Multi-modelling predictions show high uncertainty of required carbon input changes to reach a 4‰ target

Elisa Bruni, Claire Chenu, Rose Z Abramoff, Guido Baldoni, Dietmar Barkusky, Hugues Clivot, Yuanyuan Huang, Thomas Kätterer, Dorota Pikula, Heide Spiegel, et al.

Received: 14 July 2022 | Revised: 24 October 2022 | Accepted: 24 November 2022

DOI: 10.1111/ejss.13330

SPECIAL ISSUE PAPER

European Journal of Soil Science WILEY

Multi-modelling predictions show high uncertainty of required carbon input changes to reach a 4‰ target

Elisa Bruni[1,2] | **Claire Chenu**[3] | **Rose Z. Abramoff**[1,4] | **Guido Baldoni**[5] | **Dietmar Barkusky**[6] | **Hugues Clivot**[7] | **Yuanyuan Huang**[8] | **Thomas Kätterer**[9] | **Dorota Pikuła**[10] | **Heide Spiegel**[11] | **Iñigo Virto**[12] | **Bertrand Guenet**[2]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

[2]LG-ENS (Laboratoire de géologie)—CNRS UMR 8538—École normale supérieure, PSL University—IPSL, Paris, France

[3]Ecosys, INRA-AgroParisTech, Université Paris-Saclay, Campus AgroParisTech, Thiverval-Grignon, France

[4]Oak Ridge National Laboratory, Environmental Sciences Division & Climate Change Science Institute, Oak Ridge, Tennessee, USA

[5]Department of Agro Environmental Science & Technology, Alma Mater Studiorum, University of Bologna, Bologna, Italy

[6]Centre for Agricultural Landscape Research ZALF, Experimental Infrastructure Platform, Working Group "Experimental Station Müncheberg", Müncheberg, Germany

[7]Université de Reims Champagne-Ardenne, INRAE, FARE, UMR A 614, Reims, France

[8]CSIRO Oceans and Atmosphere, Aspendale, Victoria, Australia

Abstract

Soils store vast amounts of carbon (C) on land, and increasing soil organic carbon (SOC) stocks in already managed soils such as croplands may be one way to remove C from the atmosphere, thereby limiting subsequent warming. The main objective of this study was to estimate the amount of additional C input needed to annually increase SOC stocks by 4‰ at 16 long-term agricultural experiments in Europe, including exogenous organic matter (EOM) additions. We used an ensemble of six SOC models and ran them under two configurations: (1) with default parametrization and (2) with parameters calibrated site-by-site to fit the evolution of SOC stocks in the control treatments (without EOM). We compared model simulations and analysed the factors generating variability across models. The calibrated ensemble was able to reproduce the SOC stock evolution in the unfertilised control treatments. We found that, on average, the experimental sites needed an additional 1.5 ± 1.2 Mg C ha^{-1} $year^{-1}$ to increase SOC stocks by 4‰ per year over 30 years, compared to the C input in the control treatments (multi-model median ± median standard deviation across sites). That is, a 119% increase compared to the control. While mean annual temperature, initial SOC stocks and initial C input had a significant effect on the variability of the predicted C input in the default configuration (i.e., the relative standard deviation of the predicted C input from the mean), only water-related variables (i.e., mean annual precipitation and potential evapotranspiration) explained the divergence between models when calibrated. Our work highlights the challenge of increasing SOC stocks in agriculture and

[9]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden

[10]Department of Plant Nutrition and Fertilization, Institute of Soil Science and Plant Cultivation State Research Institute, Puławy, Poland

[11]Department for Soil Health and Plant Nutrition, Austrian Agency for Health and Food Safety (AGES), Vienna, Austria

[12]Departamento de Ciencias. IS-FOOD, Universidad Pública de Navarra, Pamplona, Spain

Correspondence
Elisa Bruni, LG-ENS (Laboratoire de géologie)—CNRS UMR 8538—École normale supérieure, PSL University—IPSL, Paris, France.
Email: bruni@geologie.ens.fr

Funding information
ANR-16-CONV-0003 (CLAND); GA n 101000289 (Holisoils)

accentuates the need to increasingly lean on multi-model ensembles when predicting SOC stock trends and related processes. To increase the reliability of SOC models under future climate change, we suggest model developers to better constrain the effect of water-related variables on SOC decomposition.

Highlights:

- The feasibility of the 4‰ target was studied at 16 long-term agricultural experiments.
- An ensemble of soil organic carbon models was used to estimate the uncertainty of the predictions.
- On average across the sites, carbon input had to increase by 119% compared to initial conditions.
- High uncertainty of the simulations was mainly driven by water-related variables.

KEYWORDS
4 per 1000 initiative, agriculture, carbon sequestration, climate change, European targets, multi-modelling, soil organic carbon

1 | INTRODUCTION

The latest report of the Intergovernmental Panel on Climate Change (IPCC, 2021) announced observed changes in the whole climate system in every region across the world. Although many of the changes already set in motion are irreversible over hundreds to thousands of years, strong and sustained reduction of greenhouse gas emissions (GHG) could still limit climate change (IPCC, 2021). Additional efforts to decrease the level of carbon dioxide (CO_2) and other GHGs in the atmosphere are expected from land-based mitigation solutions.

The European Commission has recently released a set of targets for European soil health (e.g., COM(2020) 381 final; COM(2021) 800 final; SWD(2021) 450 final) (European Commission, 2020, 2021a, 2021b), which includes the contribution of soils to climate change mitigation via increased atmospheric carbon (C) sequestration. These targets aim to reverse the current average decline of soil organic carbon (SOC) in European croplands (i.e., 5‰ $year^{-1}$) to a 1‰-4‰ annual increase (Veerman et al., 2020). With the same perspective, the "4 per 1000" initiative has gathered contributions from hundreds of partners across the world since 2015, to promote agricultural practices that help to maintain or enrich cultivated soils in organic C, including those which reduce the mineralization of organic C and increase its stocks in soils (Minasny et al., 2017). This will have the combined effect of improving soil quality (e.g., soil fertility and water retention) (Lal, 2008) while mitigating climate change through increased C sequestration in the soil. Despite the multiple benefits provided by increasing SOC stocks, the feasibility of a 4‰ objective with current agricultural management practices is still under debate (e.g., Chabbi et al., 2017; Rumpel et al., 2020; Soussana et al., 2019; van Groenigen et al., 2017). Recently, some studies using process-based models focused on the bio-technical feasibility of SOC stock increase targets, such as the 4‰ objective (e.g., Bruni et al., 2021; Martin et al., 2021; Riggers et al., 2021). Individual model predictions of a 4‰ increase target in Europe are relatively optimistic. That is, a required 30 to 40% C input increase in France according to Martin et al. (2021), and a 43% increase in European long-term experiments (LTEs) according to Bruni et al. (2021) under present climate conditions. A multi-modelling exercise from Riggers et al. (2021) predicted a much larger increase for German croplands, that is a 213–283% increase of C input required between 2090 and 2099, compared to 2014, under different climate change scenarios. Multi-model ensemble means are expected to provide improved estimates compared to singular model simulations, due to the relative independence of different SOC models' simulation errors (IPCC, 2007). Furthermore, simulations designed with multiple models that have underlying structural differences provide an uncertainty range of SOC projections that reflects our current understanding of SOC processes and their possible representations. The use of multi-model ensembles to predict the evolution of complex systems is a widespread practice in other disciplines, such as climate modelling (Jebeile & Crucifix, 2020; Parker, 2010; Tebaldi & Knutti, 2007). Although some efforts have been made in the soil modelling community to embrace this

practice (e.g., Farina et al., 2021; Palosuo et al., 2012; Riggers et al., 2021; Sulman et al., 2018), its use is not consolidated yet.

In the present paper, we aim to (1) use a multi-model ensemble to simulate the SOC stock evolution in long-term cropland experiments, and evaluate two multi-model ensemble configurations, one with default model parameters and the other with parameters calibrated site-by-site, (2) provide an estimate of the C input required to annually increase SOC stocks by 4‰ in 16 LTEs across Europe, and (3) identify potential factors creating uncertainty across models. With this work, we want to contribute to the understanding of the feasibility of a 4‰ SOC stock increase target in Europe, and to add a piece to the ongoing discussion about the use of multi-model ensembles in soil science.

2 | MATERIALS AND METHODS

2.1 | Experimental sites

The dataset used in this study compiles 16 long-term cropland experiments located in Europe (9 in France and 1 each in Spain, Great Britain, Sweden, Italy, Germany, Poland and Austria). Each experiment includes one or several treatments with the addition of exogenous organic material (EOM) and a control treatment without any EOM addition, but with the same crops and fertiliser inputs. In total, the database included 43 EOM treatments and 16 controls. The data consists of several measurements of SOC content and its variance across replicates, yearly crop yields and different soil characteristics (Table A1). The experiments lasted on average 26 years (median of 21 years), in the period between 1956 and 2018. EOM inputs were applied to the soil at different rates and frequencies and varied from animal manure (swine, bovine and poultry) to sewage sludge, peats, castor meal, sawdust, bio waste, green manure and household waste (i.e., residual organic material generated from residential waste). Data for the Bologna experiment were directly extracted from Triberti et al. (2008) and consisted of the average SOC stock evolution in different inorganic nitrogen (N) experiments (i.e., one treatment without any inorganic fertiliser and 3 treatments with different levels of N input).

Cropping systems (Table A2) were cereal-dominated rotations (wheat, maize, barley and oat). In particular, four were monocultures of forage crops or cereals (silage maize in Champ Noël 3, Le Rheu 1 and Le Rheu 2 and winter wheat in Broadbalk) and five sites had rotations of different cereals (winter wheat and silage or grain maize in Crécom 3 PRO, Feucherolles, La Jaillière 2 PRO, Avrillé and Bologna). The other experiments rotated cereal crops with legumes (chickpea, pea) and/or root crops (potatoes, fodder beet, fodder rape and Swedish turnip), oilseed crops (oilseed flax, sunflower, oilseed rape, and mustard), and cover crops (ryegrass). Except for Müncheberg which was irrigated in 4 out of 8 replicates between 1974 and 1981, all experiments were rainfed and managed under conventional tillage (the Ultuna trial was tilled by hand with a spade to mimic conventional tillage). Straw residues were exported from the field, except in the French and Austrian sites, where residues were partly or totally incorporated into the soil. In the French experiments Champ Noël 3, Crécom 3 PRO, La Jaillière 2 PRO, Le Rheu1 and Trévarez received optimal amounts of mineral N fertilisers both in the control and in the treatments. In the Polish experiment of Grabów, N, phosphorus (P), and potassium (K) were applied.

2.2 | Climate forcing

Daily soil surface temperature, moisture and potential evapotranspiration (PET) were simulated for each site using the land-surface model ORCHIDEE (Krinner et al., 2005). Simulations were run using a 3-hourly global climate dataset at 0.5° (GSWP3 http://hydro.iis.u-tokyo.ac.jp/GSWP3/), from which were also derived daily precipitation data. Mean annual surface temperature (MAST) during the experiments ranged between 5.7°C and 12.8°C across the sites, while mean annual precipitation (MAP) was 850 mm, with a minimum of 613 mm and a maximum of 1314 mm (Table A3). The virtual amount of C input required to increase SOC stocks was analysed over the period 1980–2010, which was the 30-yearlong interval covering the majority of the experiments.

2.3 | Soil sampling

Soil samples were collected between 0–20 and 0–40 cm depth, in 3 to 8 replicates. In Champ Noël 3, replicates were not available, and in Broadbalk, SOC was sampled using a semi-cylindrical auger, bulking together 10–20 cores from across the plot. SOC stocks were calculated using the standard formula:

$$\begin{aligned}&\mathrm{SOC}\left(\mathrm{Mg\,C\,ha^{-1}}\right)\\&\quad=\mathrm{SOC}(\%)\cdot\mathrm{BD}\left(\mathrm{g\,cm^{-3}}\right)\cdot\mathrm{sampling\,depth(cm)}\\&\qquad\cdot(1-\mathrm{coarse\,fragments\,fraction(vol.\%/100)}),\end{aligned}$$

where SOC (%) is the concentration of organic C in the soil and BD is the average bulk density of the

experimental plot. BD across the sites ranged between 1.1 and 1.7 $g\,cm^{-3}$. Its evolution over time in the EOM treatments was not taken into account due to lack of data for all experiments. In Ritzlhof, BD measurements were not available. Hence, for this site we estimated BD using the intrinsically linear pedotransfer function developed by Kaur et al. (2002), using clay, silt, and organic C content in the soil. We compared the value of BD obtained with this method to LUCAS soil maps (Ballabio et al., 2016) and found similar results. Clay content ranged from 5% to 36%, while soil pH varied from 5.8 to 8.6 across the sites. Calcium carbonate ($CaCO_3$) content was relevant in Arazuri, Colmar, Grabów and Broadbalk soils (160, 130, 77 and 20 $g_{CaCO_3} \cdot kg^{-1}{}_{soil}$ respectively), while the rest of the sites had none or negligible quantities of $CaCO_3$.

2.4 | Multi-model ensemble

Six SOC models were used for the multi-model ensemble analysis: Century (Parton et al., 1988), Roth-C (Coleman & Jenkinson, 1996), ICBM (Andrén & Kätterer, 1997), AMG (Andriulo et al., 1999), MIMICS (Wieder et al., 2015) and Millennial (Abramoff et al., 2022). All six models take as input C from plant litter and other organic material and focus on the dynamics of C within a single soil layer (0–30 cm). Four of the models (i.e., Century, Roth-C, ICBM and AMG) represent soil C dynamics using a conventional multi-compartmental structure, where C is decomposed following first-order decay rates. The number of equations (and compartments) differs from model to model. The remaining two more recent models (i.e., MIMICS and Millennial) have microbial explicit C pools, where the turnover of litter and SOC pools is governed by temperature-sensitive Michaelis–Menten kinetics. Each model was initialised with the standard modelling practice which is commonly used for the model, and using methods that reduced the running time of the spin-up (e.g., the semi-analytical spin-up for Century and Roth-C).

