NERC Open Research Archive
sharing our research

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

Article (refereed) - postprint

This is the peer reviewed version of the following article:

Cagnarini, Claudia; Renella, Giancarlo; Mayer, Jochen; Hirte, Juliane; Schulin, Rainer; Costerousse, Benjamin; Della Marta, Anna; Orlandini, Simone; Menichetti, Lorenzo. 2019. **Multi-objective calibration of RothC using measured carbon stocks and auxiliary data of a long-term experiment in Switzerland**. *European Journal of Soil Science*, 70 (4). 819-832, which has been published in final form at https://doi.org/10.1111/ejss.12802

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

© 2019 British Society of Soil Science

This version available http://nora.nerc.ac.uk/522757/

NERC has developed NORA to enable users to access research outputs wholly or partially funded by NERC. Copyright and other rights for material on this site are retained by the rights owners. Users should read the terms and conditions of use of this material at http://nora.nerc.ac.uk/policies.html#access

This document is the authors' final manuscript version of the journal article, incorporating any revisions agreed during the peer review process. There may be differences between this and the publisher's version. You are advised to consult the publisher's version if you wish to cite from this article.

The definitive version is available at http://onlinelibrary.wiley.com/

Contact CEH NORA team at
noraceh@ceh.ac.uk

The NERC and CEH trademarks and logos ('the Trademarks') are registered trademarks of NERC in the UK and other countries, and may not be used without the prior written consent of the Trademark owner.

Multi-objective calibration of RothC using measured carbon stocks and auxiliary data of a long-term experiment in Switzerland

C. CAGNARINI[a], G. RENELLA[b], J. MAYER[c], J. HIRTE[c,d], R. SCHULIN[e], B. COSTEROUSSE[f], A. DELLA MARTA[b], S. ORLANDINI[b] & L. MENICHETTI[g]

[a] *Centre for Ecology & Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd, LL57 2UW, United Kingdom,* [b] *Department of Agrifood Production and Environmental Sciences, University of Florence, P.le delle Cascine 16, 50144 Florence, Italy,* [c] *Agroscope Institute for Sustainability Sciences ISS, Reckenholzstrasse 191, 8046 Zürich, Switzerland,* [d] *University of Zurich, Department of Geography, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland,* [e] *Institute of Terrestrial Ecosystems, ETH, Universitästrasse 22, 8092 Zürich, Switzerland,* [f] *Institute of Agricultural Sciences, ETH, Eschikon 33, 8315 Lindau, Switzerland,* [g] *SLU, Ecology Department, System Ecology Unit, Inst för Ekologi, Box 7044 Uppsala, Sweden.*

Correspondence: C. Cagnarini. E-mail: ccagnar@ceh.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ejss.12802

Summary

Interactions between model parameters and low spatiotemporal resolution of available data mean that conventional soil organic carbon (SOC) models are often affected by equifinality, with consequent uncertainty in SOC forecasts. Estimation of belowground C inputs is another major source of uncertainty in SOC modelling. Models are usually calibrated on SOC stocks and fluxes from long-term experiments (LTEs), whereas other point data are not used for constraining the model parameters. We used data from an agricultural long-term (> 65 years) fertilization experiment to test a multi-objective parameter estimation approach on the RothC model, combining SOC data from different fertilization treatments with microbial biomass, basal respiration and Zimmermann’s fractions data. We also compared two methods to estimate the belowground C inputs: a conventional scaling of belowground biomass from crop harvest yield and an alternative approach based on constant belowground C for cereals measured experimentally in the field. The resulting posterior parameter distributions still suffered from some equifinality; the most stable C pool kinetic constants and composition of exogenous organic matter were the most sensitive parameters. The use of fixed belowground C inputs for cereals improved the model performance, reducing the importance of treatment-specific parameters and processes. The introduction of microbial biomass and basal respiration data was effective for increasing determination of the calibration, but also suggested a change in the model structure: the microbial biomass pool, which is proportional to the C inputs in the traditional models, could be represented by different microbial physiology functions.

Keywords: SOM dynamics, belowground C inputs, microbial biomass, basal respiration, Zimmermann's fractions.

Highlights

- Multi-objective calibration with SOC and microbial or soil fractionation data can reduce model uncertainty.
- We compared different methods to estimate belowground C inputs on a long-term trial in Switzerland.
- Fixed belowground C inputs measured in the field gave the best model performance.
- Microbial data can improve model calibration, but a change in model structure is suggested.

Introduction

Although soil organic carbon (SOC) turnover plays an important role in global warming, uncertainty in current soil organic carbon (SOC) dynamic models, and by consequence in Earth System Models (ESMs), is still very large (Bradford *et al.*, 2016). It has been claimed that such uncertainty is due to incomplete knowledge and representation of relevant processes, which has led to calls for incorporating the new understanding of SOC formation and turnover into the models (Schmidt *et al.*, 2011). Recently, various attempts have been made to account for biological properties (Allison *et al.*, 2010), soil structure (Segoli *et al.*, 2013) or a mix of biological, soil physical and chemical processes and mechanisms (Abramoff *et al.*, 2018) into new dynamic SOC models. However, evaluation of these new models is still site specific, and more research is required for successful upscaling. For practical applications in ESMs (Luo *et al*., 2016), the most used SOC models are still those based on conceptual SOC pools, which approximate SOC persistence with specific rates of first-order turnover kinetics, such as the CENTURY and RothC models (Parton *et al.*, 1988; Jenkinson & Coleman, 2008;). In these models, biotic factors like microbial physiology and enzyme activity and abiotic factors such as SOC protection by sorption or occlusion in aggregates, which make SOC persistence an ecosystem property (Schmidt *et al.*, 2011), are not represented explicitly but aggregated in the pool kinetics. Site-specific calibrations are often needed to constrain the kinetic constants of the empirical-turnover SOC pools (Ludwig *et al*., 2007), and extrapolation of the fitted parameters to other sites, even under similar land use, can give poor results. It has also been shown that most of the semi-empirical SOC models are affected by equifinality (e.g. Menichetti *et al.*, 2016): this means that different parameter values can result in diverging projections of SOC, hampering any physical interpretation of the calibrated parameters.

Data typically used for calibrating models are series of SOC stocks from long-term experiments (LTEs) across the world. In addition to the SOC stocks, other data have been measured over time in the LTEs that have been neither published nor used to constrain the SOC dynamic models. Integration of this ‘grey literature’ with the available SOC stock data can potentially reduce the uncertainty of SOC model calibration (Luo *et al.*, 2016) and reveal soil processes embedded in the calibrated parameters. For example, Fujita *et al.* (2014) found that including microbial biomass data in the calibration of CENTURY improved the simulation of soil respiration. In the last two decades, several soil physical fractionation techniques have been developed that enable us to relate operationally-defined SOC pools to the model pools (Zimmermann *et al.*, 2007; Skjemstad *et al.*, 2004). In particular, the fractionation proposed by Zimmermann defines pools that were shown to correlate well with the RothC model pools. Therefore, this fractionation scheme could be used as an additional constraint in the parameterization of the RothC model.

Another major source of uncertainty in SOC models is the estimation of belowground C inputs (Keel *et al.*, 2017). Belowground C inputs depend on multiple factors, including site-specific agronomic practices and how plant genotypes respond to them. Because direct measurements of belowground C inputs is challenging, it is difficult to produce general prediction models of this perhaps underestimated source of C. Recent research has questioned the common practice of estimating roots and exudates by scaling of relations to the aboveground plant biomass, arguing that these inaccuracies might be a major obstacle in upscaling SOC forecasts from field to global scales (Hu *et al.*, 2018; Taghizadeh-Toosi *et al.*, 2016). Thus, improving prediction of the belowground inputs on the basis of more accurate data is considered an urgent research task.

To understand the potential of combining other data from LTEs with SOC stock series, we parameterized the RothC model with measurements from the ZOFE agricultural long-term trial (> 65 years) conducted by Agroscope in Switzerland. The RothC model is a typical semi-empirical model based on conceptual C pools and first-order kinetics. It is among the most used SOC turnover model, and it is applied from field to global scales. We developed a stochastic multi-objective calibration methodology, within the GLUE framework (Beven & Binley, 2014), to integrate all the available information and deal with the equifinality of the model (Clifford *et al.*, 2014). In addition to SOC time series, other data such as microbial biomass, basal soil respiration and Zimmerman's fractions (Poeplau et al., 2013) were used to reduce the uncertainty in predictions and parameters. In addition, in ZOFE the belowground C inputs were recently measured for *Zea mays* L. and *Triticum aestivum* L. (Hirte *et al.*, 2018).

The main goals of this study were to (i) assess the effect of using different methodologies for belowground estimation of C input on SOC dynamics, (ii) parameterize the RothC model by combining the SOC data from three different treatments and assess the model equifinality and performance and (iii) evaluate the potential contribution of soil microbial biomass, soil respiration and physical fractionation data, introduced one at a time together with SOC data, on the model calibration.

