https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Beyond behavioural models: equifinality and overparameterisation undermine confidence in predictions by soil organic matter models

Marijn Van de Broek[1] and Johan Six[1]

[1]Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland

Correspondence: Marijn Van de Broek (Marijn.VandeBroek@usys.ethz.ch)

Abstract. The complexity of soil organic matter models is often not supported by sufficient data for parameter optimisation, resulting in the calibration of more parameters than can be reliably optimized with the available data. This results in equifinality, the phenomenon that multiple parameter sets generate behavioural models, i.e., similarly well-performing models that cannot be ruled out. As such trade-offs between model complexity and data availability are often overlooked for soil organic matter models, the aim of this study is to assess how equifinality affects the variability of predictions made by behavioural soil organic matter models. The results show that the number of identifiable parameters, those that do not compensate for one another, increases with the number of calibration constraints, but remained limited to five even under the most data-rich conditions. Furthermore, the size of particulate organic matter (POM) and mineral-associated organic matter (MAOM) can only be accurately simulated when data on these pool sizes are used, while the turnover rate of MAOM is reliably simulated only when $\Delta^{14}C$ data for MAOM are provided. Regardless of the type of mathematical equations used (e.g., absolute vs. relative Michaelis-Menten kinetics), or the number of optimised parameters, the tested models were able to correctly reproduce the measurements in steady state. However, different model structures led to divergent predictions upon a doubling of organic matter inputs, while the variation in the response of the behavioural models was up to eight times larger for overparameterised models compared to models for which only identifiable parameters were optimised. Our results emphasise the necessity of optimising only identifiable model parameters to avoid hidden uncertainty in model predictions.

1 Introduction

When developing environmental models, including soil organic matter (SOM) models, the mechanistic understanding of ecosystem processes is translated into mathematical equations with multiple parameters. Such models are valuable research tools, and are used to test scientific hypotheses (e.g., Laub et al., 2024; Van de Broek et al., 2024; Tang and Riley, 2015), extrapolate ecosystem properties over larger spatial and temporal scales (e.g., Tao et al., 2023; Wieder et al., 2024), support policy decisions (e.g., Campbell and Paustian, 2015), and make predictions of ecosystem properties into the future under variable external forcings (e.g., Sulman et al., 2018; Pallandt et al., 2025). Model developers need to account for the complexity and high spatial variability in ecosystems properties (e.g., Nearing et al., 1999), which often cannot be quantified due to a lack of experimental and observational data across space and time (Schindler and Hilborn, 2015; Oreskes et al., 1994). As a consequence, there is a trade-off between model complexity and data availability that determines the overall model error (Van Rompaey

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

and Govers, 2002). On the one hand, a model that does not include the most basic processes relevant to the simulated system has limited scientific and practical use. On the other hand, a model including processes of which the parameter values have not been measured or reliably estimated will be similarly unreliable. This led authors of articles presenting mechanistic SOM models to acknowledge that their proposed model structures could not be sufficiently tested because of a lack of data (Riley et al., 2014; Dwivedi et al., 2017). Thus, complex models do not necessarily outperform more simple models (Jakeman and Hornberger, 1993; Manzoni and Porporato, 2009; Perretti et al., 2013; Lawrence et al., 2009), and striking a balance between model complexity and data availability is a prerequisite to develop environmental models that can be applied reliably (Manzoni and Schimel, 2024; Lennon et al., 2024; Famiglietti et al., 2021).

Three concepts that are central to achieving this balance are identifiability, equifinality and overparameterisation. The concept of identifiability has a long history in scientific research (Rothenberg, 1971; Reiersøl, 1950), with different aspects of identifiability being defined and studied (Travis and Haddock, 1981; Delforge, 1977; Cobelli et al., 1979; Kleissen et al., 1990; DiStefano and Cobelli, 1980). The two aspects of identifiability that have been most frequently applied to environmental models are structural identifiability and practical identifiability. The concept of structural identifiability was first introduced by Bellman and Åström (1970). As a practical definition, a structural identifiability analysis assesses whether a unique set of parameter values can be obtained given a mathematical model structure without consideration of the available data (therefore also termed *a priori identifiability*). For example, the Michaelis-Menten equation (both the forward and backward formulations) to simulate the depolymerisation of organic matter or the uptake of dissolved organic matter by microbes is widely used in SOM models (Chandel et al., 2023). This equation contains one parameter in the numerator (the maximum process rate, V_{max}) and one in the denominator (the half-saturation constant, K_m). This implies that an infinite number of combinations for V_{max} and K_m can compensate for each other and result in the same output, rendering these parameters non-identifiable when optimised together (Sierra et al., 2015; Holmberg, 1982; Marschmann et al., 2019). More recently, practical identifiability analysis (Lam et al., 2022), also termed *parameter identifiability analysis* (Guillaume et al., 2019), has gained importance. This aspect of identifiability analyses assesses whether a unique set of model parameters can be found given a combination of the model structure, available (calibration) data, and data uncertainty (therefore also termed *posteriori identifiability*). It is the latter type of identifiability that will be assessed in the present study. Detailed information is present in the literature about the concepts of structural identification (Nguyen and Wood, 1982; Beck, 1987; Walter and Pronzato, 1996) and practical identification (Lam et al., 2022; Guillaume et al., 2019), or both combined (Wanika et al., 2024; Miao et al., 2011; Raue et al., 2011).

The assessment of the practical identifiability of model parameters results in the identification of sets of parameters that can be optimised together, given the available data for parameter optimisation. When identifiable model parameters are optimised together, the values of these parameters do not, by definition, compensate for each other. This results in optimised parameters that show a limited range in values. However, often no single optimal parameter value can be obtained because of measurement variability inherent to environmental systems. Hence, there is generally a range in model outcomes, originating from a range in parameter values, that cannot easily be rejected given the variability in measurements. These models are referred to as

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

behavioural models (Beven, 2006). When non-identifiable parameter sets are optimised this is referred to here as *overparameterisation* (Sierra et al., 2015; Beck, 1987).

A direct consequence of overparameterisation is *equifinality*. In the context of environmental models, this concept has first been applied to hydrology (Beven and Binley, 1992) and focuses "attention on the fact that there are many acceptable [model] representations that cannot be easily rejected and that should be considered in assessing the uncertainty associated with predictions" (Beven, 2006, p 21). In practical terms, this means that an ecosystem property can be simulated with an acceptable accuracy given the variability in observed data (i.e., behavioural models), using different combinations of parameter values. While such parameter sets provide results that cannot be rejected, this equifinality contributes to additional uncertainty when overparameterised models are used to make predictions for the future or for other geographical regions. Equifinality is generally unavoidable in the context of environmental systems, as "even if we could define the 'perfect' model, it will still be subject to equifinality if driven with non-error-free initial and boundary conditions and compared with non-error-free output measurements" (Beven, 2006, p 21). Therefore, accounting for this phenomenon is important to quantify and minimise model uncertainty. More information on the concept of equifinality can be found in Beven (1993, 2002, 2006, 2007).

Similar to all environmental models, SOM models can be subject to equifinality and overparameterisation. During the past two decades, newly developed SOM models have largely moved away from first-order, sequential compartmental type models in which the turnover time of SOM is assumed to be governed by its chemical composition. Instead, there has been a move towards the incorporation of the emerging mechanistic understanding of SOM dynamics (Manzoni and Porporato, 2009; Campbell and Paustian, 2015), as was advocated to reduce uncertainty in predictions by SOM models (Schmidt et al., 2011; Bradford et al., 2016; Blankinship et al., 2018). Most notably, the inclusion of microbial dynamics and its effect on SOM cycling has increased (Chandel et al., 2023), together with non-linear processes (Le Noë et al., 2023). However, the incorporation of microbial characteristics in SOM models generally requires the parametrisation of processes that are difficult to measure in the field (Treseder et al., 2012). While SOM models thus became more mechanistic, their parameters are often "effective parameters" that represent multiple processes and cannot be directly measured (Beven, 2002). Therefore, despite the mechanistic character of these models, multiple parameters need to be calibrated rather than being derived from measurements.

The combination of the increase in the number of model parameters and the need for calibration underlines the need to account for practical identifiability (i.e., assessing how many and which parameters can be optimised together given available data) and equifinality (i.e., avoiding that multiple parameter combinations lead to behavioural models, without it being possible for the model user to know which parameter sets are more reliable than others) during model development and application. While this has been done in the past in other scientific fields (e.g., hydrology (Sorooshian and Gupta, 1983; Kleissen et al., 1990; Kelleher et al., 2013), soil erosion modelling (Brazier et al., 2000), biological modelling (Wu et al., 2008; Browning et al., 2020; Dankwa et al., 2022), and water quality modelling (Omlin et al., 2001)), only recently did this aspect of parameter optimisation gain importance in SOM models (e.g., Marschmann et al., 2019; Sierra et al., 2015; Ahrens et al., 2014; Van de Broek

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

et al., 2025; Luo et al., 2009; Abramoff et al., 2022; Meurer et al., 2020; Guo et al., 2022, and references in the next paragraph).

The consequences of the optimisation of combinations of non-identifiable model parameters and the resulting equifinality are evident in two primary ways for SOM models. First, when models are used to simulate SOM into steady state, different combinations in the size of model pools can result in an optimal simulation of total organic matter (Braakhekke et al., 2013), a clear manifestation of equifinality. This limits the use of such models to increase mechanistic understanding of SOM dynamics, as various processes will have a varying importance in different behavioural models. Second, when such models are used to make predictions of SOM dynamics into the future under different conditions, the predictions (starting off from 'perfect' behavioural models) can widely diverge (Luo et al., 2016, 2017; Guo et al., 2022). For example, it has been shown that models generally overpredict the turnover time of soil organic carbon (SOC) and thereby overestimate the potential of soil to increase their organic carbon (OC) stocks over the coming decades, Δ^{14}C data for SOC is included as a calibration constraint (He et al., 2016; Luo et al., 2019). The optimisation of non-identifiable parameters thus increases the uncertainty of predictions made by SOM models, something that does not help with instilling confidence in SOM models (Bradford et al., 2016).

Given the increase in complexity of SOM models, combined with the lack of attention for the consequences of equifinality for model predictions, the aim of this article is to increase awareness of this concept and to provide examples of how optimising combinations of non-identifiable parameters affects the uncertainty of predictions made by SOM models. First, we used four different mathematical formulations of a rhizosphere C and N model (i.e., a model simulating only soil microbes and particulate organic matter, based on the SESAM v3.0 model by Wutzler et al. (2022)) to illustrate different concepts and consequences of practical parameter identifiability and equifinality. Next, we apply these concepts to an adaptation of the SESAM v3.0 model, to which the protection of organic matter and the simulation of the Δ^{14}C value of SOC were added. This study was guided by the following research questions: (1) How many parameters are identifiable for a rhizosphere and SOM model, given different quantities of data? (2) How much do predictions made by an overparameterised model deviate from a well-constrained model? (3) Which data are necessary to minimise equifinality in SOM models?

2 Materials and methods

2.1 Overview of the analyses

The concepts used to assesssno the effect of overparameterisation and parameter equifinality on model predictions are illustrated using four models simulating soil C and N dynamics without mineral protection of SOC (based on the SESAM v3.0 model of Wutzler et al. (2022)), termed *rhizosphere models*. Next, these concepts are applied to a microbially-driven model simulating mineral protection of SOM and the Δ^{14}C value of the simulated SOC pools, referred to as the *SOM model*. To assess the identifiability of model parameters, i.e., which parameter combinations can be optimised together without parameters compensating for each other, realistic values for all parameters need to be perturbed by a very small amount. Such values were obtained for all models by performing a frequentist calibration using a Differential Evolution Markov Chain with snooker

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

updater (DEzs) algorithm, given as many constraints on simulated pools as realistically possible. Next, to assess how different amounts of available calibration data affect model simulations in steady state, every model was calibrated using a Bayesian approach (using the DE_{zs} algorithm) for parameter sets which were either identifiable (termed the identifiable parameter model; IPM) or non-identifiable (termed the full parameter model; FPM). Last, to show the effect of parameter equifinality on model predictions, steady-state model outcomes were perturbed by doubling OC inputs for a period of 100 years. All simulations and analyses were performed in R (R Core Team, 2025).

2.2 Artificial data

In this study is, artificial SOC data were used to calibrate model parameters. We created "measurements" of the total SOC stock and fractions for one point in time, to mimic the common assumption of an SOC stock in steady state. The total SOC stock down to 0.2 m depth was calculated assuming an SOC concentration of 2 % and a bulk density of 1.2 g cm^{-3}, resulting in an SOC stock of 4,800 g C m^{-2} down to 0.2 m. It was assumed that 25 % of SOC is POM and microbes in the rhizosphere (i.e., 1,200 g C m^{-2}), which was divided into 96 % POC (1,152 g C m^{-2}) and 4 % microbes (48 g C m^{-2}). The remaining SOC was assumed to be MAOC (3,600 g C m^{-2}). The standard deviation of the SOC pools was assumed to be 10 % of their size.

The C:N ratio of plant litter inputs, microbes and enzymes were fixed at 30, 10 and 3, following Wutzler et al. (2022). Assuming POM consists of 95 % plant litter and 5 % microbial residues, the C:N ratio of POM was calculated to be 29. Similarly, MAOM was assumed to consist of equal amounts of microbial residues and unprocessed plant-derived organic matter, resulting in a C:N ratio of 20.

The $\Delta^{14}C$ values of POC and MAOC were estimated using data for density-fractionated forest soil from the ISRAD database (Lawrence et al., 2020). Selected $\Delta^{14}C$ data were limited to samples collected in the top 20 cm of the soil between 2004 and 2009 in the northern hemisphere. This resulted in $\Delta^{14}C$ values of 75.2 ‰ for POC, and 17.6 ‰ for MAOC, with the standard deviation for both values being assumed to be 10 % of their size. The average year at which these values were measured was 2007, which was taken to be the final year of the performed simulations.

