Received: 10 February 2025 | Accepted: 16 January 2026

DOI: 10.1002/saj2.70218

Soil Science Society of America Journal

ORIGINAL ARTICLE

Soil and Ecosystem Processes

ProCarbon-Soil: A dynamic model for improved model-data compatibility in carbon farming

Luís G. Barioni[1,2] | **Beatriz A. Valladão**[1] | **Vitor H. M. Mourão**[1,3] | **Robert P. Ewing**[4] | **Yusuf N. Karatay**[1,2] | **Júnior M. Damian**[1] | **Vinícius C. Melício**[1] | **Rodrigo P. A. Rejaili**[5] | **Rafael O. Silva**[2,†]

[1]Embrapa Digital Agriculture, Campinas, Brazil

[2]Global Academy of Agriculture and Food Systems, The University of Edinburgh, Edinburgh, UK

[3]Institute of Mathematics, Statistics and Scientific Computing, Campinas, Brazil

[4]Climate LLC, Seattle, USA

[5]Bayer CropScience, São Paulo, Brazil

Correspondence
Luís G. Barioni, Embrapa Digital Agriculture, Campinas, Brazil.
Email: luis.barioni@embrapa.br

[†]Deceased

Assigned to Associate Editor Girma Birru.

Funding information
Bayer Crop Science; Royal Society, Grant/Award Number: RSWVF\R2\222008

Abstract

Carbon farming is a nature-based solution to capture atmospheric CO_2 and store it as soil organic carbon (SOC). Carbon farming trading schemes (CFTS) incentivize farmers to adopt these practices. Integral to CFTS is forecasting SOC changes, typically achieved using traditional multicompartmental soil carbon models (mSCM), and monitoring total SOC stocks. However, traditional mSCM simulate unmeasurable compartments, leading to overparameterization and indeterminable partitioning among carbon compartments, suggesting a need for structural improvements. The ProCarbon-Soil (PROCS) model addresses this need by abstracting fundamental principles of mSCM, reducing SOC state variables to two (total carbon and decomposability), and employing only one stabilization parameter, compared to the four to eight state variables and 7–20 parameters typically required by mSCM. We mathematically derive methods that can use successive carbon measurements to estimate decomposability and initialize the model. PROCS can handle environmental modifiers and events such as crop rotations, tillage, and manuring events, and respond to soil characteristics and weather conditions. Tests show that PROCS can accurately reproduce synthetic SOC trajectories generated by an mSCM with perturbed parameters using short-term data (12 years) with acceptable accuracy (median root mean square error <1.03 Mg ha^{-1} and absolute median of mean bias <0.55 Mg ha^{-1}). In a cross-validation test, the mean normalized root mean square error (NRMSE) closely aligns with the coefficient of variation of white noise introduced in the synthetic data (4.15% vs. 4.00%, respectively) for augmented carbon inflow scenarios, whereas the

Abbreviations: CFTS, carbon farming trading scheme(s); CUE, soil microbial carbon use efficiency; FAST, Fourier amplitude sensitivity test; LIFS, laser-induced fluorescence spectroscopy; MAOM, mineral-associated organic matter; MB, mean bias; MDF, model data fusion; mSCM, multicompartmental soil carbon models; NRMSE, normalized root mean square error; POM, particulate organic matter; PROCS, ProCarbon-Soil; RMSE, root mean square error; SCM, soil carbon simulation model; SEM, standard error of the mean; SOC, soil organic carbon.

model exhibits higher errors for the no-carbon-inflow scenario (NRMSE = 5.48%, 7.25%, and 8.99% for 12, 24, and 50 years, respectively).

Plain Language Summary

Carbon farming rewards farmers for practices that capture carbon dioxide from the air and store it in soil, helping fight climate change. The company issuing these rewards must first accurately predict how much carbon will accumulate over time. Current prediction models are overly complex with many unmeasurable parts, making them difficult to use. We developed a simpler model called ProCarbon-Soil (PROCS) that uses only two measurable soil properties—total carbon content and how quickly it breaks down—instead of the four to eight unmeasurable parts used by traditional models. PROCS is easier to start up before making predictions and can account for farming practices like crop rotations and tillage. Testing showed that our model makes predictions of soil carbon changes that are nearly the same as those of a traditional model, making it practical for real-world applications. This advance helps make carbon farming more accessible and transparent for both farmers and carbon credit markets.

1 | INTRODUCTION

Carbon farming trading schemes (CFTS) have been proposed following the crediting mechanisms defined in Article 6 of the Paris Agreement (Simone et al., 2017) as a financial tool to incentivize carbon farming and climate change mitigation in the agricultural sector (Alexander et al., 2015; Badgery et al., 2020). In addition to removing atmospheric CO_2, increasing soil organic carbon (SOC) stocks can deliver significant co-benefits to agricultural production and food security, and enhance other socioeconomic outcomes (Chenu et al., 2019; Guillaume et al., 2022; Sharma et al., 2021; Smith et al., 2019).

Changes in SOC and biomass carbon stocks are considered good proxies for long-term net atmospheric CO_2 removals by agroecosystems and are commonly used for carbon farming accounting (Spotorno et al., 2024). Soil carbon simulation models (SCMs) can predict changes in SOC stocks and thus are widely proposed for planning and monitoring CFTS projects. SCM simulations can address a broader range of situations than those given in the experiments used for their development, hence expanding information space and time scope in a repeatable and auditable manner (Vannier et al., 2022), besides reducing the cost of planning and monitoring (Costa et al., 2020). In this context, SCM simulations are useful for (1) ex ante analysis of outcomes and target definition in CFTS projects, allowing economic evaluation of expected changes and optimization of soil management practices for high-performance carbon farming; (2) improving project performance evaluation through comparison with monitoring data and updated prediction of the project outcomes (e.g., through model data fusion [MDF]); and (3) optimizing control measures by adjusting planned actions, such as management practices and sampling frequency, to achieve project targets at the lowest possible financial cost and level of uncertainty.

Given the diverse applications, it is not surprising that dozens of SCMs have been developed (Manzoni & Porporato, 2009), addressing different requirements, data availability, scales, and purposes. For instance, SCMs can be run standalone or coupled with other component models to address problems of greater complexity. They can be integrated with biogeochemical models, crop growth models, ecosystem-scale models, and global circulation models, among others (Baveye, 2023; Campbell & Paustian, 2015; Le Noë et al., 2023; Manzoni & Porporato, 2009).

The emergence of CFTS poses new challenges to SOC modeling in agriculture because of the regional specificities of farming systems and the economic complexities of CFTS. Further development of SCMs is constrained by the limited availability of high-quality data from long-term experiments for calibration and validation, particularly in tropical regions. CFTS require that models overcome these limitations and provide reliable predictions across a wide range of specific agricultural contexts using practically and economically viable data. Furthermore, they must demonstrate robust theoretical grounding and self-consistency to be recognized as scientific models, and possess sufficient verisimilitude to meet scientific and accreditation standards (van der Voort et al., 2023; Verra, 2024).

Scientifically sound models are often first developed through systems analyses and syntheses (Walters et al., 2016). They encapsulate an understanding of less complex processes that is produced through controlled experiments and specially designed measurement methods (e.g., incubation experiments, static or dynamic flux chambers, or radiolabeled carbon). These experimental methods, while valuable for controlled studies, often lack scalability. The resulting research models often exhibit complexity and high computational demands. Furthermore, they may rely on costly measurements and be prone to overparameterization.

In contrast, an SCM's value in a CFTS depends on its ability to generate accurate predictions across a broad range of conditions and with minimal operational and monitoring costs (van der Voort et al., 2023). This ultimately translates to minimizing the number of soil measurements, efficiently collecting and verifying agronomic data at scale, and enabling the rapid, consistent, and cost-effective generation of numerous robust predictions.

Short-term CFTS requirements lean, therefore, toward a greater focus on the engineering aspects of the solution rather than on the underlying scientific representation. Longer term improvements in SCMs for CFTS will likely be driven by both improved theoretical understanding from laboratory and field experiments, particularly under controlled conditions, and the large quantity of soil carbon data that will flow from the monitoring of carbon farming initiatives. In summary, CFTS pragmatically prioritize transparency, accuracy, economic feasibility, and scalability over detailed scientific examination and theoretical understanding of the causes behind the results (Davoudabadi et al., 2021).

The engineering-driven requirements of CFTS may ultimately lead to machine learning rather than a dynamic systems modeling approach (Batjes et al., 2024; van der Voort et al., 2023). However, the current scarcity of long-term experiments across broad geographies and agricultural practices severely limits machine learning methods, particularly in tropical developing countries. Even for dynamic models, simpler rather than more complex models tend to better balance structural and parameter errors, avoiding overfitting (Villaverde et al., 2022). In the short term, therefore, hybrid modeling, data fusion, and data learning approaches are more likely to be successful, with principles from biophysics and soil science being tapped to constrain calibration (training) procedures and system dynamics (Buizza et al., 2022; Tao et al., 2023). Future scenarios where large quantities of monitored data are available may lead to more competitive applications of machine learning.

This paper describes the ProCarbon-Soil (PROCS) model, a new SCM developed through a collaborative research effort between the Brazilian Agricultural Research Corporation (Embrapa) and Bayer Crop Science. PROCS's design effectively balances model complexity with data availability for calibration and allows for improving MDF applications in SCM. It aims to combine science and engineering perspectives, especially given the requirements of CFTS and limited data availability.

Core Ideas

- ProCarbon-Soil (PROCS) model is designed for planning and monitoring carbon farming and soil carbon inventories.
- PROCS produces carbon stock trajectories similar to those of the Century model, using a reduced parameter set.
- Modeling measurable soil organic carbon (SOC) and decomposability enhances the correspondence between predictions and empirical data.
- PROCS is less prone to overparameterization and overfitting than multicompartmental models.

2 | MULTICOMPARTMENTAL MODELS IN CFTS

CFTS projects typically adopt multicompartmental soil carbon models (mSCM) as predictive tools. These models are rooted in pioneering research on carbon turnover by Hénin and Dupuis (1945), which was further refined through the development of the Century and RothC models, as documented by Coleman and Jenkinson (1996), Jenkinson and Rayner (1977), Parton (1996), Parton et al. (1988), and Parton and Rasmussen (1994).

2.1 | Multicompartment model structure

Multicompartmental models partition the carbon mass within a volume of soil (defined by a layer depth) into different compartments, according to its potential turnover rate. The flow of decomposed carbon from a compartment is partitioned into the following: (1) other—usually more stable—carbon compartments; (2) microbial carbon; and (3) gaseous carbon compounds (i.e., CO_2 and CH_4), as depicted in Figure 1.

As demonstrated by Sierra (2024), such a model constitutes a nonautonomous linear system (Equation 1).

$$\frac{d\vec{C}(t)}{dt} = \vec{B}(t)\, I(t) + \xi(t)\, A K \vec{C}(t) \tag{1}$$

in which $\vec{C}(t)$ represents carbon stock, $\vec{B}(t)$ is the partitioning of the carbon inflow rate $I(t)$, $\xi(t)$ represents environmental effects on carbon decomposition, A represents carbon

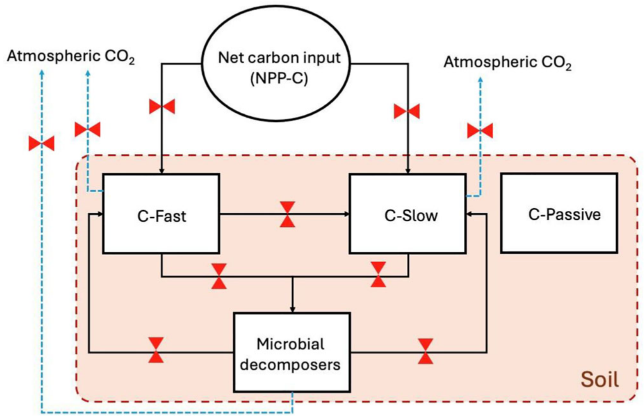

FIGURE 1 Simplified representation of the general structure of multicompartmental soil organic carbon (SOC) models depicting system dynamics with carbon flows and stocks. C-Fast, C-Slow, and C-Passive are, respectively, the carbon mass compartments for high, low, and zero potential turnover rate. NPP-C, net primary production of carbon (both aboveground and belowground).

flows between compartments, and K gives the compartments' turnover rates at a reference environment.

The general form (Equation 1) is subsequently shifted to the matrix form used in the Century model, specifically in Century's plant, soil organic matter (SOM), and environmental effects modules (Parton et al., 1994). The SOM submodel represents carbon with three compartments, the active, slow and passive pools, respectively, compartments 1, 2 and 3. The matrix components of Equation 1 take the following forms:

$$\frac{d\vec{\boldsymbol{C}}(t)}{dt} = \begin{bmatrix} \frac{dc_1(t)}{dt} \\ \frac{dc_2(t)}{dt} \\ \frac{dc_3(t)}{dt} \end{bmatrix}, \vec{\boldsymbol{C}}(\boldsymbol{t}) = \begin{bmatrix} c_1(t) \\ c_2(t) \\ c_3(t) \end{bmatrix}, \vec{\boldsymbol{B}} = \begin{bmatrix} b_1(t) \\ b_2(t) \\ 0 \end{bmatrix},$$

$$\boldsymbol{A} = \begin{bmatrix} -1 & a_{12} & a_{13} \\ a_{21} & -1 & 0 \\ a_{31} & a_{32} & -1 \end{bmatrix}$$

and

$$K = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{bmatrix}.$$

$\vec{B}(t) = (b_1(t), \ldots, b_n(t))$ contains coefficients for allocating the carbon inflow rate $I(t)$ to the n compartments, with $\sum_{i=1}^{n} b_i = 1$. In Century, the vector $\vec{B}(t)I(t)$ is given by a litter decomposition submodel, which also accounts for the spatial distribution of carbon inputs, distinguishing between aboveground and belowground compartments. No litter decomposition flows directly to the passive compartment (c_3), hence $b_3(t) = 0$.