Materials and methods

The ZOFE experiment

Zurich Organic Fertilization Experiment (ZOFE) is a long-term plot trial of the Swiss Federal Agricultural Research Institute (Agroscope) at Zurich-Reckenholz in Switzerland. It started in 1949 to compare different fertilization practices in a typical Swiss crop rotation (Oberholzer *et al.*, 2014). The site is at 420 m a.s.l., the annual precipitation is 1054 mm and the mean annual temperature is 9.4°C. The soil is a carbonate-free, loamy (14% clay) Luvisol (IUSS, 2006), with relatively small SOC content, 1.43% on average at the beginning of the experiment in 1949. Twelve different treatments are replicated in five blocks in a systematic block design. The following 8-year crop rotation has been applied to all the treatments with minor changes over time: (i) winter wheat and intercrops, (ii) maize, (iii) *Solanum tuberosum* L., (iv) winter wheat and intercrops, (v) maize, (vi) *Hordeum vulgare* L., (vii) clover grass ley and (viii) clover grass ley. The same cultivation and plant protection techniques have been used in all the treatments. The depth of ploughing has been kept constant at 20 cm at least. During the experiment the soil has had a marked loss of SOC in all treatments (Oberholzer *et al.*, 2014) and progressive acidification from a sub-neutral soil pH (H_2O) value in 1949 to around 5.5 in 2014. For the present study we used the data from the following three treatments: control (NIL) with no fertilizer or amendment application, mineral fertilizers (NPK) with 40 kg ha^{-1} of NPK at sowing and farmyard manure (FYM) with 5 t ha^{-1} of cattle organic matter every second year.

Available data

All soil properties described below refer to the ploughing depth (20 cm).

SOC and bulk density

The SOC content has been measured regularly since the beginning of the experiment by a modified K-dichromate oxidation method using hot sulphuric acid (Oberholzer *et al.*, 2014). Soil bulk density (g cm^{-3}) was measured in 2009 only in all the treatments and the following pedotransfer function was used to calculate the C stocks:

$$\text{bulk density} = 1.49 \times SOC^{-0.1261}, \qquad R^2 = 0.78,$$

where the correction SOC (%) = SOC × 1.059 was applied to account for the dichromate oxidation recovery (Leifeld *et al.*, 2009). The soil bulk density and the SOC concentration were not independent of each other, except in 2009, therefore some uncertainty in the actual SOC stock data should be acknowledged.

Microbial biomass

Microbial biomass was measured in 1988 and 1989 in spring and autumn by plate count, in 2013 before winter wheat sowing and in 2014 after winter wheat harvest by chloroform fumigation extraction (Vance *et al.*, 1987) in five replicates per treatment. Because the 1988–1989 dataset had systematically larger values than the 2013–2014 dataset, possibly because of the difference in method, we divided the microbial biomass in NPK and FYM by the corresponding values in NIL and set these microbial biomass ratios as calibration objectives. When plotted over SOC, these microbial biomass ratios collapsed into an increasing, quasi-saturating curve (Figure S1, Supporting Information).

Basal respiration

Soil basal respiration was measured in 1972, 1981, 1988, 2013 and 2014. In 2013 and 2014 the moisture content was adjusted to 50% of water holding capacity by appropriate addition of water, and the soil was closed in a bottle and pre-incubated for 7 days at 22°C. Potential water losses were determined by weighing and readjusted. Five measurements per treatment were performed. To prevent possible offsets originating from different protocols used through the years, we divided the basal respiration in NPK and FYM by the corresponding values measured in NIL.

Zimmermann's fractionation

Soil fractionation according to Zimmermann *et al.* (2007) was carried out at ZOFE by Poeplau *et al.* (2013) in an interlaboratory program for testing the reproducibility of SOC fractionation methods. The soil was collected from a treatment with mineral fertilizer in which SOC stocks and crop yield did not differ significantly from the NPK treatment considered in this study ($p < 0.05$); therefore, we assumed that the measured fractions were applicable to the NPK treatment. The Zimmermann's fractions have been found to be comparable to the DPM+RPM, HUM+BIO and IOM pools defined in the RothC model. For each fraction we sampled from a normal distribution fitted to the measurements (same treatment but different laboratories) and set these values as a calibration objective over the period 2010–2013.

The RothC model

The SOC turnover model RothC (Jenkinson & Coleman, 2008) distinguishes five organic matter pools. Two pools receive above- and below-ground inputs of plant residues and are termed decomposable plant material (DPM with turnover kinetic constant KDPM) and resistant plant material (RPM with kinetic constant KRPM), whereas a third pool is the microbial biomass (BIO with kinetic constant KBIO). The other two pools consist of soil organic matter: a pool of decomposable humified organic matter (HUM with kinetic constant KHUM) and a pool of inert organic matter (IOM), which is resistant to decomposition. In the RothC model, the non-mineralized part of the decomposed C from all the decomposable pools is partly converted into the microbial biomass BIO, whereas that remaining goes into the HUM pool. The plant inputs are defined in terms of DPM and RPM composition because the DPM fraction is added to the DPM pool and the RPM fraction to the RPM pool. The DPM/RPM ratio value of 1.44 for crop and grassland was retained as in the original RothC model. The organic amendments are defined in terms of DPM, RPM and HUM, and their composition is usually more variable than the plant litter (Mondini *et al.*, 2017). In this study, the composition of farmyard manure was calculated from two calibrated parameters, DPM_RATIO and HUM_factor, according to the following equations:

$$\mathrm{DPM}_{\mathrm{org\,amend}} = \mathrm{DPM\_RATIO},$$

$$\mathrm{HUM}_{\mathrm{org\,amend}} = (1 - \mathrm{DPM\_RATIO}) \times \mathrm{HUM\_factor},$$

$$\mathrm{RPM}_{\mathrm{org\,amend}} = 1 - \mathrm{DPM}_{\mathrm{org\,amend}} - \mathrm{HUM}_{\mathrm{org\,amend}}$$

To modify the kinetic constants, mean temperature and rainfall were collected monthly at local stations from 1949 to 2014. Evapotranspiration was calculated every month with a locally calibrated Primault equation as a function of relative humidity, sunshine, temperature and elevation. In the RothC model evapotranspiration is reported as open pan evaporation and a default conversion factor of 0.75 is used conventionally; here, a treatment-specific coefficient, TRASP, was calibrated to account for the different aboveground biomass and soil cover in the three treatments, as suggested by Herbst *et al.* (2018).

Estimation of aboveground carbon inputs

In the ZOFE experiment the main product yield of the crop, and sometimes the by-product yield, have been recorded every year since 1949. The main products have been harvested and removed, except for the intercrops, for which we considered that only half of the harvested biomass has effectively entered the soil. Aboveground by-products, such as cereal stubble, potato haulm and cover-grass stubble, which have been incorporated in the soil after harvest, were estimated to be 5% of the main product biomass. The biomass C content was taken as 0.45 for all the plant parts (Bolinder *et al.*, 2007).

Estimation of belowground carbon inputs

For estimating the belowground C inputs two approaches were tested: (i) the conventional approach proposed by Bolinder *et al.* (2007), hereafter termed 'C_{Bolinder}', in which the root biomass is calculated from the shoot biomass with a shoot:root ratio (S:R). For small-grain cereals and maize the values proposed by Bolinder *et al.* (2007) were retained, 7.4 and 5.6, respectively. For clover grass ley and intercrops an S:R ratio of 4 was used to match the measured trend in SOC in the NIL treatment when using the default model coefficients. Rhizodeposition, comprising exudates and fine dead roots, was calculated as a fixed fraction (0.65) of the estimated root biomass for all the crop species and (ii) the approach, hereafter termed 'C_{Hirte}', based on the measurements of root biomass (Hirte *et al.*, 2018) and rhizodeposition (Hirte *et al.*, 2018) that were carried out in ZOFE. During one growing season, maize was artificially enriched with ^{13}C in four treatments of ZOFE, including the NIL and FYM considered here. The total belowground C measured (roots and rhizodeposition), with an average of 930 kg ha^{-1}, did not differ significantly between treatments ($p < 0.01$). The same experiment was conducted for winter wheat and again the total belowground C was treatment-independent ($p < 0.05$), with an average of 1100 kg ha^{-1}. Therefore, in the 'C_{Hirte}' approach we set constant root biomass and rhizodeposition for maize and winter wheat, irrespective of the treatment-specific-yield, equal to the mean measured values. Belowground C inputs for summer barley were also set to constant by scaling the average belowground C inputs measured for winter wheat with the average barley yield. For clover grass ley and intercrops we kept the '$C_{Bolinder}$' approach.

In both the '$C_{Bolinder}$' and 'C_{Hirte}' methods, belowground inputs of potato were estimated as 5% of the main product yield. For all crops, only the belowground C inputs relevant to the ploughing depth were considered (for both root and rhizodeposition): the root cumulative distributions provided by Jackson *et al.* (1996) were used, with an exponent of 0.961 for annual crops and 0.943 for cover grass ley and intercrops.

Farmyard manure amendment

The farmyard manure applied was recorded throughout the experiment and C inputs were estimated by assuming a C content of 50%.

The RothC calibration method

To account for parameter interactions, RothC was calibrated within the GLUE (global likelihood uncertainty analysis) framework (Beven & Binley, 2014), according to limits of the acceptability criteria (Liu *et al.*, 2009). The calibration procedure was written in the R environment (R Core Team 2017). The parameter space was explored by sampling 100 000 parameter sets with a Latin hypercube sampling scheme (Stein, 2012). Ten parameters were calibrated simultaneously: the kinetic constants KDPM, KRPM, KBIO, KHUM (four parameters), the inert organic matter pool IOM (one parameter), the farmyard manure composition parameters DPM_RATIO and HUM_factor (two parameters) and the treatment-specific coefficients for evapotranspiration conversion TRASP_NIL, TRASP_NPK and TRASP_FYM (three parameters). Because the kinetic constants and the inert pool IOM are arbitrary definitions and, moreover, they interact with each other, we chose to calibrate all of them at the same time. This increases the posterior uncertainty, but, on the other hand, leads to more robust conclusions, not limited to the conventional values adopted in the model. The parameter priors were selected in a range between 50 and 150% of the RothC default values for the kinetic rate constants and the IOM pool, which was calculated according to Falloon *et al*. (1998), and between 0–1 for the other parameters.