2.3 Conceptual models

Two SOM models were used to assess the effect of parameter equifinality on model predictions (Fig. 1). The first is a microbially-driven model simulating coupled C and N dynamics without mineral protection of OC based on the SESAM v3.0 model (Wutzler et al., 2022), referred to as the rhizosphere model (RM). For this model, four different sets of mathematical equations were used to simulate depolymerisation of POM and microbial turnover, resulting in four rhizosphere models. The second model, termed the SOM model, is identical to the SESAM model but additionally simulates mineral protection of OC (Fig. 1). The processes are identical to the rhizosphere model, with the addition that DOM can be stabilised on soil minerals. Competition for DOM between microbes and minerals is simulated using an adapted version of the equilibrium chemistry approximation (Tang and Riley, 2013). The SESAM model was selected because it was developed with consideration for trade-offs between model complexity and data availability (Wutzler et al., 2022), the topic of this study.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Simulated inputs of organic matter (C and N) to the soil come from plant litter, with a fixed C:N ratio. In addition, OC enters the soil through rhizodeposits, which are assumed to contain no N, and inputs of mineral N come from atmospheric N deposition. The SESAM model keeps track of the size of four pools: POM consisting of microbial residues and deactivated enzymes (POM_{res}) and plant litter (POM_{lit}), microbes (MIC) and mineral N (N_{min}). In addition, dissolved organic matter (DOM) and enzymes mediating the depolymerisation of plant litter (ENZ_{lit}) and microbial residues (ENZ_{res}) are simulated without explicitly keeping track of their size, assuming they are in quasi-steady-state (i.e., their size determines the rate of depolymerisation, but they are instantly transferred to other pools upon their creation). The relative production of these enzymes is calculated in proportion to their revenue (Wutzler et al., 2022). Inputs of C and N from plant litter enter the POM_{lit} pool, while rhizodeposit C enters the DOM pool. Nitrogen leaves the system through leaching, while C leaves the system as CO_2 through microbial maintenance respiration, overflow respiration, growth respiration, and CO_2 losses upon the turnover of death microbes through grazing by predators. All model pools contain both C and N, except N_{min}, and are referred to as organic matter pools, for example POM_{lit}. When referring to the C content of a pool, they are referred to accordingly, for example POC_{lit}. Death microbes are transferred to the POM_{res} pool, while decayed enzymes are divided between the POM_{res} and DOM pools. A full description of SESAM v3.0 is provided in Wutzler et al. (2022).

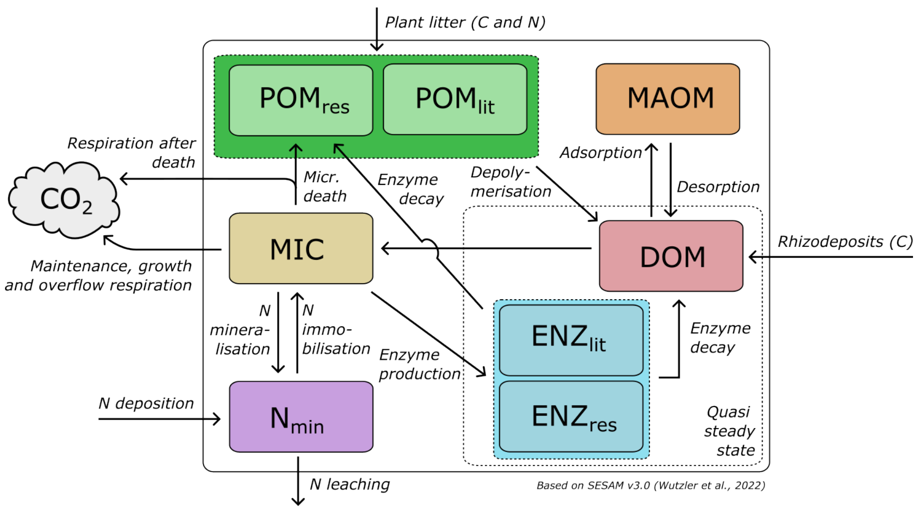

Figure 1. Illustration of the rhizosphere model and the soil organic matter (SOM) model. The rhizosphere model is the same as the SOM model, but without the simulation of the MAOM pool. In the rhizosphere model, *DOM* is not explicitly simulated, while this is done for the SOM model. The rhizosphere model is the SESAM v3.0 model (Wutzler et al., 2022).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

2.4 Mathematical models

2.4.1 The rhizosphere models

The amount of SOM was simulated down to 0.2 m for a unit surface area of 1 m^2, and expressed as g C m^{-2} or g N m^{-2}. The differential equations were solved using the *ode* solver from the deSolve package in R (Soetaert et al., 2010) with a time step of 1 day. Simulations were performed for a time span over which the model pools were shown to have reached steady state, being 500 years for the rhizosphere models, and 2000 years for the SOM model. The equations describing C and N transformations are identical to the SESAM v3.0 model, of which a detailed description can be found in Wutzler et al. (2022). Annual total OC inputs were calibrated for the SOM model (see Sect. 2.5), to obtain a correct combination of the size of the combined *POC* pools and its turnover rate, determined by optimising its Δ^{14}C value. This resulted in OC inputs of 349 g C m^{-2} yr^{-1}, of which 80 % was assumed to enter the soil as plant litter, and the remaining 20 % as rhizodeposits. Organic N enters the soil through plant litter, assuming a constant C:N ratio of 30, while rhizodeposits are assumed to only contain C. Mineral N is added to the soil through atmospheric N deposition (0.7 g N m^{-2} yr^{-1}).

For the rhizosphere models, an overview of the state variables and parameters of the rhizosphere model is presented in Tables S1 and S2, respectively. Four different sets of equations were used to assess how different mathematical formulations of the same conceptual model influence parameter identifiability and consequently the consistency in model predictions. These combined either one of two variations of (1) the Michaelis-Menten equations for depolymerisation of litter (POM_{lit} and POM_{res}) and (2) microbial turnover (Table 1). The two variations for the depolymerisation of *POM* are referred to as the absolute and relative variants. The absolute variant is formulated as:

$$Depolymerisation_{lit} = V_{max,lit} \cdot POC_{lit} \cdot \frac{\alpha_L \cdot a_E \cdot MIC}{K_{mN} + (\alpha_L \cdot a_E \cdot MIC)} \tag{1}$$

$$Depolymerisation_{res} = V_{max,res} \cdot POC_{res} \cdot \frac{\alpha_R \cdot a_E \cdot MIC}{K_{mN} + (\alpha_R \cdot a_E \cdot MIC)} \tag{2}$$

Where $V_{max,lit}$ and $V_{max,res}$ are the maximum rates of depolymerisation of POC_{lit} and POC_{res} per time step (1 day), respectively, K_{mN} is the half-saturation constant (g C m^{-3}), α_L and α_R are the proportions of total microbial investment in enzymes to depolymerise litter and residues (unitless, both values add up to 1), respectively, a_E the portion of microbial biomass invested in total enzyme production (unitless), and MIC OC in the microbial biomass pool (g C m^{-2}). These formulations, referred to as "reverse Michaelis-Menten" kinetics (Schimel and Weintraub, 2003), are identical to the original SESAM v3.0 model, and imply that the amount of depolymerisation per time step is limited by the absolute amount of extracellular enzymes. Consequently, these equations are referred to here as absolute Michaelis-Menten (MM_{abs}). They have been used in, among others, the COMISSION model (Ahrens et al., 2015, 2020), MIMICS (Wieder et al., 2015), and Millennial v2 (Abramoff et al., 2022), and have been shown to be the preferred formulation to represent depolymerisation of POM, compared to "forward Michaelis-Menten kinetics" (Tang and Riley, 2019).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

In the second variant of these equations, the rate modifier is based on ratios of model pools. These are formulated as:

$$Depolymerisation_{lit} = V_{max,lit} \cdot POC_{lit} \cdot \frac{\alpha_L \cdot a_E \cdot \frac{MIC}{POM_{lit}}}{K_{mN.lit} + \alpha_L \cdot a_E \cdot \frac{MIC}{POC_{lit}}} \tag{3}$$

$$Depolymerisation_{res} = V_{max,res} \cdot POC_{res} \cdot \frac{\alpha_R \cdot a_E \cdot \frac{MIC}{POM_{res}}}{K_{mN.res} + \alpha_R \cdot a_E \cdot \frac{MIC}{POC_{res}}} \tag{4}$$

Where $K_{mN.lit}$ is the half-saturation constant for depolymerisation of POM_{lit} (unitless, equivalent to the ratio of enzymes ($\alpha_L \cdot a_E \cdot MIC$) relative to POM_{lit}), and $K_{mN.res}$ being the half-saturation constant for depolymerisation of POM_{res} (unitless, equivalent to the ratio of enzymes ($\alpha_R \cdot a_E \cdot MIC$) relative to POM_{res}). The rate of depolymerisation (i.e., the portion of *POM* depolymerised per time step) is thus limited by the ratio of *MIC* to *POM*, rather than by the absolute amount of *MIC*. Therefore, these formulations are referred to as relative Michaelis-Menten (MM_{rel}), and imply that the more microbes (and thus extracellular enzymes) are available per unit of POM, the larger the portion of POM that will be depolymerised. Similar formulations have been used in CORPSE-N (Sulman et al., 2019), AMPSOM (Tougma et al., 2025) and SOILcarb (Van de Broek et al., 2025).

Two different approaches to simulate microbial turnover were tested: first-order decay (fo) and density-dependent turnover (DD). First-order microbial turnover per time step is simulated as:

$$Mortality = k_{mic.fo} \cdot MIC \tag{5}$$

Where $k_{mic.fo}$ is the portion of microbes turning over per time step (1 day). This formulation is common in microbially-driven SOC models (e.g., Allison et al., 2010; Riley et al., 2014; Woolf and Lehmann, 2019; Yu et al., 2020; Laub et al., 2024). Density-dependent microbial turnover (Buchkowski et al., 2017) is formulated using a logistic growth model (the Verhulst equation or Verhulst-Pearl equation) expressing the rate of change in microbial biomass per unit of time:

$$\frac{dMIC}{dt} = r \cdot MIC \cdot (1 - \frac{MIC}{K_{mic}}) \tag{6}$$

Where *MIC* is the microbial biomass (g C m^{-2}), *t* the time (d), *r* the growth rate (g C m^{2} d^{-1}) and K_{mic} the carrying capacity for *MIC* (g C m^{-2}). The latter is expressed as a portion (f_K) of total POC (the sum of POC_{lit} and POC_{res}). Equation 6 can then be reformulated as:

$$\frac{dMIC}{dt} = r \cdot P - \frac{r \cdot MIC^2}{f_k \cdot (POC_{lit} + POC_{res})} \tag{7}$$

This formulation is equivalent to the formulation of density-dependent microbial turnover in Georgiou et al. (2017), and shows that the loss of microbial biomass per unit of time (second term on the right-hand side) is a function of the square of microbial biomass. The use of this equation implies that the carrying capacity of microbial biomass (K_{mic}) needs to be known or calibrated, instead of the explicit turnover rate. Similar formulations have been shown to lead to less severe oscillations in simulated SOC stocks over time (Georgiou et al., 2017), and have been implemented in, for example, the versions of ReSOM (Tang and Riley, 2015) used in Sulman et al. (2018) and Van de Broek et al. (2024). The combinations of both approaches to simulate depolymerisation and uptake, on the one hand, and microbial turnover, on the other hand, led to four different formulations of the rhizosphere model (Table 1).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Table 1. Overview of the equations used in the four rhizosphere models (RM). MM_{abs} and MM_{rel} refer to absolute and relative Michaelis-Menten kinetics, respectively, fo and DD refer to first-order and density-dependent microbial mortality, respectively. The parameters are explained in Table S2.

RM1 (MM_{abs}_fo)	
$Depolymerisation_{lit}$	$V_{max,lit} \cdot POC_{lit} \cdot \frac{\alpha_L \cdot a_E \cdot MIC}{K_{mN}+(\alpha_L \cdot a_E \cdot MIC)}$
$Depolymerisation_{res}$	$V_{max,res} \cdot POC_{res} \cdot \frac{\alpha_R \cdot a_E \cdot MIC}{K_{mN}+(\alpha_R \cdot a_E \cdot MIC)}$
Microbial turnover	$k_{mic.fo} \cdot MIC$
RM2 (MM_{abs}_DD)	
$Depolymerisation_{lit}$	$V_{max,lit} \cdot POC_{lit} \cdot \frac{\alpha_L \cdot a_E \cdot MIC}{K_{mN}+(\alpha_L \cdot a_E \cdot MIC)}$
$Depolymerisation_{res}$	$V_{max,res} \cdot POC_{res} \cdot \frac{\alpha_R \cdot a_E \cdot MIC}{K_{mN}+(\alpha_R \cdot a_E \cdot MIC)}$
Microbial turnover	$\frac{r \cdot MIC^2}{K_{mic}}$
RM3 (MM_{rel}_fo)	
$Depolymerisation_{lit}$	$V_{max,lit} \cdot POC_{lit} \cdot \frac{\alpha_L \cdot a_E \cdot \frac{MIC}{POC_{lit}}}{K_{mN.lit}+\alpha_L \cdot a_E \cdot \frac{MIC}{POC_{lit}}}$
$Depolymerisation_{res}$	$V_{max,res} \cdot POC_{res} \cdot \frac{\alpha_R \cdot a_E \cdot \frac{MIC}{POC_{res}}}{K_{mN.res}+\alpha_R \cdot a_E \cdot \frac{MIC}{POC_{res}}}$
Microbial turnover	$k_{mic.fo} \cdot MIC$
RM4 (MM_{rel}_DD)	
$Depolymerisation_{lit}$	$V_{max,lit} \cdot POC_{lit} \cdot \frac{\alpha_L \cdot a_E \cdot \frac{MIC}{POC_{lit}}}{K_{mN.lit}+\alpha_L \cdot a_E \cdot \frac{MIC}{POC_{lit}}}$
$Depolymerisation_{res}$	$V_{max,res} \cdot POC_{res} \cdot \frac{\alpha_R \cdot a_E \cdot \frac{MIC}{POC_{res}}}{K_{mN.res}+\alpha_R \cdot a_E \cdot \frac{MIC}{POC_{res}}}$
Microbial turnover	$\frac{r \cdot MIC^2}{K_{mic}}$

2.4.2 The soil organic matter model

The inputs of C and N in the SOM model were equal to the inputs simulated in the rhizosphere models. The equations for the SOM model were chosen from the rhizosphere model with relative Michaelis-Menten depolymerisation and density-dependent microbial turnover (RM4, see Section 2.4.1). An overview of the state variables and parameters of the SOM model is presented in Table S3 and S4, respectively. Depolymerisation of POM_{lit} and POM_{res} are simulated using Eq. 3 and 4, respectively, and the mortality of microbes as the last term on the right-hand side of Eq. 7.