In the A matrix, the diagonal $a_{jj} = -1$ represents carbon outflow from compartment j, with the value -1 indicating complete allocation of decomposed carbon from i to other compartments. The off-diagonal elements a_{ij} (with $i \neq j$) represent the proportion of the decomposed carbon in compartment j transferred to compartment i over dt. In order to ensure conservation of mass, the off-diagonal elements in matrix A, representing the total proportion of decomposed carbon transferred from compartment j to the other compartments, must sum to between zero and unity: $0 \leq \sum_{i=1, i\neq j}^{n} a_{ij} \leq 1$. Enforcing $\sum_{i=1, i\neq j}^{n} a_{ij} \leq 1$, for each column, ensures that no carbon is being generated during decomposition. Since $a_{jj} = -1$, it follows that:

$$-1 \leq \sum_{i=1}^{n} a_{ij} \leq 0 \text{ for each } j = 1, \ldots, n.$$

Whenever $\sum_{i=1}^{n} a_{ij} < 0$, a proportion of the carbon decomposed from the jth compartment is lost. The proportion of mass loss, primarily as CO_2, is therefore $|\sum_{i=1}^{n} a_{ij}|$ (Huang, 2024).

The term $\xi(t)$ represents a scalar function (alternatively a matrix function), which quantifies the effect of environmental factors on the decomposition rate. Please note $\xi(t)$ is fully defined by the following notation: $\xi(\vec{\Theta}_\xi, \vec{\mathrm{X}}(\mathrm{t}), \vec{\mathrm{E}}(\mathrm{t}))$, where $\vec{\Theta}_\xi$ is the vector of parameters, $\vec{\mathrm{X}}(\mathrm{t})$ is the system's state vector, and $\vec{\mathrm{E}}(\mathrm{t})$ is the vector of exogenous environmental input variables. The vector $\vec{\mathrm{E}}(\mathrm{t})$ encompasses the trajectories of environmental factors, such as soil temperature and available soil water content. The development and usage of $\xi(t)$ is detailed in Section 3.3.

2.2 | Model development gaps and requirements for novel model design

Given the abundance of SCMs, the development of a new model should be justified by stating what currently unmet requirements will be addressed by the new model. As the PROCS model targets CFTS, we discuss its requirements accordingly.

In CFTS, it is total carbon stock $C(t)$, rather than any compartment's stock $c_i(t)$, which is of interest. This is because the variation in soil carbon stocks in a time interval Δt between stock estimates, $\Delta\mathrm{SOC}(t) = C(t) - C(t - \Delta t)$, is the metric for gross removals credited before uncertainty discounts (Gold Standard, 2020; Verra, 2023). Furthermore, most of the existing longitudinal data on SOC in agriculture and on-farm measurements of CFTS pertain to the layer mean carbon concentration and SOC stocks.

The mSCM currently applied to carbon farming can produce the trajectory of total SOC stock via $C(t) = \sum_{i=1}^{n} c_i(t)$, where $c_i(t)$ is the i^{th} element of vector $\vec{C}$. They can also generate the bulk carbon turnover time, which indicates the stability of SOC, an essential parameter for CFTS. However, mSCMs currently used in CFTS have conceptual (theoretical) rather than measurable compartments, and longitudinal total carbon stock data are insufficient for initializing conceptual compartments.

Appropriate initialization is critical to the predictive performance of dynamic models. In most mSCMs, total soil carbon stocks can be initialized from measurements, but their partition among compartments cannot be directly observed. mSCMs usually assume a system steady state at initialization based on long-term scenarios that incorporate historical information on plant productivity and litter composition (Hashimoto et al., 2011). However, such data and steady-state assumptions are typically associated with high uncertainty (Lardy, 2011; Kanari et al., 2022), although initialization from native vegetation can benefit from productivity estimated from available climate and vegetation parameters (e.g., Del Grosso et al., 2008). Also, some models require computationally demanding model spin-ups (Foereid et al., 2012; Kanari et al., 2022; Mathers et al., 2023). Although computational demand is not critical in CFTS, response time and computing costs cannot be neglected.

Options to minimize such uncertainties include using measurable carbon fractions, such as particulate- and mineral-associated SOC to constrain the partition of different pools in biogeochemical models (Dangal et al., 2022) as well as machine learning methods. Despite several initiatives, methods of mSCM initialization remain nonstandardized.

Another limitation for CFTS use is that mSCMs are overparameterized in relation to total SOC dynamics, which incurs problems in parameter identification due to low sensitivity and high collinearity (Luo et al., 2009). Likewise, using data fusion to ground-truth a model is complicated for overparameterized models (Trudinger et al., 2008; Viskari et al., 2020). Moreover, because mSCMs have no standard definitions for their theoretical compartments, they lack standard values for potential turnover rates, also known as decay rate under optimum conditions or optimum decomposition rates (Campbell & Paustian, 2015). This is likely because the concept of multiple theoretical compartments with characteristic turnover rates is not directly supported by biophysical mechanisms (Kleber & Lehmann, 2019; Lehmann & Kleber, 2015): chemical extraction protocols do not correspond to stages in microbial decomposition, nor is turnover rate a simple function of initial chemical composition. Therefore, comparing different mSCMs is largely limited to comparing their simulated carbon stock trajectories, $C(t)$, and ecosystem residence time, $(\xi(t)AK)^{-1}\vec{B}$ (Xia, 2024). Challenges in achieving correspondence between models and data may stem from a conceptual mismatch between measurements and actual soil processes, although such limitations have been addressed in the newest mSCM, as discussed below.

Current SOC modeling efforts to overcome such problems may be grouped into the following:

1. Using models that partition SOC into compartments defined by new laboratory techniques, so that compartments are rigorously initialized and their carbon stocks measurable, here denoted as mSCM+ (Robertson et al., 2019), for example, MIMICS (Kyker-Snowman et al., 2020), MEMS (Zhang et al., 2024), and Millennial (Abramoff et al., 2022).
2. Understanding how and what environmental mechanisms (i.e., chemical, biological, and physical) control SOC turnover (Todd-Brown et al., 2013; Wieder et al., 2015).
3. Developing single-compartment continuous-quality models (Bosatta & Ågren, 2003; Bruun et al., 2010), here denoted SCQM. PROCS is in this group.

Proper determination and modeling of measurable compartments of SOC could significantly improve quantification of flows and understanding of the dynamics of SOC stabilization, distinguishing mechanisms and environmental effects of the different carbon forms in soil. Indeed, the synergistic interactive evolution of scientific knowledge, sensing technology and modeling associated with the development of mSCM+ is the most promising pathway for CFTS to produce explainable, high-quality predictions from next-generation models. Techniques to assess carbon stability, like physical fractionation (Elliott et al., 1996; Six et al., 1999), mineral-specific surface area and recalcitrant compounds (Kirschbaum et al., 2019), Fourier transform infrared spectroscopy (Lei et al., 2023), and laser-induced fluorescence spectroscopy (Villas-Boas et al., 2020), have recently gained popularity and had their results applied to mSCM+. However, there are still difficulties in relating those laboratory results precisely to the partitioning among model compartments (Robertson et al., 2019).

Uncertainty is a core issue in CFTS, as current protocols penalize uncertainty in carbon crediting (Gold Standard, 2020; Verra, 2023). Meanwhile, the advanced measurements required by new generation mSCMs are currently unaffordable in CFTS. It has been argued that increasing the structural complexity of SOC dynamics is not guaranteed to reduce uncertainty (Davoudabadi et al., 2024; Manzoni & Porporato, 2009), and a more parsimonious approach offers strong advantages (Derrien et al., 2023; Shi et al., 2018).

PROCS is a parsimonious model which adopts a continuous quality approach, seeking to meet two key requirements for SCMs in CFTS applications: (1) inputs, outputs, and state variables should be affordably assessable on-farm (e.g., SOC stock and decomposability) and (2) model parameters should be statistically identifiable from total SOC stocks longitudinal data.

The first requirement allows the system to be readily initialized, validated, and corrected through MDF. Because carbon markets' main interest is to quantify long-term net balance of greenhouse gases across the soil–atmosphere interface, models and measurements must quantify changes in soil carbon stocks while considering carbon decomposability to allow evaluation of reversibility, that is, the possibility of the captured carbon being reemitted later. The associated measurements should also be cost-effective.

The second design requirement relates to the number of empirical parameters and how they are embedded into the model structure. Parameter values should be uniquely estimable from existing field experiments and readily corrected from affordable on-farm measurements.

3 | DECOMPOSABILITY IN A REDUCED-COMPARTMENT MODEL

Here, we describe how organic carbon decomposability is handled, combining mathematical derivation with considerations of compatibility with existing models, aiming to build a more parsimonious model with state variables quantifiable from measurements. Whenever possible, model inputs and most parameters are derivable and comparable with existing models. Model compatibility speeds model development and adoption by enabling direct comparison of a new model's states and parameters with previous models and allowing users of the previous models to readily learn the new model.

We start our development from the simplest differential equation SCM, the first-order kinetics carbon dynamics in which the rate of decomposition is directly proportional to the amount of SOC at any time t (Equation 2).

$$\frac{dC(t)}{dt} = I(t) - \tau_0 C(t), \tag{2}$$

where $I(t)$ is the total carbon inflow to the soil and $C(t)$ is the total carbon stock. Valladão (2022) demonstrated that it is possible to reproduce exactly the total carbon trajectory produced by an mSCM by replacing the fixed turnover rate τ_0 by a variable turnover rate $\tau(t)$:

$$\frac{dC(t)}{dt} = I(t) - \tau(t)\, C(t). \tag{3}$$

Because Equation (3) represents a single compartment, carbon outflow is $\tau(t)C(t)$ and $\tau(t)$ implicitly includes environmental decomposition. We started from the concept developed in Equation (3) and introduced a quantitative measure of decomposability, $\rho(t)$, which is a component of the turnover rate $\tau(t)$:

$$\tau(t) = \xi(t)\, \rho(t). \tag{4}$$

An essential concept in PROCS is that decomposability (ρ) indicates the (aggregate) state of vulnerability to decomposition in which carbon is found, a state which is related to the carbon's form, molecular structures, and interactions with the soil matrix. Unlike the resulting turnover rate, which is influenced by environmental factors, $\rho(t)$ is conceptually an intrinsic property of soil carbon itself. Because decomposability $\rho(t)$ is independent of environmental factors, such as soil microbiology or enzymatic activity, a given soil decomposability differs in turnover rates only under different environmental conditions. For instance, if soil with $\rho = 2E-4(day^{-1})$ was freeze-dried, its instantaneous turnover rate $\tau(t)$ would go to zero while its ρ value would remain unchanged. Dynamic environmental factors, such as soil moisture, temperature, and O_2 concentration, are accounted for in a separate decomposition modifier function $\xi(t)$, which determines the combined effects of the current soil environment on the SOC decomposition rate.

Defining $C(t)$ as a scalar quantity that represents the sum of an mSCM carbon stock vector $\vec{C}(t)$, it is possible to represent the associated carbon dynamics as $\frac{dC(t)}{dt} = I(t) - \xi(t)\rho(t)C(t)$. Therefore, making carbon losses equivalent for both models implies the following:

$$\begin{aligned} -\xi(t)\,\rho(t)\,C(t) \equiv \xi(t)\,\boldsymbol{AK}\vec{\boldsymbol{C}}(\boldsymbol{t}) \Rightarrow -\rho(t)\,C(t) \\ = \textstyle\sum_{j=i}^{n} \sum_{i=i}^{n} a_{ij}\, k_j c_j(t). \end{aligned} \tag{5}$$

Isolating $\rho(t)$ from Equation (5), we obtain (see details in Supporting Information, Section 1):

$$\rho(t) = \sum_{j=1}^{n} m_j p_j(t), \tag{6}$$

where

$$m_j = -\left(\sum_{i=1}^{n} a_{ij}\right) k_j \tag{6a}$$

in which m_j is the reference rate of carbon loss of the jth compartment, and

$$p_j(t) = \frac{c_j(t)}{\sum_{j=1}^{n} c_j(t)} \tag{6b}$$

in which p_j is the proportion of the total carbon mass in the jth compartment.

Equation (6) states that carbon decomposability is the inner product of the vector with the proportion of carbon in each compartment and the vector $\vec{M}$, which in a mSCM defines the reference rate of carbon loss of each compartment. It emphasizes that changes in carbon decomposability are actually represented in mSCMs through changes in the proportion of carbon mass in each compartment. However, while the state of an mSCM determines a unique value for $\rho(t)$, the reverse is not true. That is, multiple distinct sets of values of the vector $\vec{C}(t)$ in an mSCM can all result in the same value of $\rho(t)$, that is, equifinality. Equifinality, a consequence of the nonuniqueness of parameter combinations and a model structure built on immeasurable compartments, poses challenges for model calibration and MDF, leading to a propensity for overfitting and prediction uncertainty. The presence of equifinality also makes it difficult to initialize the model from longitudinal data, to assess an mSCM's carbon partitioning across compartments from field data, and to compare multiple different mSCMs.

Calculating $\rho(t)$ from compartment parameters of the A matrix and the carbon partition among the compartments in $\vec{C}$ allows estimating exactly the same instantaneous turnover of mSCM (assuming environmental modifier values are identical). However, to produce a whole SOC stock trajectory it is also necessary to calculate the dynamics of $\rho(t)$. Therefore, two differential equations are required: one to describe how carbon stock (mass) changes over time, and the other to model changes in turnover rate (decomposability) due to changes in chemical and biophysical properties.

3.1 | Changes in decomposability over time

Carbon's decomposability in a soil system is modified over time both by the inflow of new carbon compounds, and by transformations of carbon occurring within the soil system. Such transformations tend to produce more stabilized carbon through several mechanisms related to chemical recalcitrance, microbial stoichiometry, aggregation and aggregate disruption, organo-mineral associations, and others (Derrien et al., 2023; Mao et al., 2024).

To streamline the initial algebraic manipulation, we model soil carbon dynamics within a single soil volume defined by the top 30-cm layer and land area defined by a plot, a paddock, or a stratum. While soil three-dimensional heterogeneity does impact SOC dynamics and model results (as model inputs can also vary in space), we assume that means suffice within the vertical and horizontal bounds of our soil volume. The 30-cm depth is aligned with soil carbon dynamics calculations that are standard for inventories, such as IPCC Tier 1 (Intergovernmental Panel on Climate Change, 2006, Chap. 4), and carbon farming (Verra, 2023, 2024), and with several well-established biogeochemical models though some simulate shallower layers, for example, RothC (Jenkinson & Rayner, 1977) at 23 cm and ICBM (Andrén & Kätterer, 1997) at 25 cm.