Initialization of the model

Before 1949 the field was pasture, with the soil likely to be in equilibrium. The equilibrium conditions depend on the parameter values, therefore, we initialized the model with one ‘spin-up’ for each of the 100 000 runs. The spin-up is a long-term run, usually under the assumption of equilibrium, to determine the initial conditions of the simulation: in this case, in every spin-up the pasture C inputs were calibrated according to a deterministic optimization algorithm (limited-memory BFGS, package *lbfgs*, R Core Team, 2018) by starting from a null SOC stock and running the model over 10 000 years until we matched the SOC stock measured in 1949 when the trial was initiated. In the spin-up air temperature, rainfall and evapotranspiration were set equal to the average values between 1949 and 2014. The C pools resulting from the equilibrium spin-up runs were then used as the starting values for the simulation runs between 1949 and 2014. The effect of the equilibrium spin-up has been estimated as modest in RothC (Jenkinson & Coleman, 2008).

Calibration objectives

We aimed to study the use of multiple data to reduce uncertainty in RothC, therefore, we set four combinations of different objectives to retain or discard the parameter sets: (i) the ‘Ctot’ objective for SOC data fitting, (ii) the ‘Ctot+Biom’ objective for SOC and microbial biomass data fitting, (iii) the ‘Ctot+Rsoil’ objective for SOC and basal respiration data fitting and (iv) the ‘Ctot+Zim’ objective for SOC and Zimmermann’s fractionation data fitting. Each objective was associated with a mix of NSE (Nash–Sutcliffe efficiency) and RMSE (root mean square error) acceptance criteria, as described in Table 1. The RMSE thresholds were obtained from the standard deviation of the measured data (SOC, microbial biomass, basal respiration and Zimmermann’s fractions) and multiplied by objective-specific coefficients to obtain a statistically representative number of accepted sets (250 minimum) without compromising the parameter uncertainty.

Model sensitivity and principal component analysis

A sensitivity analysis of the model parameters was performed according to Hornberger–Spear–Young generalized sensitivity analysis (described in Beven, 2008). For each objective, the acceptance criteria used for parameter calibration were applied to sort the runs into two bins of accepted and discarded parameter sets. For each parameter, the cumulative distributions of the two bins were plotted and compared by the Kolgomorov–Smirnoff distance (*d*) statistic.

We also selected parameter sets that satisfied at least one of the four objectives described above and carried out a principal component analysis (PCA) with singular value decomposition of the matrix composed by the accepted solution sets (rows) as observations and the relative RMSE for each objective and each treatment (columns) as variables.

Results

Effect of different methodologies for belowground C input estimations

The ‘C_{Hirte}’ approach, based on ^{13}C labelling measurements, resulted in different estimates of root biomass and rhizodeposition for maize and winter wheat compared to the ‘$C_{Bolinder}$’ methodology. Averaging between 1949 and 2014, the total root biomass for maize (not limited to the top 20 cm) was greater with ‘$C_{Bolinder}$’ for all the treatments (Table 2). For winter wheat the major discrepancies were noticed in NIL and NPK, where ‘$C_{Bolinder}$’ predicted smaller and larger root biomass, respectively. With ‘C_{Hirte}’ the total rhizodeposition (not limited to the top 20 cm) was always larger than the root biomass for both maize and winter wheat, with the consequence that winter wheat rhizodeposition was greatly underestimated with ‘$C_{Bolinder}$’ (Table 2). Averages of the main product yield for maize and winter wheat are also reported in Table 2: NPK produced larger yields than NIL and FYM. Consequently, ‘$C_{Bolinder}$’ predicted very small belowground C inputs in NIL, which was particularly evident for winter wheat, and much larger inputs for NPK. Therefore, the major effect of ‘C_{Hirte}’ was to reduce the differences in C input of the NPK treatment compared to the other treatments, in particular to NIL.

Because we did not differentiate the composition of plant input between the treatments, SOC dynamics were solely controlled by the absolute value of plant C inputs. Percentage difference between the ‘C_{Hirte}’ and ‘$C_{Bolinder}$’ estimates are reported in Table 3. As noted before, ‘C_{Hirte}’ estimated smaller belowground and total C inputs for maize, e.g. by –46 and –40%, respectively, in the NPK treatment. Conversely, ‘C_{Hirte}’ estimated larger belowground and total C inputs for winter wheat, e.g. by 160 and 128%, respectively, in the NIL treatment. However, in ZOFE, maize and small grains accounted for ca. 60% of the crop rotation, so that when all the crops were considered, the differences between the two approaches attenuated (Table 3). For the full rotation, the 'C_{Hirte}' estimate of total plant C inputs was larger in NIL by 18%, smaller in NPK by –7% and similar in FYM compared to '$C_{Bolinder}$'. Again, 'C_{Hirte}' produced the effect of levelling off the plant C inputs between treatments, in particular between NIL and NPK.

Model calibration with C stock data

We first ran RothC between 1949 and 2014 with the default model parameters (Figure 1). While the measured trend in SOC was matched in the NIL treatment, RothC overestimated the SOC trends in the NPK and FYM treatments at the end of the simulations. The effect of 'C_{Hirte}' on dynamic SOC predictions was negligible compared to '$C_{Bolinder}$'. However, the effect of the 'C_{Hirte}' approach became more evident when simulating the microbial biomass and basal respiration data (Figure 1). In this case 'C_{Hirte}' predicted slightly smaller values than '$C_{Bolinder}$', which were closer to the measured ones. The two approaches did not result in appreciable differences in the simulation of the Zimmermann's fractions (data not shown). Under the 'Ctot' objective, the SOC stock data of NIL, NPK and FYM treatments were targeted at the same time by allowing simultaneous variation of the model parameters. Setting the acceptability criteria in Table 1, there were almost 2500 solutions for both '$C_{Bolinder}$' and 'C_{Hirte}' (Figures S2 and S3, Supporting Information). In both cases, the SOC density distributions slightly narrowed around the measured data in NIL, NPK and FYM (Figure 2 for

'$C_{Bolinder}$').

Parameter sensitivity analysis under the 'Ctot' objective revealed the large effect of KHUM and IOM, whose Kolgomorov–Smirnov *d* values were ~1 and 0.7, respectively, with both the '$C_{Bolinder}$' (Figure 3) and the 'C_{Hirte}' approaches. The parameters of FYM composition were also sensitive to fitting the SOC stock data. As noted in other studies (Luo *et al*., 2015), the fast-kinetic pools, such as KDPM and KBIO, did not contribute appreciably to the SOC dynamic simulations, whereas the recalcitrant plant material pool KRPM was moderately important. The conversion factors for evapotranspiration were the only treatment-dependent parameters, but their effect was moderate except for TRASP_FYM. Parameter uncertainty analysis revealed strong equifinality under the 'Ctot' objective, with clear definition for only the FYM composition (Figure 4 for the '$C_{Bolinder}$' approach). For the other parameters, the probability distributions mirrored the sensitivity analysis, with slightly sharper distributions for KHUM and IOM only.

The RothC multi-objective calibration

Under the 'Ctot' objective, the relative microbial and basal respiration data were systematically over-predicted by the '$C_{Bolinder}$' approach (Figures 5 and 6). The most severe discrepancy affected the simulation of relative microbial biomass in the FYM treatment: the simulated trend was positively related, whereas the measurements remained at a constant level. Over-prediction was still present, but less pronounced, with the 'C_{Hirte}' approach: here, the NPK relative microbial biomass data (Figure 7) and the FYM relative basal respiration data lay within the boundaries of simulation uncertainty. Adding other objectives to the SOC objective 'Ctot' did not improve fitting of the relative microbial biomass and basal respiration data in terms of RMSE (data not shown), with only a slight reduction of the simulation uncertainty boundaries (Figures 2, 6, 7). We ran the PCA of the objective-treatment specific RMSEs to investigate why the microbial data (microbial biomass and basal respiration) reduced the range of prediction uncertainty only marginally. We plotted the eigenvectors corresponding to the RMSE of each objective in the plane defined by the first two principal components. Under both '$C_{Bolinder}$' and 'C_{Hirte}' eigenvectors of the SOC RMSEs clustered together along principal component 2. They were well separated from the microbial data, which also clustered together but along the principal component 1. Therefore, the objective 'Ctot' was related to different parameter sets from 'Ctot+Biom' and 'Ctot+Rsoil'. This mismatch in the parameter sets was confirmed by the fact that with '$C_{Bolinder}$', in order to obtain the minimum number of accepted parameter sets under the 'Ctot+Biom' and 'Ctot+Rsoil' objectives, we had to relax the acceptability thresholds, while this was not true with 'C_{Hirte}'. The 'Ctot+Zim' objective had a weak relation with the leading principal components and did not show any effect in reducing the simulation uncertainty (Figures 2, 6, 7). The inclusion of the Zimmermann's fractionation data modified the probability distribution of IOM (Figure 4 for the '$C_{Bolinder}$' approach). The microbial objectives (microbial biomass and basal respiration) had a major effect on the treatment-specific evapotranspiration conversion factors TRASP with the '$C_{Bolinder}$' approach only, both in terms of sensitivity (Figure 3) and density distribution (Figure 4).