Competition for *DOM* between microbes (for uptake) and soil minerals (for adsorption) is simulated using the equilibrium chemistry approximation (ECA; Tang and Riley, 2013). This approach partitions a substrate between two enzymes using the affinity of both enzymes for the substrate. Because not all substrate is partitioned between the sinks in a single time step the DOM pool was explicitly simulated to keep track of its size, in contrast to the rhizosphere models and the original SESAM

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

v3.0 model. The implementation of ECA kinetics follows Tang and Riley (2013):

$$Uptake_{ECA} = MIC \frac{DOM}{K_{m,U}(1+\frac{SURF_{rhizo}}{K_{m,ads}}+\frac{MIC}{K_{m,U}})+DOM} \tag{8}$$

$$Adsorption_{ECA} = SURF_{rhizo} \frac{DOM}{K_{m,ads}(1+\frac{SURF_{rhizo}}{K_{m,ads}}+\frac{MIC}{K_{m,U}})+DOM} \tag{9}$$

Where $K_{m,U}$ is the affinity constant for microbial uptake of *DOM* (g C m^{-3}), $K_{m,ads}$ the affinity constant for OM adsorption (g C m^{-3}), and $SURF_{rhizo}$ the amount of available surfaces for OM adsorption in the rhizosphere (g C m^{-3}). The latter was calculated by subtracting the simulated amount of *MAOC* from the SOC stabilisation potential calculated using the clay+silt content (assumed to be 50 %) following (Georgiou et al., 2022), and multiplying this value with the volume of the rhizosphere to account for the fact that not all soil minerals are in touch with inputs of C and N from roots. This was assumed to be 10 % of total soil volume (based on calculations by Finzi et al. (2015)) at the initiation of the simulation. After the simulated OC inputs were doubled (see Sect. 2.8), also $SURF_{rhizo}$ was doubled, assuming OC inputs increase proportional to root biomass.

Desorption of OM from minerals was simulated as a first-order process:

$$Desorption = k_{des} \cdot MAOM \tag{10}$$

With k_{des} being the desorption rate (d^{-1}).

2.4.3 Simulation of Δ^{14C}

For the SOM model, the Δ^{14}C value of SOC was simulated to evaluate how including data on the Δ^{14}C values of measurable model pools (i.e., *POC* and *MAOC*) affects the identifiability of model parameters. The dataset for the annual Δ^{14}C value of atmospheric CO_2 in the northern hemisphere compiled by Van de Broek et al. (2025) (using data from Reimer et al. (2013); Hua et al. (2013) and Hammer and Levin (2017)) was used, and a lag time of 4 years between the incorporation of atmospheric $^{14}CO_2$ in plant biomass and the addition of this biomass to the SOC pool was used, following previous modelling studies (Ahrens et al., 2014; Schrumpf and Kaiser, 2015). No kinetic fractionation of ^{14}C relative to ^{12}C was assumed during the transfer of ^{14}C between simulated pools, with radioactive decay being the only mechanism leading to additional losses of ^{14}C compared to ^{12}C. Values of Δ^{14}C were calculated assuming the OC had a δ^{13}C value of -28 ‰, and that soil samples were collected in 2007 (see Section 2.2).

2.5 Deterministic parameter calibration

To evaluate the sensitivity and identifiability of model parameters, it was necessary to have a set of parameter values that leads to realistic model results with the model output being sensitive to local changes in parameter values. Therefore, a deterministic calibration was performed using a differential evolution (DE) algorithm (Mullen et al., 2011; Ardia et al., 2011). For the rhizosphere models, between 3 and 5 parameters were optimised (depending on model variant, Table S5), while for the SOC model, 9 parameters were optimised (Table S6). The parameter values were optimised by minimising the sum of squared relative errors (*SSRE*), formulated as:

$$SSRE = \sum_{i=1}^{n} \Big(\frac{meas_i - mod_i}{meas_i} \Big)^2 \tag{11}$$

Where *n* is the number of measurements and $meas_i$ and mod_i are the measured and modelled pool sizes.

For the rhizosphere models, artificial measurements of POC, MIC and the C:N ratio of total SOM were used to optimise model parameters with a population size of 180 parameter sets for 500 iterations. To make sure that the rate modifiers of the Michaelis-Menten equations for depolymerisation of microbial residues and litter were sensitive to changes in the size of model pools, their artificial measured value was assumed to be 0.5, and the *SSRE* calculated accordingly. For the rhizosphere model with first-order microbial mortality, the decay constant ($k_{mic.fo}$) was kept constant at a value of 0.01 d^{-1}, to simulate a turnover rate of the microbial biomass of 100 days. For the rhizosphere models with density-dependent microbial turnover, turnover rate of microbes was calibrated to be as close as possible to 100 days. The optimised parameter values for the rhizosphere models are shown in Table S7.

For the SOM model, artificial measurements of POC, MAOC, total OC, the C:N values of POM and MAOM, and the $\Delta^{14}C$ values of POC and MAOC were used to optimise model parameters with a population size of 180 parameter sets for 500 iterations. Similar to the rhizosphere model, the rate modifiers for the Michaelis-Menten equations simulating depolymerisation of litter and microbial residues were optimised to be as close to 0.5 as possible. The turnover rate of microbes was calibrated to be as close as possible to 100 days. The optimised parameter values for the SOM model are shown in Table S8.

2.6 Parameter identifiability analysis

To assess which parameter combinations were identifiable, given different available data sets for parameter optimisation, the methods developed by Brun et al. (2001) were applied using the *FME* package in R (Soetaert and Petzoldt, 2010). For the rhizosphere models, three scenarios of data availability were tested: (1) only data on total SOC, (2) data on total SOC and total N, and (3) data on POC, MIC and the C:N ratio of total SOC. The first scenario is one with a minimum data availability, while the second scenario is considered the most common (as C and N are often measured together). The third scenario assumes that in addition to total C and N, also microbial biomass C and POC were measured. Also for the SOM model, three scenarios of data availability were tested: (1) data on total SOC and N, (2) data on POC, MAOC and their C:N ratios, and (3) data on POC, MAOC, their C:N ratios, and the $\Delta^{14}C$ values of POC and MAOC. As the identifiability analysis needs to be performed with a parameter set that leads to an acceptable model output (Brun et al., 2001), the identifiability analysis was performed using the optimal parameter sets obtained for the different models from the deterministic DE calibration (see Sect. 2.5).

To perform the identifiability analysis, the local sensitivity of selected model state variables to variations in parameter values was quantified using a normalised, dimensionless sensitivity index (Brun et al., 2001):

$$s_{ij} = \frac{\partial y_i}{\partial \Theta_j} \frac{w_{\Theta_j}}{w_{y_i}} \tag{12}$$

Where y_i is the i^{th} state variable, Θ_j is the j^{th} model parameter, w_{Θ_j} is a weighting factor for Θ_j,and w_{y_i} is a weighting factor for y_i. The obtained sensitivity matrix (*s*) thus quantifies the rate of change in the value of model output y_i for a small

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

change in the value of parameter Θ_j. The weighting factor for the parameter value (w_{Θ_j}) was set to the optimal parameter value obtained by the DE optimisation, while the weighting factor for the state variable (w_{y_i}) is the steady-state size of this pool obtained using the optimal parameter values from the DE optimisation. The weights were thus constant for every assessed combination of state variable and parameter.

This local sensitivity analysis as implemented in the *FME* package in R (Soetaert and Petzoldt, 2010) evaluates the sensitivity for a time series of model outputs and respective measurements. As we assumed to only have steady-state measurements at one point in time, the sensitivity of the model output could only be evaluated for this time point. To do so, the sensitivity analysis from the *FME* package was run multiple times, each iteration changing the parameter values with a different relative amount (from 1×10^{-4} to 5×10^{-4} in steps of 5×10^{-5}). These results were then combined to construct the sensitivity matrix (Eq. 12), with rows containing the sensitivity indexes (s_{ij}) quantifying how varying parameters by a different relative amount affected the selected model output, and columns showing this for different parameters. One alteration made to the sensitivity analysis from the *FME* package (in the function *sensFun*) is that the tested parameter values were not restricted to be larger than the relative deviation (as was implemented by Soetaert and Petzoldt (2010)), to make sure the proposed range in parameter values was tested and not replaced by the relative deviation.

The sensitivity matrices were subsequently used to assess the parameter identifiability through the analysis of collinearity between all possible combinations of columns (i.e., parameters). As every column quantifies how every evaluated model state variable changes when a parameter value is altered, columns having a high collinearity indicate parameter combinations that are not identifiable, as a change in one parameter can be compensated by a change in one or more other model parameters. Before the collinearity was assessed, every column in the sensitivity matrix was normalized using the Euclidean norm (i.e., the square root of the sum of the squared values), following Brun et al. (2001):

$$\hat{s_j} = \frac{s_j}{\|s_j\|} \tag{13}$$

Where $\hat{s_j}$ is the normalised column for the j^{th} parameter in s_{ij}, s_j is the original column, and $\|s_j\|$ is the Euclidean norm of column s_j. Combining these columns for all parameters resulted in the normalised sensitivity matrix $\hat{S}$. For every combination of columns (i.e., parameters) in $\hat{S}$, a collinearity index was calculated (Brun et al., 2001):

$$\gamma = \frac{1}{\sqrt{min(EV[\hat{S}^T\hat{S}])}} \tag{14}$$

Where γ is the collinearity index, $\hat{S}^T\hat{S}$ is the vector product of $\hat{S}$ and the transposed vector $\hat{S}$, and $min(EV)$ is the smallest eigenvalue of this product. The interpretation of γ is that a change in one parameter can be compensated by 1 - 1/γ when one or more parameters are appropriately changed. To label a parameter set as being identifiable, we use a threshold of γ of 10, meaning that a parameter change can be undone for 90 % by changes in other parameter values. It is noted that various studies have used different values for this threshold, generally in the range $5 < \gamma < 20$ (Brun et al., 2001). A detailed description of the method presented in this section is provided in Brun et al. (2001), and examples of its application are shown in, among others, Omlin et al. (2001), Brun et al. (2002) and Sierra et al. (2015).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

2.7 Bayesian parameter calibration

Parameter optimisation was performed to obtain as many parameter combinations as possible that produced behavioural models, i.e., model outputs that cannot be readily rejected given available data. These are defined here as predictions within one standard deviation of the assumed measurements. To account for measurement uncertainty, a Bayesian calibration was performed using the Differential Evolution Markov Chain with snooker updater (DEzs) algorithm (ter Braak and Vrugt, 2008), as implemented in the *BayesianTools* package in R (Hartig et al., 2023). As no prior information on the distribution of parameter values was known, uniform priors were used within specified bounds (see Table S6) and the log-likelihood was calculated as (Vrugt, 2016):

$$\mathcal{L}(\mathbf{x}|\tilde{\mathbf{Y}},\hat{\sigma}^2) = -\frac{n}{2}\log(2\pi) - \sum_{i=1}^{n}\log(\hat{\sigma_i}) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\tilde{y_i} - y_i(\mathbf{x})}{\hat{\sigma_i}}\right)^2 \tag{15}$$

Where n is the number of observations, $\hat{\sigma_i}$ is the standard deviation of the i^{th} observation, $\tilde{y_i}$ is the i^{th} observation, and $y_i(\mathbf{x})$ is the model prediction of this observation.

For the rhizosphere models the DE_{zs} algorithm was run using 5 internal chains, higher than the three internal chains which have been shown to be sufficient to explore a high dimensional parameter space (ter Braak and Vrugt, 2008). This was done 3 times (i.e., with three independent chains), to ensure an optimal exploration of the parameter space. Each internal chain was run for 20,000 iterations. The Bayesian calibrations of the SOM model were run with a number of internal chains equal to twice the number of optimised parameters (Table 2) with 20,000 iterations each, repeated 3 times to obtain 3 independent chains. For these calibrations, the chance of a snooker jump was increased to 0.2 (from the default 0.1), to enhance exploration of the parameter space and avoid the algorithm getting stuck in a local maximum.

The four rhizosphere models were used to illustrate the concept of parameter identifiability and the consequences of equifinality on model predictions. To do so, each of the rhizosphere models was calibrated twice. In a first scenario (termed the full parameter model; FPM) all parameters for which no realistic estimates could be made or found in the literature were optimised (Table S9). These parameter sets were non-identifiable. In a second scenario (termed the Identifiable parameter model; IPM), only identifiable parameters for the assumed data were optimised: $V_{max,lit}$ and $V_{max,res}$ (see Tables S11, S13, S15 and S17). In addition to these parameter values, also the rate modifiers for the depolymerisation of POM_{lit} and POM_{res} (see Eq. 1 and 2) were optimised to be as close as possible to 0.5. For both optimisation scenarios, it was assumed that only measurements of total SOC and the C:N ratio were available, as these data are most commonly available.

For the SOM model, Bayesian parameter optimisation was performed for three assumptions on data availability: (1) data on total SOC and N (referred to as OM_{tot}), (2) data on POC, MAOC, particulate N and mineral-associated N (which can be obtained through SOM fractionation, referred to as Fractions), and (3) data on POC, MAOC, particulate N, mineral-associated N and the Δ^{14}C value of POC and MAOC (referred to as Fractions & Δ^{14}C). For each of these data sets, one full parameter model (FPM) was run by optimising eight parameters for which no reasonable estimate could be made (Table 2), and one identifiable parameter model (IPM), optimising as many parameters as could be jointly identified: two for the OM_{tot} scenario, 3 for the Fractions scenario, and 5 for the Fractions & Δ^{14}C scenario (Table 2).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Table 2. Optimised identifiable parameters during the Bayesian calibration of the identifiable parameter model (IPM) version of the soil organic matter (SOM) model for the three scenarios: OM_{tot} (calibration using only data on total organic carbon (OC) and N), Fractions (calibration based on the OC and N content of particulate (POM) and mineral-associated organic matter (MAOM)) and Fractions & $\Delta^{14}C$ (calibration based on the OC and N content and $\Delta^{14}C$ value of POM and MAOM). The columns show all parameters that were selected as requiring optimisation in the full parameter model (FPM). Information about the parameter values is presented in Table S4

Scenario	$V_{max,lit}$	$V_{max,res}$	$K_{mN.lit}$	$K_{mN.res}$	$K_{m,U}$	$K_{m,ads}$	k_{des}	f_K
IPM - OM_{tot}		X				X		
IPM - Fractions	X	X				X		
IPM - Fractions & $\Delta^{14}C$	X	X				X	X	X
Full parameter model (FPM)	X	X	X	X	X	X	X	X

2.8 Disturbing the steady state solution

All results obtained by the Bayesian calibration were within one standard deviation from the average respective measurements (behavioural models). This implies that none of these models (and their parameter sets) could readily be rejected given the variability in measurements. To assess how different sets of optimised parameter combinations, i.e., identifiable versus non-identifiable, affect model predictions starting from steady-state pool sizes, OC inputs were doubled for a period of 100 years. Subsequently, the amount of simulated total SOC, N and the $\Delta^{14}C$ of the simulated pools were analysed after 100 years. The additional uncertainty caused by parameter equifinality was assessed for the different scenarios by quantifying the absolute value and spread of model predictions upon the doubling of OC inputs.

