HAL open science

A robust initialization method for accurate soil organic carbon simulations

Eva Kanari, Lauric Cecillon, François Baudin, Hugues Clivot, Fabien Ferchaud, Sabine Houot, Florent Levavasseur, Bruno Mary, Laure Soucémarianadin, Claire Chenu, et al.

To cite this version:

Eva Kanari, Lauric Cecillon, François Baudin, Hugues Clivot, Fabien Ferchaud, et al.. A robust initialization method for accurate soil organic carbon simulations. Biogeosciences, 2022, 19 (2), pp.375-387. 10.5194/bg-19-375-2022 . hal-03546659

HAL Id: hal-03546659

https://hal.inrae.fr/hal-03546659

Submitted on 1 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Biogeosciences, 19, 375–387, 2022
https://doi.org/10.5194/bg-19-375-2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Biogeosciences
Open Access
EGU

Research article

A robust initialization method for accurate soil organic carbon simulations

Eva Kanari[1,2], Lauric Cécillon[1,3], François Baudin[2], Hugues Clivot[4], Fabien Ferchaud[5], Sabine Houot[6], Florent Levavasseur[6], Bruno Mary[5], Laure Soucémarianadin[7], Claire Chenu[6], and Pierre Barré[1]

[1]Laboratoire de Géologie, École normale supérieure, CNRS, Université PSL, IPSL, 75005 Paris, France
[2]ISTeP, UMR 7193, Sorbonne Université, CNRS, 75005 Paris, France
[3]Normandie Université, UNIROUEN, INRAE, ECODIV, 76821 Rouen, France
[4]Université de Reims Champagne Ardenne, INRAE, FARE, UMR A 614, 51097 Reims, France
[5]BioEcoAgro Joint Research Unit, INRAE, Université de Liège, Université de Lille, Université Picardie Jules Verne, 02000 Barenton-Bugny, France
[6]UMR ECOSYS, INRAE, AgroParisTech, Université Paris-Saclay, 78850 Thiverval-Grignon, France
[7]ACTA – les instituts techniques agricoles, 75012 Paris, France

Correspondence: Eva Kanari (kanari@geologie.ens.fr), Lauric Cécillon (lauric.cecillon@inrae.fr) and Pierre Barré (barre@biotite.ens.fr)

Received: 16 September 2021 – Discussion started: 23 September 2021
Revised: 25 November 2021 – Accepted: 2 December 2021 – Published: 24 January 2022

Abstract. Changes in soil organic carbon (SOC) stocks are a major source of uncertainty for the evolution of atmospheric CO_2 concentration during the 21st century. They are usually simulated by models dividing SOC into conceptual pools with contrasted turnover times. The lack of reliable methods to initialize these models, by correctly distributing soil carbon amongst their kinetic pools, strongly limits the accuracy of their simulations. Here, we demonstrate that $PARTY_{SOC}$, a machine-learning model based on Rock-Eval® thermal analysis, optimally partitions the active- and stable-SOC pools of AMG, a simple and well-validated SOC dynamics model, accounting for effects of soil management history. Furthermore, we found that initializing the SOC pool sizes of AMG using machine learning strongly improves its accuracy when reproducing the observed SOC dynamics in nine independent French long-term agricultural experiments. Our results indicate that multi-compartmental models of SOC dynamics combined with a robust initialization can simulate observed SOC stock changes with excellent precision. We recommend exploring their potential before a new generation of models of greater complexity becomes operational. The approach proposed here can be easily implemented on soil monitoring networks, paving the way towards precise predictions of SOC stock changes over the next decades.

1 Introduction

Soil organic carbon (SOC) plays an important role in sustaining soil functions and associated soil ecosystem services worldwide (IPCC, 2019). It is the largest terrestrial organic carbon reservoir, with the upper 2 m of soil storing 2400 Pg C, 3 times more carbon than the atmosphere (Jobbagy and Jackson, 2000). A mere 4 per 1000 annual decrease in SOC stocks (ca. 10 Pg C yr^{-1}) may double the global annual anthropogenic CO_2 emissions, while an equivalent increase may compensate them (Balesdent and Arrouays, 1999). This is the concept behind the 4 per 1000 initiative (Rumpel et al., 2018) that aims at increasing SOC stocks to fight global warming while ensuring food security, two Sustainable Development Goals of the United Nations (UN General Assembly, 2015). This initiative and other political headway have put the question of managing SOC stocks and assessing the global SOC sequestration potential at the top of political and scientific agendas (Vermeulen et al., 2019; FAO, 2019; Amelung et al., 2020). Despite this particular attention, the prediction of SOC stock changes remains very uncertain, which makes soils a major source of uncertainty for the evolution of atmospheric CO_2 concentration (Todd-Brown et al., 2014; He et al., 2016; Shi et al., 2018).

Published by Copernicus Publications on behalf of the European Geosciences Union.

Models of SOC dynamics can predict future SOC stock evolution by simulating carbon transfer into the soil mostly through plant organic matter inputs and microbial SOC mineralization resulting in a CO_2 flux from the soil to the atmosphere. They can have structures of various complexities, reflecting our mechanistic understanding of SOC dynamics (Shi et al., 2018; Luo et al., 2016). However, most models dedicated to prediction, including those used in Earth system models, have a simple structure dividing SOC into conceptual pools with contrasted turnover times (Manzoni and Porporato, 2009; He et al., 2016; Todd-Brown et al., 2014). These multi-compartmental models of SOC dynamics are the best option we currently have to foster science-based SOC preservation and sequestration actions, given the strong uncertainty of more complex models (Cécillon, 2021a; Dangal et al., 2021; Lee et al., 2020; Shi et al., 2018; Crowther et al., 2019). Predictions of SOC stock evolution provided by such simple models are very sensitive to the initial distribution of SOC amongst the different kinetic pools (Luo et al., 2016; Smith and Falloon, 2000; Clivot et al., 2019). This makes the question of partitioning of SOC kinetic pools a priority for improving the accuracy of multi-compartmental SOC dynamics models (Luo et al., 2016; Taghizadeh-Toosi et al., 2020).

The most commonly used method to initialize the size of SOC kinetic pools is to run spin-up simulations until a steady-state equilibrium for SOC is reached, eventually matching the initial SOC stock measurement (Wutzler and Reichstein, 2007; Taghizadeh-Toosi et al., 2020). However, this method has two well-known limitations. First, climatic SOC input and land-use or land-cover data extending over long time periods required by this approach are highly uncertain. Second, assuming steady-state equilibrium for SOC at the onset of model simulations is often unrealistic. This is due to the history of the simulated sites that often includes disturbances (e.g. fire), as well as previous changes in climate, land use, and soil management that prevent SOC pools with slow turnover times from being at equilibrium (Wutzler and Reichstein, 2007; Herbst et al., 2018; Oberholzer et al., 2014; Poeplau et al., 2011; Clivot et al., 2019). Alternative initialization procedures are needed to address these issues (Wutzler and Reichstein, 2007; Bruun and Jensen, 2002; Taghizadeh-Toosi et al., 2020).

In some models of SOC dynamics, like the AMG model (Clivot et al., 2019), a default initial SOC pool size distribution is prescribed according to basic information on land-use history (i.e. long-term cropland vs. long-term grassland; Clivot et al., 2019). This approach does not take into account the effect of recent changes in land use or historical soil management practices on SOC pool distribution. To better reflect the effect of the frequent state of non-equilibrium of SOC on its partitioning into conceptual kinetic pools, another approach has been proposed, relating results from SOC fractionation methods with SOC kinetic pool sizes (e.g. Zimmermann et al., 2007a, or Skjemstad et al., 2004, for the RothC – Rothamsted carbon – model; Dangal et al., 2021, for the DAYCENT – Daily Century – model). However, this approach also suffers from important drawbacks. First, SOC fractionation procedures are tedious and cannot be implemented on large-scale studies, though this problem may be solved by using soil infrared spectroscopy or environmental variables and machine learning (Zimmermann et al., 2007b; Viscarra Rossel et al., 2019; Sanderman et al., 2021; Cotrufo et al., 2019; Lugato et al., 2021; Baldock et al., 2013; Barthès et al., 2008; Lee et al., 2020; Dangal et al., 2021). Second, their reproducibility is questionable (Poeplau et al., 2013, 2018), and third, their use of initializing model SOC pool sizes has never been properly validated. A proper validation would require showing that (1) the size of measured SOC fractions matches the one of model kinetic pools and that (2) simulations of SOC dynamics are more accurate using this initialization strategy, compared to default simulations (on independent validation sites while other model parameters remain constant). Reasonably good correspondence between measured or soil-spectroscopy-estimated SOC fractions and modelled SOC conceptual pools has been reported in a number of studies, though with some notable discrepancies (Leifeld et al., 2009b; Herbst et al., 2018; Zimmermann et al., 2007a; Dangal et al., 2021). Conversely, the studies that attempted to initialize model SOC pool sizes using a SOC fractionation scheme generally reported no improvement in the accuracy of simulations of SOC dynamics compared to a default or a spin-up initialization approach (Leifeld et al., 2009a; Nemo et al., 2016; Cagnarini et al., 2019). Only two studies showed that an initialization based on an SOC fractionation scheme yielded more accurate simulations of observed SOC dynamics, but this was at the cost of modifying the decomposition rate of SOC kinetic pools (Skjemstad et al., 2004; Luo et al., 2014).

