Check for updates

JGR Biogeosciences

INVITED ARTICLE
10.1029/2026JG009726

Special Collection:
Biogeosciences Leaders of Tomorrow: JGR: Biogeosciences Special Collection on Emerging Scientists

Evaluating Soil Carbon Models for Sub-Saharan Africa: Revealing Knowledge Gaps in Subtropical and Tropical Soil Biogeochemistry

Sophie F. von Fromm[1,2], Katherine S. Rocci[3,4,5], Christopher O. Anuo[6], Stephen B. Asabere[7], Jeanette Kanyiri[8], Steve Kwatcho Kengdo[9], Admore Mureva[10], Kwabena A. Nketia[11,12], Lei Zhang[9], and Rose Z. Abramoff[13]

[1]Department of Biological Sciences, Dartmouth College, Hanover, NH, USA, [2]Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA, [3]University of Colorado, Boulder, CO, USA, [4]University of Michigan, Ann Arbor, MI, USA, [5]Now at University of California, Merced, CA, USA, [6]Department of Agronomy, Purdue University, West Lafayette, IN, USA, [7]Department of Physical Geography, University of Goettingen, Göttingen, Germany, [8]Department of Land Resource Management and Agricultural Technology, University of Nairobi, Nairobi, Kenya, [9]Climate and Ecosystem Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [10]Department of Natural Resources, Bindura University of Science Education, Bindura, Zimbabwe, [11]Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada, [12]Nutrient Centre for Sustainable and Digital Agriculture, University of Saskatchewan, Saskatoon, SK, Canada, [13]School of Forest Resources, University of Maine, Orono, ME, USA

Abstract Process-based soil carbon (C) models are increasingly used to project regional and global C cycle responses to climate change. However, the development and evaluation of these models has largely focused on temperate regions of North America and Europe. This geographic bias raises a critical question: Do these models capture generalizable mechanisms that can be applied to underrepresented pedological regions or encode processes specific to their developmental context? We evaluated three process-based models—Century, Millennial, and MIMICS—across 777 topsoil samples spanning the climate and pedological diversity of sub-Saharan Africa. Despite their differences in mechanistic detail, all three models performed similarly (adjusted R^2 = 0.09–0.18) in predicting soil organic carbon (SOC) stocks. Using random forest algorithms trained on observed and modeled SOC data, we identified divergences between the drivers of SOC. All three models overemphasized net primary productivity as a SOC driver and misrepresented the role of organo-mineral interactions. Bias analyses revealed that the three process-based models inadequately capture exchangeable calcium, which is increasingly recognized as an important control on SOC. Notably, increased mechanistic complexity did not improve transferability. These results have significant implications for regional C budgets and global climate projections. They underscore the importance of incorporating region-specific biogeochemistry into future soil C models in (sub-)tropical regions to enhance the precision of climate projections.

Plain Language Summary Soils store enormous amounts of carbon, and understanding how this carbon responds to climate change requires accurate computer models. However, most soil carbon models have been developed and tested in North America and Europe, raising the question of whether they can reliably predict soil carbon in other regions. We tested three widely used soil carbon models across sub-Saharan Africa using measurements from 777 soil samples spanning the region's climate and soil diversity. All three models performed poorly and similarly, explaining only 9%–18% of the observed variation in soil carbon stocks. Using machine learning, we compared what actually controls soil carbon in these soils against what the models think controls it. We found that all three models overestimate the importance of plant material inputs and fail to capture the importance of soil mineralogy. Importantly, newer and more complex models did not outperform the simpler older model. These findings suggest that current soil carbon models have fundamental structural gaps when applied outside temperate regions, which likely leads to unreliable carbon cycle projections for sub-Saharan Africa and similar tropical regions worldwide. Improving these projections will require developing models that incorporate region-specific soil processes and collecting better observational data across underrepresented regions.

Key Points:

- Process-based soil carbon models show limited transferability to sub-Saharan Africa
- Models overemphasize net primary productivity and inadequately represent organo-mineral interactions and exchangeable calcium as soil carbon controls
- Increased mechanistic complexity does not improve model performance, highlighting the need for region-specific biogeochemistry in soil carbon projections

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. F. von Fromm,
sophie.vonfromm@montana.edu

Citation:

von Fromm, S. F., Rocci, K. S., Anuo, C. O., Asabere, S. B., Kanyiri, J., Kengdo, S. K., et al. (2026). Evaluating soil carbon models for sub-Saharan Africa: Revealing knowledge gaps in subtropical and tropical soil biogeochemistry. *Journal of Geophysical Research: Biogeosciences*, *131*, e2026JG009726. https://doi.org/10.1029/2026JG009726

Received 16 JAN 2026
Accepted 3 MAY 2026

Author Contributions:

Conceptualization: Sophie F. von Fromm, Katherine S. Rocci, Christopher O. Anuo, Stephen B. Asabere, Jeanette Kanyiri, Steve Kwatcho Kengdo, Admore Mureva, Kwabena A. Nketia, Lei Zhang, Rose Z. Abramoff

© 2026. The Author(s).
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Data curation: Sophie F. von Fromm, Katherine S. Rocci, Kwabena A. Nketia, Lei Zhang, Rose Z. Abramoff
Formal analysis: Sophie F. von Fromm, Katherine S. Rocci, Rose Z. Abramoff
Funding acquisition: Sophie F. von Fromm
Investigation: Sophie F. von Fromm, Katherine S. Rocci, Christopher O. Anuo, Stephen B. Asabere, Jeanette Kanyiri, Steve Kwatcho Kengdo, Admore Mureva, Kwabena A. Nketia, Lei Zhang, Rose Z. Abramoff
Methodology: Sophie F. von Fromm, Katherine S. Rocci, Stephen B. Asabere, Jeanette Kanyiri, Steve Kwatcho Kengdo, Admore Mureva, Kwabena A. Nketia, Lei Zhang, Rose Z. Abramoff
Resources: Sophie F. von Fromm
Software: Sophie F. von Fromm, Katherine S. Rocci, Rose Z. Abramoff
Visualization: Sophie F. von Fromm
Writing – original draft: Sophie F. von Fromm, Kwabena A. Nketia
Writing – review & editing: Katherine S. Rocci, Christopher O. Anuo, Stephen B. Asabere, Jeanette Kanyiri, Steve Kwatcho Kengdo, Admore Mureva, Lei Zhang, Rose Z. Abramoff

1. Introduction

Process-based soil carbon (C) models are increasingly applied globally to simulate terrestrial carbon-climate feedbacks, yet their development and evaluation have been geographically clustered in temperate North America and Europe (Garsia et al., 2023). This creates a critical test: Do these models capture generalizable mechanisms that transfer across Earth's pedological diversity, or do they encode processes specific to their developmental context? The answer has profound implications for global C cycle projections and efforts to manage soil C in underrepresented regions. More recent process-based models explicitly simulate mechanisms, such as microbial decomposition, mineral-organic interactions, and environmental controls, rather than empirical correlations, theoretically providing greater predictive capacity under novel conditions (Bosatta & Agren, 1991; Lehmann & Kleber, 2015). However, this advantage holds only if the mechanistic representations accurately reflect the processes operating across diverse soil systems, and if the necessary data are present to accurately parameterize the models.

Pedological diversity across the globe is the result of the interaction of soil-forming factors, such as climate, parent material, time, organisms, and topography (Dokuchaev, 1883; Jenny, 1941). Different combinations of these soil-forming factors generate distinct biogeochemical environments where the dominant mechanisms controlling soil organic carbon (SOC) persistence, turnover, and environmental sensitivity may differ systematically. For example, soil mineralogy, determined by parent material and weathering stage, influences organo-mineral association mechanisms and sorption dynamics (Rasmussen et al., 2018). Highly weathered soils, dominated by iron and aluminum oxides and by low-activity clays, exhibit different stabilization processes than younger soils with 2:1 clays, higher cation exchange capacity, and higher base saturation (Six et al., 2002; Torn et al., 1997; von Fromm et al., 2021, 2023). Climate regimes similarly influence not only decomposition rates but also moisture dynamics, redox conditions, and microbial community composition and activity. Plant and microbial community effects are often related to climate regimes but can exert unique effects on soil biogeochemistry beyond those from climate (Rocci et al., 2025). These mechanistic differences have important implications for process-based model development. Mathematical representations of biogeochemical processes that successfully capture SOC dynamics in one pedological context may not directly apply to systems governed by different dominant mechanisms.

This challenge of transferability is particularly acute for tropical and subtropical soils, which differ systematically from temperate systems in terms of weathering stage, mineralogy, temperature and moisture regimes (von Fromm, Jungkunst, et al., 2025). These pedological differences result in substantially different SOC persistence and turnover rates (Shi et al., 2020; von Fromm et al., 2024; Zhang et al., 2025). While recent soil C model generations, such as the MIcrobial-MIneral Carbon Stabilization (MIMICS) and Millennial models, include more explicit representations of microbial metabolism and mineral interactions than earlier frameworks like Century (Abramoff et al., 2018, 2022; Parton et al., 1987; Wieder et al., 2014), their performance in tropical and subtropical soils largely remains untested. Sub-Saharan Africa presents a particularly critical context for evaluating process-based models. The region's highly weathered, oxide-rich soils differ substantially from the temperate conditions in which models are typically developed. Additionally, the region is facing unprecedented pressure from climate change, land use intensification, and demands for food security (Liu et al., 2025; Sakala & Santos, 2025). Evaluating model performance in sub-Saharan Africa is therefore essential for both assessing the regional reliability of SOC projections and testing whether mechanistic models can transfer successfully to pedological conditions underrepresented in their development.

