Convergent modeling of past soil organic carbon stocks but divergent projections

Z. Luo[1], E. Wang[1], H. Zheng[2], J. A. Baldock[3], O. J. Sun[4], Q. Shao[5]

[1] CSIRO Agriculture Flagship, GPO Box 1666, Canberra, ACT 2601, Australia.

[2] CSIRO Land and Water Flagship, GPO Box 1666, Canberra, ACT 2601, Australia.

[3] CSIRO Agriculture Flagship, PMB 2, Glen Osmond, SA 5064, Australia.

[4] Institute of Forestry and Climate Change Research, Beijing Forestry University, Beijing 100083, China.

[5] CSIRO Digital Productivity & Services Flagship, Private Bag 5, Wembley, WA 6913, Australia.

Correspondence to: Zhongkui Luo (zhongkui.luo@csiro.au) and Enli Wang (enli.wang@csiro.au).

Abstract

Soil carbon (C) models are important tool to understand soil C balance and project C stocks in terrestrial ecosystems, particularly under global change. The initialization and/or parameterization of soil C models can vary among studies even when the same model and dataset are used, causing potential uncertainties in projections. Although a few studies have assessed such uncertainties, it is yet unclear what these uncertainties are correlated with and how they change across varying environmental and management conditions. Here, applying a process-based biogeochemical model to 90 individual field experiments (ranging from 5 to 82 years of experimental duration) across the Australian cereal-growing regions, we demonstrated that well-designed optimization procedures enabled the model to accurately simulate changes in measured C stocks, but did not guarantee convergent forward projections (100 years). Major causes of the projection uncertainty were due to insufficient understanding of how microbial processes and soil C pool change to modulate C turnover. For a given site, the uncertainty significantly increased with the magnitude of future C input and years of the projection. Across sites, the uncertainty correlated positively with temperature, but negatively with rainfall. On average, a 331% uncertainty in projected C sequestration ability can be inferred in Australian agricultural soils. This uncertainty would increase further if projections were made for future warming and drying conditions. Future improvement in soil C modeling should focus on how microbial community and its C use efficiency change in response to environmental changes, and better conceptualization of heterogeneous soil C pools and the C transformation among those pools.

1 Introduction

Soil is the largest carbon (C) reservoir in the terrestrial biosphere and CO_2 emission from soil organic matter (SOM) decomposition accounts for ~35% of the global CO_2 emissions (Schlesinger and Andrews, 2000). Due to the large amount of soil organic carbon (SOC), carbon sequestration in soils represents a great potential for mitigating greenhouse gas emissions and climate change as well as maintaining soil fertility (Lal, 2004). Accurate projections of future change in SOC are therefore needed for C and greenhouse gas (GHG) inventories to guide the development of future policies and land management practices (Janssens et al., 2003). Due to the complex and dynamic interactions between SOC, climate, soil and land management practices, process-based SOM models have become an important tool to investigate SOC change and project SOC trends under different land uses (Jenkinson et al., 1991; Friedlingstein et al., 2006; Smith et al., 2007; Piao et al., 2009). Some studies have suggested that the uncertainties in such projections should be systematically addressed in order to judge the credibility of the underlying projections and develop appropriate polices for carbon sequestration and climate change mitigation (Friedlingstein et al., 2006; Tang et al., 2008; Todd-Brown et al., 2013; Nishina et al. 2014). Better understanding of these uncertainties and their drivers will help identify knowledge gaps and improve process-based models (Luo et al., 2014).

Uncertainty in simulation results derived from dynamic models can arise from inaccuracies in input data, deficiencies in model structure and inappropriate optimization of model parameters. For SOM models, initialization of the SOM pools can also be a major cause of divergent model projections. Most SOM models divide SOM into several conceptual pools (e.g. fast, slow and recalcitrant pools) and simulate the decomposition of each pool as a first-order decay process (Smith et al., 1997; Davidson and Janssens, 2006; Schmidt et al., 2011). In many cases, measurements are only available for total SOC, and there is no agreed-on procedure for initialization of these model pools using total SOC (Basso et al., 2011). As a result, model optimization was often conducted based on limited SOC measurements (usually at temporal scales less than decades) together with empirical initialization. The optimized model was then used to project SOC change at wider spatiotemporal scales (Friedlingstein et al., 2006; Thornton et al., 2007). Such projection is subject to unknown uncertainty (Friedlingstein et al., 2006; Tang et al., 2008; Luo et al., 2013), because it does not properly address the inaccuracies in both model initialization and model parameters, with the latter potentially caused by imperfect knowledge and model structure (Schmidt et al. 2011).

To illustrate the uncertainty propagation in SOC projections caused by initialization and parameterization and to understand what correlates to the change in the patterns of projection uncertainty, we used the Agriculture Production System sIMulator APSIM (Keating et al., 2003; Wang et al., 2002; Holzworth et al., 2014) together with data from 90 agricultural experiments at 26 sites across the Australian cereal–growing regions. The data include measurements of total SOC stock (0–30 cm), C input (i.e., amount of residue retention), crop yield, and records of management practices. The APSIM model uses a very similar SOM pool structure and first decay approach to simulate SOM dynamics to other common Earth system models (Smith et al., 1997; Friedlingstein et al., 2006; Thornton et al., 2007). We firstly conducted sensitivity analysis to identify the model parameters whose change impacted most on simulated SOC dynamics. We then used Bayesian optimization approach to derive the posterior joint distribution of the identified parameters that enabled best match between measured and observed SOC. These ensembles of parameters were used to run APSIM for each of the 90 experiments, and simulations were continued for further 100 years after the end of the experiment to produce SOC projections for uncertainty analysis. We quantified the uncertainty in SOC projections induced by both initialization of SOC pools and parameterization of algorithms for simulation of process dynamics. While the uncertainty obviously increases with years of projections, we further hypothesized that it is also influenced by site-specific climate, soil and management conditions, in addition to the impact of model initialization and parameterization. We further investigated how the projection uncertainty can be quantified by using these drivers, so that future SOC projections can become more useful with attached and well quantified uncertainties.

2 Materials and Methods

2.1 Study sites and datasets

Data from a total of 90 experimental plots located within 26 different sites (Fig. 1 in the supplement) and compiled and described by Skjemstad and Spouncer (2003) were used in this study. The experimental duration of these trials ranged from 5 to 82 years, and cover diverse climate, soil and agricultural management conditions and are representative of Australian cereal-growing regions (Table 1 in the supplement). The dataset included detailed records on crop sequence, crop yield, crop residue production (estimated according to harvest index) and agricultural management practices such as residue management (removal or retention) and fertilizer application over each year. SOC stock was determined for representative 0-30 cm soil samples at least at the beginning and end of the each experiment, with some experiments having as many as six temporal measurements. Other soil properties at the start of the experiment were also measured including total nitrogen content, bulk density, clay content and pH, and were used to initialize the APSIM model.

2.2 The APSIM model

APSIM was developed to simulate biophysical process in agricultural systems, and has been comprehensively verified and used to study productivity, nutrient cycling and environmental impacts of farming systems as influenced by climate variability and management practice (Keating et al., 2003; Wang et al., 2002; Holzworth et al., 2014). APSIM simulates crop growth and soil processes on a daily time–step in response to climate (i.e., temperature, rainfall, and radiation) and soil conditions (water availability, and nutrient status etc.). The model allows flexible specification of management options like crop and rotation type, tillage, residue management, fertilization and irrigation. The ability of APSIM to simulate SOC dynamics under different cropping and management practices has been verified (Probert et al., 1998; Luo et al., 2011).