An alternative approach using Rock-Eval® thermal analysis has recently been proposed – under the name of the $PARTY_{SOC}$ model – to estimate SOC kinetic pool sizes (Cécillon et al., 2018, 2021). $PARTY_{SOC}$ is a machine-learning model trained on Rock-Eval® data of soil samples from long-term experiments (LTEs) where the size of the centennially stable-SOC fraction can be estimated (e.g. sites including a bare fallow treatment). $PARTY_{SOC}$ incorporates recent key elements of the new understanding of SOC dynamics (Dignac et al., 2017), showing that the centennially stable-SOC fraction has specific chemical and energetic characteristics that are measurable quickly (ca. 1 h per sample) and at a reasonable cost (less than USD 60) using Rock-Eval®; it is thermally stable (i.e. high activation energy), and it is depleted in hydrogen (Barré et al., 2016; Hemingway et al., 2019; Gregorich et al., 2015; Poeplau et al., 2019; Chassé et al., 2021).

In this study, we tested if the $PARTY_{SOC}$ machine-learning model, built on a totally independent dataset from northwestern Europe, could be used to initialize the distribution of SOC pools of the simple AMG model (Clivot et al., 2019)

and improve the accuracy of its simulations. The default version of AMG is currently the most accurate model for reproducing the observed SOC stock dynamics in diverse French agricultural LTEs at the pluri-decadal scale (Martin et al., 2019). The efficient use of this model at sites covering an important pedological and climatic variability (including oceanic, continental, and tropical climate) provides further support to its robustness (Levavasseur et al., 2020; Farina et al., 2021; Saffih-Hdadi and Mary, 2008). In this model, SOC is simply divided into two pools, the "stable SOC (C_S)" that is considered inert at the timescale of the simulation and the "active SOC (C_A)" that has a mean turnover time of a few decades. A recent study (Clivot et al., 2019) determined that the optimal initial proportion of stable SOC (C_S / C_0) can deviate from the model's default value (0.65 in croplands) so that a more precise initialization of the C_S / C_0 proportion would significantly improve AMG simulations of SOC dynamics. Here, we hypothesized that the SOC pool partitioning as determined by the $PARTY_{SOC}$ machine-learning model (Cécillon et al., 2021, 2018) would be close to the mathematically optimal one for the AMG model, therefore, improving the accuracy of its SOC dynamics simulations compared to default initialization. We tested our hypothesis on 32 treatments from 9 independent French agricultural LTEs (experiment duration from 12 to 41 years with a median of 21 years) in which the AMG-optimal SOC pool partitioning could be determined by ex post optimization and for which topsoil samples collected at the onset of the experiment were available (Table 1). These LTEs were croplands established in different pedoclimates that have experienced contrasted soil management practices and land-use histories. All available initial topsoil samples were analysed with Rock-Eval®, and the results were used in the European version of the $PARTY_{SOC}$ model, $PARTY_{SOC}v2.0_{EU}$ (Cécillon et al., 2021), to compute the centennially stable-SOC proportion of each topsoil sample.

2 Materials and methods

2.1 Experimental sites

This work was conducted on nine French agricultural LTEs (Supplement Fig. S1); s 7 LTEs including 29 treatments were selected from the dataset presented in Clivot et al. (2019), from sites with available initial topsoil samples. Two additional LTEs (Colmar and Feucherolles) including a total of three treatments were obtained from the dataset published in Levavasseur et al. (2020), selecting control treatments without organic amendments and with available initial topsoil samples. Basic site and topsoil characteristics are reported in Table 1 and Supplement Table S1. Information necessary to run AMG simulations on a total of 32 treatments (initial soil physicochemical properties, detailed information on management practices, and observed climatic data during all experiments) were obtained from Clivot et al. (2019) for the 29 treatments of the 7 sites and from Levavasseur et al. (2020) for the 3 treatments of the sites of Colmar and Feucherolles.

2.2 Archive soil samples from experimental sites

Our final soil sample set included 181 topsoil samples. At each site the soil was sampled to include the whole ploughing depth (Table 1). At all sites, except Boigneville where the soil was sampled in five sublayers, the ploughing layer was sampled as one homogeneous layer. Of the final samples, 71 were from starting dates of the 9 LTEs; 24 were from LTEs intermediate dates; and 86 were from LTEs final dates. All samples were air-dried or dried at 40 °C, sieved to <2 mm, and finely ground to <250 µm using a ball mill (Retsch, Germany).

2.3 Rock-Eval® analysis of archive soil samples

All soil samples were analysed using a Rock-Eval 6® Turbo apparatus (Vinci Technologies). The samples were first pyrolysed in an inert N_2 atmosphere, then oxidized under ambient air (O_2). The heating routine applied during pyrolysis was as described in Disnar et al. (2003), starting with a 3 min isotherm at 200 °C followed by a heating ramp of 30 °C min^{-1} up to 650 °C. For the oxidation step, a 1 min isotherm was kept at 300 °C and was directly followed by a heating ramp of 20 °C min^{-1} until 850 °C was reached, followed by a 5 min isotherm at 850 °C (Baudin et al., 2015; adapted from Behar et al., 2001).

Based on 5 generated Rock-Eval® thermograms, 18 parameters were calculated for each sample and were then used as predictors by the random forest model. These include total organic carbon (TOC; in g C per kg soil) – the amount of organic C released during the analysis as a proportion of sample weight; pyrolysed organic carbon (PC; in g C per kg soil) – the sum of C released as HC, CO, and CO_2 during the pyrolysis step; the ratio of PC to TOC (PC/TOC); the S2 peak area (g C per kg soil) – the hydrocarbon gas released within the range of the pyrolysis temperature ramp; the ratio of S2 to PC (S2/PC); the PseudoS1 peak area (g C per kg soil) – the sum of C released as HC, CO, and CO_2 during the first 200 s of pyrolysis (after Khedim et al., 2021); the hydrogen index (HI; in mg HC g TOC^{-1}) – the amount of hydrocarbons released as a ratio of TOC; and the ratio of HI to the oxygen index (HI/OI_{RE6}) – where OI_{RE6} is calculated as the amount of oxygen released as CO and CO_2 gases normalized to TOC. Finally, various temperature parameters ($T70_{HC_PYR}$, $T90_{HC_PYR}$, $T30_{CO2_PYR}$, $T50_{CO2_PYR}$, $T70_{CO2_PYR}$, $T90_{CO2_PYR}$, $T70_{CO_OX}$, $T50_{CO2_OX}$, $T70_{CO2_OX}$, and $T90_{CO2_OX}$; in °C) are included in the predictor set. They describe evolution steps, namely at which temperature a specific amount (e.g. 30 %, 50 %, 70 %, or 90 %) of a given gas was released according to each thermogram (Cécillon

Table 1. Main information about the nine French agricultural long-term experiments used in this study. All sites had been croplands for at least 20 years before the onset of all experiments. Additional site information including climatic variability amongst sites and long-term history of land cover is provided in Supplement Table S1. WRB: World Reference Base.

	Auzeville	Boigneville	Colmar	Doazit	Feucherolles	Grignon-Folleville	Kerbernez	Mant	Tartas
Soil type (WRB 2014)	Luvisol	Haplic Luvisol	Calcaric Cambisol	Luvisol	Gleyic Luvisol	Luvisol	Cambisol	Dystric Luvisol	Luvic Arenosol
LTE onset	1968	1970	2000	1967	1998	1958	1978	1975	1972
Simulated period	1975–2010	1970–2011	2000–2018	1977–1989	1998–2019	1989–2008	1978–2005	1975–1992	1976–1997
Number of treatments	4	12	1	2	2	2	5	2	2
Sampling dates (number of samples)	1975 (4), 2010 (8)	1970 (29), 1998 (10), 2017 (32)	2000 (4), 2018 (6)	1977 (4), 1989 (4)	1998 (8), 2013 (8), 2018(8)	1989 (8), 2008 (8)	1978 (6), 1991 (6), 2005 (12)	1975 (4), 1992 (4)	1976 (4), 1997 (4)
Crop rotation	Annual crop rotation	Annual crop rotation	Annual crop rotation	Maize monoculture	Annual crop rotation	Annual crop rotation	Silage maize monoculture (KERB_C incl. ryegrass)	Maize monoculture	Maize monoculture
Considered depth	30	29	28	25	28	30	25	28	28
Initial SOC stock ($Mg\,C\,ha^{-1}$)	34.68	42.40	45.20	26.35	43.80	55.85	81.98	38.75	45.25
Reference	Colomb et al. (2007)	Dimassi et al. (2014)	Obriot (2016)	Lubet et al. (1993)	Noirot-Cosson et al. (2016)	Barré et al. (2008)	Vertès et al. (2007)	Messiga et al. (2010)	Morel et al. (2014)

et al., 2018). It is important to note that no pre-treatment of $CaCO_3$-containing samples was necessary before Rock-Eval® analysis. The slow pyrolysis and oxidation steps of the Rock-Eval® method allow for distinguishing carbon of organic and mineral form, since the latter is released above a given temperature. For the calculation of all of the above parameters, only the part of each thermogram corresponding to organic carbon was taken into account. For this purpose, upper temperature integration limits for Rock-Eval® temperature parameters were set to 560 °C for the CO and CO_2 pyrolysis thermograms and to 611 °C for the CO_2 oxidation thermograms (Cécillon et al., 2018; Supplement Fig. S2). R scripts used for computing Rock-Eval® parameters are available on the Zenodo platform (Cécillon, 2021b).