Here, we test model transferability by evaluating three process-based soil C models with varying mechanistic complexity—Century, MIMICS, and Millennial—across sub-Saharan Africa, using an extensive soil chemistry data set for the region (Africa Soil Information Service; Vågen et al., 2021; von Fromm et al., 2021). Our objectives are to: (a) quantify model performance and systematic biases when applied to sub-Saharan African soils, and (b) identify key drivers of SOC that are either well or poorly represented by current model structures. By comparing three process-based models, we provide insights into both the regional reliability of mechanistic SOC projections and the broader question of whether variation in mechanistic complexity alters model transferability to pedological conditions underrepresented in model development.

This paper is organized as follows: First, we conduct a comprehensive literature review to systematically assess soil C modeling efforts in sub-Saharan Africa and contextualize our evaluation within the existing knowledge

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

base. We then present our methodology, model evaluation results, and discussion of implications for model transferability and regional SOC projections.

2. Process-Based Soil Carbon Modeling in Sub-Saharan Africa: Current State and Knowledge Gaps

Soils in sub-Saharan Africa represent a substantial but poorly characterized component of global SOC dynamics (FAO, 2022). Available SOC data suggest that approximately 24 Pg of carbon (Pg C) are stored in the top five cm of soil in sub-Saharan Africa (Hengl et al., 2015), which is more than twice the current annual global fossil CO_2 emissions (ca. 10 Pg C; Friedlingstein et al., 2025). Despite the increasing development and global application of process-based models, assessment of SOC dynamics with such models in sub-Saharan Africa remains unexplored. We conducted a systematic search on Web of Science with the syntax: TS=((“soil organic carbon” OR “soil organic matter”) AND model* AND (“sub-Saharan Africa” OR “Africa south of the Sahara” OR subsahar* OR “West Africa” OR “East Africa” OR “Southern Africa” OR Sahel OR “Horn of Africa”) AND (biogeochem* OR “process-based” OR mechanistic OR “Earth system” OR “land surface” OR CENTURY OR RothC OR DNDC OR DayCent OR Yasso* OR ECOSSE OR MIMICS OR (millenni* NEAR/3 (model OR soil OR carbon)))). Records were screened in two stages (title and abstract, then full text) against the following inclusion criteria: (a) explicit use of a process-based/biogeochemical model to simulate SOC/soil organic matter (SOM) dynamics; (b) sub-Saharan Africa focus (study sites or scenarios within sub-Saharan Africa); and (c) reporting of model setup and evaluation (e.g., calibration/validation, sensitivity/uncertainty or performance metrics). Exclusions included reviews, concept-only or remote-sensing studies without SOC-process representation, global analyses without sub-Saharan Africa breakout, and purely empirical/statistical models.

The search yielded 24 publications, of which 15 studies (1991–2026) met the inclusion criteria. Retained studies were geographically clustered, with Kenya ($n = 7$) and Ghana ($n = 3$) being the most represented. Single studies included those from Zimbabwe, Burkina Faso, Senegal, and South Africa, and one multi-country study using harmonized sites across sub-Saharan Africa. The most frequently applied model families included first-generation C models, such as RothC and Century/DayCent (e.g., Couëdel et al., 2026; Kamoni et al., 2007; Ojima et al., 1993; Tan et al., 2010). These models were generally used to examine the management or climate sensitivities of SOC turnover. For instance, one study calibrated and validated process-based models for long-term Kenyan Integrated Soil Fertility Management trials using stepwise optimization and evaluation (Couëdel et al., 2026). Similarly, biogeochemical simulations in northern Ghana quantified the sensitivity of savanna-derived cropland to warming–drying and nitrogen fertilization scenarios (Tan et al., 2010). This limited and spatially uneven collection of work underscores a clear gap in the process-based representation of SOC dynamics across sub-Saharan Africa's diverse soils and management contexts.

Recent evaluations of SOC models in sub-Saharan Africa demonstrate both progress and persistent limitations. For instance, RothC applications under tropical conditions captured general trends but failed to reproduce variability linked to residue quality and soil texture, reflecting the model's fixed decomposition constants that do not account for the diverse litter inputs characteristic of sub-Saharan African management systems (Geremew et al., 2024). However, methodological advances show promise: Laub et al. (2024) applied Bayesian calibration to DayCent in Kenyan maize agroecosystems and demonstrated that parameter uncertainty declines substantially when locally constrained priors are used. Together with findings from Nyawira et al. (2021) and Ma et al. (2022), these studies indicate that temperate-derived models require region-specific re-parameterization and local validation to achieve credible performance in sub-Saharan Africa.

The limited performance of traditional soil C models in sub-Saharan Africa could reflect deficiencies in their structural design. Models such as RothC and Century use first-order kinetics where decomposition rates depend on temperature, moisture, and substrate quality, but treat microbial physiology, necromass formation, and mineral associations implicitly or ignore them entirely (Bradford et al., 2016). This is particularly problematic for sub-Saharan Africa's highly weathered soils, where Fe and Al oxides and low-activity clays exert strong controls on SOC persistence through sorptive interactions and co-precipitation processes (Rasmussen et al., 2018; von Fromm et al., 2021, von Fromm, Jungkunst, et al., 2025). Doetterl et al. (2015) demonstrated that the dominant control on SOC persistence shifts from climatic factors in younger soils to geochemical drivers (Fe/Al oxides, clay content) in highly weathered systems; precisely the conditions where traditional temperate-parameterized models may systematically misrepresent governing processes. Recent model generations such as MIMICS and

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Millennial explicitly represent microbial metabolism and mineral-organic interactions (Abramoff et al., 2018, 2022; Wieder et al., 2014), yet their performance in tropical and subtropical soils, including sub-Saharan Africa, remains largely untested.

This synthesis reveals substantial gaps in our understanding of the process-based SOC model performance in sub-Saharan Africa. First, existing studies are geographically clustered (primarily Kenya and Ghana) and focus predominantly on agricultural systems, providing limited insight into model transferability across sub-Saharan Africa's pedological and land-use diversity. Second, the models tested are mainly first-generation frameworks (RothC, Century/DayCent). Newer microbially explicit models remain unevaluated despite their theoretical advantages for capturing mineral-microbial interactions in highly weathered soils. Third, systematic comparisons of model performance across mechanistic complexity levels are absent, limiting our ability to assess whether increased mechanistic detail improves predictive reliability in (sub-)tropical systems. This study provides the first systematic evaluation of mechanistic model transferability across sub-Saharan Africa's pedological diversity, directly testing whether increased complexity improves predictive reliability in highly weathered (sub-)tropical soil systems.

3. Materials and Methods

Using an extensive soil data set from sub-Saharan Africa, we evaluated three process-based soil C models (Century, Millennial, and MIMICS). We assessed the models' performance by comparing the simulated and observed SOC stocks across 777 topsoil samples. To determine if the models capture the dominant environmental controls on SOC, we employed random forest algorithms that were trained using both observed and modeled data. This enabled us to compare the empirical relationships in the data with the mechanistic responses embedded in the process-based models. Finally, we quantified systematic biases in model predictions relative to key soil properties hypothesized to control SOC dynamics in highly weathered soils.

3.1. Study Region and Data Sets

Soil data used in this study were previously collected and analyzed as part of the African Soil Information Service (AfSIS). A total of 18,257 soil samples were collected from 60 sites (100 km^2) and from two depth intervals (0–20 and 20–50 cm) between 2010 and 2012, following the well-established Land Degradation Surveillance Framework (Vågen et al., 2016, 2021). The sites are stratified across all major Koeppen-Geiger climate zones present in sub-Saharan Africa. Within each sentinel site there were 10 plots of 1,000 m^2 randomized within 16 spatially stratified 1,000 m^2 clusters. At each sampling location, samples from four sub-plots (100 m^2) were combined into one sample for topsoils (0–20 cm) and subsoils (20–50 cm). The hierarchical sampling design allows process identification without losing the ability to understand and quantify local heterogeneity (von Fromm et al., 2021, 2023). A subset of the AfSIS data set (n = 2,002), was previously analyzed for wet chemistry, including SOC, oxalate-extractable aluminum (Al_{ox}) and iron (Fe_{ox}), exchangeable calcium (Ca_{ex}), and soil pH (Vågen et al., 2021). Clay content (<2 μm), clay + silt content (<63 μm) were quantified using laser diffraction, and clay mineralogy (1:1 and 2:1 clay minerals) was quantified using X-ray powder diffraction (von Fromm et al., 2023).

The analysis presented here is based on a subset of the AfSIS wet chemistry data set. Since none of the versions of the process-based models used in this study are depth-resolved, the analysis focuses only on topsoils (0–20 cm). We filtered the samples to include only those with complete observations. In total, we used 777 topsoil samples from various sampling locations across sub-Saharan Africa that covered a wide range of pedo-climatic conditions (Figures S1 and S2 in Supporting Information S1). The selected samples cover 14% forest, 26% cropland, 15% grassland, and 44% other (including mainly woodland, shrubland and bushland but also samples classified as other). The land cover data was collected in the field during the original sampling of the AfSIS data.