To simplify the mathematical description, we consider only one soil layer and omit erosion, leaching, and other events which disrupt carbon stability, although they also are treatable in extensions of the formulation presented herein. We establish (see details in Supporting Information, Section 2) that the effects of carbon inflow (dilution) and transformation are additive, allowing us to define

$$\frac{d\rho(t)}{dt} = \frac{df_I(t)}{dt} + \frac{df_A(t)}{dt}, \tag{7}$$

where

$$\frac{df_I(t)}{dt} = \boldsymbol{\nabla\rho(t)} \cdot \vec{\boldsymbol{B}}\boldsymbol{I}(t), \tag{7a}$$

and

$$\frac{df_A(t)}{dt} = \boldsymbol{\nabla\rho(t)} \cdot \xi(t)\,\boldsymbol{AK}\vec{\boldsymbol{C}}\boldsymbol{(t)}. \tag{7b}$$

$\frac{df_I(t)}{dt}$ represents the effect of carbon inputs flowing into the system, while $\frac{df_A(t)}{dt}$ represents the effect of alterations (transformations) of carbon forms. In mSCMs, such transformations are generally associated with carbon losses. Modeling the carbon input and output effects separately is an important feature of the PROCS model, as it enables isolating the effect of stabilization parameters due to decomposition and calibrating them using no-carbon inflow experiments; this improves parameter estimation.

The effect of carbon input on decomposability can be derived in an analytically exact form by developing Equation (7a) (see Supporting Information, Section 3), resulting in:

$$\frac{df_I(t)}{dt} = \frac{I(t)}{C(t)}\left(\rho_I(t) - \rho(t)\right) \tag{8}$$

where the decomposability of the carbon entering the soil, ρ_I, is a model input related to the quality of the inflow material, which has direct equivalence to mSCM parameters (9):

$$\rho_I(t) = \sum_{i=1}^{n} m_i\, b_i(t) \tag{9}$$

with $b_i(t)$ being the proportion of carbon entering the ith compartment at time t, and m_i the potential rate of carbon loss due to decomposition for that compartment.

By further developing Equation (7b), we find that the effect of carbon transformation (alteration) on decomposability has no general exact solution (see details in Supporting Information, Section 4):

$$\frac{df_A(t)}{dt} = \sum_{i=1}^{n} \sum_{j=1}^{n} m_i a_{ij} k_j p(t) \tag{10}$$

Note that the second term in Equation (10) has a direct dependence on the state variables of the mSCM, so it is not reducible to our PROCS state variables (C and ρ). This prevents an exact match between PROCS and a generic mSCM, except for the special case of a two-compartment model (see details in Supporting Information, Section 7).

3.2 | Single-compartment formulation for SOC stabilization

One of the simplest equations of carbon decomposition (although others could be used) is (Manzoni et al., 2012; Rovira & Rovira, 2010) as follows:

$$\frac{dC(t)}{dt} = -\left(\lambda_0 + \lambda_1 e^{-\omega t}\right) C(t) \tag{11}$$

where λ_0, λ_1, and ω are positive constants. Assuming that soil carbon decomposition can be described similarly, from Equation (11), we can derive the following (see details in Supporting Information, Section 9):

$$\frac{df_{1A}(t)}{dt} = -\omega \rho(t) \tag{12}$$

Testing indicates that dividing $\frac{df_{1A}}{dt}$ by $C(t)$ improves parameter stability and reduces uncertainty.

3.3 | Environmental drivers

Environmental factors influencing carbon turnover can be categorized into four groups:

1. continuous factors which directly modify turnover rates;
2. continuous factors which modify the speed of carbon stabilization, but with no instantaneous impact on the turnover rates;
3. events that abruptly change the value of state variables, such as tillage or manuring, and;
4. events that abruptly change model parameters, such as changing the source of carbon.

SOC instantaneous carbon turnover modifiers (1), typically soil moisture and temperature, are handled in PROCS through a scalar function $\xi(t)$ that has exogenous factors as arguments. Moreover, $\xi(t)$ also affects the rate of chemical stabilization because accelerating carbon turnover also speeds up the breakdown of labile carbon into more stable forms. This pattern is analogous to mSCM mechanisms: any mSCM following the general Equation (1) assumes implicitly that increasing carbon turnover speeds up stabilization, as accelerating turnover rates results in greater flows toward more stable compartments.

PROCS also considers factors (2) that modify the stabilization speed without instantaneous impact on decomposition through a scalar function $\mu(t) = \mu(\vec{\Theta}_\mu, \vec{X}(t), \vec{E}(t))$, where $\vec{\Theta}_\mu$ is the vector of parameters, $\vec{X}(t)$ is the system's state vector and $\vec{E}(t)$ is the vector of exogenous environmental input variables, such as soil texture.

In an mSCM, factors changing the speed of stabilization are related to changes in the partitioning of flows among compartments, that is, their balance among stabilization pathways. This would be accomplished by changing the A matrix elements or the relative value of the diagonal elements of the matrix K in Equation (1). The same effect could be achieved by making the environmental multiplier a matrix. To illustrate, we again draw on the Century soil carbon submodel, where the environmental modifier is the following multiplier to decomposition rates $AK\vec{C}(t)$:

$$f_O(t) f_W(t) f_T(t) \begin{bmatrix} f_{C_1}(t) f_X(t) & 0 & 0 \\ 0 & f_{C_2}(t) & 0 \\ 0 & 0 & f_{C_3}(t) \end{bmatrix}. \tag{13}$$

The functions $f_O(t)$, $f_W(t)$, $f_T(t)$, $f_X(t)$, $f_{C_1}(t)$, $f_{C_2}(t)$, and $f_{C_3}(t)$ are scalar functions that respectively calculate the effects of anaerobiosis, soil water, soil temperature, soil texture, and soil management in compartments 1 (fast), 2 (slow), and 3 (passive). All these functions use specific parameter sets, system state variables, and exogenous environmental input information. Since the resultant matrix (Equation 13) cannot be analytically reduced to a single scalar function, we incorporated the effects otf anaerobiosis, soil water and soil temperature into $\xi(t)$ as multipliers, the soil texture effect as $\mu(t)$ and soil management as discrete events.

Discrete events are typically associated with management interventions, which can be scheduled by the user or triggered by some internal model state (e.g., crop maturity for harvesting). Discrete events described in (3) cause instantaneous changes to state, which in PROCS means changes to $C(t)$ and $\rho(t)$. Changes in decomposability due to conventional tillage's mixing of aboveground and belowground plant carbon or manure addition are calculated as the mass weighted mean of the decomposability of the added carbon, $\rho_I(t)$, and the soil decomposability at the time of addition, $\rho(t)$; priming effects are not accounted for. In PROCS, we have the following:

$$\rho(t) \leftarrow \frac{I_q(\text{ev}, t)\,\rho_{I_q}(\text{ev}, t) + C(t)\,\rho(t)}{I_q(\text{ev}, t) + C(t)}. \tag{14}$$

where $I_q(\mathrm{ev}, t)$ represents a one-time addition of carbon (Mg ha^{-1}). The ← symbol denotes a computational assignment, not an equality. The mathematical basis for Equation (14) can be found in the Supporting Information, Section 14. Tillage likewise triggers an abrupt change in decomposability, represented in the model as follows:

$$\rho(t) \leftarrow \kappa(\mathrm{ev}, t)\,\rho(t), \tag{15}$$

where $\kappa(\mathrm{ev})$ is a scalar which depends on the properties of the event ev. For tillage, $\kappa(\mathrm{ev}) \geq 1$, and it is greater for more intense tillage (e.g., conventional) than for tillage that disturbs the soil less (e.g., no tillage). The model provides a dynamic attractor pulling ρ_T toward its non-perturbed expected state, mirroring the behavior of the biophysical dynamics in the field. The effect of erosion on SOC is set through a simple decay model of carbon concentration with soil depth when soil is under no-till management (Supporting Information, Section 15)

In contrast, the discrete events described in (4), such as sowing a new crop, change model parameters without instantaneously changing the state variables. In the case of PROCS, sowing changes the model's inflow carbon decomposability, $\rho_I(t)$, drawing on known values for various crops.

4 | MODEL FORMULATION STRENGTHS AND LIMITATIONS

PROCS, the single-compartment soil carbon model developed in the previous sections, is summarized as follows:

$$\begin{cases} \dfrac{dC(t)}{dt} = I(t) - \xi(t)\,\rho(t)\,C(t) \\ \dfrac{d\rho(t)}{dt} = \dfrac{I(t)}{C(t)}\left(\rho_I(t) - \rho(t)\right) - \xi(t)\,\mu(t)\,\dfrac{\omega\rho(t)}{C(t)} \end{cases} \tag{16}$$

where $C(t)$ is the total soil carbon stock, $I(t)$ is the carbon inflow rate, ρ is the soil carbon decomposability, and ρ_I is the decomposability of carbon inflow. ω is a non-negative parameter, and the functions $\xi(t)$ and $\mu(t)$ translate how environmental factors affect soil carbon turnover and stabilization rate, respectively (see Section 3.3 for details).

The PROCS formulation meets all the basic postulates for an SCM as defined by Sierra and Muller (2015) and therefore has the same scientific basis as other models currently applied in CFTS. Mass balance is enforced, with gas production tracked via system outflows ($\xi(t)\rho(t)C(t)$). Carbon flows are considered in terms of both mass and quality. Carbon transformations effects on the speed of decomposition are modeled by a single derivative $d\rho/dt$ rather than changes in partitioning of carbon in compartments with fixed decomposability.

Table 1 illustrates the compatibility of PROCS and multicompartmental models' formulations. Generally, only the soil endogenous stabilization function needs to be approximated in relation to traditional first-order multicompartmental models.

Internal transformations of organic matter are accounted for by ρ in a continuous manner. Although neither mSCM nor PROCS can represent measurable carbon partitions in soil, ρ has the advantage of being quantifiable from total carbon longitudinal data (see details in Supporting Information, Section 5). The model can, therefore, be initialized by using either longitudinal data or steady-state conditions (see details in Supporting Information, Section 6).

PROCS offers a more parsimonious approach to modeling carbon turnover dynamics.

In the current formulation (Equation 16), PROCS requires only the initial state defined by two state variables (carbon stock and decomposability), one intrinsic potential stabilization parameter (ω) to govern the pattern of soil carbon stabilization and two input variables (carbon inflow rate and carbon inflow quality, i.e., decomposability) to simulate the carbon stocks and carbon turnover rates in a constant environmental condition. In PROCS, stabilization is not represented through a specific mechanistic pathway, but rather through its fundamental quantitative manifestation, namely the progressive increase in soil carbon turnover time along carbon transformations. The PROCS framework is not restricted to a single formulation, as presented in Equation (16). Specifically, PROCS only assumes that the rate of stabilization increases as carbon becomes less decomposable, leading asymptotically to a steady-state turnover time. Moreover, the PROCS structure can flexibly incorporate additional stabilization factors associated with soil properties, such as texture and mineralogy, mineral-associated organic matter (MAOM), particulate organic matter (POM), soil microbial carbon use efficiency (CUE), and aggregation. For related work with DayCent, see Dangal et al. (2022).

Alternative stabilization functions with single and multiple parameters (e.g., nonlinear or quadratic forms) are fully compatible with the model structure. For example, one can use the quadratic function derived from bi-compartmental models. In that case, empirical calibration can also be constrained to a single parameter, as two of them are constrained by mathematical properties of the matrix and first principles of carbon stabilization (see details in Supporting Information, Section 7). So, at least in the case of bicompartmental models, it unequivocally demonstrates the redundancy and equifinality embodied in mSCMs.

The PROCS formulation allows drastically reducing the number of parameters. For the sake of comparison, RothC utilizes seven parameters: four parameters for compartments' turnover rates, one parameter to account for the proportion of CO_2 losses, and two parameters to estimate inert organic matter. Similarly, Century's soil carbon submodel employs eight

TABLE 1 State, flows, and environmental effects on turnover rates of a multicompartmental soil carbon model (mSCM) and its correspondence in the ProCarbon-Soil (PROCS) model.-

Model	Compartment	Exogenous Inflow	Endogenous Inflow	State (Mass)	Compartment outflow	Exogenous outflow (CO_2 Loss)	Residence time	Effect of carbon inflow on decomposability	Effect of carbon outflow on decomposability
mSCM	1	$\sum_i^{nl} L_i(t) \cdot k_i \cdot a_{i,1}$	$\sum_{i=2}^{n} \xi_i(t) \cdot C_i \cdot k_i \cdot a_{i,1}$	$c_1(t)$	$c_1(t) \cdot k_1$	$\xi_1(t) \cdot m_1 \cdot c_1(t)$	$\frac{1}{\xi_1(t) \cdot k_1}$	NA	NA
	2	$\sum_i^{nl} L_i(t) \cdot k_i \cdot a_{i,2}$	$\sum_{i=1;i\neq 2}^{n} \xi_i(t) \cdot \cdot C_i(t) \cdot k_i \cdot a_{i,2}$				$\frac{1}{\xi_2(t) \cdot k_2}$	NA	NA
	N	$\sum_i^{nl} L_i(t) \cdot k_i \cdot a_{i,n}$	$\sum_{i=1}^{n-1} \xi_i(t) \cdot C_i(t) \cdot k_i \cdot a_{i,n}$				$\frac{1}{\xi_n(t) \cdot k_n}$	NA	NA
	Total	$I(t) = \sum_{i=1}^{nl} L_i(t) \cdot (k_i - m_i)$	$\sum_{i=nl+1}^{n} \xi_i(t) \cdot C_i(t) \cdot (1 - m_i)$	$C(t) = \sum_i c_i(t)$	$\sum_i c_i(t) \cdot k_i$	$\sum_i \xi_i(t) \cdot c_i(t) \cdot m_i$	$\frac{C}{\sum_i \xi_i(t) \cdot m_i \cdot C_i}$	$\frac{I(t)}{C(t)} \cdot (\sum_{i=1}^{n} b_i m_i - \sum_{i=1}^{n} p_i(t) m_i)$	$-\xi(t) \sum_{i=1}^{n} \sum_{j=1}^{n} m_i a_{ij} k_j p(t)$
PROCS	Type of corresp.	Exact	No internal flows	Exact	–	Exact[a]	Exact[a]	Exact	Approximated[b,c]
	Total	$I(t)$	NA	$C(t)$	–	$\xi(t) \cdot \rho(t) \cdot C(t)$	$\frac{1}{\xi(t) \cdot \rho(t)}$	$\frac{I(t)}{C(t)}(\rho_I(t) - \rho(t))$	$-\xi(t) \frac{1}{C} \cdot \mu(t) \cdot \omega \cdot \rho(t)$

[a]If coupled with dynamic litter model; approximated for discrete event carbon inputs.
[b]If the environmental modifier applied to all compartments, that is, ξ scalar. Otherwise off-diagonal values and differences among diagonal elements of matrix ξ can be approximated through the μ function.
[c]It is approximated if there are three or more interacting compartments in the multicompartmental model. There is an exact generic solution for two interacting compartment models.