2.4 The $PARTY_{SOC}$ machine-learning model

The most up-to-date European version of this model, calibrated on soils from north-western Europe, used in this study, is described in detail in Cécillon et al. (2021). This model uses the 18 above-mentioned Rock-Eval® thermal analysis parameters as predictors and estimates the centennially stable-SOC proportion in a topsoil sample. The model consists of a trained non-parametric machine-learning algorithm, using the random forest approach to estimate centennially stable-SOC proportions in unknown topsoils from centred and scaled Rock-Eval® parameters. In this study the obtained centennially stable-SOC proportion of each topsoil sample was converted to centennially stable-SOC content by multiplying the predicted proportion by the total SOC content. The $PARTY_{SOC}v2.0_{EU}$ model, available on Zenodo (Cécillon, 2021b), was used without any adaptation.

2.5 The AMG model of soil organic carbon dynamics

The AMG model (Andriulo et al., 1999) was developed based on the two-compartment SOC model proposed by Hénin and Dupuis (1945). It is characterized by a simple structure consisting of three carbon pools: fresh organic matter and two SOC fractions, an active and a stable pool (Supplement Fig. S3). The model allows for the transfer of carbon from the fresh organic matter pool either to the atmosphere through microbial mineralization or into the active pool. Organic carbon from the active pool is also subject to mineralization, forming a second direct flux of CO_2 from the soil into the atmosphere. Soil organic matter (SOM) decomposition follows first-order kinetics with a rate defined by the coefficient of mineralization k (yr^{-1}), controlled by climatic conditions and soil characteristics. The h coefficient controls the yield of crop residue transformation into active carbon and depends on the type of fresh organic matter. No carbon exchange with the stable-SOC pool is possible, since it is considered inert and remains unchanged over the simulation period (here from 12 to 41 years; see Table 1). Considering the stable-SOC pool as mathematically inert at this timescale is in line with consistent observations of a significant pluridecadal persistent SOC fraction in long-term bare fallows and C_3–C_4 vegetation change chronosequences (Barré et al., 2010; Balesdent et al., 2018).

The AMG model can be mathematically described by two simple equations (Clivot et al., 2019):

$$QC = QC_S + QC_A, \quad (1)$$

$$\frac{dQC_A}{dt} = \sum_i m_i h_i - k \cdot QC_A, \quad (2)$$

where QC is the total SOC stock ($Mg\,C\,ha^{-1}$), QC_S is the stable-SOC stock ($Mg\,C\,ha^{-1}$) defined as a fraction of the initial SOC stock QC_0 (see Sect. 2.6) constant for a specific treatment, QC_A is the active-SOC stock ($Mg\,C\,ha^{-1}$), t is the time in years, m_i is the annual C input from organic residue i ($Mg\,C\,ha^{-1}\,yr^{-1}$), h is a coefficient representing the fraction of C inputs which is incorporated in SOM after 1 year related to the type of organic residue, and k is the mineralization rate constant associated with the active-C pool (yr^{-1}).

The AMG parameters (h and k) have been determined by experimental results (Clivot et al., 2019). This approach differs from most multi-compartmental SOC dynamics models for which decay rates of slower pools were calibrated indirectly, assuming an equilibrium state for SOC (Wutzler and Reichstein, 2007). The simple structure of the AMG model and the experimental determination of its decomposition rates make it less susceptible to the problem of equifinality compared to other multi-compartmental models of SOC dynamics (Clivot et al., 2019; Luo et al., 2016). Furthermore, AMG has been validated with $\delta^{13}C$ tracer data of long-term alternative sequences of C_4 and C_3 crops (Mary et al., 2020).

The version of AMG used in this study was AMGv2, described in detail in Clivot et al. (2019). Input data necessary to run simulations of SOC stocks with AMG include crop type, annual crop yields, and information regarding management of crop residues. These are used to compute annual aboveground and belowground C inputs from plants, here according to the method proposed by Bolinder et al. (2007) and adapted by Clivot et al. (2019). The coefficient of mineralization k (yr^{-1}) is calculated according to soil characteristics (clay and carbonate contents, pH, and C : N ratio) and climatic conditions (mean annual temperature, precipitation, and potential evapotranspiration; Clivot et al., 2019).

2.6 Soil organic carbon pool partitioning in the AMG model

2.6.1 Default C_S / C_0 initialization

Two default values can be used for initialization of the SOC pool distribution in AMG, depending on land-use history before the onset of simulations. The initial proportion of C_S / C_0 equals 0.65 for sites with a long-term arable land-use history. Former long-term grassland sites are expected to have lower C_S / C_0, and the value of 0.40 was used in previous studies (Saffih-Hdadi and Mary, 2008; Clivot et al., 2019). Since all sites used in this study had been under arable land for at least 20 years before the onset of the experiment, a default value of 0.65 was used.

2.6.2 $PARTY_{SOC}$-based initialization of C_S / C_0

The $PARTY_{SOC}$-based initialization of C_S / C_0 was derived from data obtained with Rock-Eval® analysis of initial topsoil samples from each LTE. Here, C_S / C_0 was estimated using the following simple four-step procedure: first, topsoil samples from the LTE's onset were analysed with Rock-Eval®, and the 18 thermal parameters described in Sect. 2.3 were calculated for each sample. Second, the thermal parameters were used as input for the $PARTY_{SOC}$ machine-learning model described in Sect. 2.4 which was run for this sample set resulting in a sample-specific prediction of the centennially stable-SOC proportion. Third, the obtained values were averaged per LTE. Fourth, the site mean of the centennially stable-SOC proportion was used (as C_S / C_0) to initialize simulations of SOC stocks for the various treatments of every site (the site standard deviation is reported in Fig. 1 and in Supplement Table S2). Supported by the evident common land-use history shared by the multiple treatments of each site before the onset of simulations and because the SOC stocks and centennially stable-SOC contents were very homogeneous amongst each site, we also performed simulations of 17 treatments for which soil samples from the onset of the LTE were not available. In these cases, we considered that the C_S / C_0 of the treatment was equal to the mean value of the respective site (Supplement Table S1 and S2).

2.6.3 Ex post optimization of C_S / C_0

Following a least-squares optimization approach, the best fit of the AMG model on observed SOC stocks time series was obtained, and the optimal initial SOC pool partitioning (C_S / C_0) was estimated accordingly for each site (Clivot et al., 2019). In sites with C_3–C_4 vegetation change chronosequences where $\delta^{13}C$ long-term monitoring data were available, the model was adapted to simultaneously match the observed evolution of C, C_3, and C_4 stocks (Clivot et al., 2019) for a given treatment.

2.7 Calculation of the centennially stable-SOC content

The content of the centennially stable-SOC pool of each LTE at initial, intermediate, and final dates was estimated through multiplication of the $PARTY_{SOC}$ estimates of the proportion of the centennially stable SOC at a given date by the corresponding total SOC content previously determined using elemental analysis (Clivot et al., 2019; Levavasseur et al., 2020). For example, for the onset of an LTE where $t = 0$, $C_S = C_S / C_0 \cdot C_0$, where C_S is the stable-SOC content (g C per kg soil), and C_0 is the total SOC content (g C per kg soil) at time $t = 0$.

2.8 Statistics

The fit between $PARTY_{SOC}$ predictions of the centennially stable-SOC proportion and ex post AMG-optimized stable-

SOC proportion was assessed by a linear regression model. The same approach was applied for the evaluation of the agreement between centennially stable-SOC content and ex post AMG-optimized stable-SOC content of initial samples. The evaluation of the performance of the AMG model, for the different SOC pool-partitioning initialization methods, was also based on simple linear regressions between simulated and observed SOC stock values. Statistical terms used to express the strength and the statistical significance of the relationships were the coefficient of determination (R^2) and the associated probability value (p value). Prediction bias and model error were expressed as the mean difference (*BIAS*) and relative mean square error (*RMSE*). The relative root mean square error (*RRMSE*) and the normalized root mean square error (*NRMSE*) were used to compare the error of different datasets (with a different range of predictions) (Smith et al., 1996; Wallach, 2006; Otto et al., 2018).

$$R^2 = \left(\frac{\sum_{i=1}^{n} \left(\left(O_i - \overline{O} \right) \cdot \left(S_i - \overline{S} \right) \right)}{\sqrt{\sum_{i=1}^{n} \left(O_i - \overline{O} \right)^2} \cdot \sqrt{\sum_{i=1}^{n} \left(S_i - \overline{S} \right)^2}} \right)^2, \quad (3)$$

$$\mathrm{BIAS} = \frac{1}{n} \sum_{i=1}^{n} \left(S_i - O_i \right), \quad (4)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(S_i - O_i \right)^2}, \quad (5)$$

$$\mathrm{RRMSE} = \frac{\mathrm{RMSE}}{\overline{O}}, \quad (6)$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{O_{\max} - O_{\min}}, \quad (7)$$

where O and S are the observed and simulated values; n is the number of observations; $\overline{O}$ and $\overline{S}$ are the means of observations and simulations, respectively; and $O_{\max}$ and $O_{\min}$ are the maximum and the minimum value observed.

The observed and simulated total SOC stock change dQC was calculated as follows for each treatment:

$$\mathrm{dQC_{obs}} = \mathrm{QC_{obs},}_{t_2} - \mathrm{QC_{obs},}_{t_1}, \quad (8)$$

$$\mathrm{dQC_{sim}} = \mathrm{QC_{sim},}_{t_2} - \mathrm{QC_{obs},}_{t_1}, \quad (9)$$

where $\mathrm{QC_{obs}}$ is the observed SOC stock at time t, $\mathrm{QC_{sim}}$ is the SOC stock at time t simulated with AMG, and t_1 indicates the start and t_2 indicates the end of simulation period.