ICBM is run at an annual time step and can be solved analytically due to the linearity of its system of equations. The model consists of two compartments: a young and an old SOC pool. Environmental factors are summarised into one coefficient (r) that affects the decomposition rates of both soil compartments equally. The response functions for the temperature and moisture used to calculate the parameter r, which had to be normalised against the Ultuna experiment, were derived from Fortin et al. (2011) and Karlsson et al. (2011). Following its standard initialization method (Saffih-Hdadi & Mary, 2008), AMG was initialised using the value of SOC during the first year of the experiments and run numerically afterwards. The model contains one fresh organic matter pool and two SOC pools (active and stable). The stable pool is considered constant throughout the simulation length, while the other pools decayed at an annual rate. Both Roth-C and Century models were solved semi-analytically, following the method described by Huang et al. (2018) and Xia et al. (2012). The method consists of (1) solving the set of differential equations by inverse calculations to determine pool sizes at steady state and (2) running the model numerically for the rest of the simulation. Century has four litter pools (structural and metabolic aboveground litter C and structural and metabolic belowground litter C) and three SOC pools (active, slow and passive), which differ in their decomposition rates. It was run at a daily time step. Roth-C simulates the SOC evolution with a monthly time step and was converted into its matrix continuous form following Parshotam (1996). The model has five pools: decomposable and resistant plant material (DPM and RPM), microbial biomass, humified organic matter (HUM) and inert organic C. This latter pool is constant through time and is calculated from the level of SOC at the beginning of the experiment (Coleman & Jenkinson, 1996). Both MIMICS and Millennial models were initialised using a Newton–Raphson approach that calculates the steady-state of the C pools analytically (stode function of the rootSolve package in R [Soetaert & Herman, 2009]). They were run numerically afterwards. MIMICS has seven SOC pools: two litter C pools that correspond to metabolic and structural litter, two microbial pools and three soil organic matter (SOM) pools (a physically protected, a bio-chemically recalcitrant and an available SOM pool). The Millennial model has five measurable pools of C: particulate organic matter (POM), low molecular weight C (LMWC), aggregate C, mineral-associated organic matter, and microbial biomass C (MIC). Both models were run at a daily time step.

2.5 | Calibration of model parameters

All models were run with two configurations: (1) using default parameters and (2) using calibrated parameters that were optimised site by site in order to fit the evolution of observed SOC stocks in the control treatments. Century, Roth-C, ICBM and AMG were coded in python (a link to the codes is provided at the end of the manuscript for all models), and the calibration of their parameters was performed using the sequential least-squares quadratic programming function in Python (SciPy v1.5.1, scipy. optimise package with method = "SLSQP"), a non-linear constrained, gradient-based optimization algorithm (Fu et al., 2019). MIMICS and Millennial were coded in R, and to optimise these models the limited-

TABLE 1 Description of the calibrated parameters related to the decomposition of soil organic carbon (SOC) in the different models

Model	Calibrated parameters	Description of the calibrated parameters	Default parameter value	Calibrated parameter value range	Unit	Reference paper
Century	$M:S_{AG}$	Metabolic: structural ratio of the aboveground litter pools	0.6916	[0.15; 0.85]		Parton et al. (1988)
	$M:S_{BG}$	Metabolic: structural ratio of the belowground litter pools	0.69	[0.15; 0.85]		
	Q_{10}	Q_{10} coefficient of the temperature response function	0.69	[0.62; 1.61]		
	T_{ref}	Reference temperature of the temperature response function	30.0	[19.80; 30.00]	°C	
Roth-C	T_{param}	Parameter of the rate modifying factor for temperature	18.27	[13.98; 29.28]		Coleman and Jenkinson (1996)
ICBM	k_1	Potential mineralization rate affecting the young and old SOC pools	0.80	[0.03;18.81]	$year^{-1}$	Andrén and Kätterer (1997)
	k_2	Potential mineralization rate affecting the old SOC pool	6.0×10^{-3}	$[1.83\times10^{-3};94.46\times10^{-3}]$	$year^{-1}$	
	r	Temperature and moisture response function parameter	1.2×10^{-1}	[0.63; 10.00]		
AMG	k_0	Potential mineralization rate of the active SOC pool	0.165	[0.03; 0.60]	$year^{-1}$	Andriulo et al. (1999)
MIMICS	f_{met}	Metabolic: structural ratio of the litter inputs	0.69	[0.32; 0.85]		Wieder et al. (2015)
	a_v	Tuning coefficient of the maximum reaction velocity of the Michaelis–Menten kinetics	8.00×10^{-6}	$[3.34\times10^{-6}; 8.00\times10^{-6}]$		
	a_k	Tuning coefficient of the half-saturation constant of the Michaelis–Menten kinetics	10.00	[6.54; 20.00]		
Millennial	Ea_{pl}	Activation energy for the maximum rate of POM decomposition	64,320	$[643.77\times10^2; 679.06\times10^2]$	J mol^{-1}	Abramoff et al. (2022)
	K_{pl}	Half-saturation constant of POM decomposition to LMWC	1.00×10^4	$[0.94\times10^4; 1.08\times10^4]$	g C m^{-2}	
	Ea_{lb}	Activation energy for the potential LMWC uptake rate	60.26×10^3	$[57.92\times10^3; 62.33\times10^3]$	J mol^{-1}	

Note: Functions where they appear are described in detail in Appendix A.

memory quasi-Newton method was used (optim function in the stats package in R, with method = "L-BFGS-B", Byrd et al., 1995). We applied a different algorithm to these two models because the SLSQP function was not available in R. The use of R for MIMICS and Millennial was imposed by the fact that they needed to be spun-up using the stode function in R each time a new set of parameters was generated for calibration.

To standardise the optimizations, we selected parameters that affect the C decomposition (see Table 1 and Appendix A). In ICBM, the young pool is multiplied by a decomposition rate (k_1) and the old pool decomposition is controlled by another decomposition rate (k_2). The input into the old pool is controlled by k_1 and treatment-specific humification fractions. Both pools are also altered by the environmental factor r. Parameters k_1, k_2 and r were optimised, following Andrén and Kätterer (1997). The active pool in AMG is decayed at a rate of *k*, which depends on environmental factors and on a potential mineralization rate (k_0). k_0 is usually optimised to fit SOC stocks (Andriulo et al., 1999; Clivot et al., 2019). In Century, C

decomposition is mostly influenced by the temperature response function, which follows the Van't Hoff relationship, based on the Q_{10} factor (van't Hoff, 1884). Following Bruni et al. (2021), we calibrated the Q_{10} and reference temperature factors (T_{ref}), after calibrating the metabolic: structural litter ratio of the aboveground ($M:S_{AG}$) and belowground ($M:S_{BG}$) litter pools. These latter parameters are used to partition the C input into the different litter pools and are a function of the nitrogen: lignin (N:L) ratio of the plants. They were optimised since no data was available on the N:L ratio of the different crops. SOC decomposition in Roth-C is also sensitive to the temperature response function, which is an empirical function initially built for the Rothamsted experiment (Jenkinson, 1990). We calibrated the temperature function parameter (T_{param}) for each experimental site. In MIMICS, we calibrated the tuning coefficients (a_v and a_k) of the temperature-sensitive kinetic parameters, on which the rates of C decomposition depend. As in Century, we also calibrated the parameter that is used to partition litter inputs into their metabolic and structural fraction (f_{met}). In Millennial, we optimised (1) the activation energy (Ea_{pl}) and (2) the half-saturation constant (K_{pl}) of the maximum rate of POM decomposition, and (3) the activation energy (Ea_{lb}) of the maximum uptake rate of the LMWC pool. Both activation energies are parameters in an Arrhenius temperature relationship and are linked to the decomposition of POM into LMWC and to the microbial uptake of LMWC (Abramoff et al., 2022).

The performance of the models was evaluated by testing the similarity between simulated and observed SOC stocks in the control treatments, in terms of (1) their correlation (ρ), (2) their standard deviation (SD) normalised against the SD of the observations, and (3) their centred root-mean-squared error (RMSE), also normalised against the SD of the observations, which is a measure of the prediction's error (Taylor, 2001). Note that the correlation between observed and simulated SOC stocks was tested with Spearman's rank correlation coefficient (ρ) because the variables did not follow a normal distribution.

Carbon input from plant material was calculated from annual crop yield measurements, following the method developed by Bolinder et al. (2007) for Canadian experiments and adapted by Clivot et al. (2019) to the same French experiments used in this study. The allometric functions used to estimate the C input and its allocation to the aboveground and belowground part of the plant can be found in Clivot et al. (2019) and have already been applied to other agricultural experiments in European temperate climates such as those in our study (Bruni et al., 2021).

2.6 | Required C input to increase SOC stocks by 4‰ per year

SOC stocks were simulated for each control treatment over the experiments' duration to evaluate the capability of the models to reproduce observed SOC stocks. The period 1980–2010 was selected to analyse the virtual amount of additional C input required to increase the SOC stocks. We simulated one scenario of SOC stock evolution, where SOC stocks increased on average by 4‰ $year^{-1}$ for 30 years, relative to the initial SOC stocks in the control treatments.

The amount of C input required to increase SOC stocks by the defined target was calculated using an inverse modelling approach that consisted in minimising the following equation:

$$J = \left| {SOC_0}^{model} \cdot target - {SOC_{30}}^{model}(\boldsymbol{I}) \right|$$

where ${SOC_0}^{model}$ and ${SOC_{30}}^{model}$ are the modelled SOC stocks at the onset of the experiment and after 30 years of simulations, respectively; $\boldsymbol{I}$ is the simulated C input required to reach the target, and $target = 1.12$ (i.e., $1 + 0.004 \times 30$) since the objective was to reach an average SOC stock increase of 4‰ $year^{-1}$ for 30 years. We used the Python function SLSQP to solve the optimization problem. Carbon input quality is accounted for differently in the different models. In Millennial, regardless of its quality 1/3 of the C input is allocated to POM and the rest 2/3 to LMWC (Abramoff et al., 2022). In Century and MIMICS, the allocation of the C input to the metabolic and structural litter pools depends on the L:N ratio of the C input material. Hence, the C input quality can be inferred by the M:S ratio of the C input. For these models, during the optimization process, we did not prescribe the quality of the C input since the optimization directly simulated the optimal allocation of C in the different litter pools to reach the 4‰ target. However, for Century, we constrained the virtual C input to have the same aboveground: belowground ratio as the initial litter inputs, assuming that crops would not change with the 4‰ implementation and that the EOM would be equally split above and below the soil surface (see Bruni et al., 2021).

In AMG and ICBM, the humification coefficient h varies according to the quality of the C input. For instance, in AMG h = 0.217 for aboveground winter wheat and h = 0.52 for cow manure (Bouthier et al., 2014). In ICBM, h = 0.125 for straw and crop residues and h = 0.31 for farmyard manure (FYM) (Andrén & Kätterer, 1997). In Roth-C, when entering the soil, 59% of litter inputs from crop plant material are allocated to DPM and 41% to the RPM

compartment, while FYM is assumed to be more decomposed and is split in the following way: 49% DPM, 49% RPM and 2% HUM (Coleman & Jenkinson, 1996). For these three models, a fraction f of the estimated C input was set to have the same quality as the litter input in the control treatment (i.e., its parametrization or its allocation to the different pools). The remaining $(1-f)$ fraction of C input was set to have the average quality of the EOM in the different treatments at the experimental site. For example, for AMG a site with initial litter input from winter wheat equal to 2 Mg C ha^{-1} $year^{-1}$ in the control treatment, and with a cow manure treatment only, a fraction $f = 2/\boldsymbol{I}$ was set to have aboveground h = 0.217 and the remaining fraction $(1-f)$ was set to have aboveground h = 0.52, where $\boldsymbol{I}$ is the estimated C input required to reach the 4‰ and where the maximum value of f was set to be 1. Both non-calibrated and calibrated models were run independently to estimate the amount of C input to reach the 4‰ target at the 16 experimental sites.

2.7 | Comparison of models outputs

Model outputs were compared using different techniques. First, we tested whether the models and the calibration affected the simulated C input needed to reach the 4‰ target across the 16 sites. This was done using a linear mixed-effect (LME) model, with fixed effects for the explanatory variables: "model", "calibration" and the interaction between the two, and including a random effect for "sites". The model was fit by maximising the log-likelihood and an analysis of variance (ANOVA) was applied to test the effect of the different explanatory variables on the simulated C input. Normality of the residuals was tested using a Shapiro–Wilk normality test, and homoscedasticity was tested using Levene's test of equality of variances. Since the residuals were not normally distributed, the data were log-transformed. Second, we looked for groups of models that behaved similarly. We created clusters based on the minimum correlation distance between models' outputs (i.e., the additional C input to reach the 4‰ target). The distance was calculated with an optimization algorithm based on a minimum spanning tree (Müller et al., 2012). To estimate which measured variables better explained the differences between the model outputs, we used a linear model and tested for the normality of the residuals using a Shapiro–Wilk test. The explanatory variables of the linear model were: MAST, MAP, PET, initial C input (C_0^{in}), clay and $CaCO_3$ content, soil C:N and pH, initial SOC stocks and N input (N_{in}). This latter was considered as a categorical variable, equal to 1 if N inputs were applied at any dose and 0 otherwise. The response variable was the relative standard deviation (RSD) among models' outputs (RSD = SD/mean · 100). To select the most parsimonious model, we performed a step-wise regression by Akaike Information Criteria (AIC). The results for the multi-model ensembles are provided as their multi-model median (MMM) and mean.