3 Results

3.1 Rhizosphere models

3.1.1 Parameter identifiability

The parameter identifiability analysis for all four rhizosphere models showed that the number of identifiable parameters was limited, and increased when more data was available for parameter optimisation (Fig. 2, Tables S11 - S18). When only total OC data were available, none of the four models had two parameters that could be simultaneously identified. For the scenario where data on both total OC and N was available, maximum two parameters could be identified together. Three parameters were jointly identifiable only when total OC, N and microbial biomass data were present. This shows that even for models developed with trade-offs between model complexity and data availability in mind (Wutzler et al., 2022), the number of parameters that can be optimised without overparameterisation is well below the total number of unknown model parameters (five or six, depending on the model formulation; Table S9).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

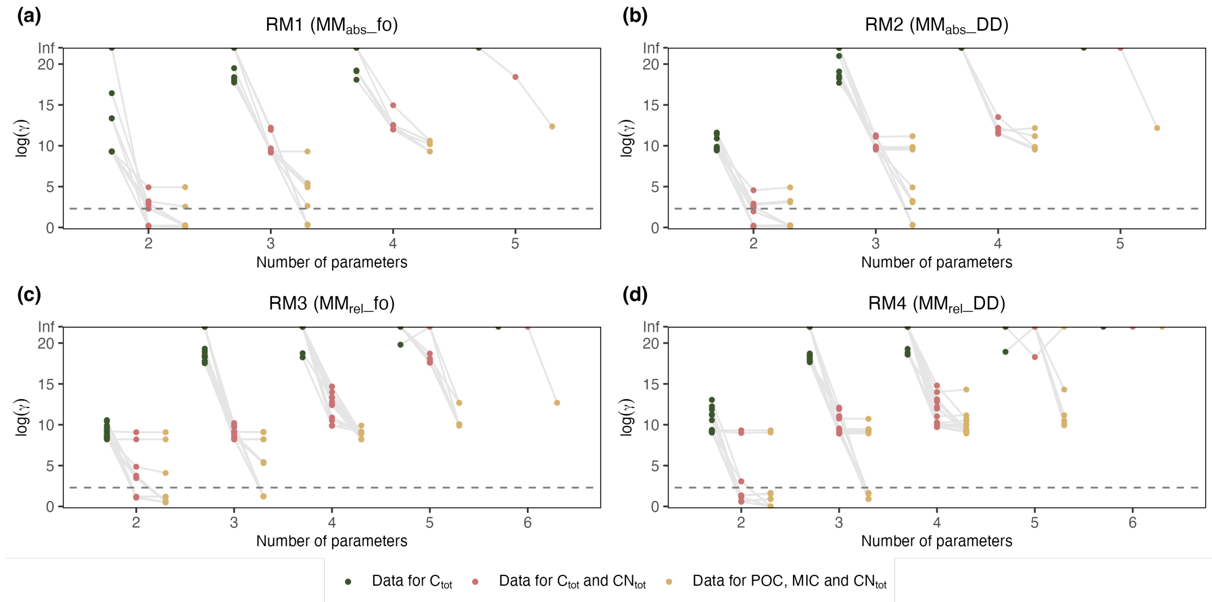

Figure 2. Logarithm of the collinearity index (γ) for all parameter combinations of the four rhizosphere models, for the scenarios where (1) only data on total SOC were available (green dots), (2) data total SOC and N were available (red dots) and (3) data on total SOC, N and microbial biomass C were available (yellow dots). The same parameter combinations are connected with grey lines. The dotted horizontal line shows the threshold in γ (log(10) = 2.3) below which parameter combinations are considered to be identifiable. The identifiable parameter combinations are shown in Tables S11 - S18.

3.1.2 The effect of overparameterisation on equifinality in rhizosphere model predictions

Disturbing the steady-state solution of the rhizosphere models by forcing a doubling of OC inputs revealed the difference in model evaluation between steady-state simulations and predictions made with the same models. The Bayesian optimisation of all rhizosphere models to steady state (the simulations at year 500 in Fig. 3) successfully produced behavioural models, of which the simulation of total OC fell within the measurement uncertainty range. If no further simulations were performed, all these model outcomes would therefore have been considered highly accurate and precise. However, predictions differed between the models when OC inputs were doubled.

Predictions by the identifiable parameter models showed that RM1 and RM2, with absolute Michaelis-Menten kinetics, predicted a lower average relative increase in OC (61.1–70.7 %; Fig. 3 c,f) compared to RM3 and RM4, with relative Michaelis-Menten kinetics (100 %; Fig. 3 i,l). This difference is due to the formulation of the rate modifiers for depolymerisation (Table 1). For RM1 and RM2, the simulated fraction of POC being depolymerised is determined by the absolute size of the microbial pool, which doubled upon a doubling of OC inputs for RM1 and increased by 67.2–70.7 % for RM2 (Fig. S1 a,c). This caused an increase in the rate modifier for depolymerisation of POC_{lit} upon a doubling of OC inputs (Fig. S1 b,d), leading to a higher fraction of POC_{lit} being depolymerised compared to the initial OC inputs. In contrast, the pools controlling the value of the

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

rate modifier for POC_{lit} in RM3 and RM4 (the ratio of MIC to POC_{lit}) remained constant upon a doubling of OC inputs, causing the rate modifier to be identical before and after a doubling of OC inputs (Fig. S1 e–h). As a result, the same fraction of POC_{lit} was depolymerised per time step upon a doubling of OC inputs. This difference shows the effect of mathematical formulations on model predictions, using either absolute or relative Michaelis-Menten rate modifiers.

Also the number of optimised parameters had an effect on predictions by the rhizosphere models. For RM1 and RM2 (with absolute Michaelis-Menten kinetics), the average relative increase in OC upon a doubling of inputs was lower for the full parameter models than for the identifiable parameter models (Fig. 3). In contrast, for RM3 and RM4 (with relative Michaelis-Mention kinetics), overparameterisation did not affect the average increase in OC, which was ca. 100 %. Moreover, compared to the identifiable parameter models, the full parameter models showed a prediction range approximately 42 times larger for RM1 and 7 times larger for RM2. These results show that the accurate prediction of steady-state stocks of SOC by behavioural models is not a sufficient criterion to evaluate the predictive capabilities of such models.

3.2 Soil organic matter model

3.2.1 Parameter identifiability

The parameter identifiability analysis for the SOM model showed that, similar to the results for the rhizosphere models, the number of identifiable parameters increased with an increasing quantity of calibration data. When only data on total SOC and N were used, at most two parameters were jointly identifiable, while three parameters were identifiable together when data on the fractions of POC, MAOC and their N content were used. The number of identifiable parameters increased to five when, in addition to C and N data on the POM and MAOM fractions, also the Δ^{14}C value of these pools were used. Also here, this analysis shows that even for the scenario with the most data, not all model parameters were identifiable.

3.2.2 The effect of overparameterisation and equifinality on steady-state models

The Bayesian parameter optimisations for the SOM model resulted in behavioural models for all calibration scenarios, as total SOC was predicted to be within one standard deviation of the average measurement at steady state (Fig. 5). However, the calibration scenario had a substantial impact on the internal dynamics of the model. For example, when model parameters were optimised using only data on total SOC and N while calibrating all model parameters, the majority of SOC could be either in *POC* or *MAOC* (Fig. 5b), clearly demonstrating equifinality resulting from overparameterisation. Further evidence for equifinality is present when looking at the simulated Δ^{14}C values. These are determined by the turnover rate of the respective pools, and provide a better understanding of the temporal dynamics of the model pools (Fig. 6). First, none of the scenarios lacking Δ^{14}C data of *POC* and *MAOC* resulted in the correct simulation of the Δ^{14}C value of these pools as measured in 2007 Fig. 6a–d). Second, the overparameterised models (FPM) without Δ^{14}C data being used for calibration showed large variations in the temporal trend of simulated Δ^{14}C values of model pools (Fig. 6b,d). This shows that the simulated turnover rate of the model pools differed substantially among the behavioural models. Third, only when Δ^{14}C data for POC and MAOC were used as a calibration constraint was the turnover time of the POC and MAOC pools correctly simulated (Fig. 6e,f).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

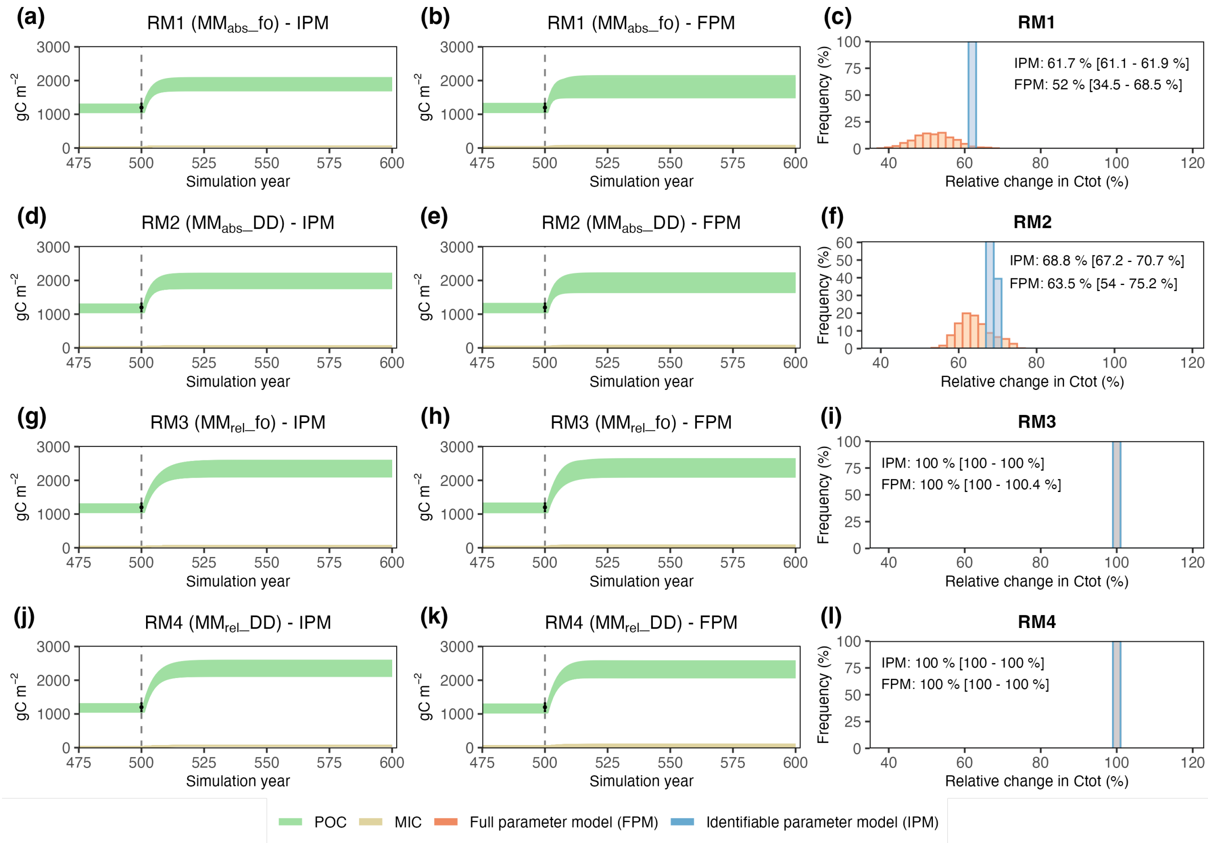

Figure 3. The effect of parameter identifiability and resulting equifinality on the prediction of SOC by the four rhizosphere models (RMs). Each row shows the results for a different model (see Table 1). The first column shows the results from a Bayesian optimisation of two identifiable parameters (the Identifiable Parameter Model: IPM), while the second column shows the results from a Bayesian optimisation of five (RM1 and RM2) or six (RM3 and RM4) parameters (the Full Parameter Model: FPM). Note that the lines for total SOC are not shown, as these overlapped with *POC*. The last column shows frequency diagrams of the relative increase in SOC after a doubling of OC inputs after simulation year 500 (as shown by the vertical dashed line). The text reports the median increase, with the range between square brackets. The black dots show the average OC measurement in simulation year 500, and vertical black bars show the standard deviation.