All data processing and statistical analyses were performed within the R programming environment (version 3.4.2) (R Core Team, 2017). For plotting, packages ggpmisc (Aphalo, 2016), reshape2 (Wickham, 2007), and ggplot2 (Wickham, 2016) were used.

3 Results

3.1 Accurate soil organic carbon pool partitioning

Centennially stable-SOC proportion values were predicted by the $\mathrm{PARTY_{SOC}}$ machine-learning model (Cécillon et al., 2021) using Rock-Eval® data measured on initial topsoil samples. The mean value for each independent site was plotted against the stable SOC proportion as determined by ex post AMG optimization (Fig. 1). The initial centennially stable-SOC proportion values predicted with $\mathrm{PARTY_{SOC}}$ ranged from 0.44 to 0.74, with a mean value of 0.59, whereas ex post AMG-optimal estimations of stable-SOC proportion covered almost the same range, from 0.45 to 0.74, with a mean value of 0.61. The two approaches were strongly correlated ($R^2 = 0.63$, significant at the $p < 0.05$ level), with a linear regression slope close to 1 ($a = 0.9$) and intercept close to 0 ($b = 0.04$), showing an unbiased relationship between $\mathrm{PARTY_{SOC}}$ estimates of the centennially stable-SOC proportion and the ex post AMG-optimized stable-SOC proportion at the onset of the nine LTEs. Although a slight discrepancy was observed for higher stable-SOC proportion values, the results validate our hypothesis showing that the centennially stable-SOC proportion determined by Rock-Eval® thermal analysis and the $\mathrm{PARTY_{SOC}}$ machine-learning model built on fully independent data provides a good estimate of the optimal stable-SOC proportion of the AMG model for unrelated French agricultural soils. When expressed as content (g C per kg soil), the fit between the $\mathrm{PARTY_{SOC}}$ predictions of the centennially stable SOC determined on initial topsoil samples and the ex post optimized stable-SOC content values was excellent ($R^2 = 0.95$; Supplement Fig. S4; optimal stable-SOC content ranged from 4.37 to 12.75 g C per kg soil across the nine sites). Furthermore, the method appears to be reliable, since additional Rock-Eval® measurements on topsoil samples from intermediate and final dates of the LTEs showed that the $\mathrm{PARTY_{SOC}}$ predictions of the centennially stable-SOC content remained remarkably constant during the experimental period at most sites (Supplement Fig. S5).

3.2 More accurate soil organic carbon simulations

In a second step, we investigated if a $\mathrm{PARTY_{SOC}}$-based initialization of the SOC pool partitioning could improve the accuracy of SOC stock simulations of the AMG model. To do so, we compared SOC stock simulations obtained with three different initializations. We first ran AMG using the default initialization method for the SOC pool partitioning ($\mathrm{C_S} / \mathrm{C_0} = 0.65$, since all LTEs were under cropland for at least 2 decades before their onset; Table 1). Then, we ran AMG simulations using the $\mathrm{PARTY_{SOC}}$-based initialization method by defining $\mathrm{C_S} / \mathrm{C_0}$ as the site-mean centennially stable-SOC proportion determined by the $\mathrm{PARTY_{SOC}}$ model. Finally, we ran AMG using the ex post optimization method to initialize the SOC pool partitioning for each

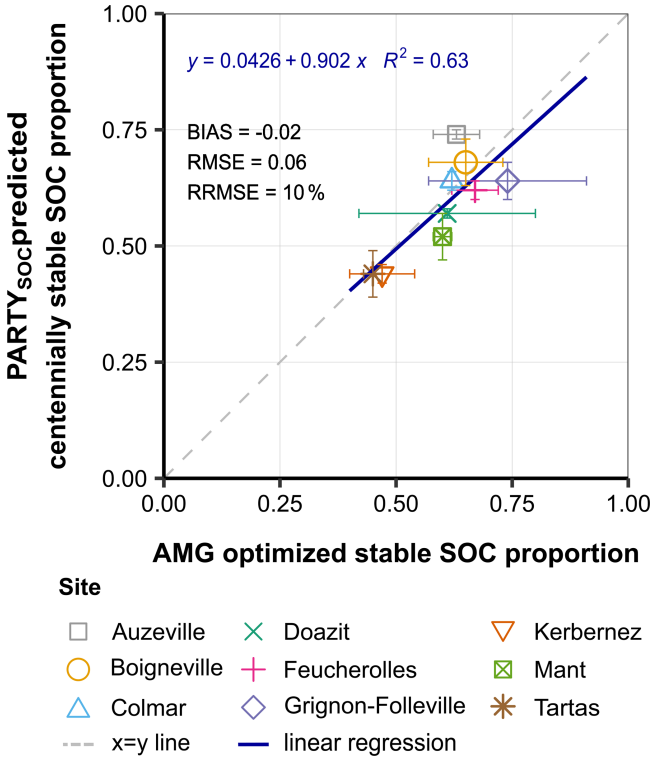

Figure 1. Performance of the $PARTY_{SOC}$ model to predict the centennially stable-SOC proportion compared to the ex post AMG-optimized stable-SOC proportion. Points represent site-mean values based on initial topsoil samples from nine independent French long-term experiments. Statistics refer to the linear regression between x and y values (blue solid line). Horizontal error bars show the uncertainty associated with the AMG-optimal stable-SOC proportion, calculated as the standard deviation of treatment-wise AMG optimizations. Vertical error bars represent the prediction error of the centennially stable-SOC proportion values, calculated as the standard deviation of the $PARTY_{SOC}$ model predictions on initial topsoil samples.

site. For all three initialization procedures, the simulated SOC stock change between the initial and last sampling date for each treatment of each site was plotted against the measured SOC stock change (Fig. 2a–c). Observed SOC stock change ranged from +6 to $-24\,Mg\,C\,ha^{-1}$ for the 32 treatments. In spite of a rather good mean agreement ($RMSE = 5.95\,Mg\,C\,ha^{-1}$), the AMG model initialized with the default procedure provided predictions of SOC stock change rather far from what was observed in two out of nine LTEs (Fig. 2a). Using the $PARTY_{SOC}$-based initialization method improved AMG simulations compared to the default method, bringing them much closer to the observed SOC stock changes ($RMSE = 3.60\,Mg\,C\,ha^{-1}$; Fig. 2a, b). $PARTY_{SOC}$-based initialization of AMG resulted in unbiased simulations ($BIAS = 0.06\,Mg\,C\,ha^{-1}$) and a strong decrease in the mean error of prediction. Unsurprisingly, AMG initialized using ex post optimized C_S / C_0 proportions predicted SOC stock changes very close to the observed ones ($RMSE = 2.12\,Mg\,C\,ha^{-1}$; Fig. 2c). AMG simulations from ex post optimized and $PARTY_{SOC}$-based initializations were remarkably comparable (Fig. 2b, c). The SOC stock simulations produced with AMG for each independent treatment are presented in Supplement Fig. S6.

It is noteworthy that the $PARTY_{SOC}$-based initialization improved the fit between observed and simulated SOC stock change, compared to AMG default initialization, especially for treatments that experienced the greatest SOC stock loss (Fig. 2a, b). In treatments that experienced no SOC stock change or a slight increase in SOC stock, the $PARTY_{SOC}$-based initialization did not improve the simulations but resulted in highly reliable predictions, similarly to AMG default or optimized initialization methods (Fig. 2a–c). This is likely explained by the history of land cover and soil management practices of the different sites. Sites presenting treatments with no change or a slight increase in SOC stocks were predominantly sites with a long cropland history (e.g. site of Boigneville; Supplement Table S1), for which the default AMG C_S / C_0 value of 0.65 is nearly optimal. Conversely, the two sites, Kerbernez and Tartas, where the ex post optimized C_S / C_0 values were far below the default value (Fig. 1) have a more complex history of land use and soil management practices. The site of Kerbernez is former grassland (during the first half of the 20th century; Supplement Table S1) that was converted into cropland 2 decades before the implementation of its arable LTE, in 1958. The site of Tartas was cultivated for a longer time before the LTE onset; however it was turned to grassland for a period in the 19th century (Supplement Table S1) and received applications of poultry manure for several years before the LTE began. In these two sites, characterized by an AMG-optimal C_S / C_0 much lower than the default value, the $PARTY_{SOC}$ machine-learning model predicted values very close to the optimal C_S / C_0 values (Fig. 1).