As input variables for the process-based models, we extracted net primary productivity (NPP), soil temperature and moisture, and estimated lignin content and "lignin to nitrogen" ratio. We used MODIS NPP data (MOD17A3HGF V6.1) as a proxy for C inputs at a given sampling location (Gottschalk et al., 2012; Smith et al., 2007). Annual NPP is derived from the sum of all 8-day Net Photosynthesis (PSN) products (MOD17A2H) covering the sampling period (2010–2012) at 500 m pixel resolution (Running & Zhao, 2021). For soil moisture and temperature, we used ERA5-Land, a global hourly reanalysis data set generated by rerunning the land component of the European Center for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System in an offline mode (Hersbach et al., 2020; Muñoz-Sabater et al., 2021). ERA5-Land data sets were

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

accessed via the Copernicus Climate Data Store Application Programming Interface using the R package "*ecmwfr* (Hufkens et al., 2019)." Hourly data sets were downloaded for the African continent for the period 2012–2025. To derive a representative topsoil layer, the upper two depth intervals (0–7 cm and 7–28 cm) were vertically aggregated into a 0–20 cm soil depth interval composite using the depth-specific weighted average of each variable. Here, numerical integration based on the trapezoidal rule was applied to maintain the vertical structure of the soil profile, following the *GlobalSoilMap* specifications (Arrouays et al., 2014). Subsequently, the 0–20 cm composite layers were aggregated annually for soil moisture and soil temperature. Bulk density was extracted from the field-level iSDA soil maps for Africa at 30 m resolution (Hengl et al., 2021) to calculate SOC stocks based on the measured SOC content in the AfSIS data set. To determine the litter lignin content and ratio with N (lignin:N), we extracted global data on litter lignin and N and plant growth from the TRY database (Kattge et al., 2020). We then determined mean litter lignin and N based on plant growth form and then determined weighted mean lignin and lignin:N values for each site based on their recorded vegetation structure.

3.2. Process-Based Soil Models

Century is a first-generation process-based soil C model that simulates organic matter decomposition using first-order kinetics, where fixed decomposition rate constants can be modified by temperature, moisture, soil texture, and substrate quality (Parton et al., 1987). The model partitions C into discrete pools (structural litter, metabolic litter, active, slow, and passive pools). The latter three pools represent SOM and are defined by their relative decomposition rates. Despite its relatively simple mathematical structure, Century has been widely applied globally to simulate long-term C dynamics under various land management and climate scenarios, including sub-Saharan Africa (Dintwe & Okin, 2018; Kamoni et al., 2007; Ojima et al., 1993). A key limitation is that Century treats microbial physiology implicitly and does not explicitly represent mineral-organic interactions.

Millennial Version 2 (hereafter "Millennial") is a mechanistically advanced soil C model that explicitly represents microbial metabolism, including microbial biomass dynamics, carbon-use-efficiency (CUE), and the formation of microbial necromass as a distinct C pool (Abramoff et al., 2018, 2022). The model incorporates temperature-dependent physiological processes through Arrhenius kinetics, with soil moisture modulating decomposition and transformation rates across multiple processes. Critically, Millennial explicitly represents mineral-organic interactions through sorption mechanisms that are further modified by soil texture and pH. The model partitions soil C into measurable pools: particulate organic matter (POM), low molecular weight C (LMWC), aggregate C, mineral-associated organic matter (MAOM), and microbial biomass. By treating both microbial necromass formation and mineral-mediated C stabilization as explicit processes, Millennial aims to better capture the mechanisms through which soil C is protected from decomposition compared to first-order kinetic models like Century.

Microbial-Mineral Carbon Stabilization is a process-based soil C model specifically designed to integrate microbial physiological and mineral controls on soil C dynamics (Wieder et al., 2014, 2015). MIMICS aims to strike a balance between the simplicity of Century and the representation of more novel soil biogeochemical mechanisms. Like Millennial, MIMICS incorporates temperature-sensitive Michaelis-Menten kinetics to govern decomposition rates, but uniquely parametrizes these kinetics by microbial functional type and litter chemical quality. Following Century, MIMICS separates litter into a metabolic and structural litter pool based on litter quality. These litter pools are then preferentially decomposed by copiotrophic and oligotrophic microbial groups, respectively, which differ in their growth strategies and substrate use patterns. Microbial turnover subsequently contributes C to physically protected and chemically protected SOM pools (comparable to MAOM and POM, respectively), as well as to the available SOM pool, from which microbes can also assimilate C. By simultaneously representing microbial community diversity, growth efficiency, and mineral-mediated stabilization, MIMICS aims to better capture the mechanisms controlling soil C persistence in mineral soils compared to traditional first-order kinetic models, while also improving predictions of soil C responses to environmental change.

We selected these three models to represent a range of mechanistic complexity and theoretical development in soil C modeling (Table S1 in Supporting Information S1). Century serves as a foundational benchmark against which many newer models are evaluated. Meanwhile, Millennial and MIMICS attempt to integrate the emerging understanding of microbial physiology and mineral stabilization. This selection allows us to test whether increased

mechanistic complexity improves the transferability of models developed primarily in temperate regions to different pedo-climatic regions, such as sub-Saharan Africa.

3.3. Process-Based Model Optimization and Parameterization

We optimized parameters for each model using approaches tailored to the model structure and available calibration methods. This approach takes into account that each model has distinct parameterization requirements and established calibration workflows. Our aim is to produce the best fit for each model, rather than compare model calibrations. To assess the effect of parameter optimization, all models were run with both default and fitted parameter values, and performance metrics were reported for both runs.

For Century, we fitted 10 out of 21 parameters using the Markov Chain Monte Carlo (MCMC) method described in Pierson et al. (2022). Parameters were selected based on both sensitivity analysis with comparable parameters in the Community Land Model (CLM; Kennedy et al. (2025) and expected importance for decomposition (e.g., soil temperature and moisture variables). Briefly, the MCMC method samples from prior ranges of each parameter and iteratively narrows the parameter space in which the model improves. Model priors were informed by parameter ranges used for comparable parameters in CLM (Kennedy et al., 2025) and model improvement was defined as decreased root mean square error (RMSE) and increased correlation for observed versus modeled soil C stocks. Default parameter values for the remaining parameters are provided in Table A2 of Abramoff et al. (2022).

For Millennial, we followed the fitting procedure described by Abramoff et al. (2022), with modifications to parameter selection. Rather than identifying parameters for fitting through formal sensitivity analysis, we selected three out of 24 parameters expected to exert the largest direct influence on decomposition rates: the activation energy and pre-exponential constant of the maximum rate of POM decomposition to LMWC, and the reference temperature controlling temperature-dependent changes in CUE. Default parameter values for the remaining parameters are provided in Table A1 of Abramoff et al. (2022). Fitting additional parameters was not pursued, as Abramoff et al. (2022) demonstrated that expanding the number of optimized parameters in Millennial produces only marginal improvements in predictive performance due to parameter equifinality and inter-parameter correlations.

For MIMICS, we fitted 8 out of 22 parameters using the same method as for Century. Parameters and priors for MIMICS were directly informed by Pierson et al. (2022), who optimized MIMICS parameter ranges for predicting soil C. Default parameter values for the remaining parameters are provided in Table 1 of Rocci et al. (2025).

Although parameters fitted and calibration workflows differ between models, for all models, we partitioned 80% of the data into a training set and 20% into a testing set. All model evaluation metrics are derived from model predictions compared to the testing set. In addition, a full list of all optimized parameters and their optimal values can be found in Table S2 in Supporting Information S1.

3.4. Statistical Modeling

To evaluate whether the simulated and measured SOC stocks have the same environmental drivers, we developed two complementary sets of random forest regression algorithms (Breiman, 2001) for each process-based model (Century, Millennial, MIMICS). This two-pronged approach enables us to distinguish which variables empirically predict SOC stocks and to which variables the process-based models mechanistically respond. Each random forest algorithm was trained using the same forcing variables specific to each process-based model and evaluated using 10-fold spatial cross-validation.

Random forest trained on observed SOC stocks were trained on the observed SOC stocks in the AfSIS data set ($n = 777$). These algorithms establish empirical relationships between environmental drivers and measured SOC stocks, revealing which variables control SOC variation in soils across sub-Saharan Africa. To account for spatial autocorrelation, samples from the same cluster were assigned together to either the training or testing sets (spatial blocking). Comparing these relationships to process-based model outputs identifies whether models capture the dominant controls on SOC.

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Random forest trained on modeled SOC stocks were trained on SOC predictions from the three process-based models forced with the AfSIS data set (n = 777), with spatial blocking applied to maintain independence between training and testing sets. By comparing the variable importance and functional relationship between the random forest models trained on the observed SOC stocks and the modeled SOC stocks, we identify which variables are mechanistically emphasized when generating predictions for process-based models. Large divergences between these two random forest approaches indicate that models respond differently to input variables than what the empirical data suggests, revealing structural biases in how models represent SOC controls.