3 | RESULTS

3.1 | Evaluation of the multi-model ensemble configurations: Prediction of the SOC stocks in the control treatments

Observed SOC stocks in the control treatments, that is, plots without EOM additions were found to

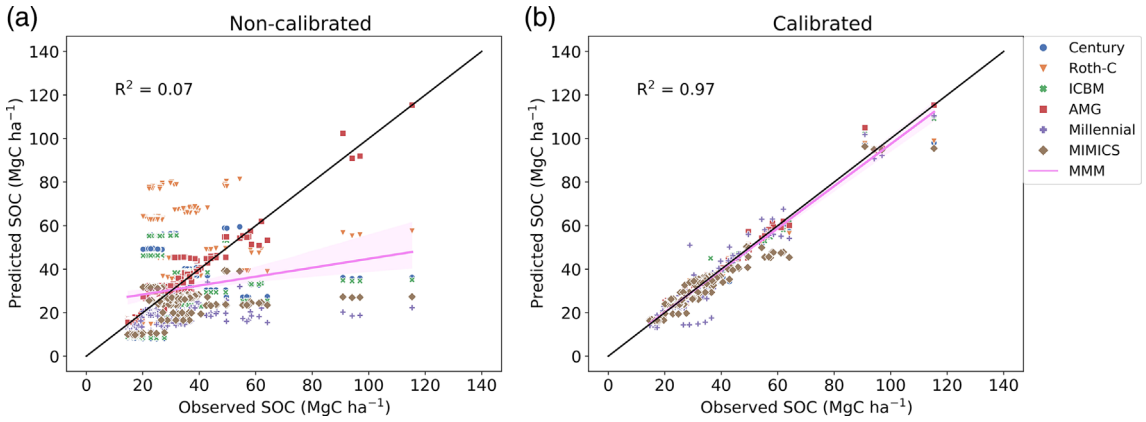

FIGURE 1 Predicted and observed soil organic carbon (SOC) stocks (Mg C ha^{-1}) in the control treatments, according to the six models (different colours and shapes) with: (a) non-calibrated and (b) calibrated parameters. The purple line represents the multi-model median (MMM) and the black line corresponds to the identity line.

decrease on average by 0.54% year^{-1} (Table A2). The calibrated models were able to correctly represent the evolution of SOC stocks in these treatments (Figure 1). In particular, the calibration of model parameters improved the representation of observed SOC stocks in the control treatments (coefficient of determination, R^2 = 0.97) (Figure 1b), compared to the non-calibrated configuration (R^2 = 0.07) (Figure 1a). The calibration also enhanced single model performances, increasing the ρ correlation coefficients to more than 0.89 in all models, approaching the normalised SDs closer to 1, and reducing the normalised centered RMSEs to less than 0.5 (Figure 2). While AMG performed better than any other model in the non-calibrated configuration (i.e., R^2 = 0.90 and RMSE = 5.46), in the calibrated configuration both ICBM and AMG outperformed the other models with the highest R^2 (0.98) and the lowest RMSE (<2.20 Mg C ha^{-1}) (Table 2). Also, the calibrated MMM (R^2 = 0.97 and RMSE = 2.4 Mg C ha^{-1}) outperformed all single models, except for ICBM and AMG.

Table 3 shows the size of the low decomposition rate pools predicted by the different models at initialization. Roth-C systematically underpredicted the size of the inert pool, compared to the other models. This latter had relatively comparable levels of C in their low decomposition rate pools, despite differences throughout the sites (Table 3).

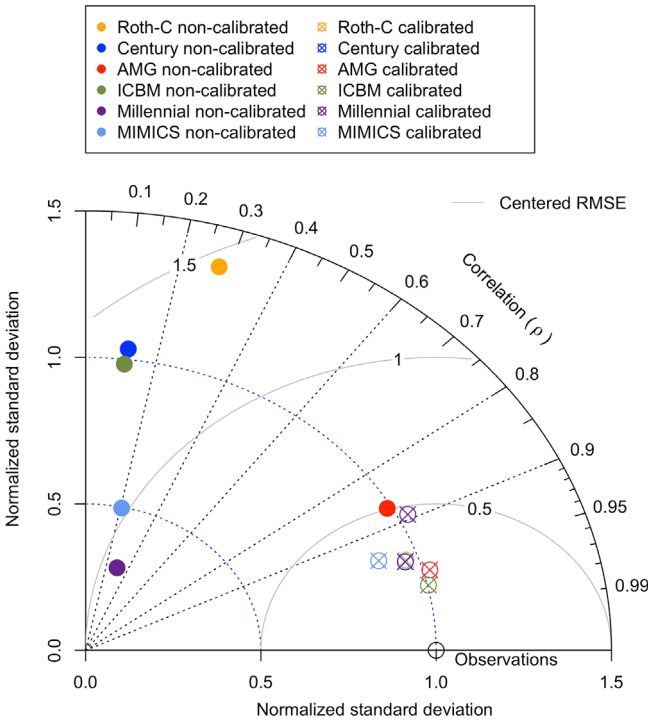

FIGURE 2 Taylor diagram showing the non-calibrated (full spots) and calibrated (crossed spots) model performances in reproducing observed soil organic carbon (SOC) stocks in the control treatments. *X*-axis and *Y*-axis show the standard deviation (SD) of simulated SOC stocks, normalised against the observed SOC stocks' SD. The circumference of the quarter circle shows the Spearman's correlation coefficient (ρ) between simulated and observed SOC stocks, and the grey arcs represent the centred normalised root-mean-squared error (RMSE).

3.2 | Evaluation of the multi-model ensemble configurations: Effect of additional C input on the SOC stock increase

In the EOM treatments, SOC stocks were found to increase by 0.28% year^{-1} on average (Table A2). The capability of the multi-model ensemble to predict the effect of additional C input on SOC stock changes is illustrated in Figure 3. The graph shows the fitted regression line between additional C input and SOC stock increase in the EOM treatments (R^2 = 0.58). For field data, the additional C input was calculated as the amount of yearly average EOM added to the soil, plus the increased crop productivity relative to the control treatment. The regression line between C input and SOC stock variation, and its CI at 95%, can be compared to the simulated additional C input required to reach a 4‰ increase of SOC

TABLE 2 Statistics of models and multi-model median performances

	R^2 non-calibrated	R^2 calibrated	RMSE non-calibrated Mg C ha^{-1}	RMSE calibrated Mg C ha^{-1}
AMG	0.90	0.98	5.46	2.16
Century	0.02	0.96	20.15	2.88
ICBM	0.02	0.98	19.63	2.04
Millennial	0.06	0.92	20.27	4.53
MIMICS	0.10	0.93	17.16	4.31
Roth-C	0.04	0.96	28.67	2.87
MMM	0.07	0.97	16.41	2.41

TABLE 3 Soil organic carbon (SOC) stock (Mg C ha^{-1}) in the low decomposition rate pool of the calibrated models at initialization, for each site

Site	AMG Stable	Century Passive	ICBM Old	Millennial Mineral-associated	MIMICS Passive	Roth-C Inert
CHNO3	26.4	17.3	40.6	13.4	9.8	3.3
COL	35.3	26.5	43.2	22.0	14.0	4.6
CREC3	40.3	27.9	59.6	22.8	13.6	5.4
FEU	25.9	19.4	30.1	15.3	9.6	3.3
LAJA2	21.1	14.9	34.6	11.3	8.1	2.6
RHEU1	23.5	15.6	5.4	13.0	9.6	2.9
RHEU2	23.7	13.9	39.4	6.2	8.4	3.0
ARAZ	36.0	26.4	6.0	19.3	15.2	4.7
ULTU	27.9	16.9	44.3	17.0	12.9	3.5
BROAD	16.1	11.2	25.7	9.3	7.0	1.9
TREV1	75.0	45.7	112.8	37.1	31.2	10.9
AVRI	30.0	21.6	48.7	16.1	11.8	3.9
BOLO	16.5	12.6	2.9	9.4	5.5	2.0
GRAB	20.2	12.9	33.3	10.2	5.6	2.5
MUNCH	12.8	7.2	12.2	7.9	4.2	1.5
RITZ	18.8	15.3	22.6	17.7	10.7	2.3

stocks (MMMs ± CI across sites). In the non-calibrated configuration, the effect of additional C input on the SOC stocks was overestimated by the MMM. On the contrary, the calibrated MMM was not significantly different from the EOM treatments' regression line. This means that, in the calibrated multi-model ensemble configuration, the predicted effect of additional C input on the SOC stocks was close to observations, when compared to the 43 EOM treatments of the 16 experimental sites. Comparing the CI of the MMMs, we can also appreciate that the predictions of the non-calibrated and calibrated configurations were significantly different from each other (Figure 3).

3.3 | Required C input to reach a 4‰ target

Table 4 shows the percentage change of C input required to reach the average 4‰ annual increase of SOC stocks, relative to the initial level of C input in the control treatment. The calibrated configuration predicted a median increase of 118.7% to reach the target and a multi-model mean of 115.6% (±56.5%) (Table 4). The variability across models was high for both non-calibrated and calibrated configurations. However, the calibration reduced the RSD across models (RSD = 48.8%) (Table 4).

Figure 4 shows the site average additional C input predicted by the single models and the MMM. The calibrated configuration predicted that C input had to increase by 1.55 ± 1.20 Mg C ha^{-1} $year^{-1}$ to reach the 4‰ target (MMM of the site average C input ± median SD), compared to the initial C input. As it was shown in Figure 3, the calibrated MMM was lower but not significantly different from the C input needed to increase SOC stocks by 4‰, inferred from EOM treatments' field data (regression line at x = 0.4 in Figure 3), that is, 1.96 ± 0.15 Mg C ha^{-1} $year^{-1}$.

The median site variability of the required additional C input was higher in the calibrated ensemble (RSD = 66.2%), compared to the non-calibrated ensemble (RSD = 40.2%) (Table 4). This means that the required amount of C input across sites had a larger variability when the models were calibrated (see also the range of values of the calibrated parameters across sites in Table 1). This was true for all models, except for AMG and Millennial where the calibration reduced the RSD of the simulated required C input across sites. Roth-C was the least sensitive to calibration, showing similar mean additional C input in both configurations (Table 4).

Table 5 shows the results of the ANOVA for the LME model. We found that both the explanatory variables "model" and "calibration" had a significant effect on the simulated C input ($p < 0.05$), that is, the difference between the C input simulated by the various models, was statistically significant and the difference between the calibrated and non-calibrated configurations was also statistically significant. We observed a non-significant interaction effect between models and calibration. This means that the effect of the calibration on the simulated C input did not depend on the model (Table 5).

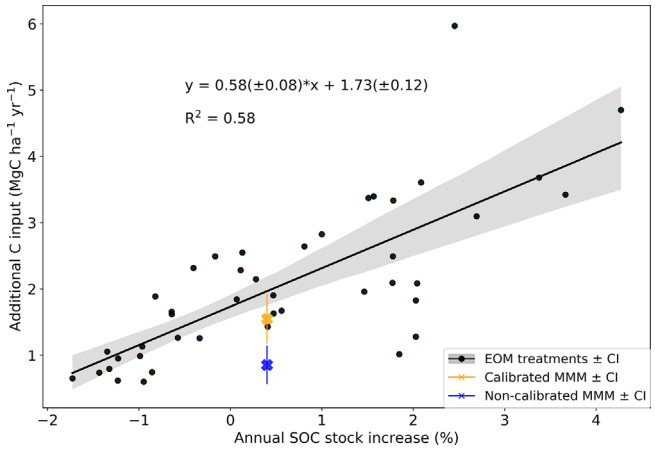

FIGURE 3 Annual soil organic carbon (SOC) stock increase (%) for different levels of additional carbon (C) input in the organic amendment treatment experiments (black spots) and additional C input required to reach the 4‰ SOC increase according to the 1) non-calibrated multi-model median (MMM) (blue cross) and the 2) calibrated MMM (orange cross). Errors are shown as confidence intervals (CI) the regression line between additional C input and SOC stock increase in the exogenous organic matter (EOM) treatments is indicated in the figure ($y = m\,(\pm SD_m)\cdot \mathrm{x} + b \pm (SD_b)$).

TABLE 4 Required percentage change of carbon (C) input to increase soil organic carbon (SOC) stocks by 4‰ per year on average over the period 1980–2010 for the non-calibrated and calibrated models' configurations

Model	Average C input change ± site RSD (%)	
	Non-calibrated	Calibrated
AMG	116.4 ± 94.9	149.2 ± 84.3
Century	29.5 ± 9.8	43.7 ± 40.6
ICBM	32.5 ± 17.9	114.7 ± 89.5
Millennial	154.9 ± 64.2	197.8 ± 57.0
MIMICS	59.6 ± 62.5	122.7 ± 75.4
Roth-C	60.7 ± 14.7	62.3 ± 51.9
Multi-model statistics		
MMM ± median RSD	60.1 ± 40.4	118.7 ± 66.2
Multi-model mean	73.4	115.6
Multi-model SD	49.8	56.5
Multi-model RSD	67.9	48.8

Note: In the table are specified the multi-model median (MMM), the multi-model mean, the standard deviation (SD) from the mean, and the relative standard deviation (RSD).

3.4 | Clusters and variability among models

The heatmap of Figure 5 shows the level of the additional simulated C input to reach the 4‰ for each site and each calibrated model. The dendrogram at the top of the figure clusters the models based on the minimal correlation distance among simulated C input. What can be appreciated from the graph is that there are two main groups of models that behave similarly when calibrated. The first cluster is formed between the AMG and ICBM models. The second cluster incorporates MIMICS, Century, Roth-C and Millennial (the correlation between the first two models being higher than between the others). The results of the stepwise AIC algorithm showed that, among the non-calibrated models, different variables had a significant effect ($p < 0.05$) on the RSD between simulated C input (i.e., initial SOC stocks, MAST and initial C input (C_0^{in})) (Table 6). When the models were calibrated, the RSD among the simulated C input was explained by water-related variables only, that is, MAP and PET (Table 6).