Discussion

Effect of different plant C input estimations

Plant C inputs are a major controlling factor in SOC dynamic models and different estimates can have a significant effect on SOC stock, i.e. at the national and global scales (Keel *et al.*, 2017). Taghizadeh-Toosi *et al.* (2016) used the long-term Broadbalk field experiment at Rothamsted, UK, to compare different methods for estimating root C input, and concluded that the traditional allometric functions based on the crop harvest yield could lead to large overestimates. Hu *et al.* (2018) carried out a meta-analysis of published aboveground and root biomass data and reported that fixed root biomass for different management and crop classes from Denmark was a better approximation than the allometric relations. Here we used experimental data of root biomass and rhizodeposition for the same species of maize and winter wheat, but under different management practices, and combined them with a long-term crop experiment. The single-season data reported by Hirte *et al.* (2018) in ZOFE suggest that the root biomass and rhizodeposition for maize and winter wheat could be decoupled from the harvest yield, confirming the findings of Hu *et al.* (2018). The root biomass estimates obtained for maize with the S:R ratio approach were markedly different from the values measured by Hirte *et al*., 2018 (Table 2), i.e. the maize shoot:root ratio of 5.6 was perhaps too large, suggesting that species variability cannot be neglected. The rhizodeposition measured by Hirte *et al.*, (2018) was independent of treatment for maize and winter wheat. Furthermore, the rhizodeposition was ca. 1.3 times the measured root biomass for both maize and winter wheat (Table 2). This rhizodeposition value was larger than the commonly used allometric function of 0.65 times the root biomass (Bolinder *et al.*, 2007) and perhaps larger than 0.5

times the plant residues. In this study we did not consider any variation in the chemical composition of the plant inputs between the treatments, although this could have major implications for the SOC models. The use of belowground inputs independent of yield can have relevant consequences for the estimated plant C inputs when compared to linear yield allometric functions, and the direction of change was not obvious: it was mainly negative for maize, but positive for winter wheat (Table 3). Because of the rotation schemes in ZOFE, we could assess only minor changes in model parameterization and performance. Future experiments are needed to assess the independence of belowground C inputs from measured yields for the most common crop cultivars.

Another important implication of the 'C_{Hirte}' approach was in the simulation of treatments with very different amounts of fertilizer, such as NIL and NPK. Mineral fertilizer application with NPK doubled crop yield compared to NIL (Table 2). When the '$C_{Bolinder}$' method was used, the larger yield with NPK resulted in proportionally larger belowground C inputs. However, the trend in SOC stock for NPK closely resembled that in the NIL treatment, with similar SOC content in 2014. This faster SOC turnover for NPK resulting from '$C_{Bolinder}$' could be explained by a positive priming effect (Kuzyakov, 2002), which could have long-term consequences on SOC stocks, as shown by Diochon *et al.* (2016) in N-deficient crop systems. Changes in C-use efficiency (CUE) between the treatments would not explain the data. In fact, CUE has been reported to increase when N is available (Manzoni *et al.*, 2012). Consequently, the NPK treatment would have larger CUE and larger SOC accumulation, which would not help to explain the faster SOM turnover with NPK under '$C_{Bolinder}$'. This is why CUE was not included in the parameterization, as was done by Luo *et al.* (2015). The

‘C_{Hirte}’ approach, with belowground C input estimates independent of yield, reduced the differences in input between NIL and NPK and improved the model fit for these two treatments without considering any eventual priming effect. The possible role of N-priming in the SOC turnover of fertilized treatments identifies the need to couple the C and N cycles into all SOC dynamic models.

Parameter calibration

It is well established that traditional SOC dynamic models based on multiple pools defined by different kinetics are not always predictive when default parameters are used (Ludwig *et al.*, 2007), in particular over a wide range of management practices (Figure 1). Widening the calibration dataset could help to constrain the models without losing any generalization capability. Here we found that combining SOC data from plots under different management practices, but with the same soil type, did not alleviate the equifinality issues (Figures 4 and S6, Supporting Information). It was possible to characterize only the composition of FYM (DPM_RATIO ~ 1), confirming that more work could be done to develop a general parameterization of exogenous organic matter from different sources (Mondini *et al.*, 2017). The humified organic matter pool KHUM, as already reported by Luo *et al.* (2015), and the IOM pool (Figures 3 and S5) were the most important in terms of model sensitivity, with some interaction between these two parameters. The density distributions of IOM changed substantially when the different objectives were set (Figures 4 and S6). Notably, the mode of IOM density distribution decreased under the ‘Ctot+Biom’ and ‘Ctot+Rsoil’ objectives. If we assume that the IOM pool effectively represents a very slowly-decomposing C pool, such as black C (Marschner *et al.*, 2008), this result would be supported by the findings of Leifeld &

Mayer (2015): in three treatments at ZOFE, including NIL, all the native C was replaced by new C in the 20-cm depth of topsoil. Conversely, when the Zimmermann's fractions were used as targets in the 'Ctot+Zim' objective, the IOM mode moved between 3 and 4 t C ha^{-1}. This value accorded with the estimate proposed by Falloon *et al.* (1998); however, the lower uncertainty of the IOM pool did not result in any improvement in the definition of KHUM. Considering the site-specific data from Leifeld & Mayer (2015) and the general argument on the existence of an inert pool (Sanderman *et al.*, 2016), we suggest that ^{14}C dating measurements could be used to initialize the IOM pool as an alternative to the fractionation techniques, such as the Zimmermann's fractions. In doing so, the IOM pool could be omitted from the parameter calibration.

The model was largely insensitive to the fast-pool kinetics KDPM, KRPM and KBIO. The microbial pool kinetic rate KBIO did not result in a better probability distribution even when the relative microbial biomass data were targeted (see discussion below on model structure). We suggest that KDPM and KRPM should be constrained separately from the calibration of other model parameters using litter decay experiments. RothC was largely insensitive to the evapotranspiration conversion coefficients under the 'Ctot' and 'Ctot+Zim' objectives (Figures 3 and S5). However, these coefficients, which were the only treatment-specific parameters, became more significant under the '$C_{Bolinder}$' approach when the relative microbial biomass and basal respiration objectives were set (Figure 4). It was expected that treatments with larger yield and aboveground biomass also had an increased transpiration demand (Herbst *et al.*, 2018), and therefore a larger conversion coefficient: TRASP_NPK > TRASP_FYM > TRASP_NIL. The parameter distributions under '$C_{Bolinder}$' confirmed the expected behaviour between NIL and NPK (Figure 4), with greater water deficit in NPK. The smaller SOC turnover in NPK, resulting from the drier soil, served to fit the relative microbial biomass and basal respiration in NPK, which were comparatively smaller than the ratio of C inputs between NPK and NIL under '$C_{Bolinder}$' (see discussion below on model structure). The fact that the TRASP coefficients were used as degrees of freedom to decouple the microbial biomass and basal respiration from the C inputs was confirmed under the 'C_{Hirte}' approach: here TRASP_NPK and TRASP_FYM were largely undefined because the plant C inputs levelled off between the treatments (Figure S6). Under 'C_{Hirte}' only TRASP_FYM was more sensitive (Figure S5) and with a larger value (Figure S6): again the transpiration coefficient was used as a lever to adjust the relative microbial biomass to the C inputs, farmyard manure in this case, without having any physical meaning.

Model uncertainty and structure

The introduction of the Zimmerman's fractions in the calibration objectives did not improve the uncertainty of prediction of SOC stocks (Figures 2 and S4), relative basal respiration (Figures 5 and S7, Supporting Information) and relative microbial biomass (Figures 6 and 7). Setting the microbial data (microbial biomass and basal respiration) as objectives slightly reduced the uncertainty of predictions of relative basal respiration (Figures 5 and S7) and predictions of relative microbial biomass (Figure 6 and 7) under '$C_{Bolinder}$', but did not improve the predictions with 'C_{Hirte}'. Furthermore, the relative microbial biomass data were overestimated under '$C_{Bolinder}$' in both the FYM and COM treatments (Figure 6) and under 'C_{Hirte}' in the COM treatment. Fujita *et al.* (2014) ran short-term soil incubations and concluded that the introduction of microbial biomass data in the CENTURY model reduced the error of predictions of respiration rate by 26%. Our analysis suggests that, when running long-term simulations, the RothC model structure might not take advantage of the microbial biomass data. In RothC, as in other first-order SOC models that include a microbial pool, the microbial biomass was linearly proportional to the rate of SOC mineralization and to the C inputs. With '$C_{Bolinder}$' the treatment NIL received much smaller C inputs than the other treatments. Consequently, RothC predicted larger microbial biomass in NPK and FYM, producing steadily increasing trends of relative microbial biomass (Figures 6 and 7). The addition of farmyard manure markedly accelerated the growth of microbial biomass in the simulations. So, while faster kinetics were needed to fit the SOC trends in NPK and FYM under '$C_{Bolinder}$', the same kinetics produced greater microbial growth, creating incompatibility between the two targets (Figures S8 and S9, Supporting Information). As already noted, 'C_{Hirte}' reduced the difference in estimates of C input between the treatments, obtaining better simulations of relative microbial biomass. Experimental data indicated that the microbial biomass was correlated with SOC stock (Insam & Domsch, 1988) in ZOFE (Figure S1), rather than linearly increasing with the C inputs. Therefore, a change in model structure is probably required to take advantage of the abundant soil microbial biomass data in the literature; for example, the linear dependence of microbial biomass to C inputs could be replaced by other non-linear, ecophysiology-controlled functions. Such a structure change is also motivated by the fact that the introduction of microbial physiology into SOC models has the potential to improve the SOC predictions (Wieder *et al.*, 2013).

Soil basal respiration measurements carried out in the laboratory are only partly related to the rates of field respiration, which respond to local environmental changes. However, the basal respiration data can indicate the potential for microbial respiration and microbial SOC decomposition. Here, because the C inputs were assigned and the SOC trends were targeted, the relative basal respiration was a check on the C balance, i.e. of the goodness of the C input estimates in NPK and FYM compared to NIL. By comparing predictions of basal respiration under '$C_{Bolinder}$' (Figure 5) with predictions under 'C_{Hirte}' (Figure S7), we conclude that the use of fixed root biomass and rhizodeposition values provided better estimates than the fixed allometric function.