For each random forest algorithm, model evaluation was performed on the testing data set, including the calculation of the R^2, mean absolute error, and RMSE. To assess the importance of each independent variable for the predictive power of the model, we calculated the "permutation feature importance." This measure can be interpreted as an explanatory variable being important if the shuffling of its values increases the model error, indicating that the model relied on that explanatory variable for prediction (Molnar, 2025). To further interpret the outcome of the random forest models, we used partial dependence plots. The partial dependence plots show the marginal effects of an explanatory variable on the predicted outcome of the random forest model (Friedman, 2001).

To evaluate systematic biases specific to each process-based model and its forcing variables, as well as important soil properties not included in the models, we calculated the bias for each sample. The bias is defined as the difference between the predicted SOC stocks by the process-based model and the observed SOC stocks in the AfSIS data set. The selected soil properties, namely soil pH, Ca_{ex}, Al_{ox}, Fe_{ox}, 1:1 and 2:1 clay minerals, have been shown to be important in explaining the variation in SOC across sub-Saharan Africa (von Fromm et al., 2021, 2023).

All statistical analyses were performed in R (version 4.4.1; R Core Team, 2024). In addition to the R packages mentioned above, the following R packages were used: "tidyverse" (Wickham et al., 2019), "ggpubr" (Kassambara, 2023), "mlr3" (Lang et al., 2019), "iml" (Molnar et al., 2018), "here" (Müller, 2020), "future" (Bengtsson, 2021), "gridExtra" (Auguie, 2017), and "FME" (Soetaert & Petzoldt, 2010). The R code and all data to reproduce all analysis can be found in (von Fromm et al., 2026).

4. Results

All three process-based models performed similarly in predicting SOC stocks across sub-Saharan Africa (adjusted R^2 = 0.09–0.18). Despite their mechanistic differences, the models diverged from empirical relationships in how they weighed environmental drivers of SOC. Most notably, they overemphasized NPP and misrepresented the roles of soil moisture and clay content. Systematic biases in the models' predictions revealed that they all inadequately represent stabilization mechanisms of SOC associated with exchangeable calcium. These findings suggest that current process-based models overlook critical geochemical processes that govern SOC dynamics in highly weathered soils, thereby limiting their transferability to sub-Saharan African pedological contexts.

4.1. Process-Based Model Performance

The three fitted process-based models (Century, Millennial, and MIMICS) show similar model performances despite their differences in model structure (Figure 1). The adjusted R^2 between observed and modeled SOC stocks ranges from 0.09 (MIMICS) to 0.18 (Millennial), with RMSE between 4.01 (Millennial) and 4.73 kg m^{-2} (Century).

Century shows the smallest improvement between its default model (adj R^2 = 0.14, RMSE = 5.06 kg m^{-2}) and its fitted model (adj R^2 = 0.16, RMSE = 4.73 kg m^{-2}; Figure 1a). This indicates that Century might be more robust to new data or that model parameters are weaker drivers of model performance than input variables (Figure S3 in Supporting Information S1; Abramoff et al., 2022; Georgiou et al., 2017). The fitted decomposition rate constants show a nuanced pattern relative to their defaults (Table S2 in Supporting Information S1): *k_active* was reduced by 43%, while *k_slow* and *k_passive* increased modestly (+10% and +21%, respectively). Additionally, litter transfer parameters shifted C pathways away from direct structural litter input to the active pool (*strlitter_to_active* −37%) and toward the slow pool (*strlitter_to_slow* +16%).

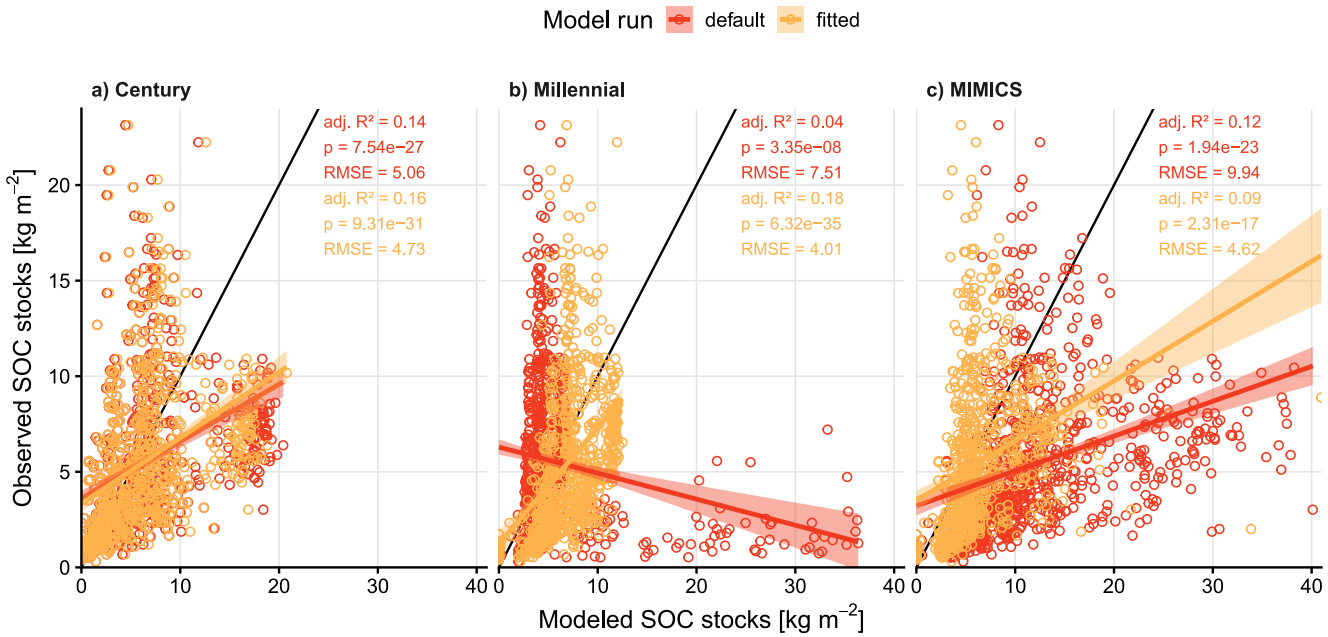


Figure 1. Comparison of the observed soil organic carbon (SOC) stocks and the modeled SOC stocks for (a) Century, (b) Millennial, and (c) MIMICS for the default and fitted model runs. The solid black lines represent the 1:1 line.

Millennial shows the largest improvement between its default (adj $R^2 = 0.04$, RMSE = 7.51 kg m^{-2}) and fitted model (adj $R^2 = 0.18$, RMSE = 4.01 kg m^{-2}; Figure 1b). This improvement was primarily driven by samples with high soil temperature values (>30°C; Figure S4 in Supporting Information S1). Of the three optimized parameters, the most consequential change was an increase in the reference temperature for temperature control on CUE (*tae_ref*: 15°C–19.6°C, +31%), with a concurrent reduction in the pre-exponential constant of the maximum POM decomposition rate (*alpha_pl*: −43%) and a negligible change in its activation energy (*eact_pl*: +2%; Table S2 in Supporting Information S1). In the default Millennial model, the low reference temperature for CUE resulted in excessive SOC accumulation under high soil temperature conditions, which the fitted parameter corrected (Figure 1b).

MIMICS showed a moderate increase in model performance between its default (adj R^2 = 0.12, RMSE = 9.94 kg m^{-2}) and fitted model (adj $R^2 = 0.09$, RMSE = 4.62 kg m^{-2}; Figure 1c). The most consequential parameter changes were a more than tripling of microbial turnover rates (*tau*, +229%), substantial reductions in the fraction of microbial necromass entering the physically protected pool (*fPHYS*, −95%), and desorption of protected C back to the available pool (*desorb*, −79%; Table S2 in Supporting Information S1). These opposing changes are aligned with relatively minimal changes in soil C predictions but altered C cycling in MIMICS, such that mineral protection is a greater mechanism for maintaining the protected C pool than microbial input. Among the three process-based model outputs, MIMICS shows the greatest overprediction in SOC stocks across the AfSIS data set, driven by sites with high clay content (Figure S5 in Supporting Information S1). Interestingly, MIMICS also has the highest proportion of SOC in the "unprotected" pool (~75% in the fitted model), while Century and Millennial store less than 10% in the "unprotected" pool (Figure S6 in Supporting Information S1). The performance of all three process-based models is relatively weak across sub-Saharan Africa but not much worse than global predictions of SOC with these models (Abramoff et al., 2022; Wang et al., 2025).