APSIM simulates the dynamics of both soil C and N stocks in each soil layer. Similar to other SOM models like RothC and Century, SOM in APSIM is divided into six conceptual pools (i.e., microbial biomass, humic organic matter and inert organic matter, together with three fresh organic matter pools, Fig. 2 in the supplement). Inert organic matter is considered to be non-susceptible to decomposition, i.e., indecomposable, due to physicochemical and/or biological protections. The amount of inert organic C is initialized at the start of the simulation and dos not change during the simulation. The decomposition of other pools is treated as a first–order decay process modified by soil temperature, moisture and nitrogen availability (for fresh organic matter pool only), leading to the release of CO_2 to the atmosphere and transfer of the remaining decomposed C to other pools. Microbial carbon use efficiency (*CUE*), i.e., the efficiency of microbial community to assimilate the decomposed SOC, determines the fraction of decomposed C transferred to other pools. The model assumes a constant *CUE* for all C pools. The flow of N depends on the C:N ratio of the receiving pool. The C:N ratio of each pool is assumed to be constant through time. The decomposition of surface residues is modified by the degree of contact of the residue with soil (Thorburn et al., 2001).

The model requires values for initial SOC content, total soil N content, bulk density, and soil hydraulic parameters for each soil layer simulated. In the Skjemstad and Spouncer (2003) dataset, measured values for SOC content, bulk density and total soil nitrogen content were provided for the 0-30 cm layer. For the deeper soil layers and hydraulic parameters in the whole soil profile, values from a measured soil profile nearest to the site were selected from the Agricultural Production Systems Research Unit (APSRU) reference sites soil database (http://www.asris.csiro.au/mapping/hyperdocs/APSRU/). Daily weather data (from 1889 to present) for each site including radiation, maximum and minimum temperatures, and rainfall was obtained from the SILO Patched Point Dataset (https://www.longpaddock.qld.gov.au/silo/).

The APSIM model was first set up for each experiment. Agricultural management including crops, residue management and fertilizer application was set according to available historical records. Crops were sown depending on rainfall (>20 mm in successive five days) and soil water content (90% of saturation water content in the top 20 cm soil). Crop cultivars were assigned according to sowing date, i.e., the earlier the sowing date, the later the maturity type of the crop cultivar. For simplification, three cultivars for each crop representing early, middle and later maturity cultivars were selected from the default cultivars in the files released with the APSIM model. For pasture, however, there was no record on the species and cultivar. Here, perennial lucerne (Medicago sativa, a commonly used species in Australian pasture) was used to represent pasture and only one cultivar–trifecta–was used in the simulation. Lucerne was sown and removed after harvesting and before sowing of annual crops in the corresponding rotations, respectively. Harvest to the height of 10 cm was assumed each time lucerne reached the flowering stage to mimic possible grazing and/or haying.

In the experiments included in this study, C from assimilation of crop growth was the only source of C input to the soil. In the APSIM model, crop growth is simulated using light interception and radiation use efficiency, modified by water and nitrogen supply. In order to achieve credible simulation of crop growth, plant available water capacity (PAWC) of the soil was adjusted. This adjusted PAWC at each site was used throughout the simulations. Despite the reliability of the APSIM model to simulate crop growth (both belowground and aboveground), we did not use the simulated aboveground C input during the simulation. Alternatively, the recorded aboveground C input (as crop residue) was manually incorporated into the model at the time of crop harvesting, whilst the simulated crop residue was removed. This manipulation eliminated the effect of imperfect match of modeled with observed crop residue on SOC dynamics.

2.3 Sensitivity analysis of SOC dynamics

A total of eight parameters (Table 2 in the supplement) that directly link to the SOC dynamics in the model were selected for sensitivity analysis in order to identify the most important ones regulating SOC dynamics. One model input for model initialization, i.e., the fraction of inert organic carbon in the total SOC at the start of the simulation (*finert*), was also included in the sensitivity analysis, due to lack of observed data of *finert* and its great effect on simulated soil C changes. To inspect the response of simulated SOC to variations of those parameters (*finert* was also called as a parameter for convenience hereafter), a univariate local sensitivity analysis was conducted by looking at the impact of one parameter at a time and fixing all other parameters. As the purpose was to identify the most influential parameter(s), a continuous wheat system with 100% residue retention (the dominant crop in the studied experiments, see Table 1 in the supplement) and a nitrogen application of 200 kg N ha^{-1} yr^{-1} were used and simulated for 100 years. The default model parameters were first used (Table 2 in the supplement), and then each parameter was sequentially increased by 10% of its default value. For each parameter, the sensitivity function (S_i) was calculated to represent the sensitivity of model output y (i.e., total 0-30 cm SOC stock) to changes in a single parameter θ_i (Soetaert and Herman, 2008):

$$S_i = \theta_i \frac{y|_{\theta_i^*} - y|_{\theta_i}}{\theta_i^{*} - \theta_i}, \quad (1)$$

where θ_i was the default parameter value, and $y|_{\theta_i}$ the model output using θ_i, θ_i^* the altered parameter value (increased by 10%) and $y|_{\theta_i^*}$ the model output using θ_i^*. Finally, the importance index of the i^{th} parameter (I_i), i.e., the overall sensitivity of the output with respect to this parameter, was calculated by summarizing the sensitivities for the 100 year outputs (n=100):

$$I_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n} {S_{ij}}^2}. \quad (2)$$

where S_{ij} was the sensitivity function for parameter i at the j^{th} year of the n (n = 100) years of each simulation. The greater the magnitude of I is, the more sensitive the model output was to the parameter (Soetaert and Herman, 2008). The importance indices were compared among the nine parameters, and the most important parameters were identified and optimized to obtain the best agreement between simulated and observed SOC dynamics for each of the 90 experiments. As the relative importance of those parameters was independent of soil and climate conditions, the typical soil and climate at Wagga Wagga (a major cropping area in Australia, and one of the 26 sites used in the main text), New South Wales of Australia were selected to conduct above analyses.

2.4 Model optimization

The differential evolution (DE) algorithm (belongs to the class of genetic algorithms) was used to optimize the most influential parameters identified. The optimization was performed in R 3.0.3 using the DEoptim function in the “DEoptim” package (Mullen et al., 2011). DE is a global optimization algorithm for continuous numerical minimization problems, which use biology-inspired operations of crossover, mutation, and selection on population in order to minimize an objective function over the course of successive generations.

To use DE, each parameter was first assumed to exhibit a uniform distribution bounded within a range (i.e., the prior distribution, see Table 2 in the supplement) that was biologically and physically possible based on previous knowledge about the process, thereby eliminating solutions in conflict with prior knowledge. The optimization performed a quasi-random walk through the multi-dimensional parameter space to find the parameter set that caused the model to generate the best match between predicted and observed SOC. The “best match” was defined as the model output that minimized the criteria selected for model evaluation (Table 3 in the supplement). Seven criteria that are commonly used in the literature were selected to assess the possible effects of criterion selection on modeling results. Using each criterion, the optimization was conducted 100 times (i.e., 100 ensembles of initial parameter values through quasi–random walk), which generated 100 ensembles of parameters (i.e., the joint posterior distribution of the most influential parameters), giving simulation results with approximately equally good matches to the observed data. Consequently, 700 ensembles of parameters (from using seven criteria) for each experiment were produced. The optimizing procedure and related simulations were operated on Bragg and Dell CPUs of CSIRO Clusters.