4 Discussion

Our study demonstrates that the $PARTY_{SOC}$ method based on Rock-Eval® thermal analysis (Cécillon et al., 2018, 2021) can estimate the initial SOC pool partitioning of the AMG model of SOC dynamics while improving its accuracy in a series of diverse and independent French LTEs. Contrary to previous studies (Skjemstad et al., 2004; Luo et al., 2014), no modifications of the decomposition rate of SOC kinetic pools were necessary to improve model predictions. The $PARTY_{SOC}$ initialization method never severely affected the model simulations, while it strongly improved them at sites where SOC stocks were far from an equilibrium state due to historical changes in soil management or land use. Areas with past changes in land use and soil management represent a large yet poorly known part of arable land in France and Europe (Fuchs et al., 2015; Erb et al., 2017) where SOC

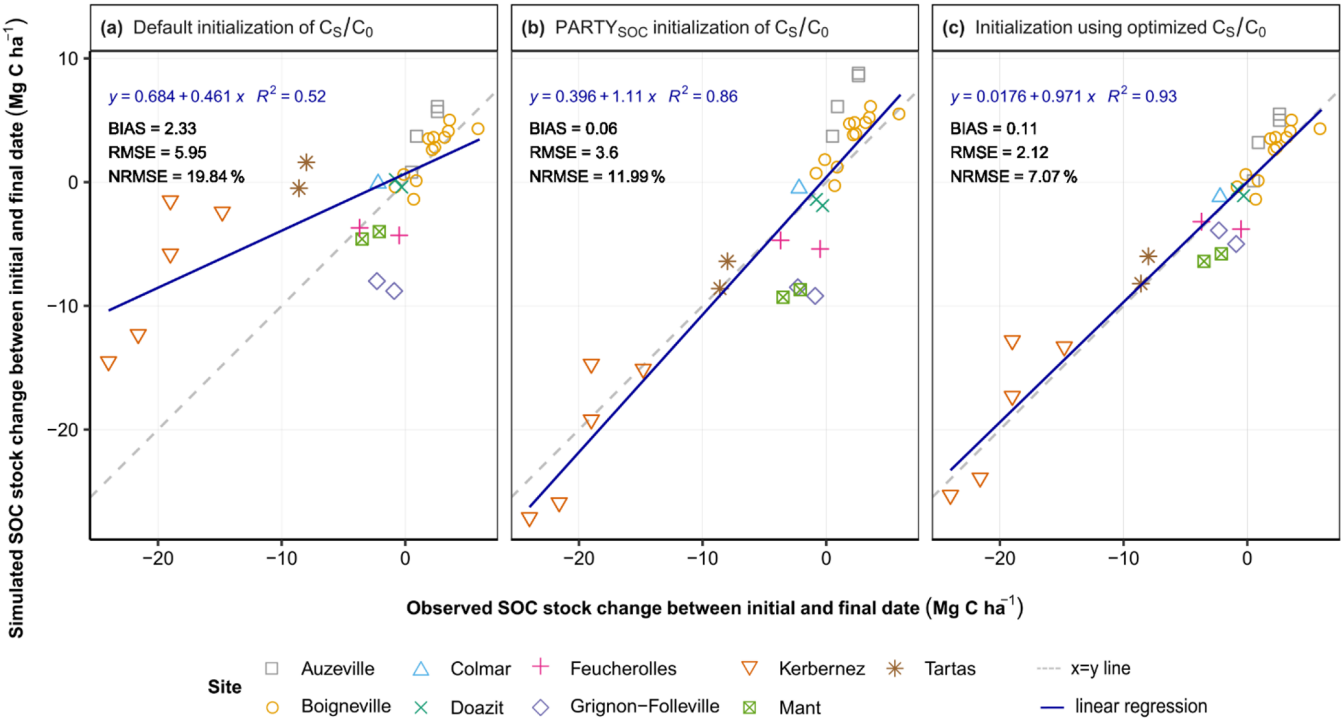

Figure 2. Observed vs. simulated change in SOC stocks between the initial and final date of 32 treatments from nine French long-term experiments. The three panels show the performance of the AMG model for three different initialization approaches. Initial SOC kinetic pool sizes were defined using **(a)** the default value for cropland ($C_S / C_0 = 0.65$), **(b)** the centennially stable-SOC proportion predicted by the $PARTY_{SOC}$ model, and **(c)** the ex post AMG-optimized C_S / C_0 proportion. Statistics refer to the linear regression between x and y values (blue solid line). Points represent the values for the 32 treatments for which AMG simulations were run.

stocks and slow-cycling SOC pools are far from equilibrium (Wutzler and Reichstein, 2007; Herbst et al., 2018; Clivot et al., 2019; Taghizadeh-Toosi et al., 2020). Therefore, by accounting for these legacy effects of site history on SOC pool partitioning, the $PARTY_{SOC}$-based initialization of the AMG model should result in more accurate simulations of SOC dynamics at a national or continental scale.

Our findings, combined with results reported in recent ensemble modelling studies (Martin et al., 2019; Farina et al., 2021), suggest that despite its simple structure and when properly initialized (e.g. using the $PARTY_{SOC}$ model) the AMG model is unsurpassed for predicting observed SOC stock changes in French agricultural LTEs and is amongst the best available modelling frameworks of SOC dynamics in European arable land (Martin et al., 2019; Farina et al., 2021). Our results demonstrate that there is still potential to increase the accuracy of simple multi-compartmental models of SOC dynamics, bringing their simulations very close to the observed values of SOC stock changes. Developing other Rock-Eval®-based initialization methods specifically designed to match the carbon pool design of other multi-compartmental SOC dynamics models such as RothC (Coleman et al., 1997) is a promising research area. More generally, we recommend that the potential of multi-compartmental SOC dynamics models be fully explored and exploited by soil biogeochemists before a new generation of models of increased complexity becomes operational. While new models including the diversity of microbial communities and related processes are emerging (Lehmann et al., 2020; Crowther et al., 2019), the uncertain structure and parametrization of more complex models is hindering their application as robust predictive tools (Shi et al., 2018). At the same time, simple conceptual models of SOC dynamics like AMG combined with novel initialization methods and data-based approaches such as $PARTY_{SOC}$ show promising improvements (Cécillon, 2021a; Dangal et al., 2021; Lee et al., 2020). The low prediction error of the AMG model when its SOC pool distribution is initialized with the $PARTY_{SOC}$ method even challenges the ability of more complex modelling approaches to achieve better performance, given the uncertainty on observed values of SOC stock changes (Schrumpf et al., 2011).

The continental or worldwide implementation of the AMG model with the $PARTY_{SOC}$-based initialization of SOC pools distribution will require additional work. First, the $PARTY_{SOC}$ machine-learning model (Cécillon et al., 2018, 2021) will have to be validated on a wider range of pedoclimates. This method, initially built on LTEs coming from north-western Europe (Cécillon et al., 2018), has now been successfully extended to new soil types and a new

climate (tropical; Cécillon et al., 2021). The good agreement between AMG-optimal stable-SOC proportion values and $PARTY_{SOC}$ predictions reported here suggests that most agricultural LTEs with accurate AMG simulations could be used as reference sites for the $PARTY_{SOC}$ model, lifting an important technical limitation to its geographical expansion (Cécillon et al., 2021). Second, the improved accuracy of model simulations using a $PARTY_{SOC}$-based initialization will also have to be demonstrated for a wider pedoclimatic range (i.e. worldwide LTEs; such as those referenced by the International Soil Carbon Network; Nave et al., 2015). Third, Rock-Eval® data from the new application areas will be required. Rock-Eval® is a high-throughput technique that is well adapted to the analysis of large soil sample sets provided by large-scale soil monitoring programmes. We recommend implementing Rock-Eval® measurements in national and continental soil monitoring networks.

5 Conclusions

Combining Rock-Eval® thermal analysis with the $PARTY_{SOC}$ machine-learning model should be considered an emerging key approach with demonstrated ability to improve the accuracy of simulations of SOC dynamics, complementary to other SOC cycling proxies (Bailey et al., 2018; Wiesmeier et al., 2019). The progressive large-scale delivery of these complementary data related to SOC dynamics will strengthen model predictions of SOC stock changes at the national to global scale, necessary for implementing efficient climate change mitigation policies (FAO, 2020).

Code availability. Code used to produce the results presented in this study, as well as the AMG model, can be made available upon reasonable request. The code of the Rock-Eval®-based $PARTY_{SOC}v2.0_{EU}$ random forest model is available on Zenodo at the permanent link of https://doi.org/10.5281/zenodo.4446138 (Cécillon, 2021b).

Data availability. Pedoclimatic and land management information used to run the AMG model are published in Clivot et al. (2019) and can be obtained from Hugues Clivot and Fabien Ferchaud upon request concerning seven sites and from Florent Levavasseur and Sabine Houot concerning the remaining two sites (Levavasseur et al., 2020). The Rock-Eval® data obtained and processed during this study are available upon request from the corresponding authors.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/bg-19-375-2022-supplement.

Author contributions. EK performed the research and produced the first version of the manuscript. PB and LC designed the study with contributions from CC and FB, whereas BM, FF, HC, FL, and SH organized the search for archive soil samples, provided data needed to run AMG simulations as well as the latest version of the AMG model, and helped with the interpretation of the results. FL performed AMG simulations for the sites of Colmar and Feucherolles. FB conducted all Rock-Eval® analyses. All authors participated in multiple scientific discussions during the preparation of this paper and finally reviewed and provided valuable feedback during the writing process.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors would like to thank ADEME (Ministry for an Ecological Transition of the Ministry for Higher Education, Research and Innovation) and the ANR (French National Research Agency) for the financial support. The authors would also like to thank Florence Savignac for her help in performing Rock-Eval® analyses and all the people involved in maintaining and monitoring the field trials used in this study over multiple decades. We would also like to thank everyone involved in the conception and creation of the AMG consortium, the AIAL database, and the SOERE PRO network for kindly providing access to their data.

Financial support. This work was supported by ADEME (French Agency for Ecological Transition, under the Ministry for an Ecological Transition and the Ministry for Higher Education, Research and Innovation; PhD grant for Eva Kanari, PROTERR project) and the ANR (French National Research Agency; PhD grant for Eva Kanari, StoreSoilC project ANR-17-CE32-0005).