4.2. Comparison of Statistical and Process-Based Models

The three process-based models (Century, Millennial, and MIMICS) show divergent patterns in how they weigh environmental and soil variables as drivers of SOC compared to empirical relationships (Figures 2 and 3, Figures S7–S9 in Supporting Information S1). For example, all three models overemphasize the importance of NPP in predicting SOC stocks relative to the empirical relationships (Figure 2). This overemphasis is most pronounced in Century and Millennial models, suggesting that these models attribute too much SOC variation to C inputs. The

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

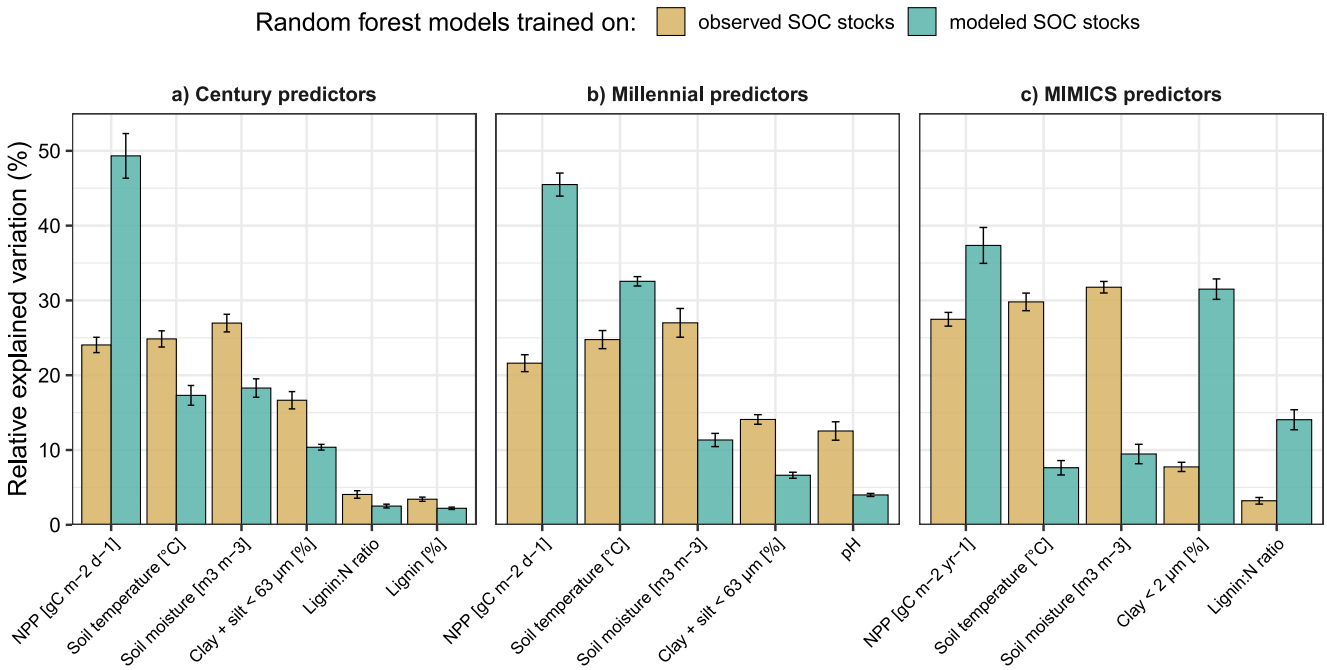


Figure 2. Variable importance plots for (a) Century predictors, (b) Millennial predictors, and (c) MIMICS predictors colored by the random forest models trained on observed soil organic carbon (SOC) stocks in the AfSIS dataset and the random forest models trained on process-based modeled SOC stocks.

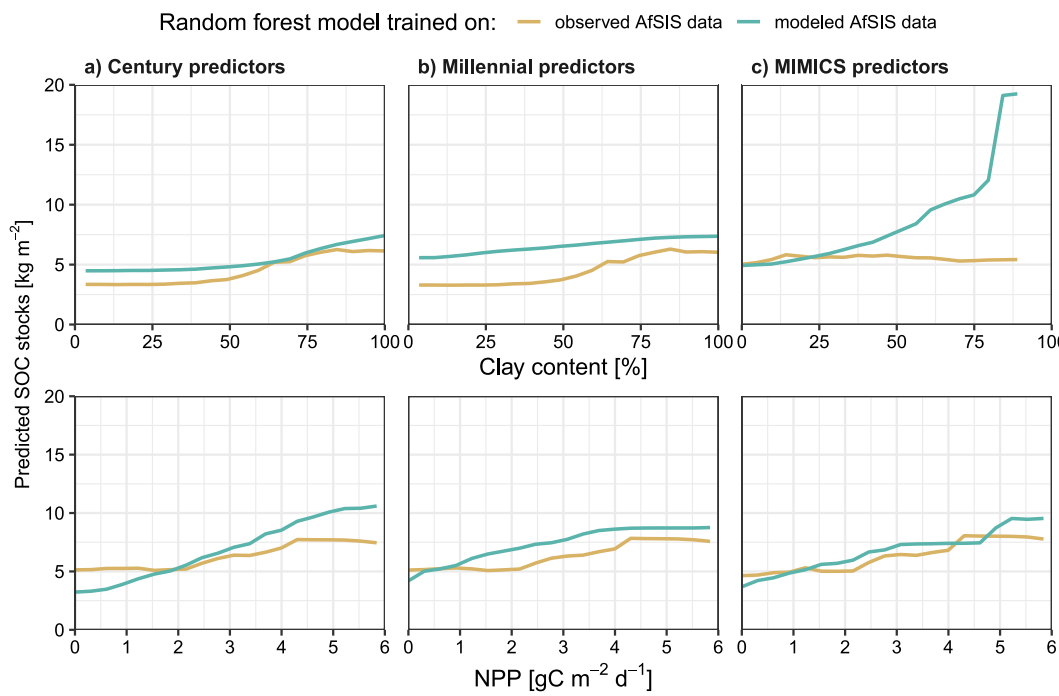


Figure 3. Partial dependence plots for (a) Century, (b) Millennial, and (c) MIMICS for clay content (top row) and net primary productivity (NPP; bottom row) colored by the random forest models trained on the observed soil organic carbon (SOC) stocks in the AfSIS dataset and the random forest models trained on process-based modeled SOC stocks. Note that the clay content for Century and Millennial is defined as all particles <63 µm, whereas for MIMICS it is defined as all particles <2 µm.

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

partial dependence plots confirm this pattern: the random forest trained on the process-based modeled SOC stocks predicts higher SOC stocks at high NPP values compared to the random forest trained on the observed SOC stocks in the AfSIS data set (Figure 3, bottom row).

Century shows the best agreement in variable importance between the random forest model trained on the observed SOC stocks in the AfSIS data set and the random forest model trained on the process-based modeled SOC stocks, with the exception of NPP (Figure 2a). However, this agreement may reflect that both approaches are missing similar important processes rather than indicating that Century adequately captures the dominant controls. In contrast, both Millennial and MIMICS show additional systematic divergences. Millennial underemphasizes the importance of soil moisture relative to the observed data (Figure 2b), indicating that the model's explicit representation of moisture-dependent physiological processes may not adequately capture how soil water availability controls SOC dynamics across sub-Saharan Africa. MIMICS exhibits the largest overall divergence from empirically derived variable importance (Figure 2c). Most notably, MIMICS overemphasizes the importance of clay content (<2 μm) in its predictions relative to empirical relationships, while simultaneously underemphasizing the importance of both soil temperature and soil moisture. The partial dependence plots reveal the severity of this divergence. The random forest model trained on the modeled SOC stocks predicts that SOC stocks increase nonlinearly with clay content. However, the random forest model trained on the observed SOC stocks shows almost no relationship between the clay content and the predicted SOC stocks (Figure 3c, top row). This pattern suggests that MIMICS's representation of mineral-organic interactions through sorption mechanisms may not adequately capture the specific mechanisms through which minerals protect C in soils across sub-Saharan Africa.

Collectively, these findings indicate that despite their mechanistic differences, all three process-based models mechanistically respond to the available input variables in ways that diverge from empirical relationships. The random forest models trained on the observed SOC stocks in the AfSIS data set differ form those trained on the process-based modeled SOC stocks (Figure S10 in Supporting Information S1), suggesting that the tested process-based models do not fully capture the environmental controls on SOC stocks in soils across sub-Saharan Africa.

4.3. Bias Estimates Based on Fitted Process-Based Models

To identify which mechanisms may be missing in the process-based models, we examined systematic biases in their predictions relative to soil geochemical properties that are either absent from or inadequately represented in model inputs. The selected geochemical predictors were chosen based on their demonstrated importance in predicting SOC abundance and persistence across sub-Saharan Africa (von Fromm et al., 2021, 2023). A systematic bias in SOC stocks with a specific geochemical soil parameter indicates that the process-based model does not adequately capture the mechanisms controlled by that property. All three fitted process-based models (Century, Millennial, and MIMICS) showed systematic biases regarding important geochemical predictors of SOC (Figure 4 and Figure S11 in Supporting Information S1) and spatially (Figure S12 in Supporting Information S1).