However, the required computing time (~2 days for one experiment and one selection criterion using 100 computer cores) has posed a significant challenge even using the high performance computing clusters (Bragg and Dell CPUs) for this multi-parameter optimization of the process-oriented APSIM model. To complete all optimizations using seven criteria for the 90 experiments, a run time of four months was expected assuming that 1000 cores could be continuously available on the clusters. For this reason, the global optimization DE was only applied for two sites, i.e., Brigalow and Tarlee, providing two cases of DE optimization as compared to an alternative and faster Bayesian sampling approach as described below.

For all the experiments, a Bayesian sampling approach was substituted for the DE optimization in order to complete the work within a reasonable time but without much sacrificing of model performance. The APSIM model was run for each experiment for 100,000 times using 100,000 ensembles of parameters that were randomly sampled from their prior distributions. The best 100 ensembles of parameters were selected as their posterior distributions through using each criterion listed in Table 3 in the supplement. At Brigalow and Tarlee, the distributions of parameters "optimized" through this Bayesian sampling approach were compared with those optimized through DE optimization. The identified parameter ensembles by Bayesian sampling approach were referred to as "optimized parameters" in the following text and used to assess the uncertainty in projected SOC.

2.5 Uncertainty in projected SOC

After obtaining the 700 ensembles of optimized parameters (i.e., after "optimization period"), the APSIM model was run continuously from the start to the end of each experiment and then for an additional 100 years after the end of each experiment using each parameter set (i.e., 700 simulations for each experiment). For the last 100-year simulations (i.e., projection period), a continuous wheat system was assumed together with 100% residue retention, which is the same as that used in sensitivity analysis. Carbon input through crop residue retention was expected to be an important factor regulating SOC dynamics in the projection period. As residue (or biomass) production is dominantly controlled by fertilizer application rates under natural rainfall condition at each site, scenarios with nitrogen application rates ranging from 0 to 300 kg N ha^{-1} yr^{-1} with increment of 20 kg N ha^{-1} yr^{-1} were modeled. These scenarios made it possible to mimic different management practices that influence C input to the soil and to assess its impact on the uncertainty of simulated SOC induced by model initialization and parameterization.

Climate data from the start year of each experiment through to 2013 was used for the corresponding simulation period. For all years from 2014 onwards, the corresponding years of the latest historical climate data were used. For example, for the possible simulations from 2014 to 2104 (91 years), the historic climate data of 91 years from 1923 to 2013 was used. As we focused on the potential uncertainty induced by model parameterization and initialization, we did not consider the uncertainty related to climate change.

SOC content in the 0-30 cm soil layer was output at the start of projection (excluding the optimization period) and at the end of each year of projection (C_i). For the i^{th} year of projection, the mean (M_{SOCi}) of C_i of the 700 estimates was calculated, and the range (R_{SOCi}) of the 95% confidence interval was calculated as the difference between 97.5^{th} and 2.5^{th} percentile of the 700 estimates. Then, the percentage uncertainty (U_{Pi}) for that year of projection was estimated based on half of the R_{SOCi} divided by the M_{SOCi}:

$$U_{P_i} = \frac{R_{SOC_i}}{2 \times M_{SOC_i}} \times 100\%, i = 1,2,3, \ldots ,100. \quad (3)$$

2.6 Attributes controlling the variability of the uncertainty

After estimating U_P, we further addressed the following question: how and why does the uncertainty (i.e., U_P) in projected SOC change across space and time? We hypothesized that U_P is associated with the management in terms of residue C inputs. At the same time, we assumed that the detailed relationship between U_P and C inputs is different not only across experiments but also across time periods of the projection. As the hierarchy of the relationships (i.e., individual-level C inputs group in experiments and time periods of the projection), a hierarchical regression model, also called multilevel model (Gelman and Hill, 2006), was fitted to estimate U_{Pi} (y_i) on C input (x_i), applied to the J = 90 experiments and K = 100 time periods of projection. The multilevel model was written as a data (the predicted U_{Pi} belonging to experiment j with k years of projection) level model, allowing the model coefficients (α and β) to vary by experiment (*j = 1, …, J*) and time period of projection (*k = 1, …, K*) (Gelman and Hill, 2006):

$$y_i \sim \mathrm{N}\left(\alpha_{j[i],k[i]} + \beta_{j[i],k[i]} x_i, \sigma_y^2\right), \text{ for } i = 1, \ldots, n, \tag{4}$$

and a decomposition of its intercepts and slopes into terms for experiment, the time period of projection and their interaction,

$$\begin{pmatrix} \alpha_{j,k} \\ \beta_{j,k} \end{pmatrix} \sim \begin{pmatrix} \alpha_j^{\text{expt}} + \alpha_k^{\text{year}} + \alpha_{j,k}^{\text{expt}\times\text{year}} \\ \beta_j^{\text{expt}} + \beta_k^{\text{year}} + \beta_{j,k}^{\text{expt}\times\text{year}} \end{pmatrix} + \begin{pmatrix} \gamma_{0j}^{\text{expt}} \\ \gamma_{1j}^{\text{expt}} \end{pmatrix} + \begin{pmatrix} \gamma_{0k}^{\text{year}} \\ \gamma_{1k}^{\text{year}} \end{pmatrix} + \begin{pmatrix} \gamma_{0jk}^{\text{expt}\times\text{year}} \\ \gamma_{1jk}^{\text{expt}\times\text{year}} \end{pmatrix}, \tag{5}$$

and models for variation,

$$\begin{pmatrix} \gamma_{0j}^{\text{expt}} \\ \gamma_{1j}^{\text{expt}} \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma^{\text{expt}}\right), \text{for } j = 1, \ldots, J \tag{6}$$

$$\begin{pmatrix} \gamma_{0k}^{\text{year}} \\ \gamma_{1k}^{\text{year}} \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma^{\text{year}}\right), \text{for } k = 1, \ldots, K \tag{7}$$

$$\begin{pmatrix} \gamma_{0jk}^{\text{expt}\times\text{year}} \\ \gamma_{1jk}^{\text{expt}\times\text{year}} \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma^{\text{expt}\times\text{year}}\right), \text{for } j = 1, \ldots, J;\ k = 1, \ldots, K. \tag{8}$$

where Σ was the 2×2 covariance matrix representing the variation of the intercepts and slopes in the population of groups (experiments and time periods of projection). In essence, there is a separate regression model for each experiment and time period combination with the coefficients estimated by the weighted average of pooled (do not consider groups) and un-pooled (consider each group separately) estimates, i.e., partial pooling. This hierarchical structure of the model allows the assessment of the variation of individual-level coefficients across groups and accounting for group-level variation in the uncertainty for individual-level coefficients.