4 | DISCUSSION

4.1 | Evaluation of the multi-model ensemble configurations

The calibration of model parameters improved the simulation of SOC stocks in the control treatments of the 16 LTEs used in this study (Figure 1). In a multi-modelling exercise, Farina et al. (2021) showed that site-specific calibration improved the simulation of SOC stocks in 7 bare-fallow LTEs in Europe. This was true both compared to a non-calibrated and to a multi-site calibration configuration (i.e., where generic parameters are optimised for all sites together). Site-specific calibration accounts for the spatial variability of model parameters across sites. However, in order to avoid overfitting, site-specific calibration requires sufficient repeated SOC measurements for each site, especially for models with many parameters.

In our study, the calibration was also validated against the effect of C input on SOC stocks. In fact, the calibrated multi-model ensemble better reproduced the effect of C input on SOC stocks in the 43 EOM treatments, compared to the non-calibrated configuration (Figure 3). In particular, we found that the MMM additional C input to reach the 4‰ simulated by the calibrated configuration was not significantly different from the observations ($p < 0.05$), whereas the non-calibrated and calibrated ensembles were different from each other at a statistically significant level (Figure 3). On the one hand, the spatial variability (i.e., across sites) was increased by the calibration (i.e., median RSD across sites were 40.4% and 66.2% in the non-calibrated and calibrated ensembles, respectively, Table 4). This is not surprising, since the parameters were

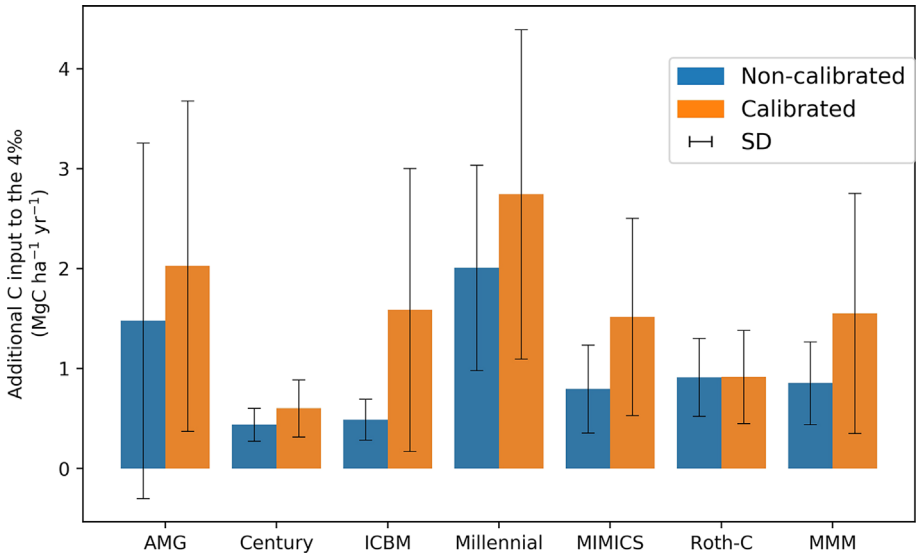

FIGURE 4 Required additional carbon (C) input (±standard deviation, SD) relative to the unfertilised control, to reach a mean annual 4‰ soil organic carbon (SOC) stock increase for 30 years across the 16 sites. The bars represent the different models and multi-model median (MMM). The non-calibrated and calibrated configurations are in blue and orange, respectively. For the MMM, the SD bar represents the median SD across models.

TABLE 5 Effect of "model" and "calibration" on the estimated carbon (C) input to reach the 4p1000 target

Variables	*p-value*
Intercept	<0.0001
Model	<0.0001
Calibration	0.0003
Model · Calibration	0.3476

Note: Results from the ANOVA of the log-transformed linear mixed-effect (LME model), with random effect of the sites.

calibrated independently at each site, whereas in the default configuration model parameters were constant across all sites (see Table 1). On the other hand, the variability of the predicted C input across models was decreased by the calibration (i.e., RSD across models was 67.9% in the non-calibrated and 48.8% in the calibrated configuration), meaning that the re-parametrization decreased the uncertainty of the predictions.

4.2 | Single model performances and MMM

Considering both configurations, AMG had the highest indices of performance (Table 2). However, AMG's comparison with the other models might be partly biased, because when the model was initially developed it was calibrated on several LTEs across France, which include many of the experiments in our database. Hence, we cannot ascertain that its application to other sites outside the European temperate zone would be as straightforward, although the model has already been evaluated on a few sites outside Europe (Andriulo et al., 1999; Saffih-Hdadi & Mary, 2008). AMG was also the only model initialised with observed initial SOC stocks, while the other models were spun-up either analytically or semi-analytically. This has enabled the model to capture the correct level of initial SOC stocks, lending it an advantage in the model comparison. In fact, AMG prescribes the initial fraction of total SOC that is considered stable (Saffih-Hdadi & Mary, 2008). That is, 65% of the total initial SOC stocks for sites with a long history of arable land use (Clivot et al., 2019). Most pool-based models do not prescribe default partitioning in the different SOC pools at the beginning of the experiment. Hence, initialization to allocate the C in their different pools is typically done by running the models with constant or repeating inputs until the C pools reach equilibrium (i.e., spin-up). The amount of C allocated to each pool at equilibrium is a function of the inputs to the model and the parameters. Spin-up assumes that soils are at equilibrium (Luo et al., 2017; Xia et al., 2012), which is often not the case, especially for the agricultural soils with changing management practices considered in this study. Hence, simulations might be started at the wrong initial values (e.g., Wutzler & Reichstein, 2007). In particular, the level of stable or low decomposition rate C pools might influence the SOC stock predictions in the long-term and could partially explain the differences in the required C input predicted by the models. Table 3 shows that the initial level of the low decomposition rate pools depends on how these pools are defined in the models. In particular, Roth-C showed the lowest proportion of C in its inert pool, which is calculated with the equation from Falloon et al. (1998) and considered constant throughout the simulation length. Roth-C and AMG are the only models with a completely inert pool. However, in Roth-C the humified C pool also has a low decomposition rate (i.e., 0.02 year^{-1}, which means that the C takes 50 years to decompose).

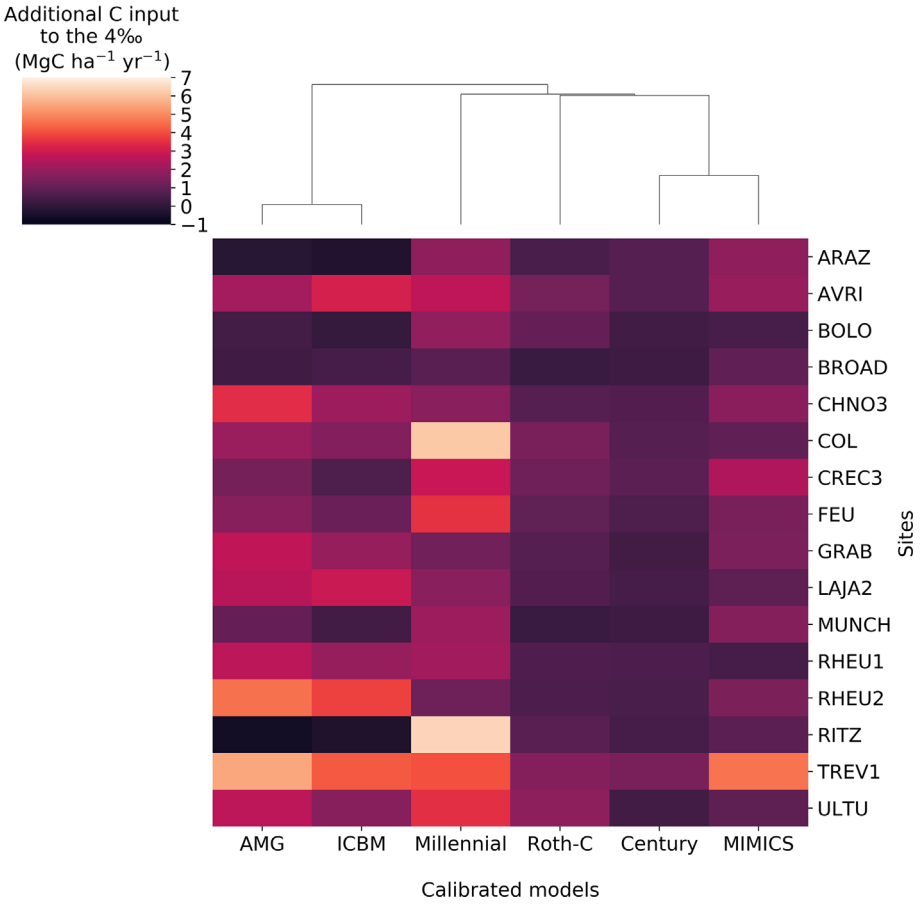

FIGURE 5 Heatmap of the simulated additional carbon (C) input to reach the 4‰, for each calibrated model and each site. Darker cells show lower C input and lighter cells represent higher C input. Dendrograms above the heatmap represent the relationship of similarity among groups of models, calculated as the minimal correlation distance.

TABLE 6 Results of the stepwise Akaike Information Criteria (AIC) model for the non-calibrated (left) and calibrated (right) configurations

Non-calibrated					**Calibrated**				
	Estimate	**Std. error**	***t* value**	***p* value**		**Estimate**	**Std. error**	***t* value**	***p* value**
Intercept	0.37	0.23	1.58	0.14	Intercept	11.64	5.90	1.97	0.07
Initial SOC stocks	0.01*	0.00	5.67	0.00	Initial SOC stocks	−0.04	0.03	−1.78	0.10
MAST	0.08*	0.02	3.90	0.00	MAP	0.01*	0.00	2.31	0.04
C_0^{in}	−0.52*	0.07	−7.61	<0.001	PET	−0.01*	0.00	−2.59	0.02
					Soil C:N	−0.80	0.43	−1.87	0.09
Residual standard error: 0.1 on 12 degrees of freedom (DF)					Residual standard error: 1.5 on 11 DF				
Multiple R^2: 0.88					Multiple R^2: 0.53				
F-statistic: 28.98 on 3 and 12 DF *p*-value <0.001					*F*-statistic: 3.16 on 4 and 11 DF *p*-value = 0.058				

Note: The linear model was originally built with the following variables: Initial soil organic carbon (SOC) stocks, mean annual surface temperature (MAST) and precipitation (MAP), potential evapotranspiration (PET), initial carbon (C) input (C_0^{in}), clay and $CaCO_3$ content, soil C:N ratio, soil pH and N input (N_{in}). This latter was provided as a categorical variable, equal to 1 if the experiment was fed with some N input, and 0 otherwise. The table shows the most significant variables selected by the stepwise algorithm to explain the relative standard deviation (RSD) of simulated C input among models. Specified at the bottom are: The residual standard error, the multiple and adjusted R^2, the *F*-statistic and the *p*-value of the selected AIC model.
*p value < 0.05.

Hence, model comparisons are difficult without considering intrinsic differences across all pools. An alternative initialization method that can be applied is called "relaxation" (Dimassi et al., 2018). This method consists of rescaling the SOC stocks allocated in the different pools through spin-up, using information about the total

observed SOC stocks. Due to the form of our problem (i.e., inverse modelling requiring us to keep the relationship between C input and SOC stock unchanged), this method could not be used. Another approach that could be tested to compare model performances using observed initial total SOC stocks would be to prescribe a priori the initial partitioning of SOC in the different pools of those models that are usually initialised with spin-up. An attempt was made for instance with Roth-C in an Australian catchment (Karunaratne et al., 2014). However, this method has proven complicated and nongeneralizable for models based on conceptual SOC pools (e.g., Sohi et al., 2001), and the scientific community has rather focused on building new modelling frameworks based on measurable pools (e.g., MEMS, Millennial) (Abramoff et al., 2022; Cotrufo et al., 2013). Initializing models with measurable SOC fractions requires additional data, which is one of the main drawbacks of such approaches.

After calibration, the models overall performed well, with an $R^2 \geq 0.92$ and an RMSE <4.6 Mg C ha^{-1}. Other than initialization and parametrization, limitations in model predictions could also be explained by the forcing data used and the processes described in the models. For example, previous land use or current management practices, which are likely to influence the level of SOC stocks at the onset of the experiment, were not prescribed in the models.

AMG not only outperformed all other models in the non-calibrated configuration, but it also performed better than the MMM in both configurations (Table 2). Tebaldi and Knutti (2007) pointed out that, while for a single given simulation the multi-model performance might not be significantly better than the single best model, improvements are more substantial when aggregated performances over many simulations are considered. Of course, as the multi-model ensemble gets larger, the estimates will be more reliable. Farina et al. (2021) suggested that the minimum number of models to obtain reliable results in SOC modelling would be ~10 models for non-calibrated multi-model ensembles, and 3 to 4 models if site-specific calibration is realisable. However, this likely depends on how much the structure varies among the multi-model ensemble.