Of the selected geochemical soil properties, soil pH, exchangeable Ca, and clay content (<2 μm) explain the largest variation in the bias in SOC stocks (adj $R^2 \geq 0.10$ for at least one of three process-based models; Figure 4). In contrast, oxalate extractable Al and Fe, clay + fine silt content (<63 μm), 1:1 clay minerals, and 2:1 clay minerals explain much less of the variation in the bias in SOC stocks (adj $R^2 \leq 0.10$ for all three process-based models; Figure S11 in Supporting Information S1). Century shows the largest bias in SOC stocks with soil pH in comparison to Millennial and MIMICS, that is Century overpredicts SOC under lower soil pH and underpredicts SOC under higher soil pH (Figure 4a). This indicates that Century is decomposing SOC too slowly under acidic soil conditions, and too quickly under basic soil conditions. The same systematic bias in SOC stocks occurs with exchangeable Ca, that is, under higher exchangeable Ca (which is usually found under higher soil pH conditions), all three process-based models are underpredicting SOC stocks (Figure 4b). This indicates that all three process-based models lack important SOC stabilization mechanisms associated with exchangeable Ca. MIMICS shows the largest bias in SOC stocks with clay content compared to Century and Millennial (Figure 4c). However, both Century and MIMICS underpredict SOC stocks under lower clay content and overpredict SOC stocks under higher clay content, whereas Millennial shows no systematic bias with clay content. This indicates that Century and MIMICS both protect too much SOC from decomposition under higher clay content. Overall, the biases in

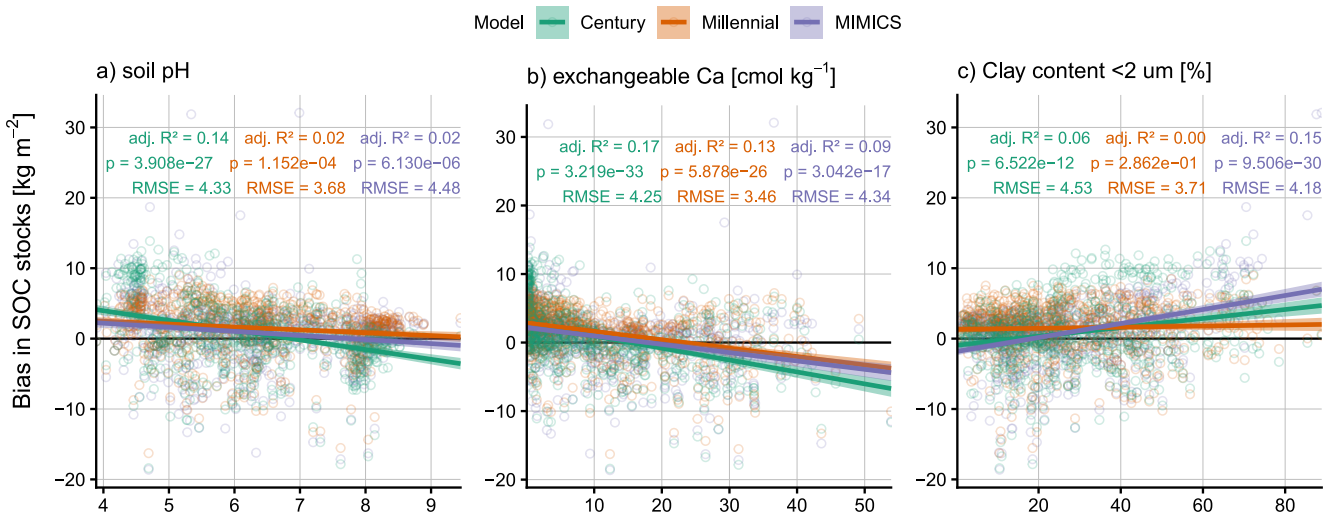


Figure 4. Estimated bias (predicted–observed) in soil organic carbon stocks colored by the fitted process-based models (Century, Millennial, MIMICS) for (a) soil pH, (b) exchangeable calcium (Ca), and (c) clay content <2 μm.

SOC stocks with the specific geochemical soil properties remain relatively low (Figure 4 and Figure S11 in Supporting Information S1). This indicates that there is substantial random noise in SOC stock bias that cannot be attributed to mechanisms controlled by geochemical soil properties.

5. Discussion

Our evaluation reveals a fundamental mismatch between the mechanistic representation of SOC controls in process-based soil C models and the dominant drivers of SOC in soils across sub-Saharan Africa. Although mechanistically complex models (Millennial and MIMICS) show promise when fitted to regional data, structural gaps limit their improvements, which cannot be remedied through parameterization alone. These findings have implications for regional C budgets and global climate projections. Our results suggest that transferring models to tropical and subtropical regions requires more than just better parameterization; it may require fundamental modifications to how models represent mineral-organic interactions and moisture-redox coupling, as well as improved input variables and data.

5.1. Tropical-Specific Soil Processes Absent From Current Model Structures

The three fitted process-based models (Century, Millennial, and MIMICS) exhibit moderate and remarkably similar performance across soils in sub-Saharan Africa (adj. R^2 = 0.09–0.18; Figure 1), despite their substantial differences in mechanistic complexity and pool structure (Figure S6 in Supporting Information S1). This similarity in model performance is striking because it contradicts the implicit assumption underlying recent model development: that explicitly representing microbial physiology and mineral-organic interactions improves predictive power. Our results suggest that mechanistic complexity does not necessarily lead to improved transferability. The value of mechanistic detail fundamentally depends on whether the chosen mechanisms accurately reflect the processes operating in the target region and if adequate and sufficient data are available.

Mineralogical Control and Weathering Stage. Soils in sub-Saharan Africa are predominantly highly weathered, with a clay mineralogy dominated by 1:1 clays (kaolinite) and iron/aluminum oxides (von Fromm et al., 2021, 2023). These minerals control SOC persistence through strong yet dynamic sorptive interactions and co-precipitation processes (Doetterl et al., 2015; Rasmussen et al., 2018). However, none of the three process-based models consider clay mineralogy as an input variable. Instead, they rely on clay content or clay + fine silt content as a proxy for mineral-mediated stabilization.

MIMICS's marked overemphasis on clay content (Figures 2 and 3c) suggests that this simplified representation inadequately captures the geochemical complexity of the region's soils. Although the biases with 1:1 and 2:1 clay minerals, as well as with oxalate-extractable Al and Fe, are small (adj $R^2 \leq 0.10$; Figure S11 in Supporting

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Information S1), the notable bias with total clay content (Figure 4c) indicates that clay mineralogy, rather than content alone, may be a more mechanistically appropriate variable for characterizing mineral-mediated SOC stabilization. However, this hypothesis has not been directly tested. Future work should evaluate whether including clay mineralogy as a model input reduces bias and improves performance. Further, investigations into efficient methods to represent clay mineralogy in process-based soil C models would be necessary and useful.

pH-mediated Stabilization and Exchangeable Cations. The systematic bias in all three models with exchangeable Ca (Figure 4b), where the models decompose SOC too rapidly under high concentrations of exchangeable Ca, reveals a missing mechanism related to cation-mediated sorption and organo-mineral interactions. Exchangeable Ca influences mineral surface charge and soil pH; these factors regulate SOC sorption onto mineral surfaces (Rasmussen et al., 2018). Century's largest bias with soil pH (Figure 4a) suggests that pH-dependent processes are inadequately represented; the model overpredicts SOC under low soil pH and underpredicts it under high soil pH. Although Millennial explicitly includes soil pH in its parameterization of mineral-organic interactions, it still shows substantial bias with respect to exchangeable Ca. This suggests that, while the mechanistic representation may be conceptually superior, it does not capture the complexity of pH-dependent sorption dynamics in these soils, potentially because of the static representation of pH in Millennial and other soil C models (Rocci et al., 2024; Zhang et al., 2021). None of the models include exchangeable cations or other ion-specific controls on mineral-organic interactions. This represents a significant structural gap for soils across sub-Saharan Africa, where such interactions are demonstrably important (Asabere et al., 2024; von Fromm et al., 2021).

Moisture Dynamics and Redox Fluctuations. Both Millennial and MIMICS underemphasize soil moisture relative to the observed empirical relationships (Figures 2b and 2c). Soils in humid and subhumid regions across sub-Saharan Africa experience intense wet-dry cycles, which induce recurrent redox oscillations and create transient anoxic microsites, even in upland soils (Keiluweit et al., 2017; Lacroix et al., 2023). These dynamics alter microbial decomposition pathways, favoring fermentation under anoxic conditions and altering mineral-organic associations through changes in the redox state of iron and manganese (Keiluweit et al., 2017). Current models, including the mechanistically advanced Millennial and MIMICS, primarily treat soil moisture as a modulator of decomposition rate, yet they do not explicitly represent redox-driven transformations of organic matter or mineral surfaces. This omission is particularly problematic for tropical soils, as moisture-driven redox cycling may be as important as steady-state moisture conditions in determining the fate of SOC.

Additional processes. Several additional processes absent from the evaluated models likely limit their applicability to sub-Saharan African soils beyond the geochemical mechanisms addressed here. Soil fauna, such as termites, earthworms, and ants, drive bioturbation, creating heterogeneous microsites that expose protected C and re-bury fresh litter (Filser et al., 2016; Jouquet et al., 2016). Similarly, recurrent fires in grasslands and savannas generate pyrogenic C (PyC) that can persist for centuries with distinct decomposition kinetics (Eckmeier et al., 2007; Saiz et al., 2015; Wei et al., 2018); however, none of the models represent PyC as a distinct pool. Finally, soils across sub-Saharan African do not necessarily operate at equilibrium due to residue removal, variable management, and rapid land-use transitions (Valbuena et al., 2015; Vanlauwe et al., 2015). However, we constrained the models to steady-state assumptions due to data limitations. These structural limitations, such as bioturbation, PyC, and non-equilibrium dynamics, are not unique to subtropical soils, but they are particularly consequential in regions that experience frequent disturbance.