Review statement. This paper was edited by Alexey V. Eliseev and reviewed by Adrián Andriulo and one anonymous referee.

References

Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Bol, R., Collins, C., Lal, R., Leifeld, J., Minasny, B., Pan, G., Paustian, K., Rumpel, C., Sanderman, J., van Groenigen, J. W., Mooney, S., van Wesemael, B., Wander, M., and Chabbi, A.: Towards a global-scale soil climate mitigation strategy, Nat. Commun., 11, 1–10, https://doi.org/10.1038/s41467-020-18887-7, 2020.

Andriulo, A. E., Mary, B., and Guerif, J.: Modelling soil carbon dynamics with various cropping sequences on the rolling pampas, Agronomie, 19, 365–377, 1999.

Aphalo, P. J.: Learn R as you learnt your mother tongue, Leanpub [code], available at: https://leanpub.com/learnr (last access: 19 January 2022), 2016.

Bailey, V. L., Bond-Lamberty, B., DeAngelis, K., Grandy, A. S., Hawkes, C. V., Heckman, K., Lajtha, K., Phillips, R. P., Sulman, B. N., Todd-Brown, K. E. O., and Wallenstein, M. D.: Soil carbon cycling proxies: Understanding their critical role in predicting climate change feedbacks, Glob. Chang. Biol., 24, 895–905, https://doi.org/10.1111/gcb.13926, 2018.

Baldock, J. A., Hawke, B., Sanderman, J., and Macdonald, L. M.: Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra, Soil Res., 51, 577–595, https://doi.org/10.1071/SR13077, 2013.

Balesdent, J. and Arrouays, D.: Usage des terres et stockage de carbone dans les sols du territoire français. Une estimation des flux nets pour la période 1900–1999, Comptes rendus l'Académie d'Agriculture Fr., 85, 265–277, 1999.

Balesdent, J., Basile-Doelsch, I., Chadoeuf, J., Cornu, S., Derrien, D., Fekiacova, Z., and Hatté, C.: Atmosphere–soil carbon transfer as a function of soil depth, Nature, 559, 599–602, https://doi.org/10.1038/s41586-018-0328-3, 2018.

Barré, P., Montagnier, C., Chenu, C., Abbadie, L., and Velde, B.: Clay minerals as a soil potassium reservoir: Observation and quantification through X-ray diffraction, Plant Soil, 302, 213–220, https://doi.org/10.1007/s11104-007-9471-6, 2008.

Barré, P., Eglin, T., Christensen, B. T., Ciais, P., Houot, S., Kätterer, T., van Oort, F., Peylin, P., Poulton, P. R., Romanenkov, V., and Chenu, C.: Quantifying and isolating stable soil organic carbon using long-term bare fallow experiments, Biogeosciences, 7, 3839–3850, https://doi.org/10.5194/bg-7-3839-2010, 2010.

Barré, P., Plante, A. F., Cécillon, L., Lutfalla, S., Baudin, F., Bernard, S., Christensen, B. T., Eglin, T., Fernandez, J. M., Houot, S., Kätterer, T., Le Guillou, C., Macdonald, A., van Oort, F., and Chenu, C.: The energetic and chemical signatures of persistent soil organic matter, Biogeochemistry, 130, 1–12, https://doi.org/10.1007/s10533-016-0246-0, 2016.

Barthès, B. G., Brunet, D., Hien, E., Enjalric, F., Conche, S., Freschet, G. T., d'Annunzio, R., and Toucet-Louri, J.: Determining the distributions of soil carbon and nitrogen in particle size fractions using near-infrared reflectance spectrum of bulk soil samples, Soil Biol. Biochem., 40, 1533–1537, https://doi.org/10.1016/j.soilbio.2007.12.023, 2008.

Baudin, F., Disnar, J. R., Aboussou, A., and Savignac, F.: Guidelines for Rock-Eval analysis of recent marine sediments, Org. Geochem., 86, 71–80, https://doi.org/10.1016/j.orggeochem.2015.06.009, 2015.

Behar, F., Beaumont, V., and De B. Penteado, H. L.: Rock-Eval 6 Technology: Performances and Developments, Oil Gas Sci. Technol., 56, 111–134, https://doi.org/10.2516/ogst:2001013, 2001.

Bolinder, M. A., Janzen, H. H., Gregorich, E. G., Angers, D. A., and VandenBygaart, A. J.: An approach for estimating net primary productivity and annual carbon inputs to soil for common agricultural crops in Canada, Agr. Ecosyst. Environ., 118, 29–42, https://doi.org/10.1016/j.agee.2006.05.013, 2007.

Bruun, S. and Jensen, L. S.: Initialisation of the soil organic matter pools of the Daisy model, Ecol. Modell., 153, 291–295, 2002.

Cagnarini, C., Renella, G., Mayer, J., Hirte, J., Schulin, R., Costerousse, B., Della Marta, A., Orlandini, S., and Menichetti, L.: Multi-objective calibration of RothC using measured carbon stocks and auxiliary data of a long-term experiment in Switzerland, Eur. J. Soil Sci., 70, 819–832, https://doi.org/10.1111/ejss.12802, 2019.

Cécillon, L.: A dual response, Nat. Geosci., 14, 262–263, https://doi.org/10.1038/s41561-021-00749-6, 2021a.

Cécillon, L.: lauric-cecillon/PARTYsoc: Second version of the $PARTY_{SOC}$ statistical model, Zenodo [code], https://doi.org/10.5281/zenodo.4446138, 2021b.

Cécillon, L., Baudin, F., Chenu, C., Houot, S., Jolivet, R., Kätterer, T., Lutfalla, S., Macdonald, A., van Oort, F., Plante, A. F., Savignac, F., Soucémarianadin, L. N., and Barré, P.: A model based on Rock-Eval thermal analysis to quantify the size of the centennially persistent organic carbon pool in temperate soils, Biogeosciences, 15, 2835–2849, https://doi.org/10.5194/bg-15-2835-2018, 2018.

Cécillon, L., Baudin, F., Chenu, C., Christensen, B. T., Franko, U., Houot, S., Kanari, E., Kätterer, T., Merbach, I., van Oort, F., Poeplau, C., Quezada, J. C., Savignac, F., Soucémarianadin, L. N., and Barré, P.: Partitioning soil organic carbon into its centennially stable and active fractions with machine-learning models based on Rock-Eval® thermal analysis ($PARTY_{SOC}$v2.0 and $PARTY_{SOC}v2.0_{EU}$), Geosci. Model Dev., 14, 3879–3898, https://doi.org/10.5194/gmd-14-3879-2021, 2021.

Chassé, M., Luftalla, S., Cécillon, L., Baudin, F., Abiven, S., Chenu, C., and Barré, P.: Long-term bare fallow soil fractions reveal thermo-chemical properties controlling soil organic carbon dynamics, Biogeosciences, 18, 1703–1718, https://doi.org/10.5194/bg-18-1703-2021, 2021.

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

Coleman, K., Jenkinson, D. S., Crocker, G. J., Grace, P. R., Klír, J., Körschens, M., Poulton, P. R., and Richter, D. D.: Simulating trends in soil organic carbon in long-term experiments using RothC-26.3, Geoderma, 81, 29–44, https://doi.org/10.1016/S0016-7061(97)00079-7, 1997.

Colomb, B., Debaeke, P., Jouany, C., and Nolot, J. M.: Phosphorus management in low input stockless cropping systems: Crop and soil responses to contrasting P regimes in a 36-year experiment in southern France, Eur. J. Agron., 26, 154–165, https://doi.org/10.1016/j.eja.2006.09.004, 2007.

Cotrufo, M. F., Ranalli, M. G., Haddix, M. L., Six, J., and Lugato, E.: Soil carbon storage informed by particulate and mineral-associated organic matter, Nat. Geosci., 12, 989–994, https://doi.org/10.1038/s41561-019-0484-6, 2019.

Crowther, T. W., van den Hoogen, J., Wan, J., Mayes, M. A., Keiser, A. D., Mo, L., Averill, C., and Maynard, D. S.: The global soil community and its influence on biogeochemistry, Science, 365, eaav0550, https://doi.org/10.1126/science.aav0550, 2019.

Dangal, S. R. S., Schwalm, C., Cavigelli, M. A., Gollany, H. T., Jin, V. L., and Sanderman, J.: Improving soil carbon es-

timates by linking conceptual pools against measurable carbon fractions in the DAYCENT Model Version 4.5, ESSOAR, https://doi.org/10.1002/essoar.10507190.1, 2021.

Dignac, M.-F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G. T., Garnier, P., Guenet, B., Hedde, M., Klumpp, K., Lashermes, G., Maron, P.-A., Nunan, N., Roumet, C., and Basile-Doelsch, I.: Increasing soil carbon storage: mechanisms, effects of agricultural practices and proxies. A review, Agron. Sustain. Dev., 37, 14, https://doi.org/10.1007/s13593-017-0421-2, 2017.

Dimassi, B., Mary, B., Wylleman, R., Labreuche, J. Ô., Couture, D., Piraux, F., and Cohan, J. P.: Long-term effect of contrasted tillage and crop management on soil carbon dynamics during 41 years, Agr. Ecosyst. Environ., 188, 134–146, https://doi.org/10.1016/j.agee.2014.02.014, 2014.

Disnar, J. R., Guillet, B., Keravis, D., Di-Giovanni, C., and Sebag, D.: Soil organic matter (SOM) characterization by Rock-Eval pyrolysis: Scope and limitations, Org. Geochem., 34, 327–343, https://doi.org/10.1016/S0146-6380(02)00239-5, 2003.