To assess the variation of individual-level coefficients (α_j^{expt} and β_j^{expt}) across different experiments, a classic linear regression was conducted to identify the effects of different sources of variation. At the experiment level, we assumed that two groups of attributes influence α_j^{expt} and β_j^{expt}: 1) uncertainty in model parameters, i.e., the three optimized parameters based on experiment-specific dataset, and 2) climate including mean annual rainfall and temperature, which are predominant factors controlling SOC dynamics during model optimization as well as during projection. The generalized variance (GV) was calculated as an indicator of the overall variation in model parameters, which is defined as the determinant of the variance-covariance matrix of the three parameters and is a scalar measure of overall multidimensional scatter. The two groups of attributes including all interactions were selected through a stepwise regression model selection by Akaike Information Criterion. Before fitting the model, GV was logarithmically transformed to satisfy additivity and linearity assumptions and then centered by subtracting the mean of the data, and rainfall and temperature were also centered. For coefficients over the time-spans of projection (α_k^{year} and β_k^{year}), their relationship with the time-span of projection were presented. All the statistical analyses including the multilevel model fitting were conducted using the R software version

3.0.3 (R Core Team, 2013).

3 Results and discussion

3.1 Sensitivity analysis and model performance

Three parameters were identified as most influential on simulated SOC (Fig. 3 in the supplement). Microbial carbon use efficiency (*CUE*) had the biggest impact. This highlights the key role of microbial process to control SOM decomposition, and the need for better capturing the dynamics and impact of microbial process in SOM models (Allison et al., 2010; Singh et al., 2010; Sinsabaugh et al., 2013; Xu et al., 2014). As *CUE* was treated as a constant in most SOM models, a framework is needed to incorporate microbial data (e.g., community, activity, and their responses and feedbacks to biotic and abiotic factors) into SOM models to provide robust estimations and predictions. Potential decomposition rate constant of humic organic matter (k_{hum}, day^{-1}) ranked the second, followed by the fraction of the humic carbon that is recalcitrant to decomposition (*finert*). This result further indicates the importance to better quantify the decomposability of the heterogeneous SOM (Schmidt et al., 2011; Sierra et al., 2011). It should be noted that the actual decomposition rate is simulated through modifying k_{hum} by a series of biotic and abiotic variables at different spatiotemporal scales, and different models simulate the responses differently (Todd-Brown et al, 2013; Exbrayat et al., 2014). Although we did not quantify the relative importance of these modifiers (e.g., soil moisture ad temperature), the results indicated that k_{hum} has to be constrained, implying the importance of determining how it responses to environmental factors. The wide distributions of *CUE*, k_{hum} and *finert* parameters (derived by constraining the model against the measurement data, Fig. 1b) imply deficiencies in our understanding of the microbial community and its activity and how they change with environmental conditions to modulate the SOM decomposition processes.

Our optimization procedure enabled accurate simulation of SOC change during the optimization period (Fig 1a) using distinct ensembles of model parameters for each experiment (Fig. 1b). Pooling together all data of the 90 experiments, the modeled average SOC of the 700 simulations could explain 96% ($P<0.001$) of the variance in observed SOC (Fig. 1a). For each experiment, model performance was nearly identical (Fig. 1a) when the simulations (using different parameter sets) were inter-compared. At the Tarlee site (Fig. 2a), for example, the RMSE between modeled and observed SOC ranged from 0.44 to 0.52 t ha^{-1}, compared with the range of 3.11 to 3.12 t ha^{-1} at Brigalow site (Fig. 2b). This high level of consistency highlights significant equifinality, i.e., different parameter ensembles leading to similar simulation results (Fig 1b, 2c and 2d), in process-based SOM models, which must be addressed in modeling studies aimed at enhanced process understanding and hypothesis testing (Tang et al., 2008; Luo et al., 2011).

3.2 Uncertainty in SOC projections

The accurate simulations of past SOC, however, do not guarantee convergent projections beyond the model optimization period. In contrast, running the model with the same parameter ensembles generated very divergent future projections (Fig. 2a and b), indicating significant uncertainty propagation with time of projection (Luo et al., 2011; Tang et al., 2008). Furthermore, the uncertainty is also related to management in terms of C input level and site conditions. At Brigalow (Fig. 2b), for example, the 95% confidence interval of projected SOC under optimal N input (i.e., no N stress for crops) ranged from 37 to 56 t ha^{-1} 10 years after the model optimization period, which increased to 26–68 t ha^{-1} for the projected SOC after 50 years. Under low N input scenario (0 kg N ha^{-1}), the uncertainty was smaller. At Tarlee (Fig. 2a), the uncertainty propagation followed a similar pattern to that at Brigalow, but the uncertainty under low N input scenario was much smaller. At Brigalow, in addition, we found that the choice of criterion (objective functions) influenced the distributions of the derived parameters (Fig. 2d) because a specific criterion only focuses on a specific aspect (e.g., mean or variance) of the data and the model results, of which the consequence for SOC simulations (e.g., the bifurcation pattern of projected SOC showed in Fig. 2b) ought to be carefully considered in future studies.

It is important to notice that the posterior distributions of model parameters were apparently different across experiments (Fig. 1b, c and d, and Fig. 4 in the supplement), confirming that model parameters are sensitive to the data constraining the model (Keenan et al., 2012; Hararuk et al., 2014; Luo et al., 2014). Our results indicate that *CUE* was likely higher for site under longer cultivation history (the Tarlee site) than for new-cleared site (the Brigalow site, Fig 2c vs 2d), implying the potential importance of land use history for constraining model parameters such as microbial carbon use efficiency because land use history has direct effect on the quantity and quality of carbon input as well as on soil properties. However, such impact needs further confirmation with more data. The distributions of the optimized model parameters were also influenced by the choice of criteria to evaluate model performance (Fig. 2d, Fig. 5 in the supplement). The differences in parameter distributions subsequently impact on the SOC projections as showed in Fig. 2b, albeit the near identical model performance in simulating historical SOC. In addition, *finert* and k_{hum} was positively related (Fig. 2c and d), implying the importance of the interactions and/or feedback between different C pools and their impacts on soil C projection. These highlight the needs for: 1) improving the science for capturing process interactions in the model such as the role of microbial processes and conceptualization of heterogeneous C pools and their transformation (Manzoni et al., 2012; Luo et al., 2014), 2) conducting model optimization conditioned on all observed data from experiments together with Bayesian inference technique, and 3) quantifying uncertainty in SOC projections with ensemble model simulations (Post et al., 2008; Weng et al., 2011; Xia et al., 2013; Hararuk et al., 2014; Luo et al., 2014).

If a continuous wheat system was practiced for 100 years after the end of each experiment at the 26 sites, optimal N management was predicted to result in an average increase in SOC (Fig. 3a), while a SOC decline under zero N input (Fig. 3b). The amount of potential SOC change depends on not only the management level (N input) and the climate and soil conditions that determine the potential productivity of crops, but also the initial SOC level at the start of the projections. Across the 90 experiments, the percentage uncertainty in the SOC projections ranged from 2% to 140% with an average of 53% under optimal N management (Fig. 3c), and from 0.8% to 108% with an average of 40% under zero N input (Fig. 3d). Applying this result to Australia's cereal-growing regions, the simulated potential SOC stock of ~7.5 Pg (Luo et al., 2013) could be subject to 53% uncertainty under no N deficient and 100% residue retention.