4.3 | Reaching a 4‰ target

Many recent works have studied the feasibility of the 4‰ target through a modelling perspective. Martin et al. (2021) estimated that a 30% to 40% increase in C input would be needed to reach a 4‰ objective in France, using an inverse Roth-C modelling approach. Bruni et al. (2021) used a similar inverse modelling approach to the Century model and applied it to 14 LTEs across Europe. They estimated that C input should increase by 43% on average, compared to the initial value of the experimental control treatments. These results are similar to our estimates from the calibrated Roth-C and Century models (Table 4). However, they are among the most optimistic ones when compared to the other models (Table 4). Furthermore, Bruni et al. (2021) showed that the level of the predicted C input was much higher under increased soil temperatures. Riggers et al. (2021) used a multi-model ensemble to predict SOC stock increase scenarios under future climate change in German croplands. They estimated an average increase of C input of 213–283% to reach an average 4‰ increase between 2090 and 2099, compared to 2014. Our results seem to be in the middle of the range of existing estimates (i.e., a median increase of 118.7% according to the calibrated ensemble), which refers to the period 1980–2010 under the current climate. Indeed, although the estimate of Riggers et al. (2021) is higher than ours, they estimated this change over a longer period (and under future changes in climate), when SOC stocks in German croplands are expected to decrease at a strong rate because of forecasted increased temperatures (Riggers et al., 2021; Wiesmeier et al., 2016).

Our findings show that the use of one single model to predict the evolution of SOC stocks and its related variables (e.g., the C input) is likely to bias the outputs of the modelling exercise. The present study raises the attention of the soil modelling community to the importance and utility of multi-model simulations and model intercomparison exercises. Multi-modelling approaches are especially necessary when models are used at new sites without previous validation. Besides, multi-modelling has been an established practice in climate projections for decades (Jebeile & Crucifix, 2020; Parker, 2010; Tebaldi & Knutti, 2007), one example being the Coupled Model Intercomparison Project (CMIP), which was created in 1995 and is nowadays the reference framework in which climate models are aggregated to predict future scenarios of climate change (Jebeile & Crucifix, 2020). These ensembles are currently used in the Intergovernmental Panel on Climate Change (IPCC) reports, considered to be the most reliable source of knowledge about climate change.

As for the feasibility of a 118.7% increase in C input, this likely depends on the reference practice against which it is compared. In fact, minerally fertilised crops might already have higher C input compared to unfertilised crops, due to higher nutrient availability that enhances net primary production (Gross & Glaser, 2021), making it harder to increase C input further. Doubling the C input where mineral fertilisers and EOM inputs are

already applied will likely require the implementation of other agricultural practices (e.g., agroforestry systems, cover cropping, improved crop rotations, and crops with high belowground biomass). This is the case for Europe, for example, where croplands are usually minerally fertilised (Eurostat, 2021) and where EOM inputs are already widely applied (Bruni et al., 2022; Foged et al., 2011; Soussana et al., 2019; Zhang et al., 2017). Our results show much lower C input requirements relative to the 10-time increase stated by Berthelin et al. (2022) to reach the 4‰. We argue that all the models used in our exercise do take into account that fresh C inputs have a fast mineralization rate in the short-term, hence it is unlikely that they are underestimating the C input requirements for this reason (Angers et al., 2022; Berthelin et al., 2022). However, our results may be over-optimistic since we did not account for future climate change (e.g., Bruni et al., 2021; Riggers et al., 2021) and because of the uncertainty brought by initialization issues, especially on the actual amount of stable SOC. Also, additional C input needs to be maintained once the SOC stocks reach a new equilibrium, that is, several decades to several hundred years (Berthelin et al., 2022). This requires transitioning to practices that permanently provide additional C input to the soil (e.g., agroforestry systems, cover cropping, and improved crop rotations) in order to maintain the goal in the long-term. In addition to that, soils that currently have higher SOC stocks are likely to require more C input to sustain a 4‰ target, compared to soils with depleted SOC stocks (Figure **S1**), due to the proportional contribution of SOC that is required. This is especially true for soils that are undergoing C losses (e.g., see SOC stock changes in the control plots, Table A2), since the SOC stocks first need to be restabilised before they might increase (Bruni et al., 2022; Soussana et al., 2019). Strategies to increase SOC stocks may take into account past and current SOC trends to set the targets, although for soils with high SOC losses this might result in zero net C gains (Bruni et al., 2022; Soussana et al., 2019). In any case, restoring degraded soils should be the priority in order to reach the objective of land degradation neutrality (LDN) of the United Nations Convention to Combat Desertification (UNCCD) (Soussana et al., 2019), and C farming strategies should aim to preserve SOC-rich soils while restoring depleting stocks. The database that was gathered for our exercise, shows that increasing SOC stocks with additional EOM is possible for soils that are undergoing SOC losses, in the middle to long-term (i.e., 8 to 54 years) (Table A2). However, it is worth noting that the use of EOM does not result in additional C sequestration but rather in locally increased SOC stocks. For example, extracting peats from former peatlands in order to apply them to crops is not itself a climate-relevant C sequestration practices because it only moves C from one place to another, with also possible losses. In our work, EOM treatments were used for methodological investigation purposes only. That is, to evaluate the multi-modelling tool with available agricultural LTEs where SOC stocks were measured after increasing C inputs to the soil. However, we did not quantify the full C cost of acquiring and applying EOM.

4.4 | Variability between calibrated models

The LME model showed that the models significantly affected the prediction of C input to reach the 4‰ target. That is, the prediction of each model was significantly different from the others (Table 5). Furthermore, there was a significant effect of the calibration (Table 5). Although the ensembles' prediction of additional C input was highly variable in both configurations (Figure 4), some models' outputs were correlated to each other once calibrated to fit the stocks (e.g., AMG and ICBM in Figure 5, as opposed to Figure S2). Many factors could be responsible for the creation of such clusters. First of all, similarities in the mathematical structure of the models, such as the number of C pools, the linearity of the system of equations, and the type of kinetics reactions. Other computational differences could have introduced this clustering behaviour. For instance, the spin-up method or the number and choice of parameters calibrated. Finally, the inherent representation of soil processes, that is, the different characterisation of pedo-climatic variables in the model functions, also known as structural uncertainty in ensemble modelling (Tebaldi & Knutti, 2007). Disregarding analogies in the mathematical structures of the models and their technical resolution, we investigated the effect of field variables on the variability of model outputs. We found that, while in the non-calibrated configuration MAST, initial SOC stocks and initial C input explained the divergence between model outputs (i.e., their RSD), only water-related variables (i.e., MAP and PET) had a significant linear effect when models were calibrated (Table 6). This means that the calibration realigned the effect of all those variables that were causing models' outputs to diverge, except for MAP and PET variables, whose effects on the RSD were significant after calibration. These results suggest that, after calibration, the high variability across models was mainly due to the representation (or non-representation) of MAP and PET in the models. The models of the ensemble do not explicitly represent soil water dynamics, but most of them consider water-related variables as forcing inputs. The version of Century that we used accounts for the effect of water on C decomposition through a

polynomial function of soil moisture, which depends on the field capacity and affects the C in the different soil pools (Bruni et al., 2021; Parton et al., 1988). Millennial considers volumetric water content effects on diffusion, matric potential, and oxygen availability, which mainly affects microbial uptake of low molecular weight C (Abramoff et al., 2022). ICBM summarises all environmental conditions into one coefficient (Andrén & Kätterer, 1997), which is calculated using functions that depend on soil moisture and temperature (Fortin et al., 2011; Karlsson et al., 2011). Roth-C and AMG are the only models that explicitly take as input MAP and PET. In AMG, the moisture function depends on MAP and PET and affects the decomposition of C in the active pool (Clivot et al., 2019). In Roth-C, monthly precipitation and PET are used to calculate a topsoil moisture deficit rate modifying factor, which affects all four dynamic C pools (Coleman & Jenkinson, 1996). MIMICS does not explicitly consider water-related variables (Wieder et al., 2015). Our finding suggests that particular attention should be given to the representation of water-related variables (e.g., soil moisture, precipitation and potential evapotranspiration) in soils, to better constrain the underlying processes that cause models to diverge (Martinez-Moyano & Richardson, 2013). Models' uncertainty around the effect of water on the SOC cycle is particularly critical in the context of climate change, where extreme events connected to precipitation are expected to increase (IPCC, 2021). Moyano et al. (2012), for instance, showed a strong effect of soil texture and SOM content on the sensitivity of C decomposition to water conditions. Our findings suggest that SOC models are likely to sense differently the effect of climate change on SOC stocks and that the uncertainty among models will be even larger when future changes in precipitation are considered. Vice versa, the uncertainty in soils' response to precipitation is likely to affect future climate change projections. This underlines the importance of multi-model ensembles, both to account for and to potentially reduce the uncertainty among SOC model predictions.

Despite the effect of water-related processes on predictions, the simulated C inputs were correlated to each other in models with similar structures (Figure 5). In particular, models with simpler structures like AMG and ICBM seemed to behave similarly when calibrated, while models with a higher number of pools clustered together (Figure 5). It is likely that the way models account for C inputs (e.g., their decomposition coefficients, their partitioning within different litter pools, and the number of litter pools itself) also affected the variability among models' outputs and created the "structural clusters" of Figure 5 (e.g., similar litter pool partitioning between Century and MIMICS). If models are correctly parametrised and simulate well the evolution of SOC stocks with time, we would expect them to converge regardless of their different mechanistic structures. However, our results suggest that the choice of the mathematical formalism and processes represented in the models significantly affects the predictions. This is particularly true for inverse modelling predictions of C input changes, where supplementary choices on the litter pool optimization have to be made.

5 | CONCLUSION

We found that the calibrated multi-model ensemble was able to correctly reproduce SOC stock changes at the 16 long-term European cropland experiments. We estimated that annual C input will have to increase by 119% (MMM) compared to the unamended controls to reach a 4‰ objective over 30 years at the experimental sites. Although still very high, we observed that the uncertainty among the different models was reduced when parameters were calibrated. The uncertainty among calibrated models was explained by MAP and PET variables, indicating that the divergence in model estimations of additional C input depended on the choice of the processes represented in the models.

We suggest that the soil modelling community should rely on multi-model ensembles to account for such uncertainty, and that model representation of water effects on SOC stocks should be further constrained in order to improve model predictions. This is particularly important since uncertainties on the effect of water availability on SOC stocks will likely affect climate change projections, due to future changes in the water cycle.

AUTHOR CONTRIBUTIONS

Elisa Bruni: Conceptualization; writing – original draft; methodology; formal analysis. **Claire Chenu:** Conceptualization; funding acquisition; supervision. **Rose Z. Abramoff:** Methodology. **Guido Baldoni:** Data curation. **Dietmar Barkusky:** Data curation. **Hugues Clivot:** Data curation. **Yuanyuan Huang:** Methodology. **Thomas Kätterer:** Data curation. **Dorota Pikuła:** Data curation. **Heide Spiegel:** Data curation. **Iñigo Virto:** Data curation. **Bertrand Guenet:** Supervision; conceptualization; funding acquisition.

ACKNOWLEDGEMENTS

The authors acknowledge Mancomunidad de la Comarca de Pamplona for maintenance and access to Arazuri site data. Research grant RTA2017-00088-C03-01 from the Instituto Nacional de Investigación Agraria y Alimentaria, INIA (Spanish Agency). The authors acknowledge Margaret Glendining, curator of the electronic Rothamsted

Archive (e-RA) for providing the data for the Broadbalk experiment. The Colmar and Feucherolles field experiments form part of the SOERE-PRO (network of long-term experiments dedicated to the study of impacts of organic waste product recycling) certified by ALLENVI (Alliance Nationale de Recherche pour l'Environnement) and integrated as a service of the "Investment for future" infrastructure AnaEE-France, overseen by the French National Research Agency (ANR-11-INBS-0001).

FUNDING INFORMATION

This work benefited from the French state aid managed by the ANR under the "Investissements d'avenir" programme with the reference ANR-16-CONV-0003 (CLAND project). EB, RZA and BG are supported by the European Union's Horizon 2020 research and innovation program under grant agreement No 101000289 (Holisoils project). RZA was also supported by the United States Department of Energy, Office of Science, and Office of Biological and Environmental Research. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the United States Department of Energy under contract DE-AC05-00OR22725.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

DATA AVAILABILITY STATEMENT

Model codes are available at https://zenodo.org/record/6835336#.YtBjNezMI0o (Bruni, 2022). Data are available under the requirement of the owners.

ORCID

Elisa Bruni https://orcid.org/0000-0001-8074-0516
Claire Chenu https://orcid.org/0000-0001-9054-0489
Iñigo Virto https://orcid.org/0000-0002-7682-4570

REFERENCES

Abramoff, R. Z., Guenet, B., Zhang, H., Georgiou, K., Xu, X., Viscarra Rossel, R. A., Yuan, W., & Ciais, P. (2022). Improved global-scale predictions of soil carbon stocks with millennial version 2. *Soil Biology and Biochemistry*, *164*, 108466. https://doi.org/10.1016/j.soilbio.2021.108466

Andrén, O., & Kätterer, T. (1997). ICBM: The introductory carbon balance model for exploration of soil carbon balances. *Ecological Applications*, *7*, 1226–1236. https://doi.org/10.1890/1051-0761(1997)007[1226:ITICBM]2.0.CO;2

Andriulo, A., Mary, B., & Guerif, J. (1999). Modelling soil carbon dynamics with various cropping sequences on the rolling pampas. *Agronomie*, *19*, 365–377. https://doi.org/10.1051/agro:19990504

Angers, D., Arrouays, D., Cardinael, R., Chenu, C., Corbeels, M., Demenois, J., Farrell, M., Martin, M., Minasny, B., Recous, S., & Six, J. (2022). A well-established fact: Rapid mineralization of organic inputs is an important factor for soil carbon sequestration. *European Journal of Soil Science*, *73*(3), 1–5. https://doi.org/10.1111/ejss.13242

Ballabio, C., Panagos, P., & Monatanarella, L. (2016). Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, *261*, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006

Berthelin, J., Laba, M., Lemaire, G., Powlson, D., Tessier, D., Wander, M., & Baveye, P. C. (2022). Soil carbon sequestration for climate change mitigation: Mineralization kinetics of organic inputs as an overlooked limitation. *European Journal of Soil Science*, *73*(1), 1351–1754. https://doi.org/10.1111/ejss.13221

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. https://doi.org/10.1016/j.agee.2006.05.013

Bouthier, A., Duparque, A., Mary, B., Sagot, S., Trochard, R., Levert, M., Houot, S., Damay, N., Denoroy, P., Dinh, J.-L., Blin, B., & Ganteil, F. (2014). Adaptation et mise en oeuvre du modèle de calcul de bilan humique à long terme AMG dans une large gamme de systèmes de grandes cultures et de polyculture-élevage. *Innovations Agronomiques*, *34*, 125–139.