Erb, K. H., Luyssaert, S., Meyfroidt, P., Pongratz, J., Don, A., Kloster, S., Kuemmerle, T., Fetzel, T., Fuchs, R., Herold, M., Haberl, H., Jones, C. D., Marín-Spiotta, E., McCallum, I., Robertson, E., Seufert, V., Fritz, S., Valade, A., Wiltshire, A., and Dolman, A. J.: Land management: data availability and process understanding for global change studies, Glob. Chang. Biol., 23, 512–533, https://doi.org/10.1111/gcb.13443, 2017.

FAO: Recarbonization of global soils: a dynamic response to offset global emissions, FAO, Rome, Italy, 8 pp., available at: https://www.fao.org/documents/card/en/c/e4839d17-c6aa-44ca-85cc-b4a6de25c143/ (last access: 19 January 2022), 2019.

FAO: Technical specifications and country guidelines for Global Soil Organic Carbon Sequestration Potential Map (GSOCseq), FAO, Rome, Italy, 48 pp., available at: https://www.fao.org/3/cb0353en/CB0353EN.pdf (last access: 19 January 2022), 2020.

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., Guenet, B., Harrison, M. T., Kirschbaum, M. U. F., Kuka, K., Kulmala, L., Liski, J., McGrath, M. J., Meier, E., Menichetti, L., Moyano, F., Nendel, C., Recous, S., Reibold, N., Shepherd, A., Smith, W. N., Smith, P., Soussana, J. F., Stella, T., Taghizadeh-Toosi, A., Tsutskikh, E., and Bellocchi, G.: Ensemble modelling, uncertainty and robust predictions of organic carbon in long-term bare-fallow soils, Glob. Chang. Biol., 27, 904–928, https://doi.org/10.1111/gcb.15441, 2021.

Fuchs, R., Herold, M., Verburg, P. H., Clevers, J. G. P. W., and Eberle, J.: Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010, Glob. Chang. Biol., 21, 299–313, https://doi.org/10.1111/gcb.12714, 2015.

Gregorich, E. G., Gillespie, A. W., Beare, M. H., Curtin, D., Sanei, H., and Yanni, S. F.: Evaluating biodegradability of soil organic matter by its thermal stability and chemical composition, Soil Biol. Biochem., 91, 182–191, https://doi.org/10.1016/j.soilbio.2015.08.032, 2015.

He, Y., Trumbore, S. E., Torn, M. S., Harden, J. W., Vaugh, L. J. S., Allison, S. D., and Randerson, J. T.: Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century, Science, 353, 1419–1424, https://doi.org/10.1126/science.aad4273, 2016.

Hemingway, J. D., Rothman, D. H., Grant, K. E., Rosengard, S. Z., Eglinton, T. I., Derry, L. A., and Galy, V. V.: Mineral protection regulates long-term global preservation of natural organic carbon, Nature, 570, 228–231, https://doi.org/10.1038/s41586-019-1280-6, 2019.

Hénin, S. and Dupuis, M.: Essai de bilan de la matière organique du sol, in: Ann. Agron. 11, Dunod (impr. de Chaix), Paris, 17–29, 1945.

Herbst, M., Welp, G., Macdonald, A., Jate, M., Hädicke, A., Scherer, H., Gaiser, T., Herrmann, F., Amelung, W., and Vanderborght, J.: Correspondence of measured soil carbon fractions and RothC pools for equilibrium and non-equilibrium states, Geoderma, 314, 37–46, https://doi.org/10.1016/j.geoderma.2017.10.047, 2018.

IPCC: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, edited by: Shukla, P. R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, E., Kissick, K., Belkacemi, M., and Malley, J., The Intergovernmental Panel on Climate Change, available at: https://www.ipcc.ch/srccl/ (last access: 19 January 2022), 2019.

Jobbagy, E. G. and Jackson, R. B.: The vertical distribution of Soil Organic Carbon and its relation to climate and vegetation, Ecol. Appl., 10, 423–436, 2000.

Khedim, N., Cécillon, L., Poulenard, J., Barré, P., Baudin, F., Marta, S., Rabatel, A., Dentant, C., Cauvy-Fraunié, S., Anthelme, F., Gielly, L., Ambrosini, R., Franzetti, A., Azzoni, R. S., Caccianiga, M. S., Compostella, C., Clague, J., Tielidze, L., Messager, E., Choler, P., and Ficetola, G. F.: Topsoil organic matter build-up in glacier forelands around the world, Glob. Change Biol., 27, 1662–1677, https://doi.org/10.1111/gcb.15496, 2021.

Lee, J., Viscarra Rossel, R. A., Zhang, M., Luo, Z., and Wang, Y.-P.: Assessing the response of soil carbon in Australia to changing inputs and climate using a consistent modelling framework, Biogeosciences, 18, 5185–5202, https://doi.org/10.5194/bg-18-5185-2021, 2021.

Lehmann, J., Hansel, C. M., Kaiser, C., Kleber, M., Maher, K., Manzoni, S., Nunan, N., Reichstein, M., Schimel, J. P., Torn, M. S., Wieder, W. R., and Kögel-Knabner, I.: Persistence of soil organic carbon caused by functional complexity, Nat. Geosci., 13, 529–534, https://doi.org/10.1038/s41561-020-0612-3, 2020.

Leifeld, J., Reiser, R., and Oberholzer, H. R.: Consequences of conventional versus organic farming on soil carbon: Results from a 27-year field experiment, Agron. J., 101, 1204–1218, https://doi.org/10.2134/agronj2009.0002, 2009a.

Leifeld, J., Zimmermann, M., Fuhrer, J., and Conen, F.: Storage and turnover of carbon in grassland soils along an elevation gradient in the Swiss Alps, Glob. Change Biol., 15, 668–679, https://doi.org/10.1111/j.1365-2486.2008.01782.x, 2009b.

Levavasseur, F., Mary, B., Christensen, B. T., Duparque, A., Ferchaud, F., Kätterer, T., Lagrange, H., Montenach, D., Resseguier, C., and Houot, S.: The simple AMG model accurately simulates organic carbon storage in soils after repeated application of ex-

ogenous organic matter, Nutr. Cycl. Agroecosyst., 117, 215–229, https://doi.org/10.1007/s10705-020-10065-x, 2020.

Lubet, E., Plénet, D., and Juste, C.: Effet à long terme de la monoculture sur le rendement en grain du maïs (Zea mays L) en conditions non irriguées, Agronomie, 13, 673–683, https://doi.org/10.1051/agro:19930801, 1993.

Lugato, E., Lavallee, J. M., Haddix, M. L., Panagos, P., and Cotrufo, M. F.: Different climate sensitivity of particulate and mineral-associated soil organic matter, Nat. Geosci., 14, 295–300, https://doi.org/10.1038/s41561-021-00744-x, 2021.

Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E. A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J. W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, Robert, B., Jones, C. D., Lara, M. J., Liang, J., McGuire, A. D., Parton, W., Peng, C., Randerson, J. T., Salazar, A., Sierra, C. A., Smith, M. J., Tian, H., Todd-Brown, K. E. O., Torn, M., Van Groenigen, K. J., Wang, Y. P., West, T. O., Wei, Y., Wieder, W. R., Xu, X., Xu, X., and Zhou, T.: Towards more realistic projections of soil carbon dynamics by Earth system models, Global Biogeochem. Cy., 30, 40–56, https://doi.org/10.1002/2015GB005239, 2016.

Luo, Z., Wang, E., Fillery, I. R. P., Macdonald, L. M., Huth, N., and Baldock, J.: Modelling soil carbon and nitrogen dynamics using measurable and conceptual soil organic matter pools in APSIM, Agr. Ecosyst. Environ., 186, 94–104, https://doi.org/10.1016/j.agee.2014.01.019, 2014.

Manzoni, S. and Porporato, A.: Soil carbon and nitrogen mineralization: Theory and models across scales, Soil Biol. Biochem., 41, 1355–1379, https://doi.org/10.1016/j.soilbio.2009.02.031, 2009.

Martin, M., Dimassi, B., Millet, F., Picaud, C., Bounoua, E.-M., Bardy, M., Bispo, A., Boulonne, L., Bouthier, A., Duparque, A., Eglin, T., Guenet, B., Huard, F., Mary, B., Mathias, E., Mignolet, C., Robert, C., Saby, N., Sagot, S., Schott, C., and Toutain, B. R.: Méthodes de comptabilisation du stockage de carbone organique des sols sous l'effet des pratiques culturales (CSopra), ADEME, available at: https://librairie.ademe.fr/produire-autrement/ (last access: 19 January 2022), 2019.

Mary, B., Clivot, H., Blaszczyk, N., Labreuche, J., and Ferchaud, F.: Soil carbon storage and mineralization rates are affected by carbon inputs rather than physical disturbance: Evidence from a 47-year tillage experiment, Agr. Ecosyst. Environ., 299, 106972, https://doi.org/10.1016/j.agee.2020.106972, 2020.

Messiga, A. J., Ziadi, N., Plénet, D., Parent, L. E., and Morel, C.: Long-term changes in soil phosphorus status related to P budgets under maize monoculture and mineral P fertilization, Soil Use Manag., 26, 354–364, https://doi.org/10.1111/j.1475-2743.2010.00287.x, 2010.