Also the simulated turnover times of the microbial pool provide indications for the presence of equifinality in the behavioural models (Fig. S2). These results show that for every scenario in which the carrying capacity of soil microbes was optimised (Fig. S2 b,d,e,f), there was a large variation in microbial turnover time for the behavioural models, ranging from ca. 7 to 500 days. Furthermore, the C:N ratio of the POM and MAOM pools was better constrained in models with identifiable parameters, compared to the full parameter models (Fig. S3). For the latter, the range in simulated C:N ratios of POM and MAOM was smaller when more calibration data was used. An incorrect simulation of the C:N ratio of POM and MAOM shows that the relative contribution of plant-derived (with a simulated C:N ratio of 30) and microbial-derived OC (with a simulated C:N ratio

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

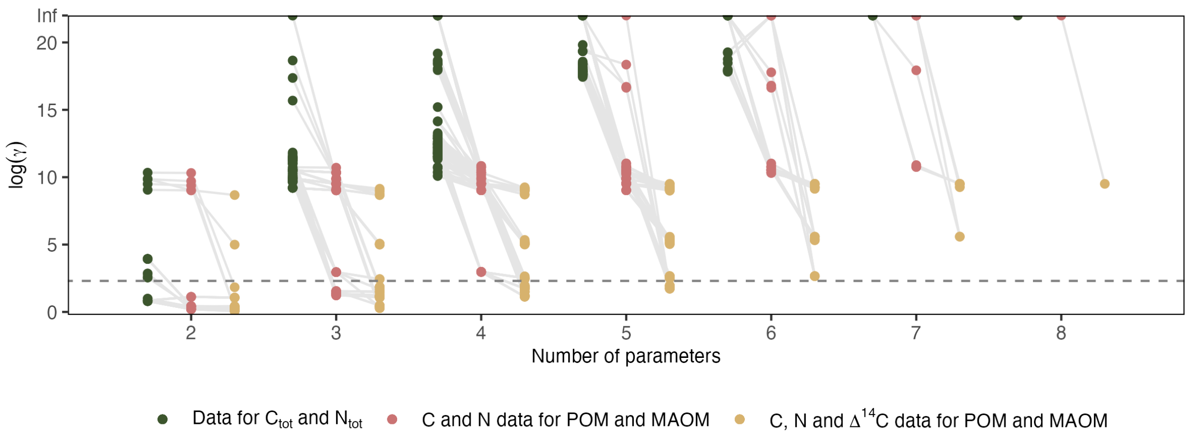

Figure 4. Logarithm of the collinearity index (γ) for all parameter combinations of the SOM model, for the scenarios where (1) only data on total SOC and N were used as calibration constraints (green dots), (2) data on the OC and N content of the POM and MAOM fractions were used (red dots), and (3) data on the OC and N content and $\Delta^{14}C$ value of the POM and MAOM fractions were used (yellow dots). The same parameter combinations are connected with grey lines. The dotted horizontal line shows the threshold in γ (log(10) = 2.3) below which parameter combinations were considered to be identifiable. The identifiable parameter combinations are shown in Tables S19–S21.

of 10) of these pools can take a range of values for the behavioural models. This is another example of the manifestation of equifinality.

3.2.3 The effect of overparameterisation and equifinality on SOM model predictions

The simulated increase in SOC stocks upon a doubling of OC inputs for the SOM model shows that overparameterisation and the resulting equifinality had a large effect on predictions (Fig. 5). All models for which only identifiable parameters were optimised (IPM) simulated a relative increase in SOC between 43.7 and 55.9 %. In contrast, overparameterised models (FPM) simulated a larger increase in SOC stocks when no data on the $\Delta^{14}C$ values of POC and MAOC were used as calibration constraints (Fig. 5c,f). This overestimation was greatest and least precise for the calibration scenario when only data on total SOC and N were available (an increase ranging from 29.4 to 104.3 %, with a median increase of 61.7 %), and slightly smaller when data on the OC and N content of the POC and MAOC pools was available (a median increase of 59.3 %). Only with $\Delta^{14}C$ data for POC and MAOC were the predictions similar between the models with identifiable parameters and the full parameter model (Fig. 5i). These results underline the importance of avoiding overparameterisation to make reliable model predictions. In addition, similar to the rhizosphere models, they show that the performance of models in steady state (i.e., the behavioural models) is not a sufficient indicator for the performance when making predictions.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

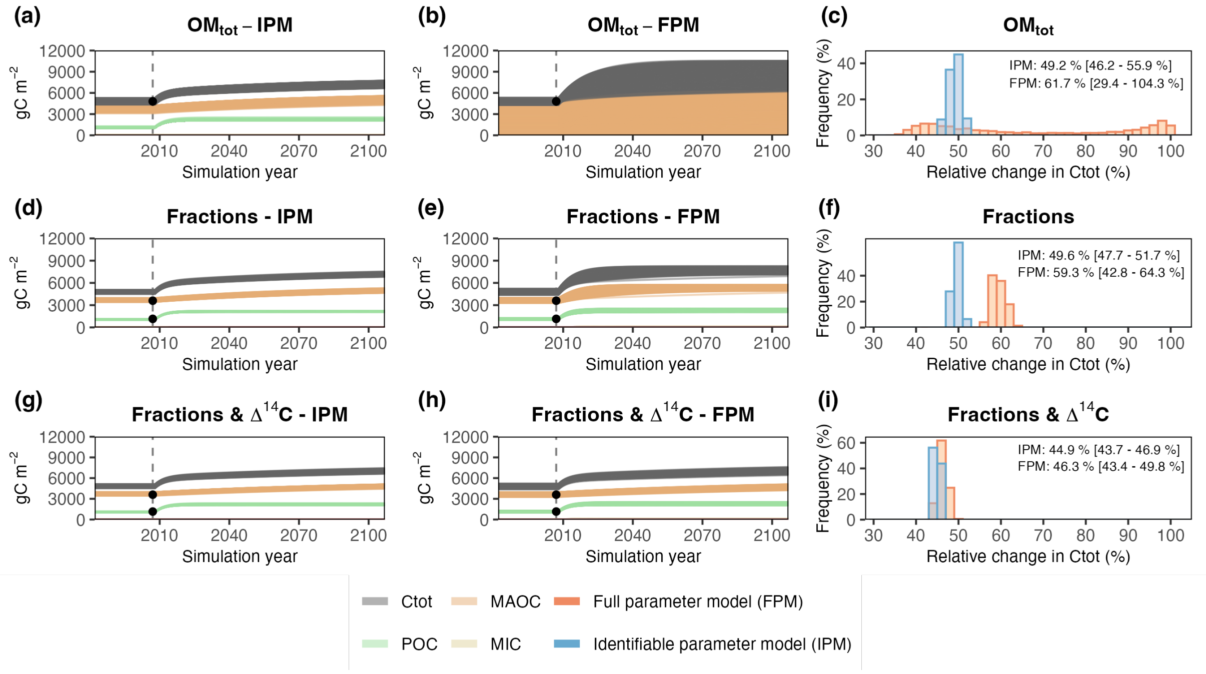

Figure 5. Results of the Bayesian calibration of the SOM model using different data constraints: (1) only total SOC and N (OM_{tot}; a–c), (2) the OC and N content of the POM and MAOM fractions (*Fractions*; d–f), and (3) the OC and N content and Δ^{14}C value of the POC and MAOC fractions (*Fractions & $\Delta^{14}C$*; g–i). The first two columns show the simulations for the different simulated pools, with the first column showing the results for the case when only identifiable parameter sets were optimised (the identifiable parameter model: IPM), while the second column shows results for the case when a non-identifiable parameter set (i.e., all model parameters) was optimised (the full parameter model: FPM). The black circles show the data used for parameter optimisation, while the vertical dashed lines show the timing of the doubling of OC inputs. The corresponding simulations of Δ^{14}C are shown in Fig. 6. The histograms on the right show the relative change in SOC for the year 2107, after a doubling of OC inputs from the year 2007 onwards. The text reports the median increase, with the range between square brackets.

4 Discussion

The discussion is structured around four main conclusions drawn from the results: (1) differences in the mathematical formulation of simulated processes led to different simulated changes in particulate organic matter, (POM) (2) although the simulated steady-state organic matter matched measurements well (the behavioural models), this alone is sufficient to evaluate model performance under an external forcing, (3) including calibration data on internal model pools and their turnover rates reduced prediction uncertainty; and (4) optimising only identifiable model parameters similarly reduced the uncertainty of predictions. The discussion concludes with making recommendations for incorporating parameter identifiability analysis into the model development and evaluation process.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

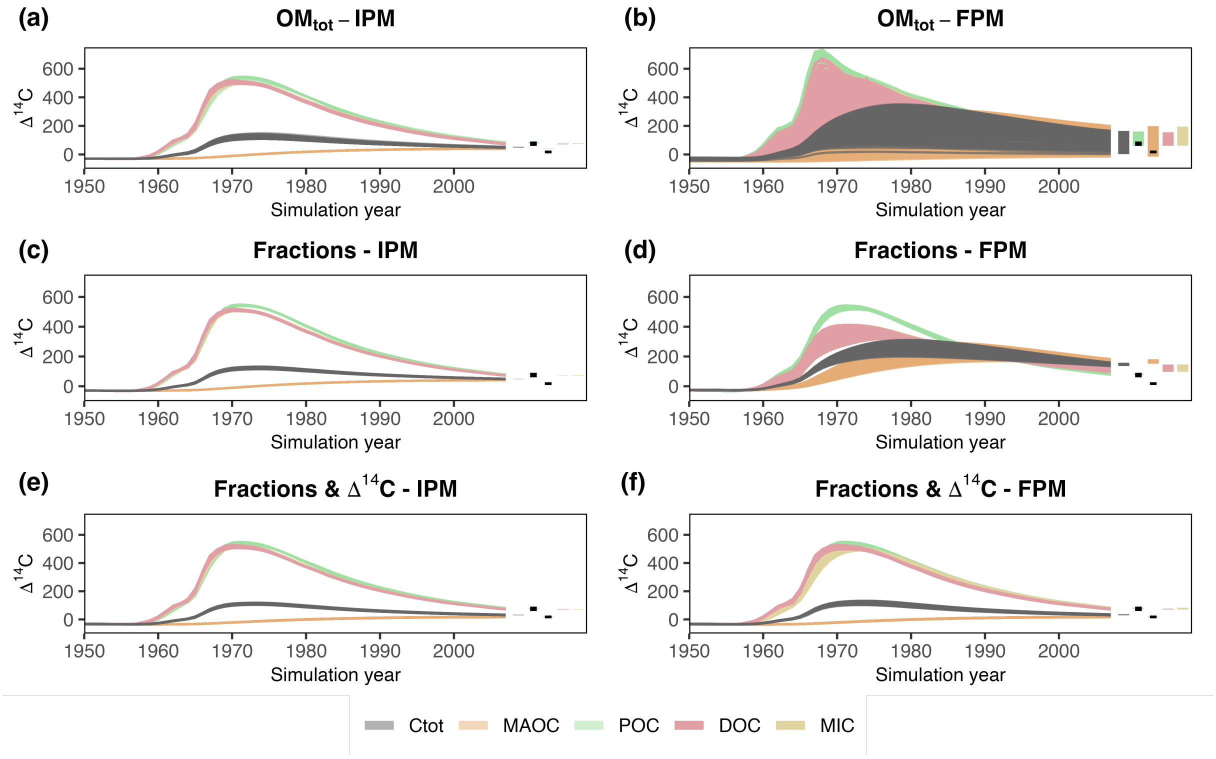

Figure 6. Simulated $\Delta^{14}C$ values of the model pools shown in Fig. 6, for the period of the 'bomb spike' in atmospheric $\Delta^{14}CO_2$ that occurred in the second half of the 20th century. The bars on the right of each graph show the range in simulated $\Delta^{14}C$ in the year 2007, together with the average ± standard deviation of assumed measurements for the POC and MAOC pools in black.

4.1 Different model structures lead to different predictions

The simulations with the rhizosphere models showed that the choice of the mathematical formulations has a large impact on predictions. Although all model formulations led to behavioural steady-state models, the use of different equations for the depolymerisation of POM (absolute versus relative Michaelis-Menten kinetics) and microbial turnover (first-order versus density-dependent) led to different predictions in SOM upon a doubling of OC inputs (Fig. 3). A similar result was obtained by Van de Broek et al. (2024), who showed that different formulations of the thermal adaptation of soil microbes resulted in large differences in predicted losses of SOC in a soil warming experiment. While many of the recently developed mechanistic SOM models use a range of mathematical formulations (Chandel et al., 2023), the quantitative evaluation of different equations on predictions made by recently-developed models received little research attention to date. Similarly, microbially-driven models forced with the same input data at the soil pedon scale have been shown to result in divergent predictions after a change in OC inputs or temperature (Sulman et al., 2018), or to result in different turnover times of SOC, POC and MAOC (Brunmayr et al., 2024). Also at the global scale, different microbially-driven models have been shown to lead to different predictions of SOC (Wieder et al., 2018). As a result, the discussion on the optimal structure of SOM models at the landscape scale is still

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

ongoing, whether on how to improve existing models (Schimel, 2023), or how to drastically change the simulated processes included in these models, and the scale at which these need to be measured (Baveye, 2023). While these discussions are vital to direct the field of soil biogeochemical modelling for the coming decades, our results show that the mathematical formulation of simulated processes should not be overlooked, and its effects on predictions evaluated more extensively.

4.2 Behavioural models do not consistently lead to well-constrained predictions

A second aspect of SOM models underlined by our results is that the correct simulation of SOM in steady state is an insufficient criterion to evaluate their predictive capabilities. This was shown by the simulations with the SOM model, for which all calibration scenarios led to the correct simulation of total SOC under steady state, while predictions made with these behavioural models showed a wide variation in responses (Fig. 5). Similarly, multiple studies have found that while overparameterised SOM models resulted in behavioural models, predictions of changes in SOM for the future diverged widely (Guo et al., 2022; Luo et al., 2016, 2017). This questions the predictions made by SOM models, and should be reduced by making better use of available data from field experiments in which environmental forcings are manipulated, independent validation of model predictions with diachronic data (Le Noë et al., 2023) or the inclusion of more data on the size and turnover time of internal model pools during parameter optimisation (see Section 4.3).

4.3 Including more calibration constraints increases trust in steady-state model simulations

Including additional data in the model calibration process offers several advantages: (1) more parameter values that cannot be measured in the field or from experiments can be optimised, while the value of fewer parameters needs to be fixed (Fig. 4), (2) the size and turnover time of different model pools is simulated more accurately (Fig. 6) and (3) model predictions are better constrained (Fig. 5). Taken together, this will increase confidence in model predictions.