Bruni, E. (2022). elisabruni/SoilC: Multimodelling_4p1000 (Version v1.0.0), Zenodo, https://zenodo.org/record/6835336#.YtBjNezMI0o

Bruni, E., Guenet, B., Clivot, H., Kätterer, T., Martin, M., Virto, I., & Chenu, C. (2022). Defining quantitative targets for topsoil organic carbon stock increase in European croplands: Case studies with exogenous organic matter inputs. *Frontiers in Environmental Science*, *10*, 824724. https://doi.org/10.3389/fenvs.2022.824724

Bruni, E., Guenet, B., Huang, Y., Clivot, H., Virto, I., Farina, R., Kätterer, T., Ciais, P., Martin, M., & Chenu, C. (2021). Additional carbon inputs to reach a 4 per 1000 objective in Europe: Feasibility and projected impacts of climate change based on century simulations of long-term arable experiments. *Biogeosciences*, *18*, 3981–4004. https://doi.org/10.5194/bg-18-3981-2021

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited-memory algorithm for bound-constrained optimization. *SIAM Journal on Scientific Computing*, *16*, 1190–1208. https://doi.org/10.1137/0916069

Chabbi, A., Lehmann, J., Ciais, P., Loescher, H. W., Cotrufo, M. F., Don, A., SanClements, M., Schipper, L., Six, J., Smith, P., & Rumpel, C. (2017). Aligning agriculture and climate policy. *Nature Climate Change*, *7*, 307–309. https://doi.org/10.1038/nclimate3286

Clivot, H., Mouny, J.-C., Duparque, A., Dinh, J.-L., Denoroy, P., Houot, S., Vertès, F., Trochard, R., Bouthier, A., Sagot, S., & Mary, B. (2019). Modeling soil organic carbon evolution in long-term arable experiments with AMG model. *Environmental Modelling & Software*, *118*, 99–113. https://doi.org/10.1016/j.envsoft.2019.04.004

Coleman, K., & Jenkinson, D. S. (1996). RothC-26.3—A model for the turnover of carbon in soil. In D. S. Powlson, P. Smith, & J. U. Smith (Eds.), *Evaluation of soil organic matter models* (pp. 237–246). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-61094-3_17

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K., & Paul, E. (2013). The microbial efficiency-matrix stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form

stable soil organic matter? *Global Change Biology*, *19*, 988–995. https://doi.org/10.1111/gcb.12113

Dimassi, B., Guenet, B., Saby, N. P. A., Munoz, F., Bardy, M., Millet, F., & Martin, M. P. (2018). The impacts of CENTURY model initialization scenarios on soil organic carbon dynamics simulation in French long-term experiments. *Geoderma*, *311*, 25–36. https://doi.org/10.1016/j.geoderma.2017.09.038

European Commission. (2020). *Communication from the commission to the European Parliament, the council, the European economic and social committee and the committee of the regions a farm to fork strategy for a fair, healthy and environmentally-friendly food system*. European Commission.

European Commission. (2021a). *Commission staff working document sustainable carbon cycles—carbon farming accompanying the Communication from the Commission to the European Parliament and the Council sustainable carbon cycles*. European Commission.

European Commission. (2021b). *Communication from the commission to the European Parliament and the Council sustainable carbon cycles*. European Commission.

Eurostat. (2021). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agri-environmental_indicator_-_mineral_fertiliser_consumption#Analysis_at_EU_level last access: November 17, 2021.

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*(8–9), 1207–1211.

Farina, R., Sándor, R., Abdalla, M., Álvaro-Fuentes, J., Bechini, L., Bolinder, M. A., Brilli, L., Chenu, C., Clivot, H., De Antoni Migliorati, M., Di Bene, C., Dorich, C. D., Ehrhardt, F., Ferchaud, F., Fitton, N., Francaviglia, R., Franko, U., Giltrap, D. L., Grant, B. B., … Bellocchi, G. (2021). Ensemble modelling, uncertainty and robust predictions of organic carbon in long-term bare-fallow soils. *Global Change Biology*, *27*, 904–928. https://doi.org/10.1111/gcb.15441

Foged, H. L., Flotats, X., Bommatí, A., & Palatsi, J. (2011). Inventory of manure processing activities in Europe. *V to the European Commission, Directorate-General Environment concerning Manure Processing Activities in Europe - Project reference* (Report No. ENV.B.1/ETU/2010/0007).

Fortin, J. G., Bolinder, M. A., Anctil, F., Kätterer, T., Andrén, O., & Parent, L. E. (2011). Effects of climatic data low-pass filtering on the ICBM temperature- and moisture-based soil biological activity factors in a cool and humid climate. *Ecological Modelling*, *222*, 3050–3060. https://doi.org/10.1016/j.ecolmodel.2011.06.011

Fu, Z., Liu, G., & Guo, L. (2019). Sequential quadratic programming method for nonlinear least squares estimation and its application. *Mathematical Problems in Engineering*, *1–8*, 2019–2018. https://doi.org/10.1155/2019/3087949

Gross, A., & Glaser, B. (2021). Meta-analysis on how manure application changes soil organic carbon storage. *Scientific Reports*, *11*, 5516. https://doi.org/10.1038/s41598-021-82739-7

Huang, Y., Lu, X., Shi, Z., Lawrence, D., Koven, C. D., Xia, J., Du, Z., Kluzek, E., & Luo, Y. (2018). Matrix approach to land carbon cycle modeling: A case study with the community land model. *Global Change Biology*, *24*, 1394–1404. https://doi.org/10.1111/gcb.13948

IPCC. (2007). Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Core Writing Team. In R. K. Pachauri, & A. Reisinger (Eds.), (pp. 104). IPCC.

IPCC. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), Cambridge University Press, In press, https://doi.org/10.1017/9781009157896

Jebeile, J., & Crucifix, M. (2020). Multi-model ensembles in climate science: Mathematical structures and expert judgements. *Studies in History and Philosophy of Science Part A*, *83*, 44–52. https://doi.org/10.1016/j.shpsa.2020.03.001

Jenkinson, D. S. (1990). The turnover of organic carbon and nitrogen in soil. Philosophical Transactions of the Royal Society of London. *Series B: Biological Sciences*, *329*(1255), 361–368.

Karlsson, T., Delin, S., Kätterer, T., Berglund, K., & Andrén, O. (2011). Simulating site-specific nitrogen mineralization dynamics in a Swedish arable field. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, *61*, 333–344. https://doi.org/10.1080/09064710.2010.490537

Karunaratne, S. B., Bishop, T. F. A., Baldock, J. A., & Odeh, I. O. A. (2014). Catchment scale mapping of measurable soil organic carbon fractions. *Geoderma*, *219–220*, 14–23. https://doi.org/10.1016/j.geoderma.2013.12.005

Kätterer, T., Bolinder, M. A., Andrén, O., Kirchmann, H., & Menichetti, L. (2011). Roots contribute more to refractory soil organic matter than above-ground crop residues, as revealed by a long-term field experiment. *Agriculture, Ecosystems & Environment*, *141*, 184–192. https://doi.org/10.1016/j.agee.2011.02.029

Kaur, R., Kumar, S., & Gurung, H. P. (2002). A pedo-transfer function (PTF) for estimating soil bulk density from basic soil data and its comparison with existing PTFs. *Soil Research*, *40*(5), 847–858. https://doi.org/10.1071/SR01023

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system: DVGM FOR COUPLED CLIMATE STUDIES. *Global Biogeochemical Cycles*, *19*, 1–33. https://doi.org/10.1029/2003GB002199

Kurzemann, F. R., Plieger, U., Probst, M., Spiegel, H., Sandén, T., Ros, M., & Insam, H. (2020). Long-term fertilization affects soil microbiota, improves yield and benefits soil. *Agronomy*, *10*, 1664. https://doi.org/10.3390/agronomy10111664

Lal, R. (2008). Carbon sequestration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 815–830. https://doi.org/10.1098/rstb.2007.2185

Levavasseur, F., Mary, B., Christensen, B. T., Duparque, A., Ferchaud, F., Kätterer, T., Lagrange, H., Montenach, D., Resseguier, C., & Houot, S. (2020). The simple AMG model accurately simulates organic carbon storage in soils after repeated application of exogenous organic matter. *Nutrient Cycling in Agroecosystems*, *117*, 215–229. https://doi.org/10.1007/s10705-020-10065-x

Luo, Y., Shi, Z., Lu, X., Xia, J., Liang, J., Jiang, J., Wang, Y., Smith, M. J., Jiang, L., Ahlström, A., Chen, B., Hararuk, O.,

Hastings, A., Hoffman, F., Medlyn, B., Niu, S., Rasmussen, M., Todd-Brown, K., & Wang, Y.-P. (2017). Transient dynamics of terrestrial carbon storage: Mathematical foundation and its applications. *Biogeosciences*, *14*, 145–161. https://doi.org/10.5194/bg-14-145-2017

Martin, M. P., Dimassi, B., Román Dobarco, M., Guenet, B., Arrouays, D., Angers, D. A., Blache, F., Huard, F., Soussana, J., & Pellerin, S. (2021). Feasibility of the 4 per 1000 aspirational target for soil carbon: A case study for France. *Global Change Biology*, *27*, 2458–2477. https://doi.org/10.1111/gcb.15547

Martinez-Moyano, I. J., & Richardson, G. P. (2013). Best practices in system dynamics modeling: I. J. Martinez-Moyano and G. P. Richardson: Best practices in SD modeling. *System Dynamics Review*, *29*, 102–123. https://doi.org/10.1002/sdr.1495

Martyniuk, S., Pikuła, D., & Kozieł, M. (2019). Soil properties and productivity in two long-term crop rotations differing with respect to organic matter management on an albic Luvisol. *Scientific Reports*, *9*, 1878. https://doi.org/10.1038/s41598-018-37087-4

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B. S., Field, D. J., Gimona, A., Hedley, C. B., Hong, S. Y., Mandal, B., Marchant, B. P., Martin, M., McConkey, B. G., Mulder, V. L., ... Winowiecki, L. (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59–86. https://doi.org/10.1016/j.geoderma.2017.01.002

Mirschel, W., Wenkel, K.-O., Wegehenkel, M., Kersebaum, K. C., Schindler, U., & Hecker, J.-M. (2007). Müncheberg field trial data set for agro-ecosystem model validation. In K. C. Kersebaum, J.-M. Hecker, W. Mirschel, & M. Wegehenkel (Eds.), *Modelling water and nutrient dynamics in soil–crop systems* (pp. 219–243). Springer. https://doi.org/10.1007/978-1-4020-4479-3_16

Moyano, F. E., Vasilyeva, N., Bouckaert, L., Cook, F., Craine, J., Curiel Yuste, J., Don, A., Epron, D., Formanek, P., Franzluebbers, A., Ilstedt, U., Kätterer, T., Orchard, V., Reichstein, M., Rey, A., Ruamps, L., Subke, J.-A., Thomsen, I. K., & Chenu, C. (2012). The moisture response of soil heterotrophic respiration: Interaction with soil properties. *Biogeosciences*, *9*, 1173–1182. https://doi.org/10.5194/bg-9-1173-2012

Müller, A. C., Nowozin, S., & Lampert, C. H. (2012). Information theoretic clustering using minimum spanning trees. In *Information theoretic clustering using minimum spanning trees* (Lecture Notes in Computer Science). Springer.

Palosuo, T., Foereid, B., Svensson, M., Shurpali, N., Lehtonen, A., Herbst, M., Linkosalo, T., Ortiz, C., Rampazzo Todorovic, G., Marcinkonis, S., Li, C., & Jandl, R. (2012). A multi-model comparison of soil carbon assessment of a coniferous forest stand. *Environmental Modelling & Software*, *35*, 38–49. https://doi.org/10.1016/j.envsoft.2012.02.004

Parker, W. S. (2010). Whose probabilities? *Predicting Climate Change with Ensembles of Models, Philosophy of Science*, *77*, 985–997. https://doi.org/10.1086/656815

Parshotam, A. (1996). The Rothamsted soil-carbon turnover model — Discrete to continuous form. *Ecological Modelling*, *86*, 283–289. https://doi.org/10.1016/0304-3800(95)00065-8

Parton, W. J., Stewart, J. W. B., Cole, C. V., & Dynamics of C, N. (1988). P and S in grassland soils: A model. *Biogeochemistry*, *5*, 109–131. https://doi.org/10.1007/BF02180320

Powlson, D. S., Bhogal, A., Chambers, B. J., Coleman, K., Macdonald, A. J., Goulding, K. W. T., & Whitmore, A. P. (2012). The potential to increase soil carbon stocks through reduced tillage or organic material additions in England and Wales: A case study. *Agriculture, Ecosystems and Environment*, *146*, 23–33. https://doi.org/10.1016/j.agee.2011.10.004

Riggers, C., Poeplau, C., Don, A., Frühauf, C., & Dechow, R. (2021). How much carbon input is required to preserve or increase projected soil organic carbon stocks in German croplands under climate change? *Plant and Soil*, *460*, 417–433. https://doi.org/10.1007/s11104-020-04806-8

Rumpel, C., Amiraslani, F., Chenu, C., Garcia Cardenas, M., Kaonga, M., Koutika, L.-S., Ladha, J., Madari, B., Shirato, Y., Smith, P., Soudi, B., Soussana, J.-F., Whitehead, D., & Wollenberg, E. (2020). The 4p1000 initiative: Opportunities, limitations and challenges for implementing soil organic carbon sequestration as a sustainable development strategy. *Ambio*, *49*, 350–360. https://doi.org/10.1007/s13280-019-01165-2

Saffih-Hdadi, K., & Mary, B. (2008). Modeling consequences of straw residues export on soil organic carbon. *Soil Biology and Biochemistry*, *40*, 594–607. https://doi.org/10.1016/j.soilbio.2007.08.022

Simoes-Mota, A., Poch, R. M., Enrique, A., Orcaray, L., & Virto, I. (2021). Soil quality assessment after 25 years of sewage sludge vs. Mineral Fertilization in a Calcareous Soil. *Land*, *10*, 727. https://doi.org/10.3390/land10070727

Soetaert, K., & Herman, P. M. J. (2009). *A practical guide to ecological modelling: Using R as a simulation platform* (p. 372). Springer.