3.3. Attributes controlling the variability of the uncertainty

The uncertainty propagation with time of prediction and across experiments could be explained using a linear model by linking the percentage uncertainty (U_P) to the C input from crop residue (C_R), i.e., $U_P = \alpha + \beta C_R$. However, both α and β changed significantly across experiments (Fig. 4a) and years of projections (Fig. 4b), and were also impacted by their interactions. Across the time periods of projection, the uncertainty increased with the number of years for projection, reflected by the linear increase in α (model intercepts) and asymptotic increase in β (model slope, Fig. 4b). The asymptotic increase in β (model slope) also implies that the relative contribution of C input to prediction uncertainty reduces with time. Across experiments, there was a marked variation in the effect of C input on U_P, indicating impact of site-specific conditions (e.g. climate and soil as described later). Across sites and years of projections, the majority of positive β implies increased uncertainty in SOC projections with increasing C input, which has not been properly addressed in previous modeling studies (Joos et al., 2001; Jones et al., 2005; Smith et al., 2005; Ogle et al., 2010). The fate of C input has direct effect on the amount of soil C. The general positive effect of C input on uncertainty would attribute to that the amount of C input ending up in the soil would be more variable and thus higher uncertainty in soil C under higher C input. These results highlight the importance of understanding the consequences of future C input changes on soil C dynamics.

The variance in model parameters (GV) across experiments had a major effect on the intercepts (positive at $P < 0.001$) and slopes (positive at $P < 0.001$) of the regression model linking U_P to C input (Table 1). As GV was logarithmically transformed when fitting the model, the increase in U_P with GV was exponential across experiments. This result highlights the crucial role to improve the representation of the sensitive microbial processes (Zhou et al., 2012; Xu et al., 2014) and the heterogeneous SOM (Sierra et al., 2011) in biogeochemical SOM models, and to constrain the space of relevant model parameters. For example, we assumed a relatively wide range of *CUE* (0.2–0.8) as the prior information for the Bayesian optimization. Sinsabaugh et al. (2013) suggested that *CUE* prediction should consider resource composition, stoichiometry constraints and biomass composition, as well as environmental drivers. A more informative prior of *CUE* could help reduce the uncertainty in soil C projections.

Rainfall and temperature, together with their interaction, had significant impact on SOC projection uncertainty through their impact on the fitted model intercepts across experiments (Table 1). α_j^{expt} increased with temperature, but tended to decrease with rainfall, implying increased uncertainty in SOC projection under future warming and drying conditions . Based on the results, the uncertainty in projected SOC will be increased by 4.95%, if average temperature is increased by 1 °C under global warming. For the slopes β_j^{expt}, rainfall and its interaction with GV had significant negative effect. These effects may reflect the impact of rainfall on both primary productivity (thus C input) and soil moisture conditions (thus microbial activity and decomposition rate of SOC), emphasizing the importance of understanding the interactions between soil processes and their responses to external drivers and management such as temperature and rainfall (Davidson and Janssens, 2006; Carvalhais et al., 2014).

4 Conclusions

Our results demonstrate that great uncertainty exists in soil C projections from process-based SOM models, due to deficiency in model initialization and parameterization to capture the process interactions, such as microbial C use efficiency and its drivers, as well as lack of detailed information to initialize the model, e.g., the heterogeneous SOM with different decomposability. The prediction uncertainty propagates with extended years of projections and C input into soil. It is also influenced by site-specific climate (temperature and rainfall) and soil conditions together with management inputs, which determine both the C input (through primary productivity) and the SOM decomposition processes. The results also suggest that C projection into warming and drying future climate will be subject to even increased uncertainty. For agricultural land uses, uncertainty caused by management practices has to be carefully considered due to its impact on microbial activity and subsequent projected SOC. For any future predictions of SOC change, ensemble simulations conditioned on total observed datasets together with a Bayesian inference technique should be used in order to quantify the uncertainties in modeling results. Based on our results, future improvement in SOM modeling should focus on how microbial community and its carbon use efficiency change in response to environmental changes, better quantification of heterogeneous SOM and the effects of its change on total soil C turnover.

Author contributions

Z.L collected data, run simulations, and performed data analysis; Z.L, E.W., J.B. designed the study; H.Z. and Q.S. was involved in statistical analysis; Z.L., E.W. and O.J.S. wrote the paper. All authors discussed the results and commented on the manuscript.

Acknowledgements

This study was supported by funding from the Australian Government Department of Agriculture, Fisheries and Forestry (DAFF) and the Grain Research and Development Corporation (GRDC). Thanks to Dr. Yiqi Luo of the University of Oklahoma, Dr. Petra Kuhnert and Dr. Jonathan Sanderman of CSIRO for their helpful comments on an earlier version of the manuscript.

References

Allison, S. D., M. D. Wallenstein, and M. A. Bradford (2010), Soil-carbon response to warming dependent on microbial physiology. Nature Geoscience, **3**, 336-340.

Basso, B., O. Gargiulo, K. Paustian, G. P. Robertson, C. Porter, P. Grace, and J. W. Jones (2011), Procedures for initializing soil organic carbon pools in the DSSAT-CENTURY model for agricultural systems. Soil Science Society of American Journal, **75**, 69-78.

Carvalhais, N., M. Forkel, M. Khomik et al. (2014), Global covariation of carbon turnover times with climate in terrestrial ecosystems. Nature, **514**, 213-217.

Davidson, E. A., and I. A. Janssens (2006), Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. Nature, **440**, 165-173.

Exbrayat, J. F., A. J. Pitman, and G. Abramowitz (2014), Response of microbial decomposition to spin-up explains CMIP5 soil carbon range until 2100 *Geosci. Model Dev*. **7,** 2683–2692

Friedlingstein, P., P. Cox, R. Betts et al. (2006), Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. Journal of Climate, **19**, 3337-3353.

Gelman, A., and J. Hill (2006), Data analysis using regression and multilevel/hierarchical models. (Cambridge University Press).

Hararuk, O., J. Y. Xia, and Y. Q. Luo (2014), Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method. Journal of Geophysical Research-Biogeoscience, **119**, 403-417.

Holzworth, D. P., N. Huth, P. G. deVoil, E. J. Zurcher, H. I. Herrmann, G. McLean, K. Chenu, E. J. van Oosterom, V. Snow, C. Murphy, A. D. Moore, H. Brown, J. P. M. Whish, S. Verrall, J. Fainges, L. W. Bell, A. S. Peake, P. L. Poulton, Z. Hochman, P. J. Thorburn, D. S. Gaydon, N. P. Dalgliesh, D. Rodriguez, H. Cox, S. Chapman, A. Doherty, E. Teixeira, J. Sharp, R. Cichota, and I. Vogeler (2014), APSIM-Evolotion towards a new generation of agricultural systems simulations. Environmental Modelling & Software, **62**, 327-350.

Janssens, I. A., A. Freibauer, P. Ciais et al. (2003), Europe's terrestrial biosphere absorbs 7 to 12% of European anthropogenic CO_2 emissions. Science, **300**, 1538-1542.

Jenkinson, D., D. Adams, and A. Wild (1991), Model estimates of CO_2 emissions from soil in response to global warming. Nature, **351**, 304-306.

Jones, C., C. McConnell, K. Coleman, P. Cox, P. Falloon, D. Jenkinson, and D. Powlson (2005), Global climate change and soil carbon stocks; predictions from two contrasting models for the turnover of organic carbon in soil. Global Change Biology, **11**, 154-166.