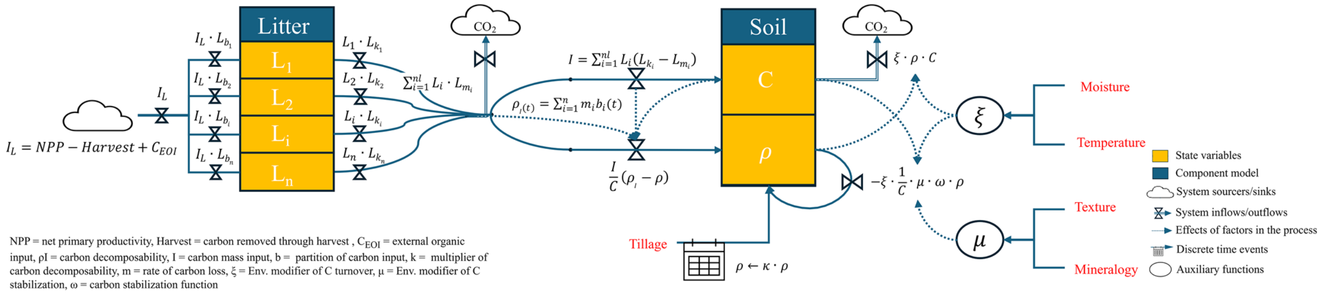

FIGURE 2 Dynamics of the ProCarbon-Soil (PROCS) soil organic carbon (SOC) model with a coupled generic litter model, including continuous environmental and management discrete event effects. In the litter model, I_L represents the total carbon inflow to the litter, $I_L = \text{NPP} - \text{Harvest} + C_{\text{EOI}}; L_{b_i}$ represents the partition of carbon among compartments (aboveground and belowground); L_i represents the carbon mass in the ith litter compartment; L_{k_i} is the reference turnover rate for the ith litter compartment, and; L_{m_i} is the reference rate of CO_2 loss of the ith litter compartment. All other symbols and equations as previously defined.

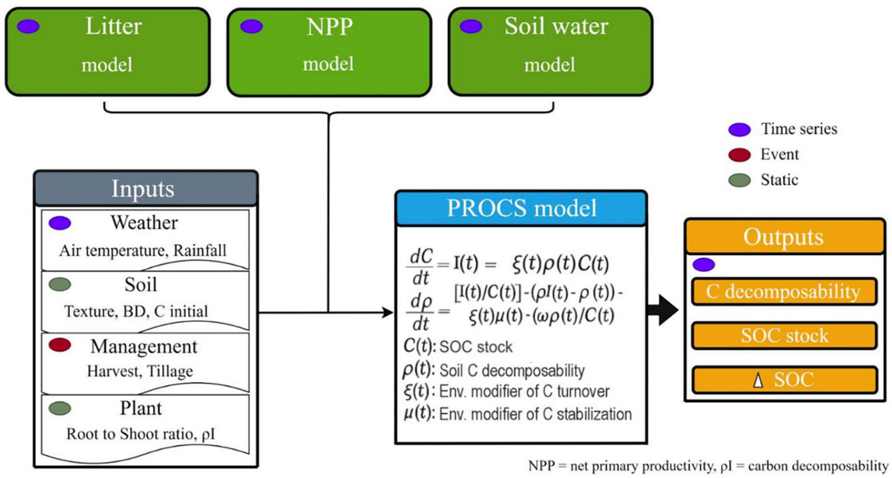

FIGURE 3 Inputs and outputs of the ProCarbon-Soil (PROCS) model and typical coupling with other models for soil carbon dynamic simulations.

parameters for turnover dynamics (three for compartments' turnover rates and five for intercompartmental flows).

Previous research has already highlighted the dominance of a single parameter: the slow compartment maximum turnover rate contributes approximately three-quarters of the variance (Mathers et al., 2023). This finding suggests that a context-specific balance should be found between model complexity and predictive performance. While complex models may allow representing a wider range of processes, appropriately parameterized simpler models may be sufficient to capture key dynamics needed to the monitoring and prediction of SOC stocks in CFTS, particularly when coupled with MDF and/or AI algorithms.

While PROCS drastically reduces the number of parameters required to account for carbon stabilization, the number of parameters required to accommodate environmental modifiers is similar to those in standard mSCM.

The PROCS model is susceptible to failure in some scenarios. Some limitations are inherited from early mSCMs from which it was derived. A limitation of the PROCS approach is that the state feedbacks of aggregation, MAOM, POM, soil microbiota, and other drivers of carbon decomposability cannot be explicit. If such measurements were considered, the dynamics of decomposability would need to be supported by measurements of indicators of carbon stability (MAOM, for instance) along the CFTS project monitoring. These could then be used to initialize decomposability, as arguments of the (t) function, and in MDF. The impact of replacing such state dynamics by periodic measurements on the prediction of carbon stocks in models for CFTS is an area for future research. At present, these measurements are not conducted in CFTS due to the same economic constraints which prevent the adoption of mSCM+.

PROCS does not explicitly represent priming effects. When fresh carbon is added, PROCS increases decomposition and turnover rates accordingly. However, as in classical mSCMs, such as Century and RothC, this response does not constitute a true priming mechanism. Priming representation would

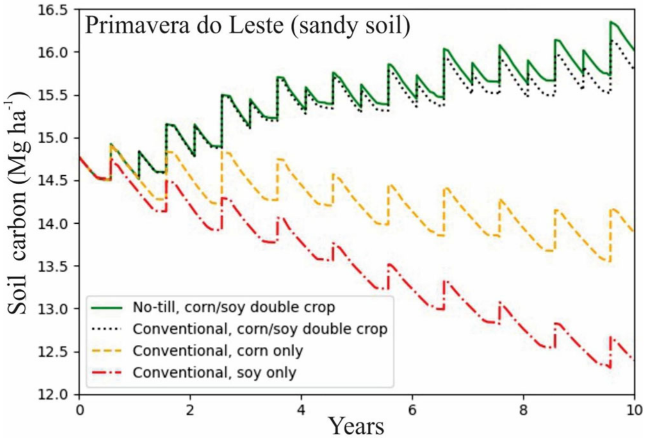

FIGURE 4 Example simulation suite comparing four cropping and tillage scenarios.

require changing maximum turnover rates of soil carbon compartments due to the addition of fresh material and microbial dynamics. This is carried out in more detailed (research-oriented) mechanistic models through explicit representations of microbial biomass dynamics, enzyme production, and carbon use efficiency regulation. As a result, PROCS may have limited ability to capture system responses under conditions where microbial growth or efficiency is strongly constrained, such as severe nutrient limitation, energy limitation, or soil contamination. In addition, PROCS shares limitations with other single-layer models; its lack of depiction of vertical environmental changes within the soil profile reduces its accuracy regarding anoxic conditions, variable rhizodeposition, and the vertical movement of soluble carbon. These constraints may reduce model accuracy in systems (e.g., systems dominated by tree components or deep-rooted grasses) where subsoil processes play a dominant role in SOC dynamics. Finally, PROCS's reliable application is constrained to the top layer of well-drained mineral soils.

Together, these limitations define the intended scope of PROCS as a parsimonious modeling framework suitable for long-term assessments of monitored agricultural systems, while highlighting conditions under which more mechanistic or depth-resolved approaches may be required.

Carbon inflows from above- and belowground litter can be treated exogenously in simplified model configurations, based on estimates derived from decay rates and CO_2 loss data obtained from literature or public datasets (SIDB; LIDET; Wu et al., 2025), or by coupling a standard litter dynamics model. We provide a generic formulation for a coupled dynamic litter model (Figure 2). A generic multicompartmental litter model with aboveground and belowground compartments can be associated with PROCS's ρ_I input variable through Equation (9). Therefore, it is possible to use an existing litter model in its original form, as demonstrated in Supporting Information, Section 16. By coupling with a litter model, the effect of aboveground litter accumulations on soil microclimate (e.g., soil temperature and moisture) may be considered by their impacts on decomposition rates through the environmental decomposition modifier function, $\xi(t)$.

5 | NUMERICAL EVALUATION AND OVERALL BEHAVIOR

5.1 | Setup, initialization, and example simulations

The model is provided as a compiled Windows dynamic linked library (.dll file) and a Linux shared object (.so file) and allows user interaction via a Python interface. This interface gives access to functions for specifying inputs and directly altering them, scheduling events, and exporting a dataset of daily values.

The model is set up for a run by supplying (1) soil texture and texture-derived parameters, (2) weather information, generally given by specifying the nearest city, and (3) the kind of carbon initialization that is desired. Soil carbon levels can be initialized any of three ways (see details in Supporting Information, Section 6): by (1) specifying the biome, such that the model solves for equilibrium values of $\boldsymbol{C}$ and ρ; (2) specifying the biome and the desired soil carbon value $\boldsymbol{C}$, such that the model attempts to solve for a decomposability value ρ that makes the given carbon value work, or (3) giving a sequence of dates and carbon values, from which the model derives a unique trajectory. Soil carbon initialization can take several seconds, as it entails solving a constrained optimization problem. Nonetheless, because only two values—soil carbon and decomposability—are being initialized, initialization is rapid compared to the lengthy spin-up needed by many mSCMs. Information about carbon inflow can be derived from user-defined crop productivity or aboveground biomass by informing harvest index and root-to-shoot ratio. It can also be provided by a coupled crop model, such as CNPP (Colmanetti et al., 2026), which was used in the simulations below, along with available soil water dynamics. Also, a simple litter model can be used to calculate CO_2 losses and carbon inflow quality through litter decomposition (Figure 3).

A comparison of four cropping and tillage scenarios (Figure 4) illustrates some of PROCS's modeling capabilities. All four scenarios model a sandy field near the Brazilian city of Primavera do Leste (15.52° S, 54.33° W) that initially supported the Cerrado biome with soil carbon at 21.3 Mg ha^{-1}. The native vegetation was burned, and the field supported summer corn with moldboard plowing for 40 years. From this point (year 0 in Figure 4), the four 10-year scenarios diverge: scenario (1) grows summer soybeans and off-season corn no-till, (2) grows the same crop sequence with moldboard plowing, (3) grows summer corn only with moldboard

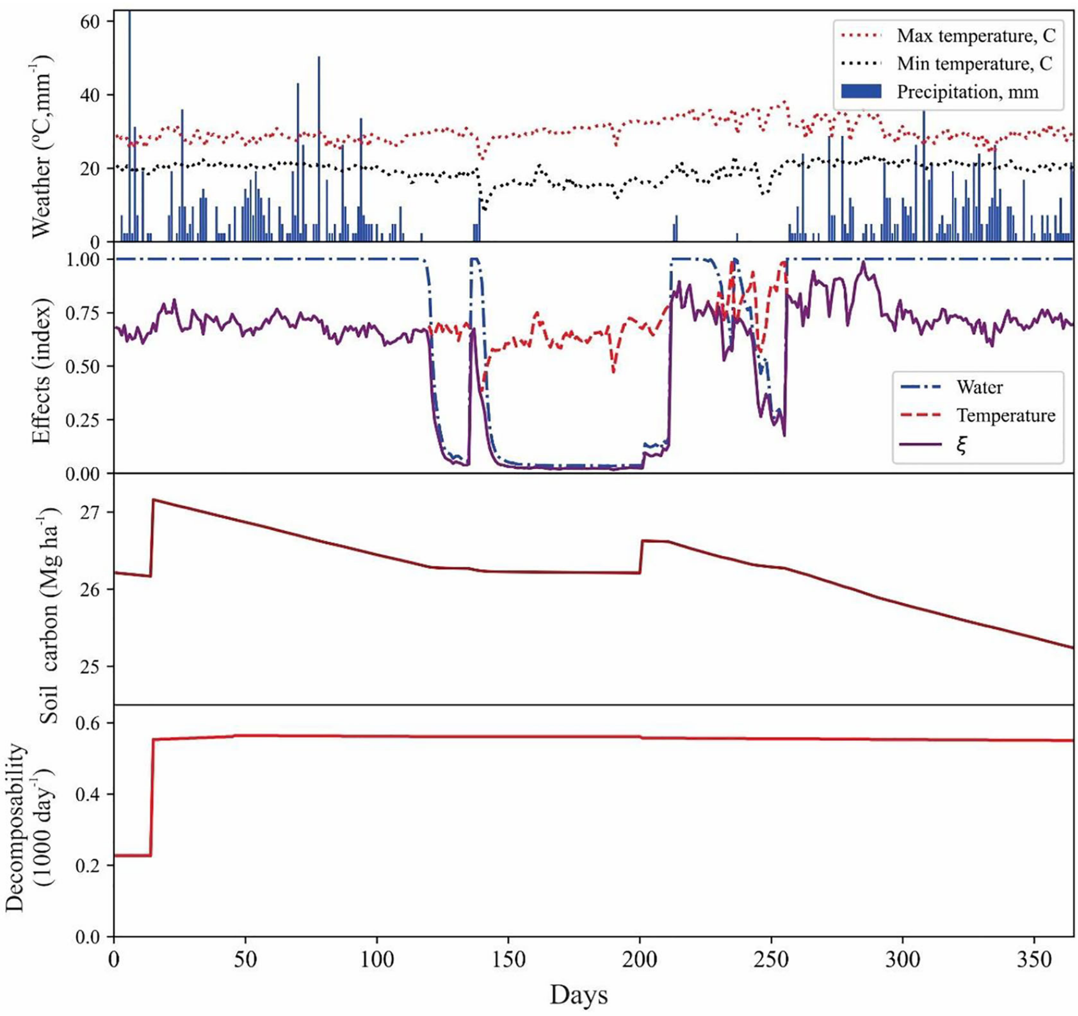

FIGURE 5 Detailed examination of a single year of simulated double-cropping. Beef manure was applied on day 15, then on day 45, the field was moldboard-plowed and corn was planted. The corn was harvested on day 200, followed by plowing and soybean planting on day 270. The top panel shows key weather variables. The second panel shows two of the key derived effects that compose $\xi(t)\xi$t: temperature and water deficit. The third panel shows soil carbon, and the fourth decomposability, multiplied by 1000 for readability.

plowing, and (4) grows summer soybeans only with moldboard plowing.