Sohi, S. P., Mahieu, N., Arah, J. R. M., Powlson, D. S., Madari, B., & Gaunt, J. L. (2001). A procedure for isolating soil organic matter fractions suitable for modeling. *Soil Science Society of America Journal*, *65*, 1121–1128. https://doi.org/10.2136/sssaj2001.6541121x

Soussana, J.-F., Lutfalla, S., Ehrhardt, F., Rosenstock, T., Lamanna, C., Havlík, P., Richards, M., Wollenberg, E. (. L.)., Chotte, J.-L., Torquebiau, E., Ciais, P., Smith, P., & Lal, R. (2019). Matching policy and science: Rationale for the '4 per 1000—Soils for food security and climate' initiative. *Soil and Tillage Research*, *188*, 3–15. https://doi.org/10.1016/j.still.2017.12.002

Sulman, B. N., Moore, J. A. M., Abramoff, R., Averill, C., Kivlin, S., Georgiou, K., Sridhar, B., Hartman, M. D., Wang, G., Wieder, W. R., Bradford, M. A., Luo, Y., Mayes, M. A., Morrison, E., Riley, W. J., Salazar, A., Schimel, J. P., Tang, J., & Classen, A. T. (2018). Multiple models and experiments underscore large uncertainty in soil carbon dynamics. *Biogeochemistry*, *141*, 109–123. https://doi.org/10.1007/s10533-018-0509-z

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*, 7183–7192. https://doi.org/10.1029/2000JD900719

Tebaldi, C., & Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A*, *365*, 2053–2075. https://doi.org/10.1098/rsta.2007.2076

Triberti, L., Nastri, A., Giordani, G., Comellini, F., Baldoni, G., & Toderi, G. (2008). Can mineral and organic fertilization help sequestrate carbon dioxide in cropland? *European Journal of Agronomy*, *29*, 13–20. https://doi.org/10.1016/j.eja.2008.01.009

van Groenigen, J. W., van Kessel, C., Hungate, B. A., Oenema, O., Powlson, D. S., & van Groenigen, K. J. (2017). Sequestering soil organic carbon: A nitrogen dilemma. *Environmental Science & Technology*, *51*, 4738–4739. https://doi.org/10.1021/acs.est.7b01427

van't Hoff, M. J. H. (1884). *Etudes de dynamique chimique*. Frederik Muller & C°.

Veerman, C., Pinto Correia, T., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B., Frison, E. A., Grabd, A., Filchew, L. H., Kriaučiūnienė, Z., Pogrzeba, M., Soussana, J.-F., Olmo, C. V., & Wittkowski, R. (2020). Caring for soil is caring for life: ensure 75% of soils are healthy by 2030 for food, people, nature and climate : report of the Mission board for Soil health and food, Publications Office. *European Commission, Directorate-General for Research and Innovation*. https://data.europa.eu/doi/10.2777/821504

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., Taylor, P. G., & Bonan, G. B. (2015). Representing life in the earth system with soil microbial functional traits in the MIMICS model. *Geoscientific Model Development*, *8*, 1789–1808. https://doi.org/10.5194/gmd-8-1789-2015

Wiesmeier, M., Poeplau, C., Sierra, C. A., Maier, H., Frühauf, C., Hübner, R., Kühnel, A., Spörlein, P., Geuß, U., Hangen, E., Schilling, B., von Lützow, M., & Kögel-Knabner, I. (2016). Projected loss of soil organic carbon in temperate agricultural soils in the 21st century: Effects of climate change and carbon input trends. *Scientific Reports*, *6*, 32525. https://doi.org/10.1038/srep32525

Wutzler, T., & Reichstein, M. (2007). Soils apart from equilibrium–consequences for soil carbon balance modelling. *Biogeosciences*, *4*(1), 125–136.

Xia, J. Y., Luo, Y. Q., Wang, Y.-P., Weng, E. S., & Hararuk, O. (2012). A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, *5*, 1259–1271. https://doi.org/10.5194/gmd-5-1259-2012

Zhang, B., Tian, H., Lu, C., Dangal, S. R. S., Yang, J., & Pan, S. (2017). Global manure nitrogen production and application in cropland during 1860–2014: A 5 arcmin gridded global dataset for earth system modeling. *Earth System Science Data*, *9*, 667–678. https://doi.org/10.5194/essd-9-667-2017

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bruni, E., Chenu, C., Abramoff, R. Z., Baldoni, G., Barkusky, D., Clivot, H., Huang, Y., Kätterer, T., Pikuła, D., Spiegel, H., Virto, I., & Guenet, B. (2022). Multi-modelling predictions show high uncertainty of required carbon input changes to reach a 4‰ target. *European Journal of Soil Science*, *73*(6), e13330. https://doi.org/10.1111/ejss.13330

APPENDIX A

The parameters calibrated in the models are linked to SOC decomposition (Table 1). Below, we detail the different functions in which they appear.

A.1. | Century

In Century, the C input is partitioned into the metabolic and structural litters according to the metabolic to a structural ratio ($M:S_{\text{ratio}}$):

$$\frac{d\text{LIT}_{S\,\text{AG,BG}}}{dt} = f_{S\,\text{AG,BG}} \cdot I_{\text{AG,BG}} - F_{\text{LS AG,BG}} \quad \text{(A1)}$$

$$\frac{d\text{LIT}_{M\,\text{AG,BG}}}{dt} = f_{M\,\text{AG,BG}} \cdot I_{\text{AG,BG}} - F_{\text{LM AG,BG}} \quad \text{(A2)}$$

where AG = aboveground and BG = belowground, $\frac{f_{M\,\text{AG,BG}}}{f_{S\,\text{AG,BG}}} = M:S_{\text{ratio}}$ is the metabolic:structural ratio of the litter inputs, $\text{LIT}_S(t)$ and $\text{LIT}_M(t)$ are the state variables of the structural and metabolic litter pools, respectively (g C m^{-2}), I is the C input (g C m^{-2} d^{-1}), F_{LS} is the outflux from the structural litter pool (g C m^{-2} d^{-1}), and F_{LM} is the outflux from the metabolic litter pool (g C m^{-2} d^{-1}):

$$F_{\text{LS AG,BG}} = \text{LIT}_{S\,\text{AG,BG}}(t) \cdot k_{LS} \cdot f(T) \cdot f(W) \cdot e^{-3 \cdot \text{lignin}_{S\,\text{AG,BG}}} \quad \text{(A3)}$$

$$F_{\text{LM AG,BG}} = \text{LIT}_{M\,\text{AG,BG}}(t) \cdot k_{\text{LM}} \cdot f(T) \cdot f(W) \quad \text{(A4)}$$

where $k_{LS} = 0.01$ and $k_{LM} = 0.041$ (d^{-1}) are the turnover rates of the structural and metabolic litter pools, respectively, $f(T)$ and $f(W)$ are the temperature and moisture response functions, $\text{lignin}_{S\,\text{AG}} = 0.76$ and $\text{lignin}_{S\,\text{BG}} = 0.72$ are the lignin fractions in the aboveground and belowground structural litter pools, respectively.

And where the temperature response function is defined as:

$$f(T) = Q_{10}^{\frac{(T(t)-T_{\text{ref}})}{10}} \quad \text{(A5)}$$

where Q_{10} is the temperature coefficient of the Van't Hoff equation (van't Hoff, 1884), T_{ref} is the reference temperature (°C), and $T(t)$ is temperature (°C).

A.2. | Roth-C

In Roth-C, the temperature response function takes the form:

$$f(T) = \frac{47.91}{1 + e^{\left(\frac{106.06}{T(t)+T_{\text{param}}}\right)}} \quad \text{(A6)}$$

where $T(t)$ is temperature (°C) and T_{param} is a parameter.

A.3. | ICBM

In ICBM, the ordinary differential equations of the young and old SOC pools are:

$$\frac{dY}{dt} = i - k_1 \cdot r \cdot Y(t) \quad \text{(A7)}$$

$$\frac{dO}{dt} = h \cdot k_1 \cdot r \cdot Y(t) - k_2 \cdot r \cdot O(t) \quad \text{(A8)}$$

where $Y(t)$ is the state variable of the young SOC pool (kg C m^{-2}), $O(t)$ is the state variable of the old SOC pool (kg C m^{-2}), i is the C input (kg C m^{-2} $year^{-1}$), k_1 is the potential mineralization rate affecting both the young and the old SOC pools ($year^{-1}$), k_2 is the potential mineralization rate affecting the old SOC pool ($year^{-1}$), r is the environmental parameter, and h is the "humification coefficient", that is, the fraction of the annual outflux from the young to the old pool. The environmental parameter r was calculated using the temperature and moisture response functions described in Fortin et al. (2011) and Karlsson et al. (2011) and normalised against the Ultuna experiment.

A.4. | AMG

In AMG, the mineralization rate constant k of the active pool ($year^{-1}$) depends on:

$$k = k_0 \cdot f(T) \cdot f(W) \cdot f(A) \cdot f(\text{CaCO}_3) \quad \text{(A9)}$$

where k_0 is the potential mineralization rate of the active SOC pool ($year^{-1}$), $f(T)$ is the temperature response function, $f(W)$ is the water response function, and $f(A)$ and $f(CaCO_3)$ are functions describing the effect of clay and $CaCO_3$ soil content on SOC mineralization.

A.5. | MIMICS

In MIMICS, the ordinary differential equations of the structural and metabolic litter pools are:

$$\frac{d\text{LIT}_S}{dt} = (1 - f_{\text{MET}}) \cdot I - F_{\text{LS}} \quad \text{(A11)}$$

$$\frac{d\mathrm{LIT}_M}{dt} = f_{\mathrm{MET}} \cdot I - F_{\mathrm{LM}} \tag{A12}$$

$$F_{\mathrm{LS}} = \mathrm{MIC}_r(t) \cdot V_{\max} \cdot \frac{\mathrm{LIT}_S(t)}{K_m + \mathrm{LIT}_S(t)} + \mathrm{MIC}_k(t) \cdot V_{\max} \cdot \frac{\mathrm{LIT}_S(t)}{K_m + \mathrm{LIT}_S(t)} \tag{A13}$$

$$F_{\mathrm{LM}} = \mathrm{MIC}_r(t) \cdot V_{\max} \cdot \frac{\mathrm{LIT}_M(t)}{K_m + \mathrm{LIT}_{M(t)}} + \mathrm{MIC}_k(t) \cdot V_{\max} \cdot \frac{\mathrm{LIT}_M(t)}{K_m + \mathrm{LIT}_{M(t)}} \tag{A14}$$

where $\mathrm{LIT}_s(t)$ and $\mathrm{LIT}_M(t)$ are the state variables of the structural and metabolic litter pools, respectively (mg C cm^{-3}), $\mathrm{MIC}_r(t)$ and $\mathrm{MIC}_k(t)$ are the state variables of the copiotrophic and oligotrophic microbial biomass pools, respectively (mg C cm^{-3}), I is the C input (mg C cm^{-3} d^{-1}), f_{MET} is the fraction of the C input that goes to the metabolic litter pool, and F_{LS} and F_{LM} are the outfluxes from the two litter pools (mg C cm^{-3} d^{-1}), and with temperature-sensitive maximum reaction velocities $V_{\max}$ (mg C [mg MIC] $-^{1}$ d^{-1}) and half-saturation constants K_m (mg C cm^{-3}) of the Michaelis–Menten kinetics:

$$V_{\max} = e^{V_{\mathrm{slope}} \cdot T(t) + V_{\mathrm{int}}} \cdot a_v \cdot V_{\mathrm{mod}} \tag{A15}$$

$$K_m = e^{K_{\mathrm{slope}} \cdot T(t) + K_{\mathrm{int}}} \cdot a_k \cdot K_{\mathrm{mod}} \tag{A16}$$

where V_{slope} (ln (mg C [mg MIC] $-^{1}$ d-1)°C^{-1}) and K_{slope} (ln (mg C cm^{-3})°C-1) are regression coefficients (ln (mg C [mg MIC] $-^{1}$ d^{-1})°C^{-1}), V_{int} (ln (mg C [mg MIC] $-^{1}$ d^{-1})) and K_{int} (ln (mg C cm^{-3})) are regression intercepts, a_v and a_k are tuning coefficients, V_{mod} and K_{mod} are coefficients modifying $V_{\max}$ and K_m for fluxes into the microbial pools, and $T(t)$ is temperature.