Conclusion

After calibrating the RothC model by combining the control, NPK and FYM amended soils from the ZOFE long-term trial in Switzerland, we found that the model parameters were not determined well because of the many parameter interactions, with consequent equifinality. The humified organic matter pool kinetic rate KHUM and the inert organic matter pool IOM were the most sensitive parameters and they interacted with each other. The distribution of IOM improved when the microbial biomass and basal respiration data were included in the calibration, converging towards a null value. The fast turnover pools DPM and RPM were not significant for long-term SOC dynamics and we recommend the calibration of their kinetic rates with litter decomposition experiments.

This equifinality was partly alleviated by the introduction of multiple objectives in the calibration. When the Zimmermann's fractions data were included, the IOM converged towards a larger value that was comparable to the estimate of the traditional pedotransfer function used for RothC, but it contrasted with ^{14}C measures indicating that at ZOFE native C was replaced by new C. It was possible to determine the composition of the FYM amendment (mostly fast-decomposable), confirming the importance of characterizing the quality of organic matter inputs. Addition of microbial biomass data could be advantageous in terms of simulation uncertainties and model parameterization, but the RothC model structure prevented the use of this additional information. We can expect advantages in introducing more realistic microbial physiology functions and in constraining the microbial biomass with field data. Addition of the Zimmermann's fractions did not improve the model uncertainty. The use of fixed root biomass and rhizodeposition values for maize and winter wheat

produced a better fit than use of the conventional allometric function based on shoot:root ratio values. When belowground C inputs were set to constant, microbial biomass and basal respiration data matched better without introducing any treatment-specific parameter or invoking any N-driven priming effect. This indicates the critical need for measurements of belowground C inputs for different crops, cultivars and treatments. The proposed method for RothC multi-objective parameter estimation could be extended with the support of more standardized measurement campaigns designed *ad-hoc* to include other data than SOC.

Supporting Information

Figure S1: Relative microbial biomass over SOC (%) in the ZOFE trial including the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments.

Figure S2: Number of accepted parameter sets using the C input estimate approach '$C_{Bolinder}$'.

Figure S3: Number of accepted parameter sets using the C input estimate approach 'C_{Hirte}'.

Figure S4: Simulation uncertainty of relative microbial biomass ($Biom/Biom_{NON}$) under the four objectives described in this study (total C, 'Ctot', total C and microbial biomass 'Ctot+Biom', total C and soil basal respiration 'Ctot+Rsoil', total C and results from a Zimmermann fractionation 'Ctot+Zim') using the C input estimate approach 'C_{Hirte}' for the (a) control (NIL), (b) mineral fertilizer (NPK) and (c) farmyard manure (FYM) treatments. Crosses are measured values.

Figure S5: Kolgomorov–Smirnov *d* statistics for parameter sensitivity analysis under the four objectives described in this study (total C, 'Ctot', total C and microbial biomass 'Ctot+Biom', total C and soil basal respiration 'Ctot+Rsoil', total C and results from a Zimmermann fractionation 'Ctot+Zim') using the C input estimate approach 'C_{Hirte}'.

Figure S6: Parameter probability distribution under the four objectives described in this study (total C, 'Ctot', total C and microbial biomass 'Ctot+Biom', total C and soil basal respiration 'Ctot+Rsoil', total C and results from a Zimmermann fractionation 'Ctot+Zim') using the C input estimate approach 'C_{Hirte}'.

Figure S7: Simulation uncertainty of microbial biomass relative to the NIL treatment

($Biom/Biom_{NIL}$) under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘C_{Hirte}’ for the (a) mineral fertilizer (NPK) and (b) farmyard manure (FYM) treatments. Crosses are measured values.

Figure S8: Principal component analysis (PCA) of RMSEs of the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘$C_{Bolinder}$‘ for the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments. (a) eigenvectors (arrows) and PC scores (grey dots) plotted in the plane of the first two principal components. (b) eigenvalues of the first eight principal components.

Figure S9: Principal component analysis (PCA) of RMSEs of the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘C_{Hirte}‘ for the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments. (a) eigenvectors (arrows) and PC scores (grey dots) plotted in the plane of the first two principal components. (b) eigenvalues of the first eight principal components.

Acknowledgements

We are very grateful to those who shared the data with us: Lucie Gunst from Agroscope for providing the SOC series, Hans-Rudolf Oberholzer from Agroscope for the basal respiration data in 2013–2014, Dr. Jens Leifeld from Agroscope for basal respiration and microbial biomass data before 2013, Dr. Christopher Poepleau from Thünen Institute for the Zimmermann's fractionation data. We also thank two anonymous reviewers who greatly contributed to improvement of the manuscript. This research has been partially funded by the Doctoral Program of the DISPAA of the University of Florence.

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

Abramoff, R., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E.A., Finzi, A., Moorhead, D., Schimel, J., Torn, M. & Mayes, M.A. 2018. The Millennial model: in search of measurable pools and transformations for modeling soil carbon in the new century. *Biogeochemistry*, **137**, 51–71.

Allison, S. D., Wallenstein, M.D. & Bradford, M.A. 2010. Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, **3**, 336–40.

Beven, K. & Binley, A. 2014. GLUE: 20 Years on. *Hydrological Processes*, **28**, 5897–5918.

Beven, K. J. 2008. *Environmental Modelling: An Uncertain Future*? London: Routledge.

Bolinder, M. A., Janzen, H.H., Gregorich, E.G., Angers, D.A. & VandenBygaart, A.J. 2007. An approach for estimating net primary productivity and annual carbon inputs to soil for common agricultural crops in Canada. *Agriculture, Ecosystems & Environment*, **118**, 29–42.

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

Clifford, D., Pagendam, D., Baldock, J., Cressie, N., Farquharson, R., Farrell, M., Macdonald, L. and Murray, L. 2014. Rethinking soil carbon modelling: a stochastic approach to quantify uncertainties. *Environmetrics*, **25**, 265–78.

Diochon, A., Gregorich, E. G., Kellman, L., Morrison, M. & Ma, B.L. 2016. Greater soil C inputs accelerate loss of C in cropping systems with low N input. *Plant and Soil*, **400**, 93–105.

Falloon, P., Smith, P., Coleman, K. & Marshall, S. (1998). Estimating the size of the inert organic matter pool from total soil organic carbon content for use in the Rothamsted carbon model. *Soil Biology & Biochemistry*, **30**, 1207–1211.

Fujita, Y., Witte, J.P.M. & Van Bodegom, P.M. 2014. Incorporating microbial ecology concepts into global soil mineralization models to improve predictions of carbon and nitrogen fluxes. *Global Biogeochemical Cycles*, **28**, 223–38.

Herbst, M., Welp G., Macdonald A., Jate M., Hädicke A., Scherer H., Gaiser T., Herrmann F., Amelung W. & Vanderborght J. 2018. Correspondence of measured soil carbon fractions and RothC pools for equilibrium and non-equilibrium states. *Geoderma*, **314**, 37-46.

Hirte, J., Leifeld, J., Abiven, S., Oberholzer, H. R., & Mayer, J. (2018). Below ground carbon inputs to soil via root biomass and rhizodeposition of field-grown maize and wheat at harvest are independent of net primary productivity. *Agriculture, Ecosystems & Environment*, **265**, 556-566.

Hirte, J., Leifeld, J., Abiven, S. & Mayer J. 2018. Maize and wheat root biomass, vertical distribution, and size class as affected by fertilization intensity in two long-term field trials. *Field Crops Research*, **216**, 197–208.

Hu, T., Sørensen, P., et al. 2018. Root biomass in cereals, catch crops and weeds can be reliably estimated without considering aboveground biomass. *Agriculture, Ecosystems & Environment*, **251**, 141–48.

Insam, H. & Domsch, K.H. 1988. Relationship between soil organic carbon and microbial biomass on chronosequences of reclamation sites. *Microbial Ecology*, **15**, 177–188.

IUSS Working Group WRB. 2006. *World Reference Base for Soil Resources 2006*. World Soil Resources Reports No 103. FAO, Rome.

Jackson, R. B., Canadell, J., Ehleringer, J.R., Mooney H.A., Sala O.E. & Schulze E.D. 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia*, **10**8, 389–411.

Jenkinson, D. S. & Coleman, K. 2008. The turnover of organic carbon in subsoils. Part 2. modelling carbon turnover. *European Journal of Soil Science*, **59**, 400–413.

Keel, S. G., Leifeld, J., Mayer, J., Taghizadeh-Toosi, A. & Olesen, J. E. 2017. Large uncertainty in soil carbon modelling related to method of calculation of plant carbon input in agricultural systems. *European Journal of Soil Science*, **68**, 953–963.

Kuzyakov, Y. 2002. Review: factors affecting rhizosphere priming effects. *Journal of Plant Nutrition and Soil Science*, **165**, 382–96.

Leifeld, J. & Mayer J. 2015. ^{14}C in cropland soil of a long-term field trial–experimental variability and implications for estimating carbon turnover, *SOIL*, **1**, 537–542.

Leifeld, J., Reiser, R., Oberholzer, H. R. 2009. Consequences of conventional versus organic farming on soil carbon: Results from a 27-year field experiment. *Agronomy Journal*,

101, 1204–1218.

Liu, Y., Freer, J., Beven, K. & Matgen, P. 2009. Towards a limits of acceptability approach to the calibration of hydrological models: extending observation error. *Journal of Hydrology*, **36**7, 93–103.