Joos, F., I. C. Prentice, S. Sitch, R. Meyer, G. Hooss, G. K. Plattner, S. Gerber, and K. Hasselmann (2001), Global warming feedbacks on terrestrial carbon uptake under the Intergovernmental Panel on Climate Change (IPCC) emission scenarios. Global Biogeochemical Cycles **15**, 891-907.

Keating, B. A., P. S. Carberry, G. L. Hammer et al. (2003), An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy, **18**, 267-288.

Keenan, T. F., E. Davidson, A. M. Moffat, W. Munger, and A. D. Richardson (2012) Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. Global Change Biology, **18**, 2555-2569 (2012).

Lal, R. (2004), Soil Carbon sequestration impacts on global blimate bhange and food security. Science, **304**, 1623-1627.

Luo, Y., K. Ogle, C. Tucker, S. Fei, C. Gao, S. LaDeau, J. Clark, and D. S. Schimel (2011), Ecological forecasting and data assimilation in a data-rich era. Ecological Applications, **21**, 1429-1442.

Luo, Y., T. F. Keenan, and M. Smith (2014), Predictability of the terrestrial carbon cycle. Global Change Biology, DOI: 10.1111/gcb.12766.

Luo, Z., E. Wang, B. A. Bryan, D. King, G. Zhao, X. Pan, U. Bende-Michl (2013), Meta-modeling soil organic carbon sequestration potential and its application at regional scale. Ecological Applications, **23**, 408-420.

Luo, Z., E. Wang, O. J. Sun, C. J. Smith, and M. E. Probert (2011), Modeling long-term soil carbon dynamics and sequestration potential in semi-arid agro-ecosystems. Agricultural and Forest Meteorology, **151**, 1529-1544.

Luo, Z., E. Wang, I. R. P. Fillery, L. M. Macdonald, N. Huth, and J. Baldock (2014), Modelling soil carbon and nitrogen dynamics using measurable and conceptual soil organic matter pools in APSIM. Agriculture, Ecosystems & Environment, **186**, 94-104.

Manzoni, S., P. Taylor, A. Richter, A. Porporato, and G. I. Ågren (2012), Environmental and stoichiometric controls on microbial carbon use efficiency in soils. New Phytol., **196,**79-91.

Mullen, K.M., D. Ardia, D. L. Gil, D. Windover, and J. Cline (2011), DEoptim: An R package for global optimization by differential evolution. Journal of Statistical Software, **40**, 6.

Nishina, K., Ito, A., Beerling, D. J., Cadule, P., Ciais, P., Clark, D. B., Falloon, P., Friend, A. D., Kahana, R., Kato, E., Keribin, R., Lucht, W., Lomas, M., Rademacher, T. T., Pavlick, R., Schaphoff, S., Vuichard, N., Warszwaski, L., and Yokohata, T. (2014) Quantifying uncertainties in soil carbon responses to changes in global mean temperature and precipitation. Earth System Dynamics, **5**, 197–209.

Ogle, S. M., F. J. Breidt, M. Easter, S. Williams, K. Killian, and K. Pautian (2010), Scale and uncertainty in modeled soil organic carbon stock changes for US croplands using a process-based model. Global Change Biology, **16**, 810-822.

Piao, S. L., J. Y. Fang, P. Ciais, P. Peylin, Y. Huang, S. Sitch, and T. Wang (2009), The carbon balance of terrestrial ecosystems in China. Nature, **458**, 1009-1013.

Post, J., F. F. Hattermann, V. Krysanova, and F. Suckow (2008), Parameter and input data uncertainty estimation for the assessment of long-term soil organic carbon dynamics. Environmental Modelling and Software, **23**, 125-138.

Probert, M., J. Dimes, B. Keating, R. Dalal, W. Strong (1998), APSIM's water and nitrogen modules and simulation of the dynamics of water and nitrogen in fallow systems. Agricultural Systems, **56**, 1-28.

Qian, S. S., T. F. Cuffney, I. Alameddine, G. McMahon, and K. H. Reckhow (2010), On the application of multilevel modeling in environmental and ecological studies. Ecology, 91, 355-361.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Schlesinger, W., and J. Andrews (2000), Soil respiration and the global carbon cycle. Biogeo chemistry, **48**, 7-20.

Schmidt, M. W. I., M. S. Torn, S. Abiven et al. (2011), Persistence of soil organic matter as an ecosystem property. Nature **478**, 49-56.

Sierra, C. A., M. E. Harmon, and S. S. Perakis (2011), Decomposition of heterogeneous organic matter and its long-term stabilization in soils. Ecological Monographs, **81**, 619-634.

Singh, B. K., R. D. Bardgett, P. Smith, and D. Reay (2010), Microorganisms and climate change: terrestrial feedbacks and mitigation options. Nature Reviews Microbiology **8**, 779-790.

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

Skjemstad, J., and L. R. Spouncer (2003), Integrated Soils Modelling for the National Carbon Accounting System:(estimating Changes in Soil Carbon Resulting from Changes in Land Use). (Australian Greenhouse Office).

Smith, J., P. Smith, M. Wattenbach et al. (2005), Projected changes in mineral soil carbon of European croplands and grasslands, 1990-2080. Global Change Biology, **11**, 2141-2152.

Smith, J., P. Smith, M. Wattenbach et al. (2007), Projected changes in the organic carbon stocks of cropland mineral soils of European Russia and the Ukraine, 1990–2070. Global Change Biology, **13**, 342-356.

Smith, P., J. U. Smith, D. S. Powlson et al. (1997), A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. Geoderma, **81**, 153-225.

Soetaert, K., and P. M. Herman PM (2008) A practical guide to ecological modelling: using R as a simulation platform. (Springer).

Tang, J., and Q. Zhuang (2008), Equifinality in parameterization of process-based biogeochemistry models: A significant uncertainty source to the estimation of regional carbon dynamics. Jounal of Geophysical Research-Biogeosciences, **113**, G04010.

Thorburn, P. J., M. E. Probert, and F. A. Robertson (2001), Modelling decomposition of sugar cane surface residues with APSIM-residue. Field Crop Research **70**, 223-232.

Thornton, P.E., J. F. Lamarque, N. A. Rosenbloom, and N. M. Mahowald (2007), Influence of carbon-nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. Global Biogeochemical Cycles, **21**, GB4018.

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

Wang. E., Robertson, M.J., Hammer, G.L., Carberry, P.S., Holzworth, D., Meinke, H., Chapman, S.C., Hargreaves, J.N.G., Huth, N.I., McLean, G., 2002. Development of a generic crop model template in the cropping system model APSIM. European Journal of Agronomy, **18**, 121-140.

Weng, E., and Y. Luo (2011), Relative information contributions of model vs. data to short- and long-term forecasts of forest carbon dynamics. Ecological Applications **21**, 1490-1505.

Xia, J., Y. Luo, Y. P. Wang, O. Hararuk (2013), Traceable components of terrestrial carbon storage capacity in biogeochemical models. Global Change Biology, **19,** 2104-2116.

Xu, X., J. P. Schimel, P. E. Thornton, X. Song, F. M. Yuan, and S. Goswami (2014), Substrate and environmental controls on microbial assimilation of soil organic carbon: a framework for Earth system models. Ecology Letters, **17**, 547-555.