5.2. Do Newer Models Address Gaps in Soil Carbon Controls? A Mixed Answer

MIMICS and Millennial represent a significant advancement in mechanistic representation by explicitly including microbial functional traits and mineral-organic interactions, which are absent in Century. However, our results show that these improvements only address some of the structural gaps that limit model performance in sub-Saharan Africa.

Both Millennial and MIMICS substantially improved when fitted to the AfSIS data, with Millennial showing the largest improvement (Figure 1b). Millennial's improvement was primarily driven by correcting the reference temperature for temperature-dependent CUE from 15°C to 19.6°C. This finding suggests that explicit representations of microbial physiology, when parameterized for a given region, can capture important controls. Despite its greater mechanistic complexity, Millennial only required three parameters for fitting, compared to 10 for Century. This suggests that the model's explicit physiological structure provided a better foundation for regional adaptation.

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Despite these improvements, Millennial and MIMICS still show systematic biases with geochemical soil properties (Figure 4). MIMICS' overemphasis on clay content is particularly striking, as it explains over 30% of modeled SOC variance, whereas observations explain less than 10% (Figure 2c). This suggests that, although the framework for representing microbial groups and mineral interactions is conceptually sound, the specific parameterization of sorption mechanisms does not reflect how minerals stabilize C in soils across sub-Saharan Africa. The problem is not that the model includes mineral interactions but rather that the model neglects additional controlling factors, such as clay mineralogy, pH-dependent sorption dynamics, and redox-mediated changes in mineral properties.

Similarly, both Millennial and MIMICS underemphasize soil moisture despite its evident importance in observations. This indicates that the mechanistic representation of moisture-dependent processes, while more explicit than in Century, does not sufficiently capture the influence of moisture in tropical regions. This is likely due to the absence of mechanistic representation of redox dynamics, bioturbation, and other moisture-coupled processes.

5.3. Input Variable Uncertainty as a Model Constraint

Beyond the structural limitations of how models represent processes, the input variables themselves carry inherent uncertainties that likely contribute to model error. The climate and NPP data used to force the models (ERA5-Land reanalysis and MODIS NPP) are generated by intermediate models that involve spatial interpolation and temporal aggregation. This introduces potential biases in regions such as sub-Saharan Africa with sparse observational networks. For example, the NPP data derive from a light-use efficiency model and represent 8-day composites aggregated to annual totals, which can obscure actual temporal variability in C inputs. Similarly, the ERA5-Land soil temperature and moisture data are outputs of an offline land surface model rather than direct observations. In addition, it is challenging to quantify at the site-level how much C is actually entering the soil.

Assuming balanced organic matter inputs and decomposition when running models at steady state is unlikely to hold in systems experiencing global change (von Fromm, Olson, et al., 2025). The discrepancy between the steady-state assumption in our tested models and the dynamic nature of soils across sub-Saharan African likely contributes to the systematic errors in model predictions. Addressing this limitation requires long-term observations of soil C dynamics and input variables at the site-level across sub-Saharan Africa to constrain models in dynamic simulation modes (Le Noë et al., 2023).

A related and underappreciated constraint is that increasing mechanistic complexity only improves predictions if the additional parameters introduced can be adequately constrained by observations. In data-rich temperate systems, decades of monitoring networks and long-term experimental sites have supported model calibration and validation. Sub-Saharan Africa lacks equivalent infrastructure, meaning that added complexity risks amplifying prediction uncertainty rather than reducing it. This creates a compounding challenge: the soils most in need of improved representation are those for which the fewest constraining data exist, suggesting that investments in observational infrastructure and model development must proceed in parallel.

5.4. Implications for Global Carbon Budgets and Climate Projections

Soils across sub-Saharan Africa store approximately 24 Pg C in the top 5 cm alone (Hengl et al., 2015). The moderate performance and systematic biases of all three process-based models have important implications for projections of the continental and global C cycles.

Uncertainty in Regional Carbon Budgets. The adjusted R^2 values ranging from 0.09 to 0.18 across the models indicate substantial prediction uncertainty at regional scales. For a given site, predictions could deviate from observations by 4–5 kg m^{-2} (one RMSE). Scaled across sub-Saharan Africa's ∼1.3 billion hectares, this uncertainty represents billions of tons of C. This uncertainty is problematic because sub-Saharan Africa is targeted for soil C sequestration initiatives and climate mitigation strategies that rely on process-based model projections. Without improved model performance, such initiatives risk overestimating C gains or misallocating resources to regions where models perform poorly.

Systematic Biases in Climate Change Projections. The biases identified here, particularly the overemphasis on NPP and the underrepresentation of mineralogical controls, are not random errors that average out at large scales. Rather, they are directional biases that will systematically affect climate projections. For example, models that

overweight C input sensitivity and underweight mineral-mediated stabilization may overestimate SOC losses during droughts, which reduce NPP, while underestimating SOC responses to management practices that alter soil mineral properties or pH. These biases could lead to incorrect projections of climate-carbon feedback with consequences for global climate change assessments and policy decisions based on them. In addition, our analysis was necessarily limited to topsoils (0–20 cm), the layer most directly influenced by recent plant inputs. In these deeply weathered systems, however, a substantial and often dominant portion of the total profile C stock resides in subsoils, where stabilization is governed even more strongly by the mineralogical and geochemical controls that our study identifies as poorly represented in current models (von Fromm, Jungkunst, et al., 2025). Therefore, the structural biases we report, particularly the overemphasis on NPP and inadequate representation of oxide and cation-mediated stabilization, are likely to be amplified when predicting full-profile C stocks. This suggests that improving models for tropical C budgets will require not only the incorporation of region-specific mechanisms but also their depth-explicit application and evaluation against deeper soil data sets/whole-profile soil data sets.

Implications Beyond Sub-Saharan Africa. Although this study focuses on sub-Saharan Africa, the findings likely apply to tropical and subtropical soils worldwide. Process-based models were primarily developed in temperate regions and applied globally with minimal regional adaptation. The structural mismatches identified here (i.e., the inadequate representation of clay mineralogy, and pH-dependent sorption) are not unique to sub-Saharan Africa, but rather, characterize highly weathered tropical soils worldwide. Therefore, similar model inadequacies likely affect projections for the Amazon, Southeast Asia, and other tropical regions. This suggests that current assessments may systematically underestimate global C cycle uncertainty.

Opportunities for Improved Regional Representation. Despite these limitations, the finding that mechanistically complex models (Millennial and MIMICS) can substantially improve when fitted to regional data is hopeful. Systematic re-parameterization efforts leveraging regional soil and microbial data sets, as demonstrated by recent work in Kenya (Laub et al., 2024), could substantially improve model performance. However, this requires a coordinated effort to: (a) expand regional soil property data sets, (b) develop region-specific parameterizations, and (c) validate models against independent data. These efforts would improve regional projections and provide valuable insights into the critical mechanisms for SOC dynamics across diverse soils. Alternatively, parameterizing at smaller scales, such as the soil profile scale, may significantly improve soil C predictions (Tao et al., 2024) but is limited by reduced data availability in tropical and subtropical regions.

6. Conclusions

This study evaluated three process-based soil C models (Century, Millennial, and MIMICS) with increasing degrees of mechanistic complexity across soils in sub-Saharan Africa. Despite their substantial differences in complexity, all three models exhibited similar performance (adj. R^2 = 0.09–0.18) and converged on similar biases. This indicates that increased mechanistic complexity does not necessarily lead to improved transferability. Rather, model performance fundamentally depends on whether the chosen mechanisms accurately represent the processes operating in the target region. In other words, mechanistic complexity will likely only be beneficial if the right mechanisms are included.

We identified discrepancies between current process-based models and the pedogenic and biogeochemical conditions of soils in sub-Saharan Africa. The tested models overemphasize C inputs and inadequately represent mineral-mediated stabilization mechanisms, which include clay mineralogy and pH-dependent sorption dynamics. Additionally, the models omit important processes, including exchangeable cations, redox-driven C transformations, bioturbation-mediated C cycling, pyC persistence, and the socio-ecological drivers of non-equilibrium soil conditions. These gaps are not simple parameterization issues, but rather, they reflect fundamental structural limitations in how current models conceptualize soil C cycling.

Newer-generation models, such as Millennial and MIMICS, that explicitly represent microbial physiology and mineral-organic interactions, represent conceptual advances and may show promise for regional adaptation when parameterized with local data. However, our results demonstrate that mechanistic detail is valuable only when the chosen mechanisms are relevant and accurately parameterized for the target system. Despite its sophisticated mineral-interaction framework, MIMICS's overemphasis on clay content illustrates this principle: a more complex mechanism is not inherently superior if it misrepresents regional soil geochemistry. Therefore, advancing soil C modeling in sub-Saharan Africa and other underrepresented pedological regions requires model

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

development tailored to regionally important soil C mechanisms, as well as target field sampling and laboratory analysis to ensure the necessary data is present to constrain the (additional) model parameters.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

All data and code needed to reproduce the presented results can be accessed via github (https://github.com/katierocci/SSA_soils_workshop_models) and is published in (von Fromm et al., 2026).

Acknowledgments
This research was supported by a workshop held at Dartmouth College, which was funded by the William H. Neukom 1964 Institute for Computational Science, the Department of Geography, the Department of Biological Sciences, and the Associate Dean of the Faculty, Science Division.