Ludwig, B., Schulz, E., Rethemeyer, J., Merbach, I. & Fless, H. 2007. Predictive modelling of C dynamics in the long-term fertilization experiment at Bad Lauchstädt with the Rothamsted Carbon Model. *European Journal of Soil Science*, **58**, 1155–1163.

Luo, Y. *et al*. 2016. Toward more realistic projections of soil carbon dynamic by earth system models. *Global Biogeochemical Cycles*, **30**, 40–56.

Luo, Z. Wang, E., Zheng, H., Baldock, J.A., Sun, O.J. & Shao Q.5 2015. Convergent modelling of past soil organic carbon stocks but divergent projections. *Biogeosciences*, **12**, 4373–83.

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

Marschner, B., Brodowski, S., Dreves, A., Gleixner, G. et al. 2008. How relevant is recalcitrance for the stabilization of organic matter in soils? *Journal of Plant Nutrition and Soil Science*, **171**, 91–110.

Menichetti, L., Kätterer, T. & Leifeld, J. 2016. Parametrization consequences of constraining soil organic matter models by total carbon and radiocarbon using long-term field data. *Biogeosciences*, **13**, 3003–19.

Mondini, M., Cayuela, M.L., Sinicco, T., Fornasier, F., Galvez, A. & Sánchez-Monedero, M.A. 2017. Modification of the RothC model to simulate soil C mineralization of exogenous organic matter. *Biogeosciences*, **14**, 3253–3274.

Oberholzer, H. R., Leifeld, J. & Mayer, J. 2014. Changes in soil carbon and crop yield over 60 years in the Zurich organic fertilization experiment, following land-use change from grassland to cropland. *Journal of Plant Nutrition and Soil Science*, **177**, 696–704.

Parton, W. J., Stewart, J. W. B. & Cole, C. V. 1988. Dynamics of C, N, P and S in grassland soils: A model. *Biogeochemistry*, **5**, 109–31.

Poeplau, C., Don A., Dondini M., Leifeld J., Nemo R., Schumacher J., Senapati N. & Wiesmeier M. 2013. Reproducibility of a soil organic carbon fractionation method to derive RothC carbon pools. *European Journal of Soil Science*, **64**, 735–46.

R Core Team. *R: A Language and Environment for Statistical Computing*. 2018. R Foundation for Statistical Computing, Vienna, Austria.

Sanderman, J., Baisden, W. T. & Fallon, S. (2016). Redefining the inert organic carbon pool. *Soil Biology & Biochemistry*, **92**, 149-152.

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

Segoli, M., De Gryze S., Doua, F., Lee, J., Post, W.M., Denef, K. & Six, J. 2013. AggModel: A soil organic matter model with measurable pools for use in incubation studies. *Ecological Modelling*, **263**, 1–9.

Skjemstad, J.O., Spouncer, L.R.,Cowie, B. & Swift, R.S. 2004. Calibration of the Rothamsted organic carbon turnover model (RothC ver.26.3), using measurable soil organic carbon pools. *Australian Journal of Soil Resarch*, **42**, 79-88.

Stein, M. 2012. Large sample properties of simulations using latin hypercube sampling. *Technometrics*, **29**, 37–41.

Taghizadeh-Toosi, A., Christensen,B.T., Glendining, M. & Olesen J.E. 2016. Consolidating soil carbon turnover models by improved estimates of belowground carbon input. *Scientific Reports*, **6**:32568.

Vance, E. D., Brookes P.C. & Jenkinson D.D. 1987. An extraction method for measuring soil microbial biomass C. *Soil Biology & Biochemistry*, **19**, 703–707.

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

Zimmermann, M., Leifeld J., Schmidt M.W.I., Smith P. & Fuhrer J. 2007. Measured soil organic matter fractions can be related to pools in the RothC model. *European Journal of Soil Science*, **58**, 658–667.

TABLE CAPTIONS

Table 1: Acceptability criteria under the four objectives described in this study (total C,'Ctot', total C and microbial biomass 'Ctot+Biom', total C and soil basal respiration 'Ctot+Rsoil', total C and results from a Zimmermann fractionation 'Ctot+Zim') for the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments.

Table 2: Estimates of total root biomass, rhizodeposition and main product yield for maize and winter wheat using the two approaches for C input estimates described in this study ('C_{Hirte}' and '$C_{Bolinder}$') in the ZOFE trial between 1949–2014 for the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments.

Table 3: Percentage difference of belowground and total C inputs estimates for maize, winter wheat and all crops calculated with the two approaches for C input estimates described in this study (('C_{Hirte}' – '$C_{Bolinder}$') / '$C_{Bolinder}$' × 100) in the ZOFE trial between 1949–2014 for the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments.

Objectives	Acceptability Criteria
'Ctot'	Mean (NSE_NIL, NSE_NPK, NSE_FYM) > 0 & RMSE_CTOT_NIL < SD_CTOT x 2.5 & RMSE_CTOT_NPK < SD_CTOT x 2.5 & RMSE_CTOT_FYM < SD_CTOT x 2.5
'Ctot+Biom'	'Ctot' criteria & RMSE_BIOM_NIL < SD_BIOM x 1.7 & RMSE_BIOM_NPK < SD_BIOM x 1.7 & RMSE_BIOM_FYM < SD_BIOM x 1.7
'Ctot+Rsoil'	'Ctot' criteria & RMSE_RSOIL_NIL < SD_RSOIL x 1.7 & RMSE_RSOIL_NPK < SD_RSOIL x 1.7 & RMSE_RSOIL_FYM < SD_RSOIL x 1.7
'Ctot+Zim'	'Ctot' criteria & RMSE_ZIM < SD_ZIM

SD refers to the standard deviation of the measured data, RMSE refers to the root mean squared error of the modelled and measured data, NSE refers to the Nash–Sutcliffe efficiency, an alternative model error measurement (refer to the main text for details).

Table 1

	Roots/ t C ha^{-1}			
	‘$C_{Bolinder}$’			**‘C_{Hirte}’**
Treatments	**NIL**	**NPK**	**FYM**	**All**
Maize mean	0.52 (0.39)	1.03 (0.31)	0.76 (0.34)	0.4 (0.12)
Winter wheat mean	0.26 (0.11)	0.63 (0.14)	0.43 (0.1)	0.47 (0.13)
	Rhizodeposition/ t C ha^{-1}			
	‘$C_{Bolinder}$’			**‘C_{Hirte}’**
Treatments	**NIL**	**NPK**	**FYM**	**All**
Maize mean	0.34 (0.25)	0.67 (0.2)	0.49 (0.22)	0.53 (0.13)
Winter wheat mean	0.17 (0.07)	0.41 (0.09)	0.28 (0.07)	0.63 (0.13)
	Product Yield/ t C ha^{-1}			
Treatments	**NIL**	**NPK**	**FYM**	
Maize mean	1.17 (0.87)	2.91 (0.94)	1.99 (0.95)	
Winter wheat mean	0.77 (0.34)	1.88 (0.54)	1.37 (0.37)	

Mean, geometric mean; SD, standard deviation, in brackets.

Table 2

	Belowground C input difference/ %			**Total plant C input difference/ %**		
Treatments	**NIL**	**NPK**	**FYM**	**NIL**	**NPK**	**FYM**
Maize mean	7	–46	–26	6	–40	–23
Winter wheat mean	160	6	56	128	5	44
All species mean	27	–10	4	18	–7	2

Mean = geometric mean.

Table 3

FIGURE CAPTIONS

Figure 1: Measurements and simulations of SOC, relative microbial biomass and relative basal respiration with RothC default parameters using the two approaches for C input estimates described in this study (‘C_{Hirte}’ and ‘$C_{Bolinder}$’) in the ZOFE trial between 1949–2014 for the control (NIL), mineral fertilizer (NPK) and farmyard manure (FYM) treatments.

Figure 2: Simulation uncertainty for SOC of accepted parameter sets under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘$C_{Bolinder}$’ for the (a) control (NIL), (b) mineral fertilizer (NPK) and (c) farmyard manure (FYM) treatments.

Figure 3: Kolgomorov–Smirnov *d* statistics for parameter sensitivity analysis under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘$C_{Bolinder}$’.

Figure 4: Parameter probability distribution under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘$C_{Bolinder}$’.

Figure 5: Simulation uncertainty of basal respiration relative to the NIL treatment ($Resp/Resp_{NIL}$) under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and

results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘$C_{Bolinder}$‘ for the (a) mineral fertilizer (NPK) and (b) farmyard manure (FYM) treatments. Crosses are measured values.

Figure 6: Simulation uncertainty of microbial biomass relative to the NIL treatment ($Biom/Biom_{NIL}$) under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘$C_{Bolinder}$‘ for the (a) mineral fertilizer (NPK) and (b) farmyard manure (FYM) treatments. Crosses are measured values.

Figure 7: Simulation uncertainty of microbial biomass relative to the NIL treatment ($Biom/Biom_{NIL}$) under the four objectives described in this study (total C, ‘Ctot’, total C and microbial biomass ‘Ctot+Biom’, total C and soil basal respiration ‘Ctot+Rsoil’, total C and results from a Zimmermann fractionation ‘Ctot+Zim’) using the C input estimate approach ‘C_{Hirte}’ for the (a) mineral fertilizer (NPK) and (b) farmyard manure (FYM) treatments. Crosses are measured values.