Zhou, J., K. Xue, J. P. Xie et al. (2012), Microbial mediation of carbon-cycle feedbacks to climate warming. Nature Climate Change, 106-110.

Table 1. The effects of experiment-specific variance of model parameters and climate on individual-level coefficients (i.e., α_j^{expt} and β_j^{expt} in Fig. 4a).

Factor[†]	α^{expt}				β^{expt}			
	Estimate	SE	t value	P	Estimate	SE	t value	P
Model intercept	26.35	2.14	12.30	***	1.62	0.33	4.89	***
GV	3.15	0.55	5.69	***	0.17	0.088	1.97	•
R	–0.059	0.016	–3.63	***	–0.0055	0.0026	–2.15	*
T	4.95	1.35	3.66	***	–0.16	0.21	–0.77	0.44
GV × R	–	–	–	–	–0.0018	0.00061	–2.87	**
GV × T	–0.57	0.33	–1.74	•	–	–	–	–
R × T	–0.046	0.010	–4.49	***	0.0021	0.0014	1.46	0.15
Whole model R^2		**0.44**		*******		**0.21**		*******

***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$; •, $P < 0.1$.

[†]GV, generalized variance of the identified three model parameters including microbial carbon use efficiency, decomposition rate of humic organic carbon and the fraction of inert organic carbon; R, the annual average rainfall; T, the annual average temperature. GV was logarithmically transformed and centered, and R and T were also centered when fitting the model.

Figure legends

Figure 1. Model performance in simulating soil organic carbon (SOC) dynamics (a) and the corresponding optimized model parameters (b) across the studied 90 experiments. Circles and bars (a) indicate the average and 95% confidence interval of the simulations for each experiment using different parameter ensembles. Red and blue symbols in (a) highlight the data at Tarlee and Brigalow respectively, corresponding to the data in Fig. 2. Dashed line is the 1:1 line in (a). The parameter ensembles at Tarlee and Brigalow are highlighted in (b). See Fig.2 for the means of the colorful symbols in (b), showing the different ranges of optimized fraction of inert organic carbon (*finert*).

Figure 2. Projected soil organic carbon dynamics at two case sites Tarlee (a) and Brigalow (b) and the correspondingly used parameter ensembles (c and d). Black symbols show the observations. Seven criteria (RMSE, MAE, pMAE, IoA, rIoA, NSE and rNSE, see Table 3 in the supplement for details) are used to derive the posterior joint distribution of model parameters (*CUE*, k_{hum} and finert). *CUE*, microbial carbon use efficiency; k_{hum}, the potential decomposition rate of humic organic carbon; *finert*, the fraction of inert organic carbon.

Figure 3. Projected SOC (a and b) and its percentage uncertainty (c and d) under high (a and c) and low (b and d) carbon input scenarios after 100-year simulations in 90 experiments across 26 sites. Concentric circles show the different experiments at the same site. The sizes of the pies correspond to the projected average of SOC content (a and b) and the corresponding percentage uncertainty (c and d). Blue and red circles indicate that the average of the 700 simulations is increased and decreased, respectively, compared with the SOC content at the start of the projection. Blue and red sectors of the pies in (c) and (d) indicate the fraction of 700 bootstrapping simulations that shows an increase and a decrease of the projected SOC, respectively, compared with the SOC content at the start of the projection

period.

Figure 4. Coefficients (estimate ± standard deviation) for the regression model: $U_P = \alpha + \beta C_R$. The model is fitted to estimate the effects of carbon input (C_R) on the percentage uncertainty (U_P) in soil organic carbon projections, applied to 90 experiments (a) and 100 time-spans of projection (b). $\hat{\alpha}$, $\hat{\beta}$ and σ show the data-level coefficients (i.e., averaging over experiments and time-spans of projection) and errors, respectively. In (a), experiments are sorted according to α_j^{expt}. The coefficients at the experiment × time-span level are not shown. See more details in the Methods for the regression model.

Figure 1

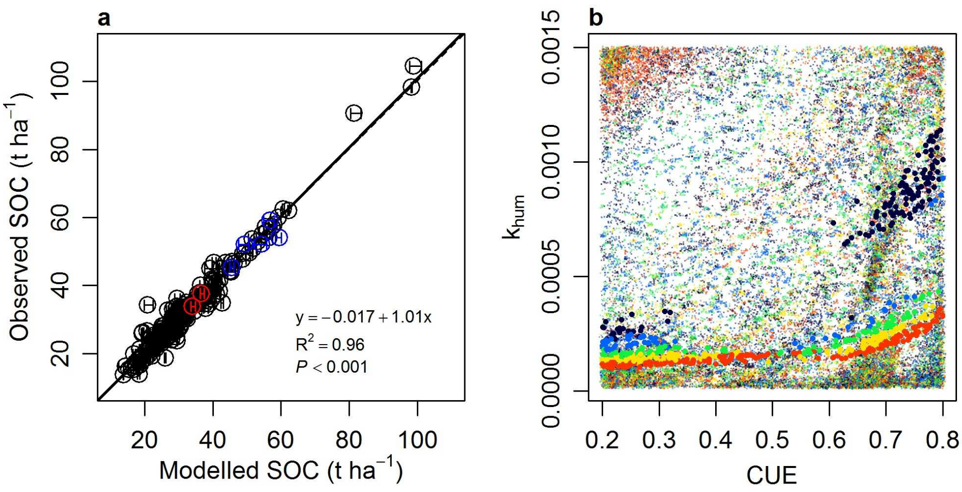

a
b
y = −0.017 + 1.01x
R2 = 0.96
P < 0.001
Observed SOC (t ha−1)
Modelled SOC (t ha−1)
20
40
60
80
100
khum
CUE
0.0000
0.0005
0.0010
0.0015
0.2
0.3
0.4
0.5
0.6
0.7
0.8