Differences between these scenarios are apparent from the first harvest, with biomass inputs being the primary driver of differences, and tillage being secondary. Differences between years are due to weather, which drives day-to-day changes in decomposition. The sawtooth pattern in the soil carbon plots (Figure 4) is an artifact related to how the current version of the soil model is coupled with the crop model. In the current coupling, all non-harvested carbon, including estimates of root exudates during plant growth, goes into the soil at harvest rather than giving root exudates and senesced material to the soil throughout the growing season. Given the typical 5-year timeframe of CFTS models, within-season imbalances can be considered negligible (Brummitt et al., 2024; Oldfield et al., 2022; Smith et al., 2020). Future implementations of the model will have an improved crop model coupling, giving smoother curves.

A close examination of a single year of simulation (Figure 5) illustrates how the weather is abstracted to form the $\xi(t)$ adjustment to decomposability. This second simulation includes applying manure, moldboard plowing, and planting and harvesting a crop, showing how different processes affect the quantity and decomposability of soil carbon.

The temperature effect reflects the degree to which a given day's temperature is below the optimum temperature. Likewise, a water deficit effect reflects decomposition being hindered by a lack of available water. Under flooded conditions, lack of oxygen would also limit decomposition, but this effect is omitted from the figure as the example site is not prone to flooding. As described in Section 3.3, $\xi(t)$ is the product of these three dynamic adjustments to decomposition.

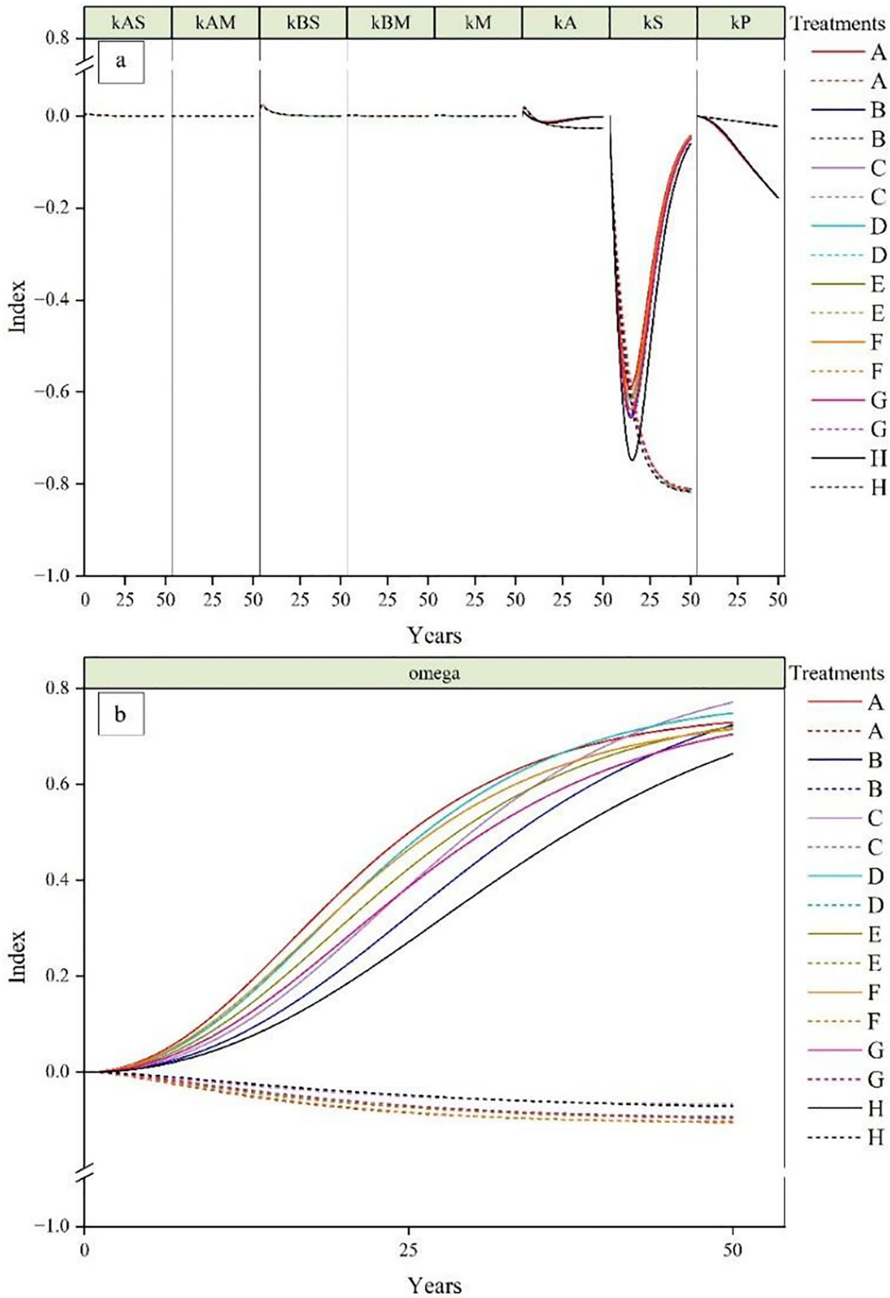

FIGURE 6 Local sensitivity analysis of parameters for Century (a) and ProCarbon-Soil (PROCS) (b). Solid lines show the "no input" condition, and dashed lines the "double input." Letters correspond to treatments tested, described in more detail in Supporting Information, Section 10. In (a), *k* represents the turnover rate for the pool named: *k*AS, aboveground structural; *k*AM, aboveground metabolic; *k*BS, belowground structural; *k*BM, belowground metabolic; *k*M, microbial; *k*A, active; *k*S, slow, and *k*P, passive.

5.2 | Synthetic data and in-silico experiments

PROCS was systematically evaluated through in silico experiments using synthetic datasets. While rarely used in soil carbon modeling, synthetic datasets are potent tools for mimicking the statistical properties of real-world data while offering complete control over data attributes. They provide a methodical and thorough means of testing model assumptions and performance, and are especially useful when field data are limited (Tekchandani et al., 2024). Following synthetic data generation, we first evaluated PROCS parameter sensitivity through standard procedures and compared it with a classic mSCM, the Century soil carbon submodel. In a second experiment, we evaluated PROCS's ability to deterministically forecast carbon stocks based on early trajectory information. Finally, the level of uncertainty associated with PROCS estimates, when calibrated with noisy synthetic data, was evaluated through *k*-fold cross-validation.

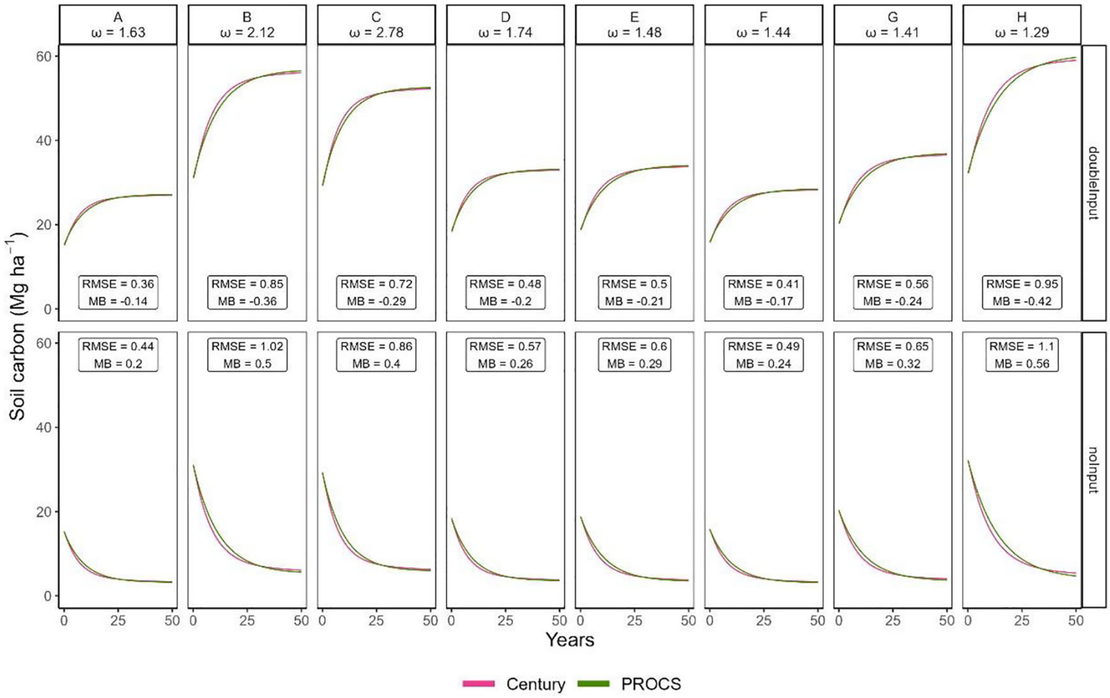

FIGURE 7 ProCarbon-Soil (PROCS) calibration curves compared with reference trajectories generated by the Century model, for eight treatments and two input scenarios ("no input" and "double input"). The respective values for estimated parameter ω, root mean square error (RMSE), and mean bias error (MB) are exhibited in the legend boxes.

5.2.1 | Synthetic data and scenario building

Synthetic data procedures generated (1) 50-year deterministic reference trajectories for SOC stocks, which were assumed to represent the true system state, and (2) synthetic time series datasets to represent SOC stock measurements.

The reference trajectories were generated in a factorial design using the Century carbon submodel. The first factor involved generating eight different parameter sets (labeled A–H), and the second factor pertained to carbon inflow quantities. The parameter sets were generated using a log-normal distribution of Century's standard decomposition parameters, based on the mean values proposed by Parton et al. (1994) for the aboveground structural (kAS), aboveground metabolic (kAM), belowground structural (kBS), belowground metabolic (kBM), microbial (kM), active (kA), slow (kS), and passive (kP) compartments. Stochastic parameter perturbation was used to account for model parameter uncertainty and site-specific system behavior. For each perturbed parameter set, two contrasting scenarios were designed. The first scenario ("no-input") imposes zero soil carbon inflow ($I = 0$), while the second scenario ("double-input") simulates soil carbon inflow that is twice that needed for SOC stock steady state. All trajectories started from identical initial SOC stocks and steady state conditions. The steady state was achieved for Century by adjusting the soil and litter compartments' carbon partitioning, and for PROCS by adjusting carbon decomposability. Century and PROCS outputs were made compatible by calculating the total carbon influx as the sum of Century litter to soil flows, and the soil carbon stock from Century as the sum of the carbon masses in the active, slow, and passive soil compartments. Litter dynamics were simulated in Century, but litter carbon was not included in SOC stocks. Details on the settings used to generate synthetic data can be found in Supporting Information, Section 10.

Synthetic time series representing SOC stock measurements were produced by adding Gaussian noise with a coefficient of variation of 4% to the true system state, that is, the reference trajectory evaluated at the sampling time t. Note that 100 time series were generated for each scenario.

5.2.2 | Parameter selection (identifiability analysis)

Parameter identification analysis followed the method outlined by Brun et al. (2001) and Omlin and Reichert (1999), who suggest selecting parameters for calibration based on local sensitivity and collinearity analysis. The influential and independent parameters are then calibrated, while the others remain fixed, reducing calibration complexity and improving computational efficiency.

Local sensitivity analysis evaluates the effect of small perturbations in parameter values on the model output at specific

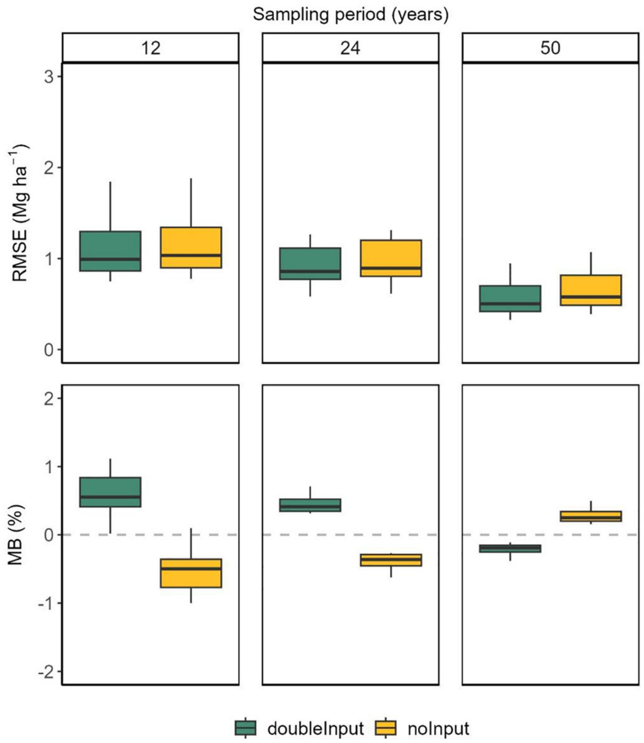

FIGURE 8 Distribution of the root mean square error (RMSE) (top panel) and mean bias error (MB) (bottom panel) between PROCS predictions for the "no input" and "double input" for 50 years over all reference trajectories, using different sampling time spans and a time interval of one year. Distributions represented by the box-plots are related to the eight different sets of perturbed trajectories. See text and Supporting Information, Section 10 for details.

points in time, calculating a sensitivity index from the partial derivative of the output with respect to each parameter, scaled appropriately. The higher the absolute sensitivity value, the higher its influence over a selected output variable.

Collinearity analysis was applied to identify how parameters may vary in concert to fit to data. A collinearity index of γ means that the output change caused by perturbing one parameter can be compensated by $(1 - 1/\gamma)$% by changing the other parameters appropriately. If the collinearity index exceeds a chosen value, typically between 15 and 20, then the parameter set is considered poorly identifiable (Brun et al., 2001), giving evidence of redundancy of parameters in the model structure, that is, overparameterization (Mason & Perreault, 1991). As the current PROCS formulation has only one parameter, it inherently circumvents these concerns about collinearity.