Morel, C., Ziadi, N., Messiga, A., Bélanger, G., Denoroy, P., Jeangros, B., Jouany, C., Fardeau, J. C., Mollier, A., Parent, L. E., Proix, N., Rabeharisoa, L., and Sinaj, S.: Modeling of phosphorus dynamics in contrasting agroecosystems using long-term field experiments, Can. J. Soil Sci., 94, 377–387, https://doi.org/10.4141/CJSS2013-024, 2014.

Nave, L., Johnson, K., Ingen, C. van, Agarwal, D., Humphrey, M., and Beekwilder, N.: International Soil Carbon Network (ISCN) Database V3-1, International Soil Carbon Network, https://doi.org/10.17040/ISCN/1305039, 2015.

Nemo, K., Coleman, K., Dondini, M., Goulding, K., Hastings, A., Jones, M. B., Leifeld, J., Osborne, B., Saunders, M., Scott, T., Teh, Y. A., and Smith, P.: Soil Organic Carbon (SOC) Equilibrium and Model Initialisation Methods: an Application to the Rothamsted Carbon (RothC) Model, Environ. Model. Assess., 22, 215–229, https://doi.org/10.1007/s10666-016-9536-0, 2016.

Noirot-Cosson, P. E., Vaudour, E., Gilliot, J. M., Gabrielle, B., and Houot, S.: Modelling the long-term effect of urban waste compost applications on carbon and nitrogen dynamics in temperate cropland, Soil Biol. Biochem., 94, 138–153, https://doi.org/10.1016/j.soilbio.2015.11.014, 2016.

Oberholzer, H. R., Leifeld, J., and Mayer, J.: Changes in soil carbon and crop yield over 60 years in the Zurich Organic Fertilization Experiment, following land-use change from grassland to cropland, J. Plant Nutr. Soil Sci., 177, 696–704, https://doi.org/10.1002/jpln.201300385, 2014.

Obriot, F.: Epandage de produits résiduaires organiques et fonctionnement biologique des sols: De la quantification des impacts sur les cycles carbone et azote à l'évaluation multicritère de la pratique à l'échelle de la parcelle (PhD), AgroParisTech, 456 pp., available at: https://hal.inrae.fr/tel-02801385 (last access: 19 January 2022), 2016.

Otto, S. A., Kadin, M., Casini, M., Torres, M. A., and Blenckner, T.: A quantitative framework for selecting and validating food web indicators, Ecol. Indic., 84, 619–631, https://doi.org/10.1016/j.ecolind.2017.05.045, 2018.

Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., and Gensior, A.: Temporal dynamics of soil organic carbon after land-use change in the temperate zone – carbon response functions as a model approach, Glob. Change Biol., 17, 2415–2427, https://doi.org/10.1111/j.1365-2486.2011.02408.x, 2011.

Poeplau, C., Don, A., Dondini, M., Leifeld, J., Nemo, R., Schumacher, J., Senapati, N., and Wiesmeier, M.: Reproducibility of a soil organic carbon fractionation method to derive RothC carbon pools, Eur. J. Soil Sci., 64, 735–746, https://doi.org/10.1111/ejss.12088, 2013.

Poeplau, C., Don, A., Six, J., Kaiser, M., Benbi, D., Chenu, C., Cotrufo, M. F., Derrien, D., Gioacchini, P., Grand, S., Gregorich, E., Griepentrog, M., Gunina, A., Haddix, M., Kuzyakov, Y., Kühnel, A., Macdonald, L. M., Soong, J., Trigalet, S., Vermeire, M. L., Rovira, P., van Wesemael, B., Wiesmeier, M., Yeasmin, S., Yevdokimov, I., and Nieder, R.: Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison, Soil Biol. Biochem., 125, 10–26, https://doi.org/10.1016/j.soilbio.2018.06.025, 2018.

Poeplau, C., Barré, P., Cécillon, L., Baudin, F., and Sigurdsson, B. D.: Changes in the Rock-Eval signature of soil organic carbon upon extreme soil warming and chemical oxidation – A comparison, Geoderma, 337, 181–190, https://doi.org/10.1016/j.geoderma.2018.09.025, 2019.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, available at: https://www.R-project.org/ (last access: 19 January 2022), 2017.

Rumpel, C., Amiraslani, F., Koutika, L. S., Smith, P., Whitehead, D., and Wollenberg, E.: Put more carbon in soils to meet Paris climate pledges, Nature, 564, 32–34, https://doi.org/10.1038/d41586-018-07587-4, 2018.

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

Sanderman, J., Baldock, J. A., Dangal, S. R. S., Ludwig, S., Potter, S., Rivard, C., and Savage, K.: Soil organic carbon fractions in the Great Plains of the United States: an application of mid-infrared spectroscopy, Biogeochemistry, 156, 97–114, https://doi.org/10.1007/s10533-021-00755-1, 2021.

Schrumpf, M., Schulze, E. D., Kaiser, K., and Schumacher, J.: How accurately can soil organic carbon stocks and stock changes be quantified by soil inventories?, Biogeosciences, 8, 1193–1212, https://doi.org/10.5194/bg-8-1193-2011, 2011.

Shi, Z., Crowell, S., Luo, Y., and Moore, B.: Model structures amplify uncertainty in predicted soil carbon responses to climate change, Nat. Commun., 9, 2171, https://doi.org/10.1038/s41467-018-04526-9, 2018.

Skjemstad, J. O., Spouncer, L. R., Cowie, B., and Swift, R. S.: Calibration of the Rothamsted organic carbon turnover model (RothC ver. 26.3), using measurable soil organic carbon pools, Aust. J. Soil Res., 42, 79–88, https://doi.org/10.1071/SR03013, 2004.

Smith, J. U., Smith, P., and Addiscot, T.: Quantitative methods to evaluate and compare soil organic matter (SOM) models, in: Evaluation of Soil Organic Matter Models Using existing Long-Term Datasets, edited by: Powlson, D. S., Smith, J. U., and Smith, P., Springer Berlin, Heidelberg, 181–199, 1996.

Smith, P. and Falloon, P. D.: Modelling refractory soil organic matter, Biol. Fertil. Soils, 30, 388–398, https://doi.org/10.1007/s003740050019, 2000.

Taghizadeh-Toosi, A., Cong, W. F., Eriksen, J., Mayer, J., Olesen, J. E., Keel, S. G., Glendining, M., Kätterer, T., and Christensen, B. T.: Visiting dark sides of model simulation of carbon stocks in European temperate agricultural soils: allometric function and model initialization, Plant Soil, 450, 255–272, https://doi.org/10.1007/s11104-020-04500-9, 2020.

Todd-Brown, K. E. O., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., Zhang, Q., and Allison, S. D.: Changes in soil organic carbon storage predicted by Earth system models during the 21st century, Biogeosciences, 11, 2341–2356, https://doi.org/10.5194/bg-11-2341-2014, 2014.

UN General Assembly: UN General Assembly, Transforming our world: the 2030 Agenda for Sustainable Development, 21 October 2015, A/RES/70/1, General assembly seventieth session, New York, available at: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E (last access: 19 January 2022), 2015.

Vermeulen, S., Bossio, D., Lehmann, J., Luu, P., Paustian, K., Webb, C., Augé, F., Bacudo, I., Baedeker, T., Havemann, T., Jones, C., King, R., Reddy, M., Sunga, I., Von Unger, M., and Warnken, M.: A global agenda for collective action on soil carbon, Nat. Sustain., 2, 2–4, https://doi.org/10.1038/s41893-018-0212-z, 2019.

Vertès, F., Simon, J.-C., Laurent, F., and Besnard, A.: Prairies et qualité de l'eau. Evaluation des risques de lixiviation d'azote et optimisation des pratiques, Fourrag., 192, 423–440, 2007.

Viscarra Rossel, R. A., Lee, J., Behrens, T., Luo, Z., Baldock, J., and Richards, A.: Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls, Nat. Geosci., 12, 547–552, https://doi.org/10.1038/s41561-019-0373-z, 2019.

Wallach, D.: Evaluating Crop Models, in: Working with Dynamic Crop Models Evaluation, Analysis, Parameterization, and Applications, edited by: Wallach, D., Makowski, D., and Jones, J. W., Elsevier, Amsterdam, 11–54, 2006.

Wickham, H.: Reshaping Data with the reshape Package, J. Stat. Softw., 21, 1–20, https://doi.org/10.18637/jss.v021.i12, 2007.

Wickham, H.: ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York, ISBN 978-3-319-24277-4, 2016.

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H. J., and Kögel-Knabner, I.: Soil organic carbon storage as a key function of soils – A review of drivers and indicators at various scales, Geoderma, 333, 149–162, https://doi.org/10.1016/j.geoderma.2018.07.026, 2019.

Wutzler, T. and Reichstein, M.: Soils apart from equilibrium – Consequences for soil carbon balance modelling, Biogeosciences, 4, 125–136, https://doi.org/10.5194/bg-4-125-2007, 2007.

Zimmermann, M., Leifeld, J., Schmidt, M. W. I., Smith, P., and Fuhrer, J.: Measured soil organic matter fractions can be related to pools in the RothC model, Eur. J. Soil Sci., 58, 658–667, https://doi.org/10.1111/j.1365-2389.2006.00855.x, 2007a.

Zimmermann, M., Leifeld, J., and Fuhrer, J.: Quantifying soil organic carbon fractions by infrared-spectroscopy, Soil Biol. Biochem., 39, 224–231, https://doi.org/10.1016/j.soilbio.2006.07.010, 2007b.