When the parameters of a SOM model with internal pools are optimised without sufficient data to constrain the size of all pools, several combinations of internal pool sizes can lead to behavioural models that accurately simulate the combined size of all pools (i.e., the total amount of SOM). This manifestation of equifinality has been shown to occur in various SOM models (Braakhekke et al., 2013, 2014; Guo et al., 2022). In addition, this may lead to behavioural models with an incorrect simulation of the turnover time of the internal model pools, and therefore total SOM (Braakhekke et al., 2014; He et al., 2016; Brunmayr et al., 2024; Van de Broek et al., 2025). To alleviate these issues, several studies have used additional data besides data on the total amount of SOM in the parameter optimisation process, such as data on stable carbon isotopes ($\delta^{13}C$; e.g., van Dam et al., 1997; Poage and Feng, 2004), radiocarbon isotopes ($\Delta^{14}C$; e.g., Ahrens et al., 2014; Braakhekke et al., 2014; Tifafi et al., 2018; Yu et al., 2020; Tougma et al., 2025), a combination of both isotopes (e.g., Wang et al., 2020; Van de Broek et al., 2025), and data on the size of internal model pools (e.g., Ahrens et al., 2015; Guo et al., 2022; Laub et al., 2024; Van de Broek et al., 2024). Similar to the results presented here, studies devoted to the topic of parameter optimisation under different data constraints consistently found that including more data during the calibration process led to (1) parameter ranges that were better constrained (Ahrens et al., 2014; Braakhekke et al., 2014; Van de Broek et al., 2025) and (2) the distribution of simulated organic matter among model pools better matching measurements (Guo et al., 2022). When this exercise was done

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

in combination with a parameter identifiable analysis, similar conclusions were drawn, combined with the observation that although additional data were used, not all parameters could be optimised (Sierra et al., 2015).

Our results and previous research thus show that model parameter values can be better constrained when including more data in the parameter optimisation process. However, this has to be accompanied by an identifiability analysis to confidently determine how many and which parameters can be optimised together. For our SOM model, only three parameters could be optimised when data on total SOM, POM and MAOM were present. This increased to five parameters when this was amended with Δ^{14}C values of MAOC and POC. Similarly, Sierra et al. (2015) found that no more than 4 parameters of linear-pool SOC models were jointly identifiable. As a consequence, even when data on SOM fractions and Δ^{14}C were used for our SOM model, three parameters needed to be fixed, while this increased to five parameters when only data on POM and MAOM were available, and six parameters when only data on total SOM was present. As the values of these parameters are generally not based on data, but rather estimates, this implies that there is a hidden uncertainty in the quality of these predictions. This can be quantified by assessing how the values of fixed parameters affect predictions (Brun et al., 2001).

Only when data on the size and Δ^{14}C value of POC and MAOC were used during parameter optimisation were the sizes of these pools, and their turnover times, correctly simulated (Figs. 5g,h and 6e,f). Based on these results, it is recommended to use data on the Δ^{14}C value of internal model pools in the calibration process, in line with Van de Broek et al. (2025). When this data is not available, it is worthwhile to report the Δ^{14}C value or turnover time of simulated model pools, as this may serve as an indication of whether the simulated turnover times are in line with observations (Mathieu et al., 2015; Balesdent et al., 2018; Lawrence et al., 2020; Sierra et al., 2024; von Fromm et al., 2024). Whether Δ^{14}C data on total SOC is sufficient to achieve similar results as when data on the Δ^{14}C value of internal model pools is present is a topic for future research. The recommendations mentioned above are equally relevant for the validation of SOM models, which has been shown to be often applied insufficiently and inadequately (Garsia et al., 2023; Le Noë et al., 2023).

4.4 Reducing uncertainty in model predictions through identifiability analysis

The importance of practical identifiability and equifinality has long been recognised across environmental disciplines that use simulation models, as outlined in the introduction. Only more recently have these concepts received attention in the field of SOM modelling (e.g. Sierra et al., 2015; Marschmann et al., 2019; Guo et al., 2022), and are they mentioned in review articles (Schimel, 2023; Baveye, 2023; Le Noë et al., 2023; Manzoni and Schimel, 2024). In line with these studies, our results show that optimising identifiable parameters of a SOM model leads to better constrained predictions of SOM upon a doubling of C inputs (note that only the precision of predictions was evaluated, as we did not have data to assess the accuracy). From this, it is clear that knowledge about which parameters can be jointly identified, given the quantity and type of data available for calibration, is necessary to reliably apply a model. However, results from an identifiability analysis are rarely reported in articles describing (new) SOM models. A possible explanation is that this concept has been neglected so far, given its historical underrepresentation in the SOM modelling literature. Other reasons are related to the fact that "there is still a great need for development of methods, software and training to ensure all modellers are able to assess and react appropriately to non-identifiability" (Guillaume et al., 2019, p. 428). A last reason can be that although many techniques for parameter identifiability

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

have been developed, their description in the literature can be very technical and difficult to implement by non-expert modellers. One way to solve this issue would be to develop accessible software packages that non-expert modellers can use to perform an identifiability analysis on their model of choice. The potential of developing such tools to make identifiability analysis more accessible is evidenced by the incorporation of the sensitivity-based identifiability analysis presented by Brun et al. (2001) into the FME package in R (Soetaert and Petzoldt, 2010), which has been used by most studies assessing the identifiability of parameters in SOM models (e.g., Sierra et al., 2015; Abramoff et al., 2022; Guo et al., 2022). In addition, developing such tools will make it possible to apply and compare the results of different methods of parameter identifiability analyses. While an overview of the different methods available to perform structural and practical identifiability analyses is beyond the scope of this discussion, the interested reader is referred to Walter and Pronzato (1996); Raue et al. (2011); Miao et al. (2011); Raue et al. (2014); Guillaume et al. (2019); Lam et al. (2022) and Wanika et al. (2024).

4.5 Including parameter identifiability analysis in the model development process

Based on our results, it is argued here that a parameter identifiability analysis should be an integral part of the model calibration and validation process, together with previously described and equally important aspects to be taken into account (e.g., Jakeman et al., 2006; Mai, 2023). We therefore concur with Guillaume et al. (2019), who recommended that "any modeling study should document whether a model is non-identifiably, the source of potential non-identifiability and how this affects intended project outcomes". Based on the results presented in this study, we suggest the following:

1. When developing a novel SOM model, information on the identifiability of model parameters for different representative scenarios on data availability should be provided, in order for model users to know which parameters can be jointly calibrated, given available data.

2. This should be accompanied by results of a sensitivity analysis, in order for users to know which parameters, to which the model output is not sensitive, should be avoided during parameter optimisation.

3. For non-identifiable parameters, reference values should be suggested based on observations, experiments or the scientific literature.

4. The results of a newly developed model should not only be shown for a steady-state simulation, but complemented by how model outputs change when environmental forcings (e.g., temperature, soil moisture or C inputs) are varied. This way, it can be evaluated if the model behaves as desired under changing environmental conditions.

5 Conclusions

This study assessed how (1) equifinality, arising from overparameterisation, and (2) the choice of mathematical formulations impact the variability of predictions made by SOM models. The key findings are summarised as follows. (1) The accurate simulation of total SOM in steady state is not a sufficient criterion to evaluate model performance. This was demonstrated by the

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

diverging predictions of SOM upon a doubling of OM inputs for models using different mathematical equations (e.g., absolute versus relative Michaelis-Menten kinetics) and overparameterised models, versus models being optimised using identifiable parameters. Specifically, the variation in model response for overparameterised models was up to eight times larger compared to when only identifiable parameters were optimised. (2) The amount of calibration data determines how many model parameters are identifiable, and can thus be jointly optimised without their values compensating for each other. Our results confirmed previous studies showing that the number of identifiable model parameters is generally lower than the number of unknown parameters. (3) The type of calibration data is equally important, as it dictates which pools can have their size and turnover rate constrained. With only total SOC data, the distribution of simulated OM among the model pools cannot be evaluated, while data on $\Delta^{14}C$ is necessary to correctly simulate the turnover rate of OM pools. This implies that a reliable application of SOM models requires measurements of the size of model pools and data on their turnover rate. Without such data, predictions by SOM models will not be reliable. Based on our findings, we urge that parameter identifiability analysis becomes a standard procedure when developing and applying SOM models. This will remove the hidden uncertainty in model predictions caused by equifinality, thereby increasing confidence in SOM models.

Code availability. The exact version of the codes used to produce the results used in this paper is archived on Zenodo under https://doi.org/10.5281/zenodo.17974745 (Van de Broek, 2025) under the GPL-2 license, as are input data and scripts to run the model and produce the plots for all the simulations presented in this paper.

Author contributions. MVdB conceived and designed the study, developed the model code, performed the model simulations, and took the lead in writing the original draft of the manuscript. JS contributed to the interpretation of the results and editing of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research has been financially supported by the Swiss National Science Foundation (SNSF; Ambizione grant number PZ00P2_193617/1, granted to Marijn Van de Broek).

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

References

Abramoff, R. Z., Guenet, B., Zhang, H., Georgiou, K., Xu, X., Viscarra Rossel, R. A., Yuan, W., and Ciais, P.: Improved global-scale predictions of soil carbon stocks with Millennial Version 2, Soil Biology and Biochemistry, 164, 108 466, https://doi.org/10.1016/j.soilbio.2021.108466, 2022.

Ahrens, B., Reichstein, M., Borken, W., Muhr, J., Trumbore, S. E., and Wutzler, T.: Bayesian calibration of a soil organic carbon model using Δ^{14}C measurements of soil organic carbon and heterotrophic respiration as joint constraints, Biogeosciences, 11, 2147–2168, https://doi.org/10.5194/bg-11-2147-2014, publisher: Copernicus GmbH, 2014.

Ahrens, B., Braakhekke, M. C., Guggenberger, G., Schrumpf, M., and Reichstein, M.: Contribution of sorption, DOC transport and microbial interactions to the 14C age of a soil organic carbon profile: Insights from a calibrated process model, Soil Biology and Biochemistry, 88, 390–402, https://doi.org/10.1016/j.soilbio.2015.06.008, 2015.

Ahrens, B., Guggenberger, G., Rethemeyer, J., John, S., Marschner, B., Heinze, S., Angst, G., Mueller, C. W., Kögel-Knabner, I., Leuschner, C., Hertel, D., Bachmann, J., Reichstein, M., and Schrumpf, M.: Combination of energy limitation and sorption capacity explains 14C depth gradients, Soil Biology and Biochemistry, 148, 107 912, https://doi.org/10.1016/j.soilbio.2020.107912, 2020.

Allison, S. D., Wallenstein, M. D., and Bradford, M. A.: Soil-carbon response to warming dependent on microbial physiology, Nature Geoscience, 3, 336–340, https://doi.org/10.1038/ngeo846, iSBN: 1752-0894 Publisher: Nature Publishing Group, 2010.

Ardia, D., Boudt, K., Carl, P., Mullen, K. M., and Peterson, B. G.: Differential Evolution with DEoptim: An Application to Non-Convex Portfolio Optimization, R Journal, 3, 27–34, https://doi.org/10.32614/RJ-2011-005, 2011.

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.

Baveye, P. C.: Ecosystem-scale modelling of soil carbon dynamics: Time for a radical shift of perspective?, Soil Biology and Biochemistry, 184, 109 112, https://doi.org/10.1016/j.soilbio.2023.109112, 2023.

Beck, M. B.: Water quality modeling: A review of the analysis of uncertainty, Water Resources Research, 23, 1393–1442, https://doi.org/10.1029/WR023i008p01393, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/WR023i008p01393, 1987.

Bellman, R. and Åström, K. J.: On structural identifiability, Mathematical Biosciences, 7, 329–339, https://doi.org/10.1016/0025-5564(70)90132-X, 1970.

Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, Advances in Water Resources, 16, 41–51, https://doi.org/10.1016/0309-1708(93)90028-E, 1993.

Beven, K.: Towards a coherent philosophy for modelling the environment, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 458, 2465–2484, https://doi.org/10.1098/rspa.2002.0986, publisher: Royal Society, 2002.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K.: Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process, Hydrology and Earth System Sciences, 11, 460–467, https://doi.org/10.5194/hess-11-460-2007, publisher: Copernicus GmbH, 2007.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279–298, https://doi.org/10.1002/hyp.3360060305, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.3360060305, 1992.

Blankinship, J. C., Berhe, A. A., Crow, S. E., Druhan, J. L., Heckman, K. A., Keiluweit, M., Lawrence, C. R., Marín-Spiotta, E., Plante, A. F., Rasmussen, C., Schädel, C., Schimel, J. P., Sierra, C. A., Thompson, A., Wagai, R., and Wieder, W. R.: Improving understanding of soil or-

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

ganic matter dynamics by triangulating theories, measurements, and models, Biogeochemistry, 140, 1–13, https://doi.org/10.1007/s10533-018-0478-2, iSBN: 1053301804, 2018.

Braakhekke, M. C., Wutzler, T., Beer, C., Kattge, J., Schrumpf, M., Ahrens, B., Schöning, I., Hoosbeek, M. R., Kruijt, B., Kabat, P., and Reichstein, M.: Modeling the vertical soil organic matter profile using Bayesian parameter estimation, Biogeosciences, 10, 399–420, https://doi.org/10.5194/bg-10-399-2013, 2013.

Braakhekke, M. C., Beer, C., Schrumpf, M., Ekici, A., Ahrens, B., Hoosbeek, M. R., Kruijt, B., Kabat, P., and Reichstein, M.: The use of radiocarbon to constrain current and future soil organic matter turnover and transport in a temperate forest, Journal of Geophysical Research: Biogeosciences, 119, 372–391, https://doi.org/10.1002/2013JG002420, 2014.

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A., and Crowther, T. W.: Managing uncertainty in soil carbon feedbacks to climate change, Nature Climate Change, 6, 751–758, https://doi.org/10.1038/nclimate3071, publisher: Nature Publishing Group, 2016.

Brazier, R. E., Beven, K. J., Freer, J., and Rowan, J. S.: Equifinality and uncertainty in physically based soil erosion models: application of the GLUE methodology to WEPP–the Water Erosion Prediction Project–for sites in the UK and USA, Earth Surface Processes and Landforms, 25, 825–845, https://doi.org/10.1002/1096-9837(200008)25:8<825::AID-ESP101>3.0.CO;2-3, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1096-9837%28200008%2925%3A8%3C825%3A%3AAID-ESP101%3E3.0.CO%3B2-3, 2000.

Browning, A. P., Warne, D. J., Burrage, K., Baker, R. E., and Simpson, M. J.: Identifiability analysis for stochastic differential equation models in systems biology, Journal of The Royal Society Interface, 17, 20200 652, https://doi.org/10.1098/rsif.2020.0652, publisher: Royal Society, 2020.

Brun, R., Reichert, P., and Künsch, H. R.: Practical identifiability analysis of large environmental simulation models, Water Resources Research, 37, 1015–1030, https://doi.org/10.1029/2000WR900350, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2000WR900350, 2001.

Brun, R., Kühni, M., Siegrist, H., Gujer, W., and Reichert, P.: Practical identifiability of ASM2d parameters—systematic selection and tuning of parameter subsets, Water Research, 36, 4113–4127, https://doi.org/10.1016/S0043-1354(02)00104-5, 2002.