Figure 6 shows the results of sensitivity indices for Century and PROCS models over the 50 years considered in the study. For the Century model carbon turnover, k_S is by far the most sensitive parameter (Figures S1 and S2). Sensitivity indexes differed considerably between "no-input" and "double-input" scenarios and varied with data time span, suggesting that the period of data covered in the calibration can influence which parameter sets should be selected for calibration. Under the "no-input" scenario, Century had two or three identifiable parameters (depending on the duration of the sampling period), while under the "double-input" scenario, the number of identifiable parameters varied from two to four (see details in Supporting Information, Section 11). Additionally, the near-zero sensitivity of several Century parameters (Figure 6a) suggests that the model is overparameterized for total SOC trajectories. Meanwhile, the influence of PROCS's single parameter ω also increases over time (Figure 6b) but is much less sensitive for the "double-input" condition than it is for "no-input."

This result indicates that the model's complexity exceeds what the carbon trajectory data allow us to determine for the chosen variable set. PROCS's ω was much more sensitive in the "no-input" scenario as carbon stabilization effects are isolated. In the "double-input" scenario, sensitivity was very low for short time series because of the strong influence of ρ_I in the dynamics. Given this high dependance on ρ_I in the short term, accurate ρ_I estimates for different plant residues and organic fertilizers will be essential for predicting initial trajectories that have high carbon inflows. Moreover, although ρ_I has been designated as a model input, its high short-term sensitivity indicates the reasonableness of reestimating its value through MDF for individual or grouped fields. Future research should address ρ_I estimation and its use in MDF.

A global sensitivity analysis was conducted to determine how variations in the model's parameter ω influence predicted soil carbon dynamics. Unlike local sensitivity analysis, which evaluates small perturbations around a specific point, the global approach considers the entire possible range of variation of the parameters to evaluate nonlinear interactions and joint effects. We employed the Fourier amplitude sensitivity test (FAST), as described by Cukier et al. (1978) and Saltelli et al. (1999). The FAST method quantifies first-order (direct) and total-order (including interactive) contributions of the parameter to output variance. Since the PROCS model includes only one parameter, the first-order and total-order sensitivity indices coincide. A sampling using 100 values for parameter ω, uniformly distributed in the interval 0.001–5, was used to perform the FAST method.

The results (see details in Supporting Information, Section 12) demonstrate that ω consistently dominates the model's output variability, with sensitivity indices approaching unity under all examined conditions (Figure S3). The variability of the sensitivity index increases with longer sampling periods, suggesting that the temporal scale of observation and the applied treatment regime ("no-input" vs. "double-input") can influence the degree to which the parameter's dominance is expressed. Despite these minor fluctuations, the parameter remains the primary source of uncertainty in the model, confirming its critical role in driving system behavior over multiple temporal and management scenarios.

Soil Science Society of America Journal

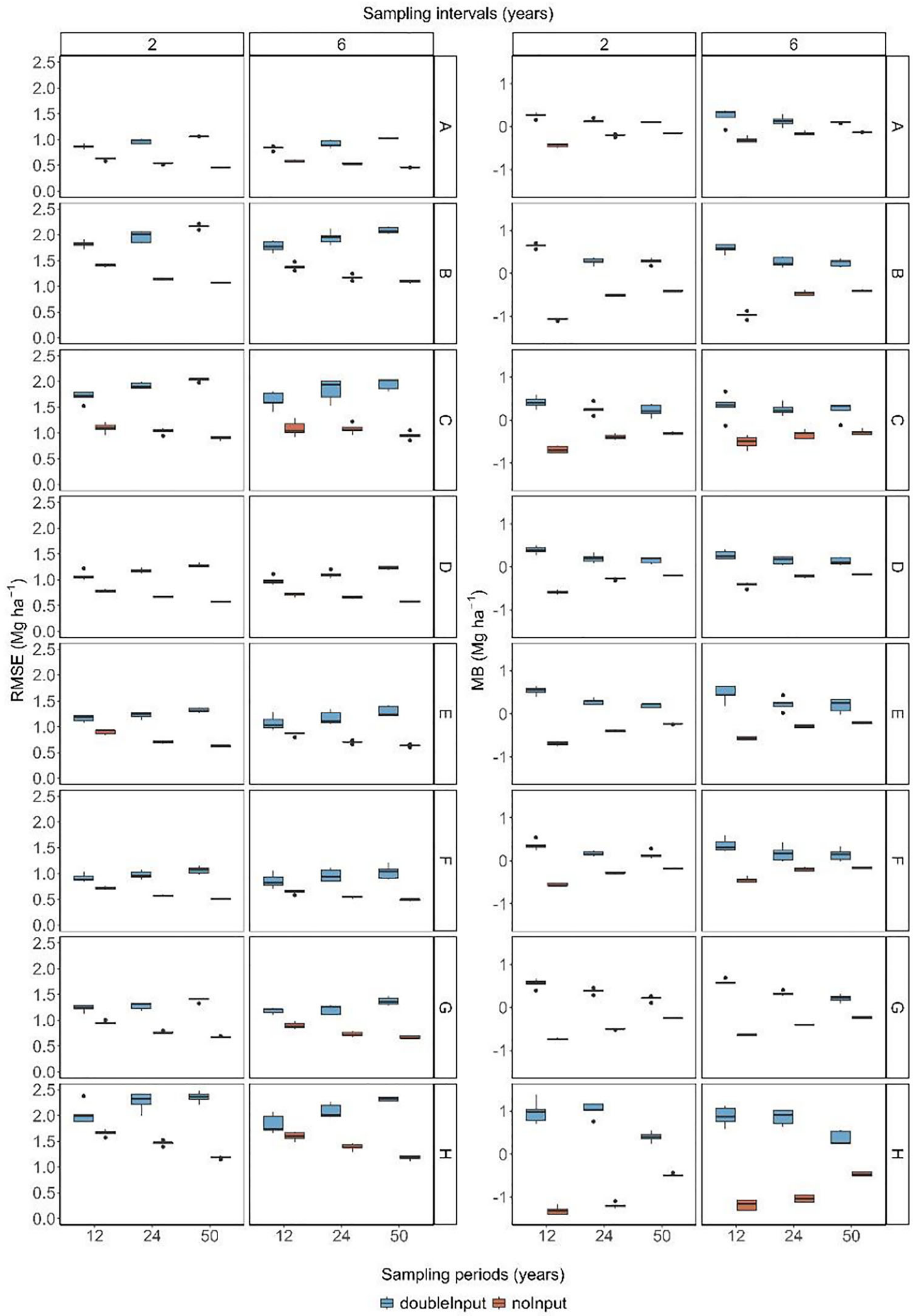

FIGURE 9 Root mean square error (RMSE) and mean bias error (MB) for *k*-fold ($k = 5$) cross-validation of the ProCarbon-Soil (PROCS) model using noisy synthetic data for different sampling periods and intervals, for the "no input" and "double input" scenarios.

5.2.3 | Model trajectory comparison and flexibility testing

While reducing the number of parameters is desirable to avoid equifinality and overfitting, it may have the side effect of decreasing a model's ability to fit a wide range of data patterns. To test against this possible drawback, we examined PROCS's ability to reproduce the Century-simulated reference trajectories, considering sampling intervals fixed in 1 year. The PROCS ω parameter was calibrated to minimize the residual sum of squares simultaneously for the "no-input" and "double-input" treatments. Century reference trajectories and PROCS calibrated trajectories are presented in Figure 7.

Output dissimilarity between PROCS and Century was evaluated through the root mean square error (RMSE) and maximum absolute C stock difference, for each set of

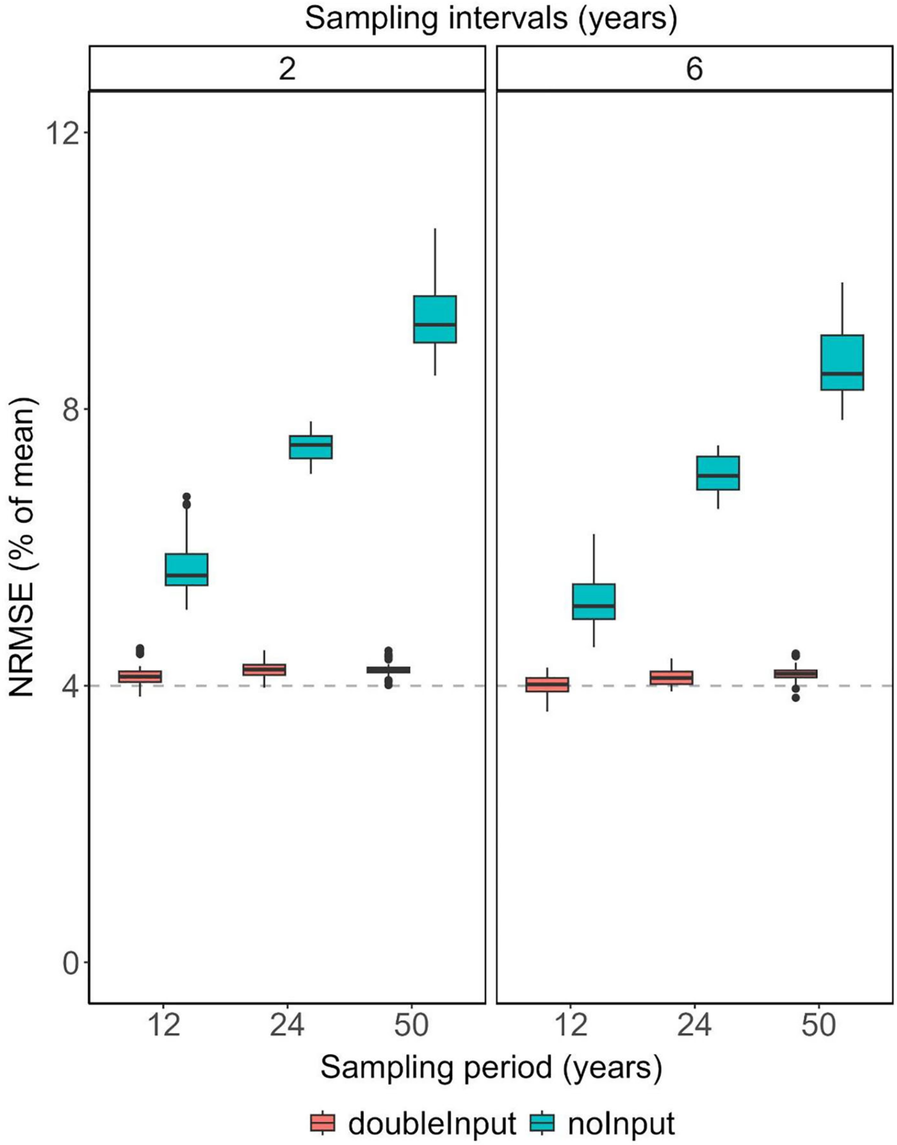

FIGURE 10 Normalized root mean square error (NRMSE) expressed as a percentage of the mean for two different input scenarios: "double input" (red) and "no input" (blue). Results are shown across sampling periods of 12, 24, and 50 years, under two sampling interval scenarios of 2 years (left) and 6 years (right). The dashed line represents the input data noise level (4% coefficient of variation, CV).

perturbed parameters and carbon input levels. Different values for different reference trajectories are expected as they reflect the different parameter sets used to generate the reference trajectories. Mean bias (MB) tended to be positive for the "no-input" scenario (+0.2 to + 0.56 Mg C ha^{-1}) and negative for the "double-input" scenario (−0.42 to −0.14 Mg C ha^{-1}). For comparison, data analyzed in the context of the Bayer's PROCarbono project in Brazilian farms with twin cropland plots of around 30 ha each and eight trenches produced standard error of the mean (SEM) for carbon stocks around 3.8 Mg ha^{-1}, and an SEM of stock change around 5.4 Mg ha^{-1}. In other words, PROCS can follow disturbed parameter Century trajectories with lower uncertainty than is usually found in SOC stock measurements on farms. Further discussion on model errors reported in field studies can be found in Section 5.2.4.

The single parameter formulation of PROCS facilitates constraining values within a well-defined expected range, which allows for good projections even from short time series (Figure 8). As expected, however, as the sampling period increased and more information was made available to the model, RMSEs decreased from 1.03 to 0.57 Mg ha^{-1} for "no-input" and from 0.99 to 0.50 Mg ha^{-1} for "double-input" treatments.

While the Century trajectories cannot be assumed to be the the system's true state, the reported bias and error values suggest a potential for refining the carbon stabilization function. However, the tests reported here show no major structural limitations of PROCS in representing Century-generated total SOC stocks over carbon farming projects' expected duration (up to 50 years), despite PROCS relying on only two state variables and a single empirical SOC stabilization parameter.

The slightly higher errors in the "no-input" treatments are expected, as in that scenario only the stabilization process (dF_A) operates, and it was approximated by an arbitrary function (Equation 11) which does not mimic exactly an mSCM's behavior. Meanwhile, in the "double-input" scenario, the analytically exact (dF_I) dilutes the dF_A error.

5.2.4 | Calibration and cross-validation

The accuracy of a predictive model largely relies on the precise estimation of its parameters, which can be uncertain due to measurement errors, spatial variability, and limited representation of available observational data, including the duration of experiments and their geographic coverage. The PROCS model's SOC estimate uncertainty was evaluated through a k-fold cross-validation procedure on the synthetic data generated as described in Section 5.2.1 and in Supporting Information, Section 10. PROCS's environmental decomposition modifier, $\xi(t)$, and stabilization modifier, $\mu(t)$ were set to values compatible with those used in Century to generate the curves. The ω values were calibrated independently for each of the eight treatments (A–H), sampling periods, and sampling intervals in a fivefold cross-validation with 100 measurement time series generated for each scenario (Section 5.2.1). The distribution of parameter values, RMSE and MB, were then analyzed jointly (Figure 9).