A.6. | Millennial

In Millennial V2, the decomposition of particulate organic matter (POM) into low molecular weight carbon (LMWC) is:

$$F_{pl} = V_{pl} S_{w,D} P \frac{B}{K_{pl} + B} \tag{A17}$$

where V_{pl} is the maximum rate of POM decomposition to LMWC (d^{-1}):

$$V_{pl} = \alpha_{pl} e^{-Ea_{pl}/(R(T(t)+273.15))} \tag{A18}$$

$V_{lb}$is the maximum uptake rate of LMWC (d^{-1}):

$$V_{lb} = \alpha_{lb} e^{-Ea_{lb}/(R(T(t)+273.15))} \tag{A19}$$

And where $S_{w,D}$ is the diffusion limitation of substrates, P is the POM pool, B is the microbial biomass pool, K_{pl} is the half-saturation constant of POM decomposition to LMWC (g C m^{-2}), α_{pl}(g C m^{-2} (g C m^{-2})$^{-1}$ d^{-1}) and α_{lb} (g C m^{-2} (g C m^{-2})$^{-1}$ d^{-1}) are the pre-exponential constants of V_{pl} and V_{lb}, respectively, Ea_{pl} (J mol^{-1}) and Ea_{lb} (J mol^{-1}) are the activation energies of V_{pl} and V_{lb}, respectively, R is the gas constant (J K^{-1} mol^{-1}), $T(t)$ is temperature (°C).

TABLE A1 Information on the soil of the control treatments at the onset of the 16 long-term experiments

Site	Sampling depth cm	Bulk density g cm^{-3}	Carbonate $gCaCO_3$ kg^{-1}	Clay %	Initial SOC stock Mg C ha^{-1}	Soil C:N	pH	Reference paper
Champ Noël 3 (CHNO3)	0–30	1.35	0	15	40.57	8.96	6.30	Clivot et al. (2019)
Colmar (COL)	0–28	1.30	129.57	23	54.33	10.52	8.33	Levavasseur et al. (2020)
Crécom 3 PRO (CREC3)	0–30	1.36	0	15	62.00	10.17	6.15	Clivot et al. (2019)
Feucherolles (FEU)	0–29	1.32	0	16	39.78	9.89	6.73	Levavasseur et al. (2020)
La Jaillière 2 PRO (LAJA2)	0–25	1.37	0	21	32.42	9.01	6.80	Levavasseur et al. (2020)
Le Rheu 1 (RHEU1)	0–30	1.27	0	16	36.23	10.05	5.85	Clivot et al. (2019)
Le Rheu 2 (RHEU2)	0–30	1.28	0	14	36.53	8.22	6.05	Clivot et al. (2019)
Arazuri (ARAZ)	0–30	1.67	160.00	28	55.39	6.44	8.60	Simoes-Mota et al. (2021)
Ultuna (ULTU)	0–20	1.40	0	36	42.91	8.82	6.23	Kätterer et al. (2011)
Broadbalk (BROAD)	0–23	1.25	20.00	25	24.84	8.95	7.25	Powlson et al. (2012)
Trévarez (TREV1)	0–30	1.48	0	19	115.33	9.49	6.01	Clivot et al. (2019)
Avrillé (AVRI)	0–30	1.40	0	18	46.20	8.91	6.59	Clivot et al. (2019)
Bologna (BOLO)	0–40	1.16	0	28	25.41	7.00	6.90	Triberti et al. (2008)
Grabów (GRAB)	0–25	1.40	76.66	5	31.08	10.76	5.87	Martyniuk et al. (2019)
Müncheberg (MUNCH)	0–25	1.47	0	5	19.66	10.00	5.95	Mirschel et al. (2007)
Ritzlhof (RITZ)	0–25	1.1	0.03	23	28.88	9.42	6.88	Kurzemann et al. (2020)
Mean		1.35	24.14	19.19	43.22	9.61	6.66	
Median		1.35	0	19.00	39.78	9.16	6.59	
Minimum		1.10	0	5.00	19.66	6.44	5.85	
Maximum		1.67	160	36.00	115.33	10.76	8.60	

TABLE A2 Agronomic information on the experimental sites

Site	Experiment length	Treatment name	Rotations*	C input from plants Mg ha^{-1} year^{-1}	C input from EOM Mg ha^{-1} year^{-1}	Treatment type	Initial SOC stocks Mg ha^{-1}	SOC stock change** % year^{-1}
CHNO3	19	Min	sM	1.29	0.00	Reference + N***	40.57	−0.92
		LP	sM	1.49	0.79	Pig Manure	43.30	−0.89
COL	14	T0	wW/Mg/sB/S	2.79	0.00	Reference	54.33	−0.78
		BIO1	wW/Mg/sB/S	3.93	1.01	Biowaste	54.78	0.15
		BOUE1	wW/Mg/sB/S	3.96	0.49	Sewage sludge	54.33	−0.61
		CFB1	wW/Mg/sB/S	4.04	1.07	Cow manure	51.42	−0.01
		DVB1	wW/Mg/sB/S	4.00	1.08	Green manure + sewage sludge	53.69	0.18
		FB1	wW/Mg/sB/S	3.93	1.36	Cow manure	53.69	−0.01
CREC3	23	Min	wW/sM	1.84	0.00	Reference + N***	62.00	−0.06
		FB2	wW/sM	1.92	1.82	Cow manure	61.27	0.49
		FV	wW/sM	1.96	0.47	Poultry manure	64.07	−1.46
FEU	16	T0	wW/Mg	2.22	0.00	Reference	39.78	−0.66
		BIO1	wW/Mg	3.44	2.21	Biowaste	41.23	3.60
		DVB1	wW/Mg	3.45	2.45	Green manure + sewage sludge	40.52	3.69
		FB1	wW/Mg	3.55	2.28	Cow Manure	42.99	1.36
		OMR1	wW/Mg	3.45	2.11	Household waste	39.68	1.72
LAJA2	15	Min	sM/wW	1.59	0.00	Reference + N***	32.42	−1.43
		CFB	sM/wW	1.25	1.14	Cow manure	31.79	−0.88
		CFP	sM/wW	1.21	1.00	Pig manure	31.36	−1.09
		CFV	sM/wW	1.31	0.94	Poultry manure	31.36	−1.60
		FB	sM/wW	1.29	1.44	Cow manure	31.01	−0.64
		FP	sM/wW	1.27	1.07	Pig manure	33.05	−1.03
		FV	sM/wW	1.40	0.93	Poultry manure	33.40	−1.59
RHEU1	16	Min	sM	1.31	0.00	Reference + N***	36.23	−1.51
		CFB1	sM	1.31	1.06	Cow manure	36.23	−1.21
RHEU1	16	T0	sM	1.03	0.00	Reference	36.53	−1.72
		CFP1	sM	1.20	0.78	Pig manure	36.53	−1.28
		FP	sM	1.30	1.63	Pig manure	36.53	−0.74

(Continues)

TABLE A2 (Continued)

Site	Experiment length	Treatment name	Rotations*	C input from plants Mg ha^{-1} year^{-1}	C input from EOM Mg ha^{-1} year^{-1}	Treatment type	Initial SOC stocks Mg ha^{-1}	SOC stock change** % year^{-1}
ARAZ	17	D0_N0	W/B/Sf/O	1.34	0.00	Reference	55.39	1.00
		D1_F1	W/B/Sf/O	1.92	2.79	Sewage sludge	62.17	0.40
		D1_F2	W/B/Sf/O	1.87	1.30	Sewage sludge	63.19	1.22
		D1_F3	W/B/Sf/O	1.95	0.68	Sewage sludge	63.19	1.22
		D2_F1	W/B/Sf/O	1.75	5.56	Sewage sludge	74.02	0.22
		D2_F2	W/B/Sf/O	1.84	2.60	Sewage sludge	57.53	2.32
		D2_F3	W/B/Sf/O	1.98	1.32	Sewage sludge	63.18	0.93
ULTU	53	P0_B	O/sT/Mu/sB/Fb/OsR/FR/M	1.03	0.00	Reference	41.72	−0.52
		S_F	O/sT/Mu/sB/Fb/OsR/FR/M	1.10	1.77	Straw	42.28	−0.09
		GM_H	O/sT/Mu/sB/Fb/OsR/FR/M	1.82	1.76	Green manure	40.6	0.11
		PEAT_I	O/sT/Mu/sB/Fb/OsR/FR/M	1.14	1.97	Peat	41.16	2.17
		FYM_J	O/sT/Mu/sB/Fb/OsR/FR/M	1.76	1.91	Farmyard manure	41.72	0.69
		SD_L	O/sT/Mu/sB/Fb/OsR/FR/M	0.82	1.84	Sawdust	40.88	0.56
		SS_O	O/sT/Mu/sB/Fb/OsR/FR/M	2.59	1.84	Sewage sludge	43.12	1.36
BROAD	48	3_Nill	wW	0.36	0.00	Reference	24.84	−0.09
		19_Cast	wW	0.95	0.43	Castor meal	32.74	0.42
		22_FYM	wW	2.07	2.99	Farmyard manure	61.49	0.38
TREV1	23	Min	RG/Mg/wW/sM	1.94	0.00	Reference + N***	115.33	−0.66
		FB	RG/Mg/wW/sM	2.04	1.52	Cow manure	110.67	−0.39
		FP	RG/Mg/wW/sM	2.02	1.18	Pig manure	109.50	−0.18
AVRI	8	T1TR	wW/sM	1.62	0.00	Reference	46.20	−1.18
		T2TR	wW/sM	1.71	1.58	Cow manure	47.13	−0.76
BOLO	29	T0	M/wW	1.96	0.00	Reference + N*** (b)	25.41	0.41
		CM	M/wW	2.21	2.24	Cow manure	28.63	1.18
		CS	M/wW	2.15	2.64	Cow slurry	26.70	0.93
GRAB	33	CP	P/B/wW/Mu/sM	2.10	0.00	Reference + N*** (c)	31.08	−0.88
		T1	P/B/wW/Mu/sM	2.39	0.98	Farmyard manure	33.18	−0.54
MUNCH	54	CP	gM/wR/P/sW/S/sB/wW/ Oflax/fPea/sM/	0.47	0.00	Reference	19.66	−0.29
		FYM2		0.50	1.40	Farmyard manure	20.48	0.18

TABLE A2 (Continued)

Site	Experiment length	Treatment name	Rotations*	C input from plants Mg ha^{-1} year^{-1}	C input from EOM Mg ha^{-1} year^{-1}	Treatment type	Initial SOC stocks Mg ha^{-1}	SOC stock change** % year^{-1}
			gM/wR/P/sW/S/sB/wW/ Oflax/fPea/sM/					
RITZ	28	CP	M/sW/wB/Pea/wW/wB	1.52	0.00	Reference	28.88	0.59
		BW	M/sW/wB/Pea/wW/wB	1.88	1.73	Biowaste compost	28.88	1.39

*Rotations legend: M = maize / wM = winter maize / sM = silage maize / Mg = maize grain/gM = green maize /W = wheat / wW = winter wheat / sW = spring wheat /B = barley / wB = winter barley / sB = spring barley/ O = Oats / P = potato/ S = sugar beet /Sf = sunflower/sT = Swedish turnip/Mu = mustard/ Fb = fodder beet/ OsR = oilseed rape /FR = fodder rape /RG = ray grass /wR = winter rye/Oflax = oil flax/fPea = fodder peas/Pea = peas.

**Calculated by approximating the SOC stock evolution with a linear regression of the form: SOC = mt + b, where t = the number of the year, m is the slope and b is the intercept. The relative change was estimated as SOC change (% year^{-1}) = $m/b \cdot 100$.

***(a) Optimal amount of N inputs in both the reference and the treatments; (b) in Bologna, data represent the mean of several treatments with different inorganic fertilisation rates (see Triberti et al. (2008)); (c) in Grabów N was applied as ammonium nitrate (34% N), phosphorus (P) as triple superphosphate (45% P_2O_5) and potassium (K) as potassium chloride (60% K_2O).

TABLE A3 Mean annual climate variables extracted from the GSWP3 climate dataset (i.e., mean annual precipitation (MAP)) or simulated by the ORCHIDEE model at each site (i.e., mean annual potential evapotranspiration (PET), mean annual surface temperature (MAST), and mean annual soil water content (SWC))

Site	Coordinates	Experiment duration	PET mm	MAP mm	MAST °C	SWC $kg_{H2O}\ m^{-2}_{soil}$
CHNO3	48.09°N, 1.78°W	1990–2008	1107.5	818.1	12.2	21.6
COL	48.11°N, 7.38°E	2000–2013	866.6	1126.7	9.7	24.6
CREC3	48.32°N, 3.16°W	1986–2008	1131.3	1150.1	11.8	22.9
FEU	48.88°N, 1.96°E	1998–2013	1049.9	707.3	11.9	21.2
LAJA2	47.44°N, 0.98°W	1995–2009	1314.7	794.7	12.8	20.5
RHEU1	48.09°N, 1.78°W	1994–2009	1106.6	841.2	12.3	21.8
RHEU2	48.09°N, 1.78°W	1994–2009	1106.6	841.2	12.3	21.8
ARAZ	42.81°N, 1.72°W	2002–2018	1416.4	866.0	12.6	20.3
ULTU	59.82°N, 17.65°E	1956–2008	824.5	613.4	5.7	22.6
BROAD	51.81°N, 0.37°W	1968–2015	872.0	665.6	10.3	21.5
TREV1	48.15°N, 3.76°W	1986–2008	1139.5	1314.5	11.9	23.4
AVRI	47.50°N, 0.60°W	1984–1991	1170.1	680.7	12.0	20.0
BOLO	44.55°N, 11.35°E	1972–2000	1474.3	890.9	11.3	19.4
GRAB	51.35°N, 21.66°E	1980–2012	974.8	638.1	8.5	13.5
MUNCH	14.11°N, 52.51°E	1963–2016	938.3	639.9	9.2	20.9
RITZ	48.18°N, 14.25°E	1991–2018	675.5	1010.5	9.1	25.4
Mean			1073.0	849.9	10.9	21.3
Median			1106.6	829.6	11.9	21.5
Minimum			675.5	613.4	5.7	13.5
Maximum			1474.3	1314.5	12.8	25.4

Note: The table also specifies the geographical coordinates of the experiments and the years of the simulations.