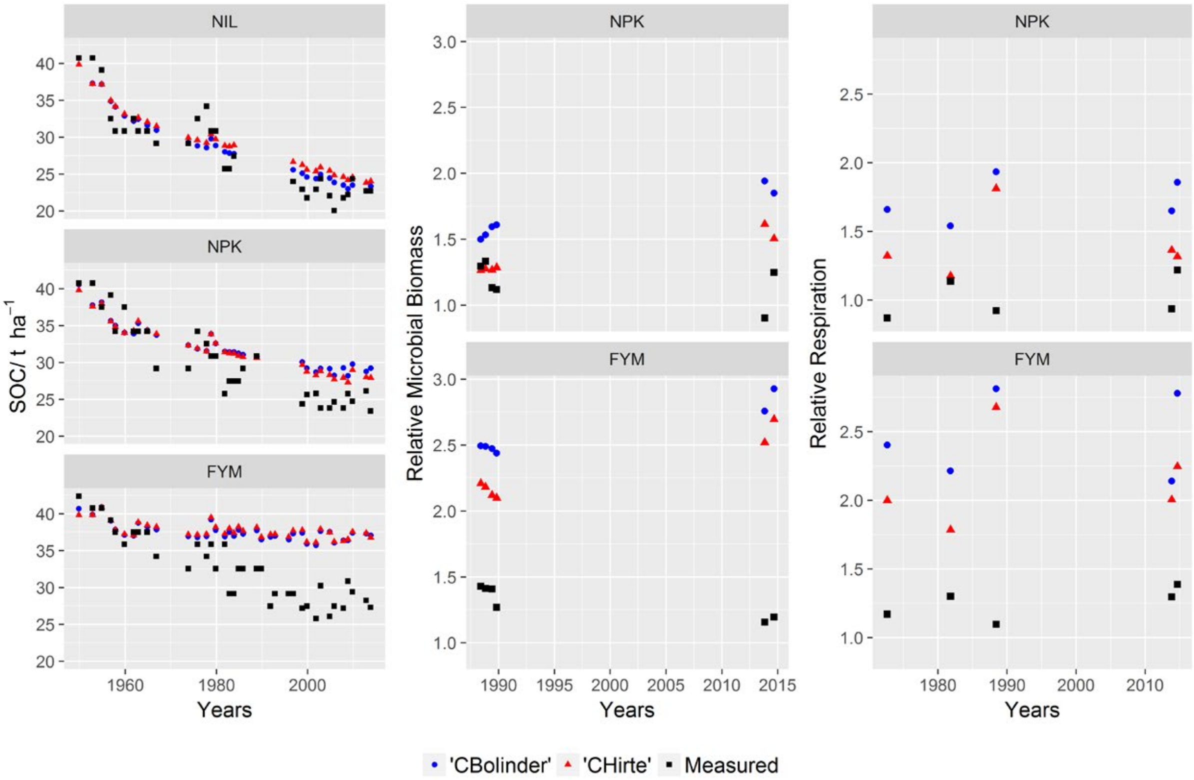
NIL
NPK
FYM
SOC/ t ha^{-1}
40
35
30
25
20
1960
1980
2000
Years
Relative Microbial Biomass
3.0
2.5
2.0
1.5
1.0
1990
1995
2000
2005
2010
2015
Years
Relative Respiration
2.5
2.0
1.5
1.0
1980
1990
2000
2010
Years
● 'CBolinder'
▲ 'CHirte'
■ Measured

Figure 1

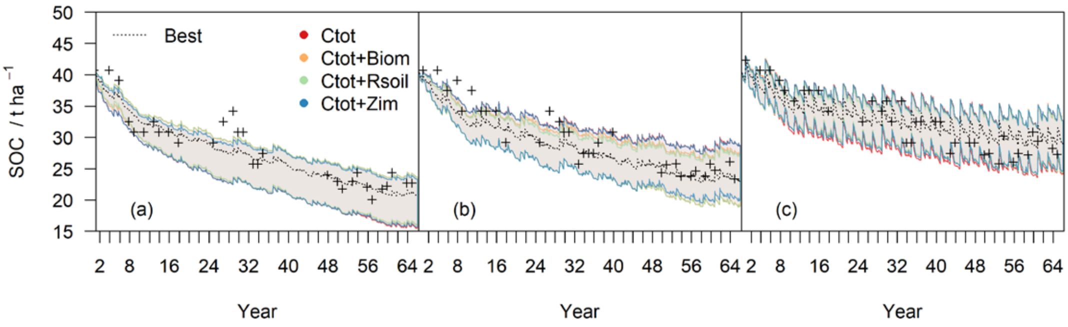
Best
● Ctot
● Ctot+Biom
● Ctot+Rsoil
● Ctot+Zim
SOC / t ha $^{-1}$
50
45
40
35
30
25
20
15
(a)
(b)
(c)
2
8
16
24
32
40
48
56
64
Year
Year
Year

Figure 2

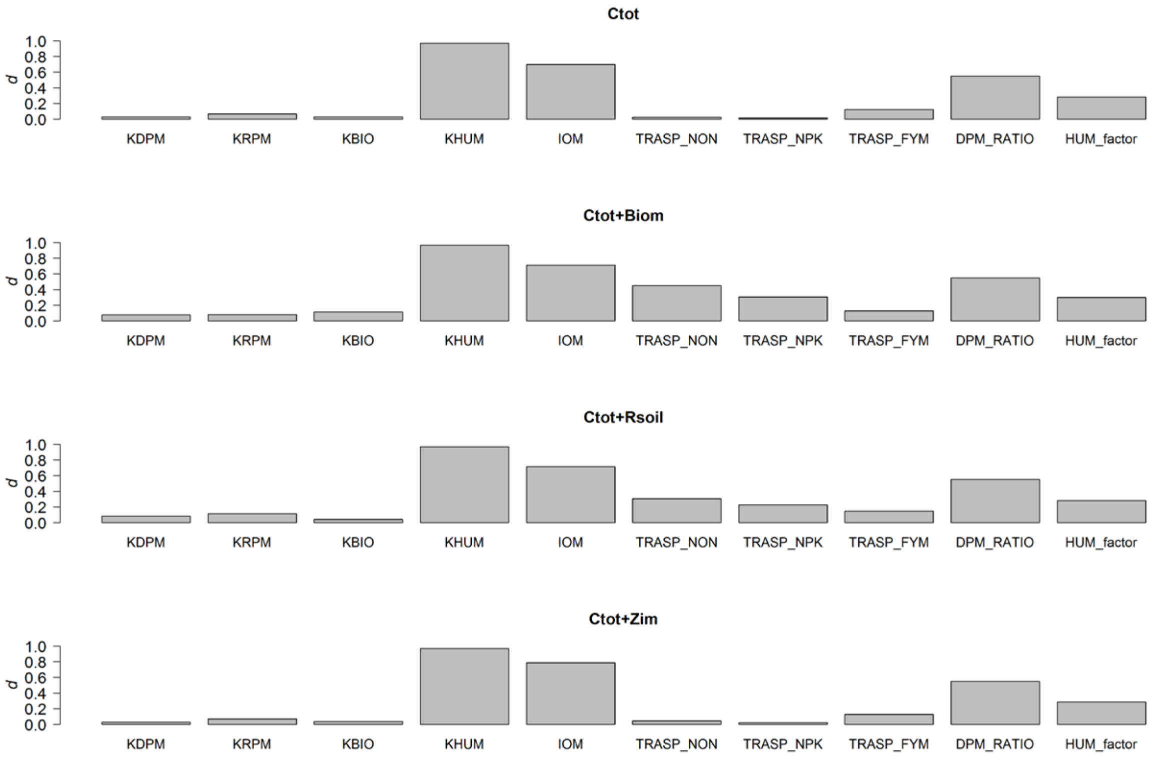
Ctot
d
1.0
0.8
0.6
0.4
0.2
0.0
KDPM
KRPM
KBIO
KHUM
IOM
TRASP_NON
TRASP_NPK
TRASP_FYM
DPM_RATIO
HUM_factor
Ctot+Biom
Ctot+Rsoil
Ctot+Zim