Our synthetic data experiments found uncertainty in total carbon stocks usually between 1 and 2 Mg ha^{-1}. Those values are lower than what is typically reported for field studies reported elsewhere. In a study evaluating five long-term experimental sites in Austria, Bernardini et al. (2024) found RMSE of 4.71 $\pm$ 0.09 Mg ha^{-1} for the C-TOOL model and 7.00 $\pm$ 0.15 Mg ha^{-1} for RothC in their cross-validation process. In a study conducted in Brazil using the Century model, Bortolon et al. (2011) reported an RMSE of 2.60 Mg ha^{-1}, which was 7.6% of the measured values. In a national study across the United States, Gautam et al. (2020) used the DayCent model for simulating 654 observed points, finding an RMSE of 19 Mg ha^{-1}. The lower uncertainty reported in this study may be due to the comparisons being made with data generated for each individual reference trajectory. While this approach is well-suited to study model structural adequacy, the disparity raises a compelling hypothesis: much of the uncertainty in field studies may arise from high variability in behavior observed across different fields. This challenge could be effectively addressed through MDF techniques.

Figure 10 shows that the input uncertainty for the double-input treatment, expressed as the normalized root mean square error (NRMSE), was only marginally greater than the Gaussian noise coefficient of variation of 4%. It is unlikely that other model structures would be able to reduce such errors, as they already approach the variability of the reference trajectory. In the absence of carbon input ("no-input" scenario), the errors were larger and increased with increasing duration, suggesting a pronounced nonlinear behavior of $\frac{df_A}{dt}$. This highlights a potential area for model refinement in such extreme scenarios, which are, however, unlikely to be required in typical CFTS projects. Despite being proportionally higher due to low carbon stocks in the "no-input" treatment, the absolute values of RMSE were comparable between the "no-input" and "double-input" scenarios. Results from Section 5.2.3 and this section suggest that adding more complexity to the carbon turnover calculations is unlikely to significantly improve SOC stock projections in relation to the PROCS model in typical CFTS conditions.

6 | SUMMARY AND CONCLUSIONS

This study proposes decomposability as a property of the soil carbon system latent in mSCM, enabling the modeling of SOC dynamics via two linked differential equations tracking carbon mass and decomposability changes. This approach mitigates equifinality, grounding soil model states firmly in measurable data while maintaining compatibility with mSCM states and their parameters. Decomposability can be derived from total carbon longitudinal data, improving model initialization and interaction.

Analytical tests demonstrate that PROCS effectively reproduces carbon trajectories of a classic mSCM (Century) using a more generic structure with only one empirical carbon stabilization rate tuning parameter. Cross-validation results under increased carbon inflow reveal a minimal increase in uncertainty, suggesting that the complexity of multiple compartments may not significantly improve SOC stock projections compared to the parsimonious PROCS approach based on total SOC data.

PROCS exhibits promising potential for CFTS applications, and it may expedite Tier 3 development to support national UNFCCC inventories. However, it still requires comprehensive calibration and rigorous field data evaluation before it can be confidently implemented. Ongoing research within the Embrapa and Bayer Crop Science cooperation is focusing on crop model coupling, field data calibration, inclusion of vertical heterogeneity in carbon concentration and environmental conditions, and the development of MDF algorithms informed by on-farm data, to further assess and advance PROCS's ability to predict real-world SOC trajectories.

AUTHOR CONTRIBUTIONS

Luís G. Barioni: Conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; validation; writing—original draft; writing—review and editing. **Beatriz A. Valladão**: Conceptualization; data curation; formal analysis; investigation; methodology; software; visualization; writing—review and editing. **Vitor H. M. Mourão**: Formal analysis; software; validation; writing—review and editing. **Robert P. Ewing**: Investigation; validation; writing—original draft; writing—review and editing. **Yusuf N. Karatay**: Conceptualization; formal analysis; investigation; writing—original draft; writing—review and editing. **Júnior M. Damian**: Conceptualization; formal analysis; investigation; methodology; writing—original draft; writing—review and editing. **Vinícius C. Melício**: Conceptualization; formal analysis; software; visualization. **Rodrigo P. A. Rejaili**: Formal analysis; resources; visualization. **Rafael O. Silva**: Writing—original draft; writing—review and editing.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Luís G. Barioni https://orcid.org/0000-0003-1716-1428
Beatriz A. Valladão https://orcid.org/0000-0001-5125-8392
Vitor H. M. Mourão https://orcid.org/0000-0001-7460-8527
Robert P. Ewing https://orcid.org/0000-0002-5466-9303
Yusuf N. Karatay https://orcid.org/0000-0001-8163-4950
Júnior M. Damian https://orcid.org/0000-0002-1692-6751
Vinícius C. Melício https://orcid.org/0009-0005-4100-2189
Rodrigo P. A. Rejaili https://orcid.org/0009-0005-4586-1909
Rafael O. Silva https://orcid.org/0000-0003-3440-580X

REFERENCES

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

Alexander, P., Paustian, K., Smith, P., & Moran, D. (2015). The economics of soil C sequestration and agricultural emissions abatement. *Soilless*, *1*(1), 331–339. https://doi.org/10.5194/SOIL-1-331-2015

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

Badgery, W., Murphy, B., Cowie, A., Orgill, S., Rawson, A., Simmons, A., & Crean, J. (2020). Soil carbon market-based instrument pilot—The sequestration of soil organic carbon for the purpose of obtaining carbon credits. *Soil Research*, *59*(1), 12–23. https://doi.org/10.1071/SR19331

Batjes, N. H., Ceschia, E., Heuvelink, G. B. M., Demenois, J., le Maire, G., Cardinael, R., Arias-Navarro, C., & van Egmond, F. (2024). Towards a modular, multi-ecosystem monitoring, reporting and verification (MRV) framework for soil organic carbon stock change assessment. *Carbon Management*, *15*(1), 2410812. https://doi.org/10.1080/17583004.2024.2410812

Baveye, P. C. (2023). Ecosystem-scale modelling of soil carbon dynamics: Time for a radical shift of perspective?. *Soil Biology and Biochemistry*, *184*, 109112. https://doi.org/10.1016/J.SOILBIO.2023.109112

Bernardini, L. G., Rosinger, C., Bodner, G., Keiblinger, K. M., Izquierdo-Verdiguier, E., Spiegel, H., Retzlaff, C. O., & Holzinger, A. (2024). Learning vs. understanding: When does artificial intelligence outperform process-based modeling in soil organic carbon prediction?. *New Biotechnology*, *81*, 20–31. https://doi.org/10.1016/J.NBT.2024.03.001

Bortolon, E. S. O., Mielniczuk, J., Tornquist, C. G., Lopes, F., & Bergamaschi, H. (2011). Validation of the century model to estimate the impact of agriculture on soil organic carbon in Southern Brazil. *Geoderma*, *167–168*, 156–166. https://doi.org/10.1016/J.GEODERMA.2011.08.008

Bosatta, E., & Ågren, G. I. (2003). Exact solutions to the continuous-quality equation for soil organic matter turnover. *Journal of Theoretical Biology*, *224*(1), 97–105. https://doi.org/10.1016/S0022-5193(03)00147-4

Brummitt, C. D., Mathers, C. A., Keating, R. A., O'Leary, K., Easter, M., Friedl, M. A., DuBuisson, M., Campbell, E. E., Pape, R., Peters, S. J. W., & Kumar, A. A. (2024). Solutions and insights for agricultural monitoring, reporting, and verification (MRV) from three consecutive issuances of soil carbon credits. *Journal of Environmental Management*, *369*, 122284. https://doi.org/10.1016/j.jenvman.2024.122284

Brun, R., Reichert, P., & Künsch, H. R. (2001). Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, *37*(4), 1015–1030. https://doi.org/10.1029/2000WR900350

Bruun, S., Ågren, G. I., Christensen, B. T., & Jensen, L. S. (2010). Measuring and modeling continuous quality distributions of soil organic matter. *Biogeosciences*, *7*(1), 27–41. https://doi.org/10.5194/BG-7-27-2010

Buizza, C., Quilodrán Casas, C., Nadler, P., Mack, J., Marrone, S., Titus, Z., Le Cornec, C., Heylen, E., Dur, T., Baca Ruiz, L., Heaney, C., Díaz Lopez, J. A., Kumar, K. S. S., & Arcucci, R. (2022). Data learning: Integrating data assimilation and machine learning. *Journal of Computational Science*, *58*, 101525. https://doi.org/10.1016/J.JOCS.2021.101525

Campbell, E. E., & Paustian, K. (2015). Current developments in soil organic matter modeling and the expansion of model applications: A review. *Environmental Research Letters*, *10*(12), 123004. https://doi.org/10.1088/1748-9326/10/12/123004

Chenu, C., Angers, D. A., Barré, P., Derrien, D., Arrouays, D., & Balesdent, J. (2019). Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. *Soil and Tillage Research*, *188*, 41–52. https://doi.org/10.1016/J.STILL.2018.04.011

Coleman, K., & Jenkinson, D. S. (1996). RothC-26.3—A Model for the turnover of carbon in soil. *Evaluation of Soil Organic Matter Models*, *38*, 237–246. https://doi.org/10.1007/978-3-642-61094-3_17

Colmanetti, M., Oldoni, H., Silva, L., Wolff, W., Rejaili, R., Barioni, L., de Almeida, I., Ewing, R., & Cuadra, S. (2026). Calibration and evaluation of a carbon net primary productivity (CNPP) module based on the simple model. *Agronomy Journal*. https://doi.org/10.22541/au.176402479.99853267/v1

Costa, C., Dittmer, K. M., Shelton, S. W., Bossio, D. A., Zinyengere, N., Luu, P., Heinz, S., Egenolf, K., Rowland, B., Zuluaga, A., Klemme, J., Mealey, T., Smith, M., & Wollenberg, E. K. (2020). *How soil carbon accounting can improve to support investment-oriented actions promoting soil carbon storage*. CGIAR Research Program on Climate Change, Agriculture and Food Security. https://hdl.handle.net/10568/110284

Cukier, R. I., Levine, H. B., & Shuler, K. E. (1978). Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, *26*(1), 1–42. https://doi.org/10.1016/0021-9991(78)90097-9

Dangal, S. R. S., Schwalm, C., Cavigelli, M. A., Gollany, H. T., Jin, V. L., & Sanderman, J. (2022). Improving soil carbon estimates by linking conceptual pools against measurable carbon fractions in the DAYCENT model version 4.5. *Journal of Advances in Modeling Earth Systems*, *14*(5), e2021MS002622. https://doi.org/10.1029/2021MS002622

Davoudabadi, M. J., Pagendam, D., Drovandi, C., Baldock, J., & White, G. (2021). Advanced Bayesian approaches for state-space models with a case study on soil carbon sequestration. *Environmental Modelling & Software*, *136*, 104919. https://doi.org/10.1016/J.ENVSOFT.2020.104919

Davoudabadi, M. J., Pagendam, D., Drovandi, C., Baldock, J., & White, G. (2024). Innovative approaches in soil carbon sequestration modelling for better prediction with limited data. *Scientific Reports*, *14*(1), 1–13. https://doi.org/10.1038/s41598-024-53516-z

Del Grosso, S., Parton, W., Stohlgren, T., Zheng, D., Bachelet, D., Prince, S., Hibbard, K., & Olson, R. (2008). Global potential net primary production predicted from vegetation class, precipitation, and temperature. *Ecology*, *89*(8), 2117–2126. https://doi.org/10.1890/07-0850.1

Derrien, D., Barré, P., Basile-Doelsch, I., Cécillon, L., Chabbi, A., Crème, A., Fontaine, S., Henneron, L., Janot, N., Lashermes, G., Quénéa, K., Rees, F., & Dignac, M. F. (2023). Current controversies on mechanisms controlling soil carbon storage: Implications for interactions with practitioners and policy-makers. A review. *Agronomy for Sustainable Development*, *43*(1), 1–28. https://doi.org/10.1007/S13593-023-00876-X

Elliott, E. T., Paustian, K., & Frey, S. D. (1996). Modeling the measurable or measuring the modelable: A hierarchical approach to isolating meaningful soil organic matter fractionations. *Evaluation of Soil Organic Matter Models*, *38*, 161–179. https://doi.org/10.1007/978-3-642-61094-3_12

Foereid, B., Bellamy, P. H., Holden, A., & Kirk, G. J. D. (2012). On the initialization of soil carbon models and its effects on model predictions for England and Wales. *European Journal of Soil Science*, *63*(1), 32–41. https://doi.org/10.1111/J.1365-2389.2011.01407.X

Gautam, S., Mishra, U., Scown, C. D., & Zhang, Y. (2020). Sorghum biomass production in the continental United States and its potential impacts on soil organic carbon and nitrous oxide emissions. *Global Change Biology Bioenergy*, *12*(10), 878–890. https://doi.org/10.1111/GCBB.12736

Gold Standard. (2020). *Soil organic carbon framework methodology*. https://globalgoals.goldstandard.org/402-luf-agr-fm-soil-organic-carbon-framework-methodolgy/

Guillaume, T., Makowski, D., Libohova, Z., Bragazza, L., Sallaku, F., & Sinaj, S. (2022). Soil organic carbon saturation in cropland–grassland systems: Storage potential and soil quality. *Geoderma*, *406*, 115529. https://doi.org/10.1016/J.GEODERMA.2021.115529

Hashimoto, S., Wattenbach, M., & Smith, P. (2011). A new scheme for initializing process-based ecosystem models by scaling soil carbon pools. *Ecological Modelling*, *222*(19), 3598–3602. https://doi.org/10.1016/j.ecolmodel.2011.08.011

Hénin, S., & Dupuis, M. (1945). Essai de bilan de la matière organique du sol. *Annales Agronomiques*, *11*, 17–29.