Brunmayr, A. S., Hagedorn, F., Moreno Duborgel, M., Minich, L. I., and Graven, H. D.: Radiocarbon analysis reveals underestimation of soil organic carbon persistence in new-generation soil models, Geoscientific Model Development, 17, 5961–5985, https://doi.org/10.5194/gmd-17-5961-2024, publisher: Copernicus GmbH, 2024.

Buchkowski, R. W., Bradford, M. A., Grandy, A. S., Schmitz, O. J., and Wieder, W. R.: Applying population and community ecology theory to advance understanding of belowground biogeochemistry, Ecology Letters, 20, 231–245, https://doi.org/10.1111/ele.12712, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ele.12712, 2017.

Campbell, E. E. and Paustian, K.: Current developments in soil organic matter modeling and the expansion of model applications: a review, Environmental Research Letters, 10, 123 004, https://doi.org/10.1088/1748-9326/10/12/123004, publisher: IOP Publishing, 2015.

Chandel, A. K., Jiang, L., and Luo, Y.: Microbial Models for Simulating Soil Carbon Dynamics: A Review, Journal of Geophysical Research: Biogeosciences, 128, e2023JG007 436, https://doi.org/10.1029/2023JG007436, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023JG007436, 2023.

Cobelli, C., Lepschy, A., and Jacur, G. R.: Identifiability of compartmental systems and related structural properties, Mathematical Biosciences, 44, 1–18, https://doi.org/10.1016/0025-5564(79)90026-9, 1979.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Dankwa, E. A., Brouwer, A. F., and Donnelly, C. A.: Structural identifiability of compartmental models for infectious disease transmission is influenced by data type, Epidemics, 41, 100 643, https://doi.org/10.1016/j.epidem.2022.100643, 2022.

Delforge, J.: The problem of structural identifiability of a linear compartmental system: Solved or not?, Mathematical Biosciences, 36, 119–125, https://doi.org/10.1016/0025-5564(77)90019-0, 1977.

DiStefano, J. and Cobelli, C.: On parameter and structural identifiability: Nonunique observability/reconstructibility for identifiable systems, other ambiguities, and new definitions, IEEE Transactions on Automatic Control, 25, 830–833, https://doi.org/10.1109/TAC.1980.1102439, conference Name: IEEE Transactions on Automatic Control, 1980.

Dwivedi, D., Riley, W. J., Torn, M. S., Spycher, N., Maggi, F., and Tang, J. Y.: Mineral properties, microbes, transport, and plant-input profiles control vertical distribution and age of soil carbon stocks, Soil Biology and Biochemistry, 107, 244–259, https://doi.org/10.1016/j.soilbio.2016.12.019, iSBN: 0038-0717 Publisher: Elsevier Ltd, 2017.

Famiglietti, C. A., Smallman, T. L., Levine, P. A., Flack-Prain, S., Quetin, G. R., Meyer, V., Parazoo, N. C., Stettz, S. G., Yang, Y., Bonal, D., Bloom, A. A., Williams, M., and Konings, A. G.: Optimal model complexity for terrestrial carbon cycle prediction, Biogeosciences, 18, 2727–2754, https://doi.org/10.5194/bg-18-2727-2021, publisher: Copernicus GmbH, 2021.

Finzi, A. C., Abramoff, R. Z., Spiller, K. S., Brzostek, E. R., Darby, B. A., Kramer, M. A., and Phillips, R. P.: Rhizosphere processes are quantitatively important components of terrestrial carbon and nutrient cycles, Global Change Biology, 21, 2082–2094, https://doi.org/10.1111/gcb.12816, 2015.

Garsia, A., Moinet, A., Vazquez, C., Creamer, R. E., and Moinet, G. Y. K.: The challenge of selecting an appropriate soil organic carbon simulation model: A comprehensive global review and validation assessment, Global Change Biology, 29, 5760–5774, https://doi.org/10.1111/gcb.16896, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.16896, 2023.

Georgiou, K., Abramoff, R. Z., Harte, J., Riley, W. J., and Torn, M. S.: Microbial community-level regulation explains soil carbon responses to long-term litter manipulations, Nature Communications, 8, 1223, https://doi.org/10.1038/s41467-017-01116-z, publisher: Nature Publishing Group, 2017.

Georgiou, K., Jackson, R. B., Vindušková, O., Abramoff, R. Z., Ahlström, A., Feng, W., Harden, J. W., Pellegrini, A. F. A., Polley, H. W., Soong, J. L., Riley, W. J., and Torn, M. S.: Global stocks and capacity of mineral-associated soil organic carbon, Nature Communications, 13, 3797, https://doi.org/10.1038/s41467-022-31540-9, publisher: Nature Publishing Group, 2022.

Guillaume, J. H. A., Jakeman, J. D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M. C., Jakeman, A. J., Keesman, K. J., Razavi, S., and Stigter, J. D.: Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose, Environmental Modelling & Software, 119, 418–432, https://doi.org/10.1016/j.envsoft.2019.07.007, 2019.

Guo, X., Viscarra Rossel, R. A., Wang, G., Xiao, L., Wang, M., Zhang, S., and Luo, Z.: Particulate and mineral-associated organic carbon turnover revealed by modelling their long-term dynamics, Soil Biology and Biochemistry, 173, 108 780, https://doi.org/10.1016/j.soilbio.2022.108780, 2022.

Hammer, S. and Levin, I.: Monthly mean atmospheric d14CO2 at Jungfraujoch and Schauinsland from 1986 to 2016 [dataset], https://doi.org/10.11588/data/10100, 2017.

Hartig, F., Minunno, F., and Paul, S.: BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics, https://CRAN.R-project.org/package=BayesianTools, r package version 0.1.8, 2023.

He, Y., Trumbore, S. E., Torn, M. S., Harden, J. W., Vaughn, 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, iSBN: 1095-9203 (Electronic) 0036-8075 (Linking), 2016.

Holmberg, A.: On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities, Mathematical Biosciences, 62, 23–43, https://doi.org/10.1016/0025-5564(82)90061-X, 1982.

Hua, Q., Barbetti, M., and Rakowski, A. Z.: Atmospheric radiocarbon for the period 1950-2010, Radiocarbon, 55, 2059–2072, https://doi.org/10.2458/azu_js_rc.v55i2.16177, iSBN: 0033-8222|escape}, 2013.

Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, Water Resources Research, 29, 2637–2649, https://doi.org/10.1029/93WR00877, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/93WR00877, 1993.

Jakeman, A. J., Letcher, R. A., and Norton, J. P.: Ten iterative steps in development and evaluation of environmental models, Environmental Modelling & Software, 21, 602–614, https://doi.org/10.1016/j.envsoft.2006.01.004, 2006.

Kelleher, C., Wagener, T., McGlynn, B., Ward, A. S., Gooseff, M. N., and Payn, R. A.: Identifiability of transient storage model parameters along a mountain stream, Water Resources Research, 49, 5290–5306, https://doi.org/10.1002/wrcr.20413, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wrcr.20413, 2013.

Kleissen, F. M., Beck, M. B., and Wheater, H. S.: The Identifiability of Conceptual Hydrochemical Models, Water Resources Research, 26, 2979–2992, https://doi.org/10.1029/WR026i012p02979, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/WR026i012p02979, 1990.

Lam, N. N., Docherty, P. D., and Murray, R.: Practical identifiability of parametrised models: A review of benefits and limitations of various approaches, Mathematics and Computers in Simulation, 199, 202–216, https://doi.org/10.1016/j.matcom.2022.03.020, 2022.

Laub, M., Blagodatsky, S., Van de Broek, M., Schlichenmaier, S., Kunlanit, B., Six, J., Vityakon, P., and Cadisch, G.: SAMM version 1.0: a numerical model for microbial- mediated soil aggregate formation, Geoscientific Model Development, 17, 931–956, https://doi.org/10.5194/gmd-17-931-2024, publisher: Copernicus GmbH, 2024.

Lawrence, C. R., Neff, J. C., and Schimel, J. P.: Does adding microbial mechanisms of decomposition improve soil organic matter models? A comparison of four models using data from a pulsed rewetting experiment, Soil Biology and Biochemistry, 41, 1923–1934, https://doi.org/10.1016/j.soilbio.2009.06.016, 2009.

Lawrence, C. R., Beem-Miller, J., Hoyt, A. M., Monroe, G., Sierra, C. A., Stoner, S., Heckman, K., Blankinship, J. C., Crow, S. E., McNicol, G., Trumbore, S., Levine, P. A., Vindušková, O., Todd-Brown, K., Rasmussen, C., Hicks Pries, C. E., Schädel, C., McFarlane, K., Doetterl, S., Hatté, C., He, Y., Treat, C., Harden, J. W., Torn, M. S., Estop-Aragonés, C., Asefaw Berhe, A., Keiluweit, M., Della Rosa Kuhnen, , Marin-Spiotta, E., Plante, A. F., Thompson, A., Shi, Z., Schimel, J. P., Vaughn, L. J. S., von Fromm, S. F., and Wagai, R.: An open-source database for the synthesis of soil radiocarbon data: International Soil Radiocarbon Database (ISRaD) version 1.0, Earth System Science Data, 12, 61–76, https://doi.org/10.5194/essd-12-61-2020, publisher: Copernicus GmbH, 2020.

Le Noë, J., Manzoni, S., Abramoff, R., Bölscher, T., Bruni, E., Cardinael, R., Ciais, P., Chenu, C., Clivot, H., Derrien, D., Ferchaud, F., Garnier, P., Goll, D., Lashermes, G., Martin, M., Rasse, D., Rees, F., Sainte-Marie, J., Salmon, E., Schiedung, M., Schimel, J., Wieder, W., Abiven, S., Barré, P., Cécillon, L., and Guenet, B.: Soil organic carbon models need independent time-series validation for reliable prediction, Communications Earth & Environment, 4, 1–8, https://doi.org/10.1038/s43247-023-00830-5, publisher: Nature Publishing Group, 2023.

Lennon, J. T., Abramoff, R. Z., Allison, S. D., Burckhardt, R. M., DeAngelis, K. M., Dunne, J. P., Frey, S. D., Friedlingstein, P., Hawkes, C. V., Hungate, B. A., Khurana, S., Kivlin, S. N., Levine, N. M., Manzoni, S., Martiny, A. C., Martiny, J. B. H., Nguyen, N. K., Rawat, M., Talmy, D., Todd-Brown, K., Vogt, M., Wieder, W. R., and Zakem, E. J.: Priorities, opportunities, and challenges for integrating microorganisms into Earth system models for climate change prediction, mBio, 15, e00 455–24, https://doi.org/10.1128/mbio.00455-24, publisher: American Society for Microbiology, 2024.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X., and Zhang, L.: Parameter Identifiability, Constraint, and Equifinality in Data Assimilation with Ecosystem Models, Ecological Applications, 19, 571–574, https://www.jstor.org/stable/27645995, publisher: [Wiley, Ecological Society of America], 2009.

Luo, Z., Wang, E., Shao, Q., Conyers, M. K., and Liu, D. L.: Confidence in soil carbon predictions undermined by the uncertainties in observations and model parameterisation, Environmental Modelling & Software, 80, 26–32, https://doi.org/10.1016/j.envsoft.2016.02.013, 2016.

Luo, Z., Wang, E., and Sun, O. J.: Uncertain future soil carbon dynamics under global change predicted by models constrained by total carbon measurements, Ecological Applications, 27, 1001–1009, https://www.jstor.org/stable/26155933, publisher: [Wiley, Ecological Society of America], 2017.

Luo, Z., Wang, G., and Wang, E.: Global subsoil organic carbon turnover times dominantly controlled by soil properties rather than climate, Nature Communications, 10, 3688, https://doi.org/10.1038/s41467-019-11597-9, publisher: Springer US, 2019.

Mai, J.: Ten strategies towards successful calibration of environmental models, Journal of Hydrology, 620, 129 414, https://doi.org/10.1016/j.jhydrol.2023.129414, 2023.

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

Manzoni, S. and Schimel, J. P.: Advances in modelling soil microbial dynamics, Soil Biology and Biochemistry, 197, 109 535, https://doi.org/10.1016/j.soilbio.2024.109535, 2024.

Marschmann, G. L., Pagel, H., Kügler, P., and Streck, T.: Equifinality, sloppiness, and emergent structures of mechanistic soil biogeochemical models, Environmental Modelling & Software, 122, 104 518, https://doi.org/10.1016/j.envsoft.2019.104518, 2019.

Mathieu, J. A., Hatté, C., Balesdent, J., and Parent, : Deep soil carbon dynamics are driven more by soil type than by climate: a worldwide meta-analysis of radiocarbon profiles, Global Change Biology, 21, 4278–4292, https://doi.org/10.1111/gcb.13012, iSBN: 1365-2486, 2015.

Meurer, K. H. E., Chenu, C., Coucheney, E., Herrmann, A. M., Keller, T., Kätterer, T., Nimblad Svensson, D., and Jarvis, N.: Modelling dynamic interactions between soil structure and the storage and turnover of soil organic matter, Biogeosciences, 17, 5025–5042, https://doi.org/10.5194/bg-17-5025-2020, publisher: Copernicus GmbH, 2020.

Miao, H., Xia, X., Perelson, A. S., and Wu, H.: On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics, SIAM Review, 53, 3–39, https://doi.org/10.1137/090757009, 2011.

Mullen, K., Ardia, D., Gil, D., Windover, D., and Cline, J.: DEoptim: An R Package for Global Optimization by Differential Evolution, Journal of Statistical Software, 40, 1–26, https://doi.org/10.18637/jss.v040.i06, 2011.

Nearing, M. A., Govers, G., and Norton, L. D.: Variability in Soil Erosion Data from Replicated Plots, Soil Science Society of America Journal, 63, 1829–1835, https://doi.org/10.2136/sssaj1999.6361829x, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.2136/sssaj1999.6361829x, 1999.

Nguyen, V. V. and Wood, E. F.: Review and Unification of Linear Identifiability Concepts, SIAM Review, 24, 34–51, https://www.jstor.org/stable/2029430, publisher: Society for Industrial and Applied Mathematics, 1982.

Omlin, M., Brun, R., and Reichert, P.: Biogeochemical model of Lake Zürich: sensitivity, identifiability and uncertainty analysis, Ecological Modelling, 141, 105–123, https://doi.org/10.1016/S0304-3800(01)00257-5, 2001.