Figure 3

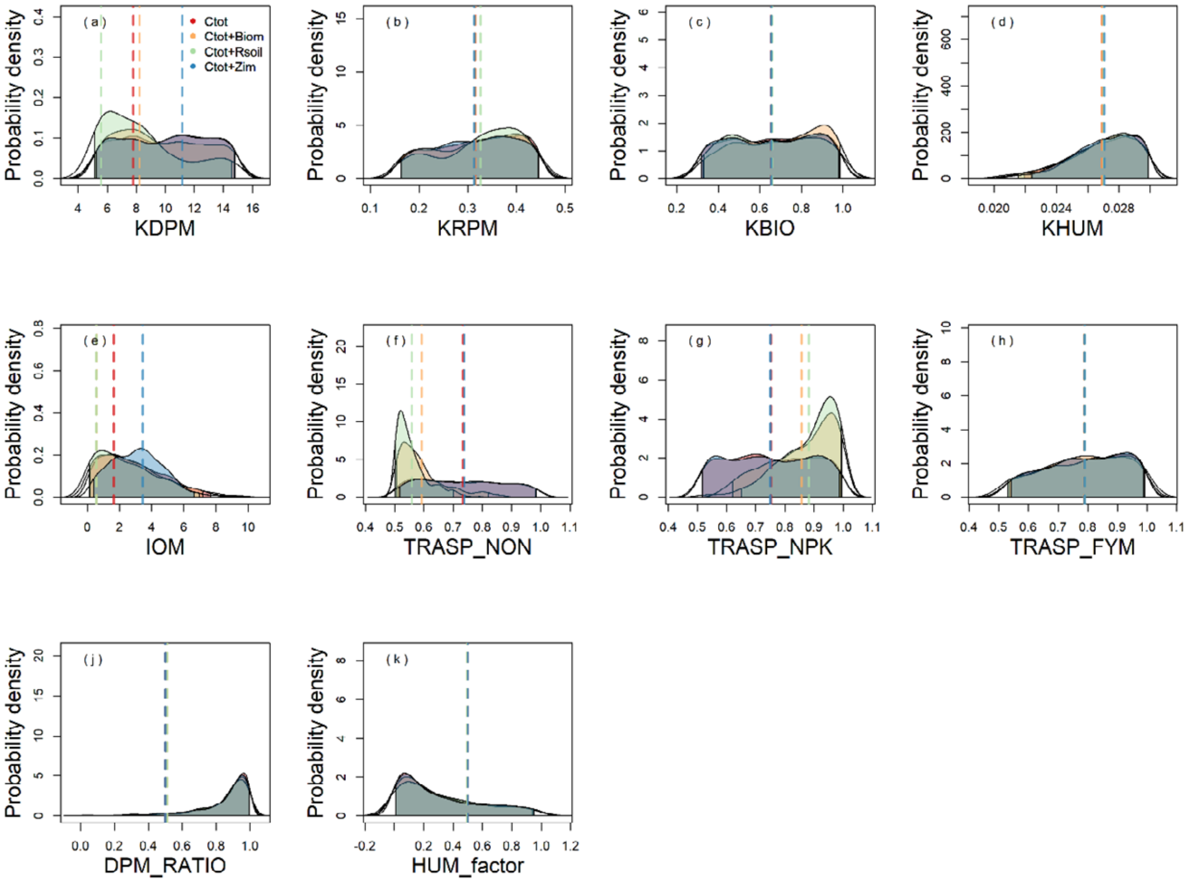
(a)
Ctot
Ctot+Biom
Ctot+Rsoil
Ctot+Zim
Probability density
0.0 0.1 0.2 0.3 0.4
4 6 8 10 12 14 16
KDPM
(b)
Probability density
0 5 10 15
0.1 0.2 0.3 0.4 0.5
KRPM
(c)
Probability density
0 1 2 3 4 5 6
0.2 0.4 0.6 0.8 1.0
KBIO
(d)
Probability density
0 200 400 600
0.020 0.024 0.028
KHUM
(e)
Probability density
0.0 0.2 0.4 0.6 0.8
0 2 4 6 8 10
IOM
(f)
Probability density
0 5 10 15 20
0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1
TRASP_NON
(g)
Probability density
0 2 4 6 8
0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1
TRASP_NPK
(h)
Probability density
0 2 4 6 8 10
0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1
TRASP_FYM
(j)
Probability density
0 5 10 15 20
0.0 0.2 0.4 0.6 0.8 1.0
DPM_RATIO
(k)
Probability density
0 2 4 6 8
-0.2 0.2 0.4 0.6 0.8 1.0 1.2
HUM_factor

Figure 4

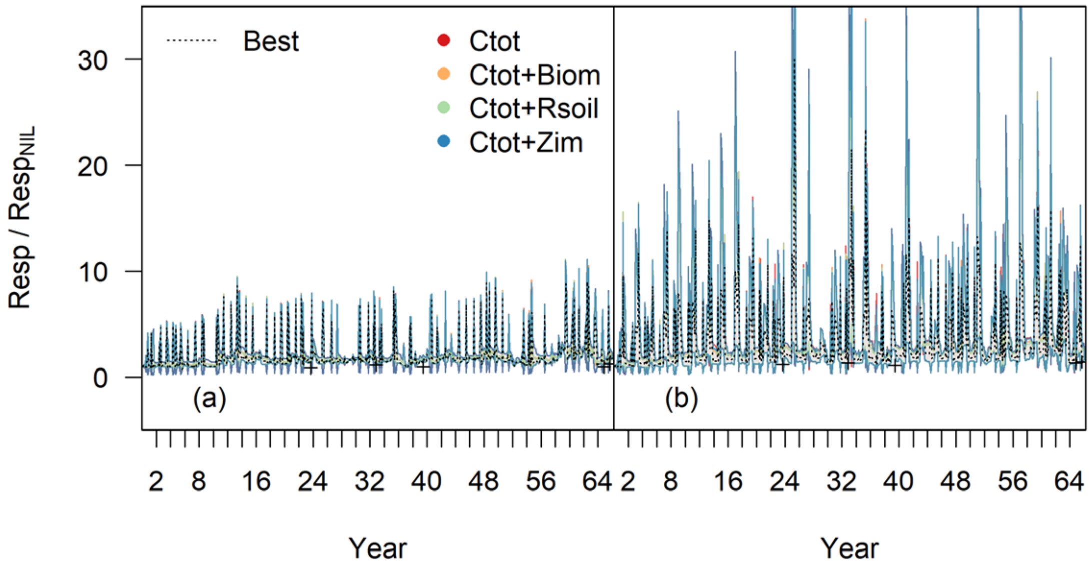
Best
● Ctot
● Ctot+Biom
● Ctot+Rsoil
● Ctot+Zim
Resp / Resp_{NIL}
30
20
10
0
(a)
(b)
2
8
16
24
32
40
48
56
64
Year

Figure 5

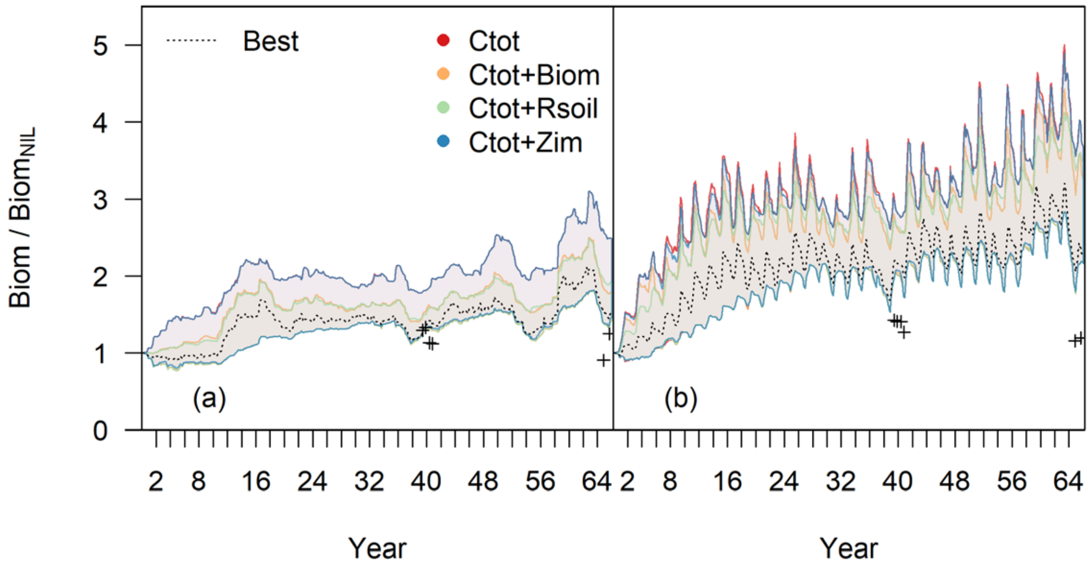
Best
Ctot
Ctot+Biom
Ctot+Rsoil
Ctot+Zim
Biom / Biom$_{NIL}$
0
1
2
3
4
5
(a)
(b)
2
8
16
24
32
40
48
56
64
Year

Figure 6

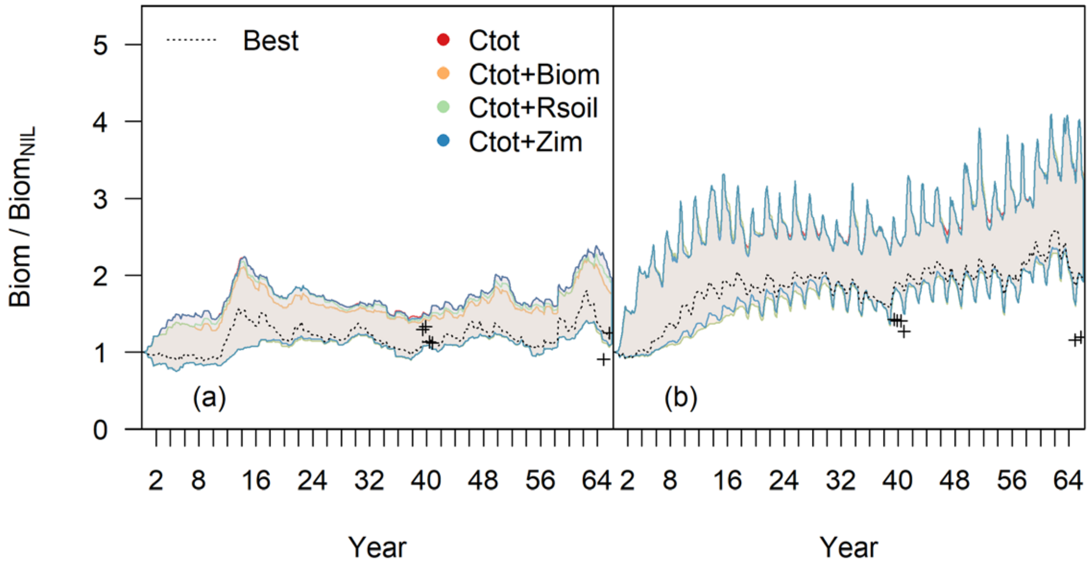
Best
Ctot
Ctot+Biom
Ctot+Rsoil
Ctot+Zim
Biom / Biom$_{NIL}$
0
1
2
3
4
5
(a)
(b)
2
8
16
24
32
40
48
56
64
Year
Year

Figure 7