Huang, Y. (2024). Developing matrix models for land carbon models. In Y. Luo & B. Smith (Eds.), *Land carbon cycle modeling: Matrix approach, data assimilation, ecological forecasting, and machine learning* (2nd ed., 29–33). CRC Press. https://doi.org/10.1201/9781032711126-7

Intergovernmental Panel on Climate Change. (2006). *2006 IPCC guidelines for national greenhouse gas inventories: Volume 4. Agriculture, forestry and other land use*. Institute for Global Environmental Strategies. https://www.ipcc-nggip.iges.or.jp/public/2006gl/vol4.html

Jenkinson, D. S., & Rayner, J. H. (1977). The turnover of soil organic matter in some of the Rothamsted classical experiments. *Soil Science*, *123*(5), 298–305. https://doi.org/10.1097/00010694-197705000-00005

Kanari, E., Cécillon, L., Baudin, F., Clivot, H., Ferchaud, F., Houot, S., Levavasseur, F., Mary, B., Soucémarianadin, L., Chenu, C., & Barré, P. (2022). A robust initialization method for accurate soil organic carbon simulations. *Biogeosciences*, *19*(2), 375–387. https://doi.org/10.5194/BG-19-375-2022

Kirschbaum, M. U. F., Giltrap, D. L., McNally, S. R., Liáng, L. L., Hedley, C. B., Moinet, G. Y. K., Blaschek, M., Beare, M. H., Theng, B. K. G., Hunt, J. E., & Whitehead, D. (2019). Estimating the mineral surface area of soils by measured water adsorption. Adjusting for the confounding effect of water adsorption by soil organic carbon. *European Journal of Soil Science*, *71*(3), 382–391. https://doi.org/10.1111/ejss.12892

Kleber, M., & Lehmann, J. (2019). Humic substances extracted by alkali are invalid proxies for the dynamics and functions of organic matter in terrestrial and aquatic ecosystems. *Journal of Environmental Quality*, *48*(2), 207–216. https://doi.org/10.2134/JEQ2019.01.0036

Kyker-Snowman, E., Wieder, W. R., Frey, S. D., & Grandy, A. S. (2020). Stoichiometrically coupled carbon and nitrogen cycling in the MIcrobial-MIneral Carbon Stabilization model version 1.0 (MIMICS-CN v1.0). *Geoscientific Model Development*, *13*(9), 4413–4434. https://doi.org/10.5194/GMD-13-4413-2020

Lardy, R., Bellocchi, G., & Soussana, J.-F. (2011). A new method to determine soil organic carbon equilibrium. *Environmental Modelling & Software*, 016. https://doi.org/10.1016/j.envsoft.2011.05.016

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., … Guenet, B. (2023). Soil organic carbon models need independent time-series validation for reliable prediction. *Communications Earth & Environment*, *4*(1), 1–8. https://doi.org/10.1038/s43247-023-00830-5

Lehmann, J., & Kleber, M. (2015). The contentious nature of soil organic matter. *Nature*, *528*(7580), 60–68. https://doi.org/10.1038/nature16069

Lei, W., Pan, Q., Teng, P., Yu, J., & Li, N. (2023). How does soil organic matter stabilize with soil and environmental variables along a black soil belt in Northeast China? An explanation using FTIR spectroscopy data. *Catena*, *228*, 107152. https://doi.org/10.1016/J.CATENA.2023.107152

Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X., & Zhang, L. (2009). Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. *Ecological Applications*, *19*(3), 571–574. https://doi.org/10.1890/08-0561.1

Manzoni, S., Piñeiro, G., Jackson, R. B., Jobbágy, E. G., Kim, J. H., & Porporato, A. (2012). Analytical models of soil and litter decomposition: Solutions for mass loss and time-dependent decay rates. *Soil Biology and Biochemistry*, *50*, 66–76. https://doi.org/10.1016/J.SOILBIO.2012.02.029

Manzoni, S., & Porporato, A. (2009). Soil carbon and nitrogen mineralization: Theory and models across scales. *Soil Biology and Biochemistry*, *41*(7), 1355–1379. https://doi.org/10.1016/J.SOILBIO.2009.02.031

Mao, H. R., Cotrufo, M. F., Hart, S. C., Sullivan, B. W., Zhu, X., Zhang, J., Liang, C., & Zhu, M. (2024). Dual role of silt and clay in the formation and accrual of stabilized soil organic carbon. *Soil Biology and Biochemistry*, *192*, 109390. https://doi.org/10.1016/J.SOILBIO.2024.109390

Mason, C. H., & Perreault, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research*, *28*(3), 268. https://doi.org/10.2307/3172863

Mathers, C., Black, C. K., Segal, B. D., Gurung, R. B., Zhang, Y., Easter, M. J., Williams, S., Motew, M., Campbell, E. E., Brummitt, C. D., Paustian, K., & Kumar, A. A. (2023). Validating DayCent-CR for cropland soil carbon offset reporting at a national scale. *Geoderma*, *438*, 116647. https://doi.org/10.1016/j.geoderma.2023.116647

Oldfield, E. E., Eagle, A. J., Rubin, R. L., Rudek, J., Sanderman, J., & Gordon, D. R. (2022). Crediting agricultural soil carbon sequestration. *Science*, *375*(6586), 1222–1225. https://doi.org/10.1126/science.abl7991

Omlin, M., & Reichert, P. (1999). A comparison of techniques for the estimation of model prediction uncertainty. *Ecological Modelling*, *115*(1), 45–59. https://doi.org/10.1016/S0304-3800(98)00174-4

Parton, W. J. (1996). The CENTURY model. *Evaluation of Soil Organic Matter Models*, *38*, 283–291. https://doi.org/10.1007/978-3-642-61094-3_23

Parton, W. J., Ojima, D. S., Cole, C. V., & Schimel, D. S. (1994). A general model for soil organic matter dynamics: Sensitivity to litter chemistry, texture and management. In R. B. Bryant & R. W. Arnold (Eds.), *Quantitative modeling of soil forming processes* (pp. 147–167). ASA, CSSA, SSSA. https://doi.org/10.2136/SSSASPECPUB39.C9

Parton, W. J., & Rasmussen, P. E. (1994). Long-term effects of crop management in wheat-fallow: II. CENTURY model simulations. *Soil Science Society of America Journal*, *58*(2), 530–536. https://doi.org/10.2136/SSSAJ1994.03615995005800020040X

Parton, W. J., Stewart, J. W. B., & Cole, C. V. (1988). Dynamics of C, N, P and S in grassland soils: A model. *Biogeochemistry*, *5*(1), 109–131. https://doi.org/10.1007/BF02180320

Robertson, A. D., Paustian, K., Ogle, S., Wallenstein, M. D., Lugato, E., & Cotrufo, M. F. (2019). Unifying soil organic matter formation and persistence frameworks: The MEMS model. *Biogeosciences*, *16*(6), 1225–1248. https://doi.org/10.5194/BG-16-1225-2019

Rovira, P., & Rovira, R. (2010). Fitting litter decomposition datasets to mathematical curves: Towards a generalised exponential approach. *Geoderma*, *155*(3–4), 329–343. https://doi.org/10.1016/J.GEODERMA.2009.11.033

Saltelli, A., Tarantola, S., & Chan, K. P. S. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, *41*(1), 39–56. https://doi.org/10.1080/00401706.1999.10485594

Sharma, M., Kaushal, R., Kaushik, P., & Ramakrishna, S. (2021). Carbon farming: Prospects and challenges. *Sustainability*, *13*(19), 11122. https://doi.org/10.3390/SU131911122

Shi, Z., Crowell, S., Luo, Y., & Moore, B. (2018). Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nature Communications*, *9*(1), 1–11. https://doi.org/10.1038/s41467-018-04526-9

Sierra, C. A. (2024). Compartmental dynamical systems and carbon cycle models. In Y. Luo & B. Smith (Eds.), *Land carbon cycle modeling: Matrix approach, data assimilation, ecological forecasting, and machine learning* (2nd ed., pp. 45–50). CRC Press. https://doi.org/10.1201/9781032711126-9

Sierra, C. A., & Müller, M. (2015). A general mathematical framework for representing soil organic matter dynamics. *Ecological Monographs*, *85*(4), 505–524. https://doi.org/10.1890/15-0361.1

Simone, T. E., Lambert, D. M., Cuvaca, I., & Eash, N. S. (2017). Soil carbon sequestration, carbon markets, and conservation agriculture practices: A hypothetical examination in Mozambique. *International Soil and Water Conservation Research*, *5*(3), 167–179. https://doi.org/10.1016/J.ISWCR.2017.06.001

Six, J., Elliott, E. T., & Paustian, K. (1999). Aggregate and soil organic matter dynamics under conventional and no-tillage systems. *Soil Science Society of America Journal*, *63*(5), 1350–1358. https://doi.org/10.2136/SSSAJ1999.6351350X

Smith, P., Adams, J., Beerling, D. J., Beringer, T., Calvin, K. V., Fuss, S., Griscom, B., Hagemann, N., Kammann, C., Kraxner, F., Minx, J. C., Popp, A., Renforth, P., Vicente Vicente, J. L., & Keesstra, S. (2019). Land-management options for greenhouse gas removal and their impacts on ecosystem services and the sustainable development goals. *Annual Review of Environment and Resources*, *44*, 255–286. https://doi.org/10.1146/annurev-environ-101718-033129

Smith, P., Soussana, J.-F., Angers, D., Schipper, L., Chenu, C., Rasse, D. P., Batjes, N. H., Van Egmond, F., Mcneill, S., Kuhnert, M., Arias-Navarro, C., Olesen, J. E., Chirinda, N., Fornara, D., Wollenberg, E., Álvaro-Fuentes, J., Sanz-Cobena, A., & Klumpp, K. (2020). How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Global Change Biology*, *26*(1), 219–241. https://doi.org/10.1111/GCB.14815

Spotorno, S., Gobin, A., Vanongeval, F., Del Borghi, A., & Gallo, M. (2024). Carbon farming practices assessment: Modelling spatial changes of soil organic carbon in Flanders, Belgium. *Science of The Total Environment*, *922*, 171267. https://doi.org/10.1016/J.SCITOTENV.2024.171267

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., … Luo, Y. (2023). Microbial carbon use efficiency promotes global soil carbon storage. *Nature*, *618*(7967), 981–985. https://doi.org/10.1038/s41586-023-06042-3

Tekchandani, P., Das, A. K., & Kumar, N. (2024). Large-scale secure model learning and inference using synthetic data for IoT-based big data analytics. *Computers and Electrical Engineering*, *119*, 109565. https://doi.org/10.1016/J.COMPELECENG.2024.109565

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., & Allison, S. D. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, *10*(3), 1717–1736. https://doi.org/10.5194/BG-10-1717-2013

Trudinger, C. M., Raupach, M. R., Rayner, P. J., & Enting, I. G. (2008). Using the Kalman filter for parameter estimation in biogeochemical models. *Environmetrics*, *19*(8), 849–870. https://doi.org/10.1002/ENV.910

Valladão, B. A. (2022). *Avaliação de modelos de dinâmica de carbono do solo baseada na identificação e estimação de parâmetros*. Unicamp—Universidade Estadual de Campinas.

van der Voort, T. S., Verweij, S., Fujita, Y., & Ros, G. H. (2023). Enabling soil carbon farming: Presentation of a robust, affordable, and scalable method for soil carbon stock assessment. *Agronomy for Sustainable Development*, *43*(1), 1–12. https://doi.org/10.1007/s13593-022-00856-7

Vannier, C., Cochrane, T. A., Zawar Reza, P., & Bellamy, L. (2022). An analysis of agricultural systems modelling approaches and examples to support future policy development under disruptive changes in New Zealand. *Applied Sciences*, *12*(5), 2746. https://doi.org/10.3390/APP12052746

Verra. (2023). *VM0042 improved agricultural land management, v2.1*. https://verra.org/methodologies/vm0042-improved-agricultural-land-management-v2-1/

Verra. (2024). *VMD0053 model calibration, validation, and uncertainty guidance for the methodology for improved agricultural land management, v2.0*. Verra. https://verra.org/methodologies/vmd0053-model-calibration-validation-and-uncertainty-guidance-for-the-methodology-for-improved-agricultural-land-management-v2-0/

Villas-Boas, P. R., Franco, M. A., Martin-Neto, L., Gollany, H. T., & Milori, D. M. B. P. (2020). Applications of laser-induced breakdown spectroscopy for soil analysis, part I: Review of fundamentals and chemical and physical properties. *European Journal of Soil Science*, *71*(5), 789–804. https://doi.org/10.1111/EJSS.12888

Villaverde, A. F., Pathirana, D., Fröhlich, F., Hasenauer, J., & Banga, J. R. (2022). A protocol for dynamic model calibration. *Briefings in Bioinformatics*, *23*(1), 1–19. https://doi.org/10.1093/BIB/BBAB387

Viskari, T., Laine, M., Kulmala, L., Mäkelä, J., Fer, I., & Liski, J. (2020). Improving Yasso15 soil carbon model estimates with ensemble adjustment Kalman filter state data assimilation. *Geoscientific Model Development*, *13*(12), 5959–5971. https://doi.org/10.5194/GMD-13-5959-2020

Walters, J. P., Archer, D. W., Sassenrath, G. F., Hendrickson, J. R., Hanson, J. D., Halloran, J. M., Vadas, P., & Alarcon, V. J. (2016). Exploring agricultural production systems and their fundamental components with system dynamics modelling. *Ecological Modelling*, *333*, 51–65. https://doi.org/10.1016/J.ECOLMODEL.2016.04.015

Wieder, W. R., Allison, S. D., Davidson, E. A., Georgiou, K., Hararuk, O., He, Y., Hopkins, F., Luo, Y., Smith, M. J., Sulman, B., Todd-Brown, K., Wang, Y.-P., Xia, J., & Xu, X. (2015). Explicitly representing soil microbial processes in Earth system models. *Global Biogeochemical Cycles*, *29*(10), 1782–1800. https://doi.org/10.1002/2015GB005188

Wu, Q., Ni, X., Sun, X., Chen, Z., Hong, S., Berg, B., Zheng, M., Chen, J., Zhu, J., Ai, L., Zhang, Y., & Wu, F. (2025). Substrate and climate determine terrestrial litter decomposition. *Proceedings of the National Academy of Sciences*, *122*(7), 2420664122. https://doi.org/10.1073/pnas.2420664122

Xia, J. (2024). Overview of traceability analysis. In Y. Luo & B. Smith (Eds.), *Land Carbon cycle modeling: Matrix approach, data assimilation, ecological forecasting, and machine learning* (2nd ed., 105–111). https://doi.org/10.1201/9781032711126-22

Zhang, Y., King, A. E., Hamilton, E., & Cotrufo, M. F. (2024). Representing cropping systems with the MEMS 2 ecosystem model. *Agronomy Journal*, *116*(5), 2328–2345. https://doi.org/10.1002/AGJ2.21611

SUPPORTING INFORMATION

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

How to cite this article: Barioni, L. G., Valladão, B. A., Mourão, V. H. M., Ewing, R. P., Karatay, Y. N., Damian, J. M., Melício, V. C., Rejaili, R. P. A., & Silva, R. O. (2026). ProCarbon-Soil: A dynamic model for improved model-data compatibility in carbon farming. *Soil Science Society of America Journal*, *90*, e70218. https://doi.org/10.1002/saj2.70218