Oreskes, N., Shrader-Frechette, K., and Belitz, K.: Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences, Science, 263, 641–646, https://doi.org/10.1126/science.263.5147.641, iSBN: 00368075, 1994.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Pallandt, M., Schrumpf, M., Lange, H., Reichstein, M., Yu, L., and Ahrens, B.: Modelling the effect of climate–substrate interactions on soil organic matter decomposition with the Jena Soil Model, Biogeosciences, 22, 1907–1928, https://doi.org/10.5194/bg-22-1907-2025, publisher: Copernicus GmbH, 2025.

Perretti, C. T., Munch, S. B., and Sugihara, G.: Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data, Proceedings of the National Academy of Sciences, 110, 5253–5257, https://doi.org/10.1073/pnas.1216076110, publisher: Proceedings of the National Academy of Sciences, 2013.

Poage, M. A. and Feng, X.: A theoretical analysis of steady state δ 13 C profiles of soil organic matter, Global Biogeochemical Cycles, 18, 1–13, http://doi.wiley.com/10.1029/2003GB002195, 2004.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2025.

Raue, A., Kreutz, C., Maiwald, T., Klingmüller, U., and Timmer, J.: Addressing parameter identifiability by model-based experimentation, IET Systems Biology, 5, 120–130, https://doi.org/10.1049/iet-syb.2010.0061, publisher: The Institution of Engineering and Technology, 2011.

Raue, A., Karlsson, J., Saccomani, M. P., Jirstrand, M., and Timmer, J.: Comparison of approaches for parameter identifiability analysis of biological systems, Bioinformatics, 30, 1440–1448, https://doi.org/10.1093/bioinformatics/btu006, 2014.

Reiersøl, O.: Identifiability of a Linear Relation between Variables Which Are Subject to Error, Econometrica, 18, 375–389, https://doi.org/10.2307/1907835, publisher: [Wiley, Econometric Society], 1950.

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., Buck, C. E., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Haflidason, H., Hajdas, I., Hatté, C., Heaton, T. J., Hoffmann, D. L., Hogg, A. G., Hughen, K. A., Kaiser, K. F., Kromer, B., Manning, S. W., Niu, M., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Staff, R. A., Turney, C. S. M., and Plicht, J. v. d.: Intcal13 and marine13 radiocarbon age calibration curves 0 – 50,000 years cal bp, Radiocarbon, 55, 1869–1887, https://doi.org/10.2458/rc.v51i4.3569, iSBN: 0033-8222, 2013.

Riley, W. J., Maggi, F., Kleber, M., Torn, M. S., Tang, J. Y., Dwivedi, D., and Guerry, N.: Long residence times of rapidly decomposable soil organic matter: application of a multi-phase, multi-component, and vertically resolved model (BAMS1) to soil carbon dynamics, Geoscientific Model Development, 7, 1335–1355, https://doi.org/10.5194/gmd-7-1335-2014, iSBN: 1991-959X, 2014.

Rothenberg, T. J.: Identification in Parametric Models, Econometrica, 39, 577–591, https://doi.org/10.2307/1913267, publisher: [Wiley, The Econometric Society], 1971.

Schimel, J.: Modeling ecosystem-scale carbon dynamics in soil: The microbial dimension, Soil Biology and Biochemistry, 178, 108 948, https://doi.org/10.1016/j.soilbio.2023.108948, 2023.

Schimel, J. P. and Weintraub, M. N.: The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model, Soil Biology and Biochemistry, 35, 549–563, https://doi.org/10.1016/S0038-0717(03)00015-4, 2003.

Schindler, D. E. and Hilborn, R.: Prediction, precaution, and policy under global change, Science, 347, 953–954, https://doi.org/10.1126/science.1261824, publisher: American Association for the Advancement of Science, 2015.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. a., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. a. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property., Nature, 478, 49–56, https://doi.org/10.1038/nature10386, 2011.

Schrumpf, M. and Kaiser, K.: Large differences in estimates of soil organic carbon turnover in density fractions by using single and repeated radiocarbon inventories, Geoderma, 239-240, 168–178, https://doi.org/10.1016/j.geoderma.2014.09.025, 2015.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Sierra, C. A., Malghani, S., and Müller, M.: Model structure and parameter identification of soil organic matter models, Soil Biology and Biochemistry, 90, 197–203, https://doi.org/10.1016/j.soilbio.2015.08.012, 2015.

Sierra, C. A., Ahrens, B., Bolinder, M. A., Braakhekke, M. C., von Fromm, S., Kätterer, T., Luo, Z., Parvin, N., and Wang, G.: Carbon sequestration in the subsoil and the time required to stabilize carbon for climate change mitigation, Global Change Biology, 30, e17 153, https://doi.org/10.1111/gcb.17153, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.17153, 2024.

Soetaert, K. and Petzoldt, T.: Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME, Journal of Statistical Software, 33, 1–28, https://doi.org/10.18637/jss.v033.i03, 2010.

Soetaert, K., Petzoldt, T., and Setzer, R. W.: Solving Differential Equations in R: Package deSolve, Journal of Statistical Software, 33, 1–25, https://doi.org/10.18637/jss.v033.i09, 2010.

Sorooshian, S. and Gupta, V. K.: Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness, Water Resources Research, 19, 260–268, https://doi.org/10.1029/WR019i001p00260, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/WR019i001p00260, 1983.

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., and Classen, A. T.: Multiple models and experiments underscore large uncertainty in soil carbon dynamics, Biogeochemistry, 141, 109–123, https://doi.org/10.1007/s10533-018-0509-z, 2018.

Sulman, B. N., Shevliakova, E., Brzostek, E. R., Kivlin, S. N., Malyshev, S., Menge, D. N., and Zhang, X.: Diverse Mycorrhizal Associations Enhance Terrestrial C Storage in a Global Model, Global Biogeochemical Cycles, 33, 501–523, https://doi.org/10.1029/2018GB005973, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GB005973, 2019.

Tang, J. and Riley, W. J.: Weaker soil carbon–climate feedbacks resulting from microbial and abiotic interactions, Nature Climate Change, 5, 56–60, https://doi.org/10.1038/nclimate2438, publisher: Nature Publishing Group, 2015.

Tang, J. and Riley, W. J.: Competitor and substrate sizes and diffusion together define enzymatic depolymerization and microbial substrate uptake rates, Soil Biology and Biochemistry, 139, 107 624, https://doi.org/10.1016/j.soilbio.2019.107624, 2019.

Tang, J. Y. and Riley, W. J.: A total quasi-steady-state formulation of substrate uptake kinetics in complex networks and an example application to microbial litter decomposition, Biogeosciences, 10, 8329–8351, https://doi.org/10.5194/bg-10-8329-2013, publisher: Copernicus GmbH, 2013.

Tao, F., Huang, Y., Hungate, B. A., Manzoni, S., Frey, S. D., Schmidt, M. W. I., Reichstein, M., Carvalhais, N., Ciais, P., Jiang, L., Lehmann, J., Wang, Y.-P., Houlton, B. Z., Ahrens, B., Mishra, U., Hugelius, G., Hocking, T. D., Lu, X., Shi, Z., Viatkin, K., Vargas, R., Yigini, Y., Omuto, C., Malik, A. A., Peralta, G., Cuevas-Corona, R., Di Paolo, L. E., Luotto, I., Liao, C., Liang, Y.-S., Saynes, V. S., Huang, X., and Luo, Y.: Microbial carbon use efficiency promotes global soil carbon storage, Nature, 618, 981–985, https://doi.org/10.1038/s41586-023-06042-3, publisher: Nature Publishing Group, 2023.

ter Braak, C. J. F. and Vrugt, J. A.: Differential Evolution Markov Chain with snooker updater and fewer chains, Statistics and Computing, 18, 435–446, https://doi.org/10.1007/s11222-008-9104-9, 2008.

Tifafi, M., Camino-Serrano, M., Hatté, C., Morras, H., Moretti, L., Barbaro, S., Cornu, S., and Guenet, B.: The use of radiocarbon 14 C to constrain carbon dynamics in the soil module of the land surface model ORCHIDEE (SVN r5165), Geoscientific Model Development, 11, 4711–4726, https://doi.org/10.5194/gmd-11-4711-2018, 2018.

Tougma, I. A., Van de Broek, M., Six, J., Gaiser, T., Holz, M., Zentgraf, I., and Webber, H.: AMPSOM: A measureable pool soil organic carbon and nitrogen model for arable cropping systems, Environmental Modelling & Software, 185, 106 291, https://doi.org/10.1016/j.envsoft.2024.106291, 2025.

Travis, C. C. and Haddock, G.: On structural identification, Mathematical Biosciences, 56, 157–173, https://doi.org/10.1016/0025-5564(81)90052-3, 1981.

Treseder, K. K., Balser, T. C., Bradford, M. A., Brodie, E. L., Dubinsky, E. A., Eviner, V. T., Hofmockel, K. S., Lennon, J. T., Levine, U. Y., MacGregor, B. J., Pett-Ridge, J., and Waldrop, M. P.: Integrating microbial ecology into ecosystem models: challenges and priorities, Biogeochemistry, 109, 7–18, https://www.jstor.org/stable/41490541, publisher: Springer, 2012.

van Dam, D., van Breemen, N., and Veldkamp, E.: Soil organic carbon dynamics: variability with depth in forested and deforested soils under pasture in Costa Rica, Biogeochemistry, 39, 343–375, https://doi.org/10.1023/A:1005880031579, 1997.

Van de Broek, M.: Van de Broek and Six - GMD2026 - for peer review only, https://doi.org/10.5281/zenodo.17974745, 2025.

Van de Broek, M., Riley, W. J., Tang, J., Frey, S. D., and Schmidt, M. W. I.: Thermal Adaptation of Enzyme-Mediated Processes Reduces Simulated Soil CO2 Fluxes Upon Soil Warming, Journal of Geophysical Research: Biogeosciences, 129, e2024JG008 619, https://doi.org/10.1029/2024JG008619, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2024JG008619, 2024.

Van de Broek, M., Govers, G., Schrumpf, M., and Six, J.: A microbially driven and depth-explicit soil organic carbon model constrained by carbon isotopes to reduce parameter equifinality, Biogeosciences, 22, 1427–1446, https://doi.org/10.5194/bg-22-1427-2025, publisher: Copernicus GmbH, 2025.

Van Rompaey, A. J. J. and Govers, G.: Data quality and model complexity for regional scale soil erosion prediction, International Journal of Geographical Information Science, 16, 663–680, https://doi.org/10.1080/13658810210148561, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/13658810210148561, 2002.

von Fromm, S. F., Hoyt, A. M., Sierra, C. A., Georgiou, K., Doetterl, S., and Trumbore, S. E.: Controls and relationships of soil organic carbon abundance and persistence vary across pedo-climatic regions, Global Change Biology, 30, e17 320, https://doi.org/10.1111/gcb.17320, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.17320, 2024.

Vrugt, J. A.: Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation, Environmental Modelling & Software, 75, 273–316, https://doi.org/10.1016/j.envsoft.2015.08.013, 2016.

Walter, E. and Pronzato, L.: On the identifiability and distinguishability of nonlinear parametric models, Mathematics and Computers in Simulation, 42, 125–134, https://doi.org/10.1016/0378-4754(95)00123-9, 1996.

Wang, Z., Qiu, J., and Van Oost, K.: A multi-isotope model for simulating soil organic carbon cycling in eroding landscapes (WATEM_C v1.0), Geoscientific Model Development, 13, 4977–4992, https://doi.org/10.5194/gmd-13-4977-2020, publisher: Copernicus GmbH, 2020.

Wanika, L., Egan, J. R., Swaminathan, N., Duran-Villalobos, C. A., Branke, J., Goldrick, S., and Chappell, M.: Structural and practical identifiability analysis in bioengineering: a beginner's guide, Journal of Biological Engineering, 18, 20, https://doi.org/10.1186/s13036-024-00410-x, 2024.

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., Taylor, P. G., and Bonan, G. B.: 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, publisher: Copernicus GmbH, 2015.

https://doi.org/10.5194/egusphere-2025-6297
Preprint. Discussion started: 3 March 2026

Wieder, W. R., Hartman, M. D., Sulman, B. N., Wang, Y.-P., Koven, C. D., and Bonan, G. B.: Carbon cycle confidence and uncertainty: Exploring variation among soil biogeochemical models, Global Change Biology, 24, 1563–1579, https://doi.org/10.1111/gcb.13979, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.13979, 2018.

Wieder, W. R., Hartman, M. D., Kyker-Snowman, E., Eastman, B., Georgiou, K., Pierson, D., Rocci, K. S., and Grandy, A. S.: Simulating Global Terrestrial Carbon and Nitrogen Biogeochemical Cycles With Implicit and Explicit Representations of Soil Microbial Activity, Journal of Advances in Modeling Earth Systems, 16, e2023MS004 156, https://doi.org/10.1029/2023MS004156, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023MS004156, 2024.

Woolf, D. and Lehmann, J.: Microbial models with minimal mineral protection can explain long-term soil organic carbon persistence, Scientific Reports, 9, 6522, https://doi.org/10.1038/s41598-019-43026-8, publisher: Nature Publishing Group, 2019.

Wu, H., Zhu, H., Miao, H., and Perelson, A. S.: Parameter Identifiability and Estimation of HIV/AIDS Dynamic Models, Bulletin of Mathematical Biology, 70, 785–799, https://doi.org/10.1007/s11538-007-9279-9, 2008.

Wutzler, T., Yu, L., Schrumpf, M., and Zaehle, S.: Simulating long-term responses of soil organic matter turnover to substrate stoichiometry by abstracting fast and small-scale microbial processes: the Soil Enzyme Steady Allocation Model (SESAM; v3.0), Geoscientific Model Development, 15, 8377–8393, https://doi.org/10.5194/gmd-15-8377-2022, publisher: Copernicus GmbH, 2022.

Yu, L., Ahrens, B., Wutzler, T., Schrumpf, M., and Zaehle, S.: Jena Soil Model (JSM v1.0; revision 1934): a microbial soil organic carbon model integrated with nitrogen and phosphorus processes, Geoscientific Model Development, 13, 783–803, https://doi.org/10.5194/gmd-13-783-2020, publisher: Copernicus GmbH, 2020.