Figure 2

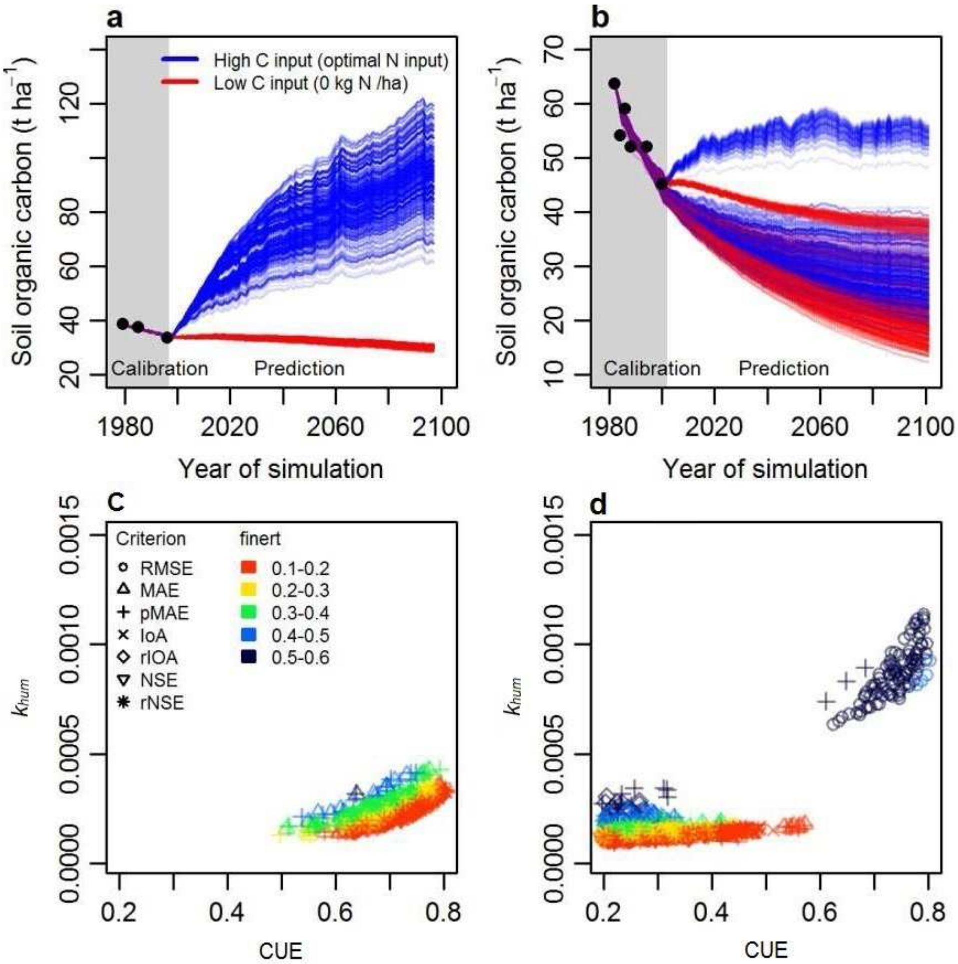
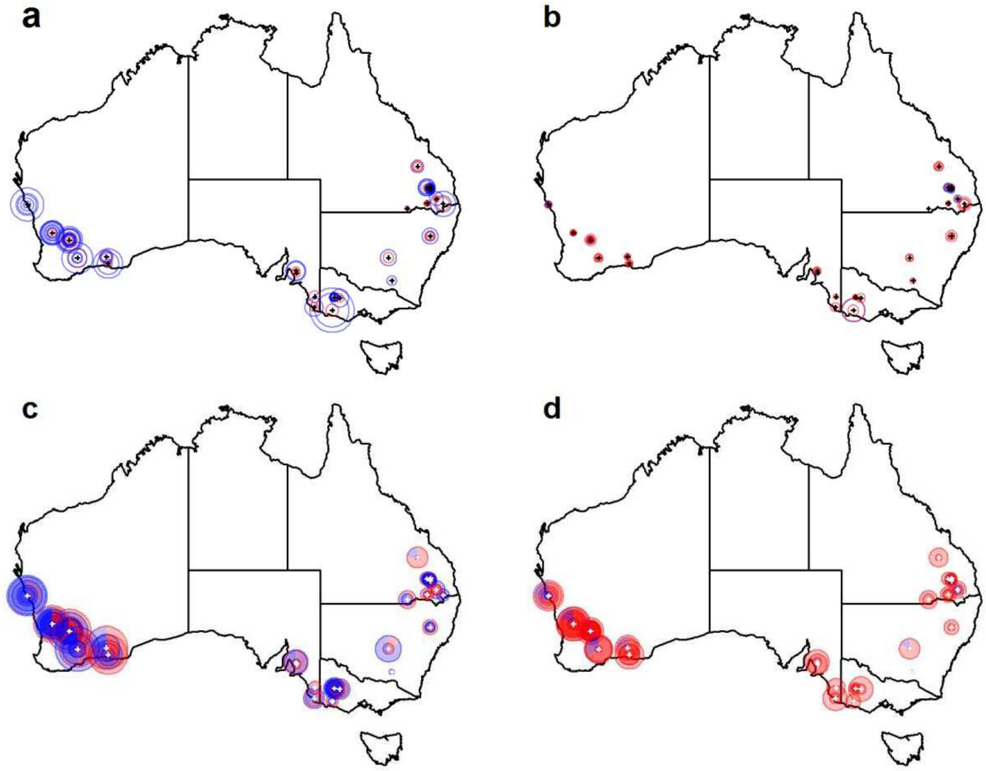
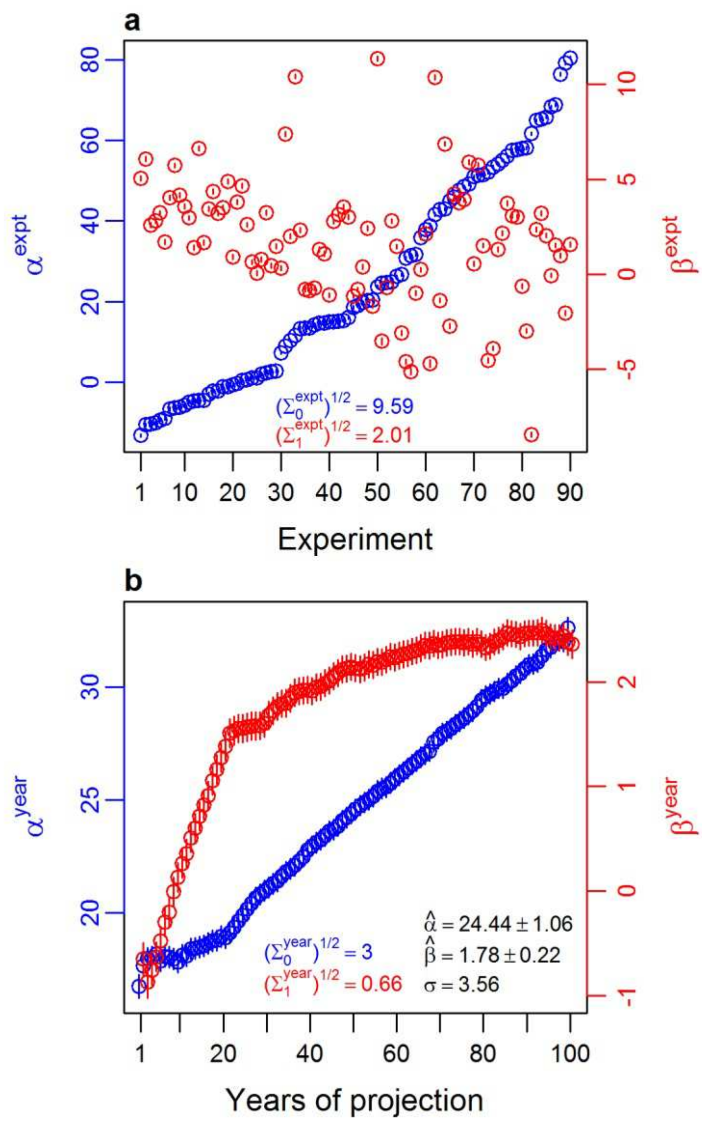

a
b
High C input (optimal N input)
Low C input (0 kg N /ha)
Soil organic carbon (t ha-1)
Calibration
Prediction
Year of simulation
c
d
Criterion
RMSE
MAE
pMAE
IoA
rIOA
NSE
rNSE
finert
0.1-0.2
0.2-0.3
0.3-0.4
0.4-0.5
0.5-0.6
khum
CUE

Figure 3



648 Figure 4



649