References

Abramoff, R. Z., Guenet, B., Zhang, H., Georgiou, K., Xu, X., Rossel, R. A. V., et al. (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

Abramoff, R. Z., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E., et al. (2018). The millennial model: In search of measurable pools and transformations for modeling soil carbon in the new century. *Biogeochemistry*, *137*(1), 51–71. https://doi.org/10.1007/s10533-017-0409-7

Arrouays, D., Grundy, M. G., Hartemink, A. E., Hempel, J. W., Heuvelink, G. B. M., Hong, S. Y., et al. (2014). Chapter Three—GlobalSoilMap: Toward a fine-resolution global grid of soil properties. In D. L. Sparks (Ed.), *Advances in agronomy* (Vol. 125, pp. 93–134). Academic Press. https://doi.org/10.1016/B978-0-12-800137-0.00003-0

Asabere, S. B., Abrefa Nketia, K., Iddris, N. A., Ashiagbor, G., & Sauer, D. (2024). Nitrogen enrichment does not adversely affect exchangeable bases in rainfed urban arable soils of Kumasi, a tropical West African city. *Geoderma*, *444*, 116864. https://doi.org/10.1016/j.geoderma.2024.116864

Auguie, B. (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics (Version 2.3) [Computer software]. https://CRAN.R-project.org/package=gridExtra

Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in R using futures. *The R Journal*, *13*(2), 208. https://doi.org/10.32614/RJ-2021-048

Bosatta, E., & Agren, G. I. (1991). Dynamics of carbon and nitrogen in the organic matter of the soil: A generic theory. *The American Naturalist*, *138*(1), 227–245. https://doi.org/10.1086/285213

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A., & Crowther, T. W. (2016). Managing uncertainty in soil carbon feedbacks to climate change. *Nature Climate Change*, *6*(8), 751–758. https://doi.org/10.1038/nclimate3071

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Couëdel, A., Laub, M., Ranaivomanana, R., Falconnier, G. N., Cardinael, R., Mucheru-Muna, M. W., et al. (2026). Evaluating DayCent and STICS in simulating the long-term impact of contrasting organic resource amendments on soil organic carbon and maize yields in Sub-Saharan Africa. *Field Crops Research*, *335*, 110169. https://doi.org/10.1016/j.fcr.2025.110169

Dintwe, K., & Okin, G. S. (2018). Soil organic carbon in savannas decreases with anthropogenic climate change. *Geoderma*, *309*, 7–16. https://doi.org/10.1016/j.geoderma.2017.08.030

Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Casanova Pinto, M., et al. (2015). Soil carbon storage controlled by interactions between geochemistry and climate. *Nature Geoscience*, *8*(10), 780–783. https://doi.org/10.1038/ngeo2516

Dokuchaev, V. V. (1883). Russkii chernozem. *Free Economic Society*.

Eckmeier, E., Gerlach, R., Skjemstad, J. O., Ehrmann, O., & Schmidt, M. W. I. (2007). Minor changes in soil organic carbon and charcoal concentrations detected in a temperate deciduous forest a year after an experimental slash-and-burn. *Biogeosciences*, *4*(3), 377–383. https://doi.org/10.5194/bg-4-377-2007

FAO. (2022). Global soil organic carbon sequestration potential map – GSOCseq v.1.1. https://doi.org/10.4060/cb9002en

Filser, J., Faber, J. H., Tiunov, A. V., Brussaard, L., Frouz, J., De Deyn, G., et al. (2016). Soil fauna: Key to new carbon models. *SOIL*, *2*(4), 565–582. https://doi.org/10.5194/soil-2-565-2016

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., et al. (2025). Global Carbon Budget 2024. *Earth System Science Data*, *17*(3), 965–1039. https://doi.org/10.5194/essd-17-965-2025

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1189–1232. https://doi.org/10.1214/aos/1013203451

Garsia, A., Moinet, A., Vazquez, C., Creamer, R. E., & Moinet, G. Y. K. (2023). The challenge of selecting an appropriate soil organic carbon simulation model: A comprehensive global review and validation assessment. *Global Change Biology*, *29*(20), 5760–5774. https://doi.org/10.1111/gcb.16896

Georgiou, K., Abramoff, R. Z., Harte, J., Riley, W. J., & Torn, M. S. (2017). Microbial community-level regulation explains soil carbon responses to long-term litter manipulations. *Nature Communications*, *8*(1), 1223. https://doi.org/10.1038/s41467-017-01116-z

Geremew, B., Tadesse, T., Bedadi, B., Gollany, H. T., Tesfaye, K., Aschalew, A., et al. (2024). Evaluation of RothC model for predicting soil organic carbon stock in north-west Ethiopia. *Environmental Challenges*, *15*, 100909. https://doi.org/10.1016/j.envc.2024.100909

Gottschalk, P., Smith, J. U., Wattenbach, M., Bellarby, J., Stehfest, E., Arnell, N., et al. (2012). How will organic carbon stocks in mineral soils evolve under future climate? Global projections using RothC for a range of climate change scenarios. *Biogeosciences*, *9*(8), 3151–3171. https://doi.org/10.5194/bg-9-3151-2012

Hengl, T., Heuvelink, G. B. M., Kempen, B., Leenaars, J. G. B., Walsh, M. G., Shepherd, K. D., et al. (2015). Mapping soil properties of Africa at 250 m resolution: Random forests significantly improve Current predictions. *PLoS One*, *10*(6), e0125814. https://doi.org/10.1371/journal.pone.0125814

Hengl, T., Miller, M. A. E., Križan, J., Shepherd, K. D., Sila, A., Kilibarda, M., et al. (2021). African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Scientific Reports*, *11*(1), 6130. https://doi.org/10.1038/s41598-021-85639-y

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hufkens, K., Stauffer, R., & Campitelli, E. (2019). The ecwmfr package: An interface to ECMWF API endpoints [Computer software]. https://bluegreen-labs.github.io/ecmwfr/

Jenny, H. (1941). *Factors of soil Formation: A System of quantitative pedology*. McGraw-Hill. Retrieved from https://books.google.com/books?id=h-dIAAAAMAAJ

Jouquet, P., Chintakunta, S., Bottinelli, N., Subramanian, S., & Caner, L. (2016). The influence of fungus-growing termites on soil macro and micro-aggregates stability varies with soil type. *Applied Soil Ecology*, *101*, 117–123. https://doi.org/10.1016/j.apsoil.2016.02.001

Kamoni, P. T., Gicheru, P. T., Wokabi, S. M., Easter, M., Milne, E., Coleman, K., et al. (2007). Evaluation of two soil carbon models using two Kenyan long term experimental datasets. *Soil Carbon Stocks at Regional Scales*, *122*(1), 95–104. https://doi.org/10.1016/j.agee.2007.01.011

Kassambara, A. (2023). Ggpubr: "ggplot2" Based Publication Ready Plots (Version R package version 0.6.0) [Computer software]. https://CRAN.R-project.org/package=ggpubr

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., et al. (2020). TRY plant trait database – Enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. https://doi.org/10.1111/gcb.14904

Keiluweit, M., Wanzek, T., Kleber, M., Nico, P., & Fendorf, S. (2017). Anaerobic microsites have an unaccounted role in soil carbon stabilization. *Nature Communications*, *8*(1), 1771. https://doi.org/10.1038/s41467-017-01406-6

Kennedy, D., Dagon, K., Lawrence, D. M., Fisher, R. A., Sanderson, B. M., Collier, N., et al. (2025). One-at-a-Time parameter perturbation ensemble of the community land model, version 5.1. *Journal of Advances in Modeling Earth Systems*, *17*(8), e2024MS004715. https://doi.org/10.1029/2024MS004715

Lacroix, E. M., Aeppli, M., Boye, K., Brodie, E., Fendorf, S., Keiluweit, M., et al. (2023). Consider the anoxic microsite: Acknowledging and appreciating spatiotemporal redox heterogeneity in soils and sediments. *ACS Earth and Space Chemistry*, *7*(9), 1592–1609. https://doi.org/10.1021/acsearthspacechem.3c00032

Lang, M., Binder, M., Richter, J., Schratz, P., Pfisterer, F., Coors, S., et al. (2019). mlr3: A modern object-oriented machine learning framework in R. *Journal of Open Source Software*, *4*(44), 1903. https://doi.org/10.21105/joss.01903

Laub, M., Necpalova, M., Van de Broek, M., Corbeels, M., Ndungu, S. M., Mucheru-Muna, M. W., et al. (2024). Modeling integrated soil fertility management for maize production in Kenya using a Bayesian calibration of the DayCent model. *Biogeosciences*, *21*(16), 3691–3716. https://doi.org/10.5194/bg-21-3691-2024

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

Le Noë, J., Manzoni, S., Abramoff, R., Bölscher, T., Bruni, E., Cardinael, R., et al. (2023). Soil organic carbon models need independent time-series validation for reliable prediction. *Communications Earth and Environment*, *4*(1), 158. https://doi.org/10.1038/s43247-023-00830-5

Liu, M., Zheng, S., Pendall, E., Smith, P., Liu, J., Li, J., et al. (2025). Unprotected carbon dominates decadal soil carbon increase. *Nature Communications*, *16*(1), 2008. https://doi.org/10.1038/s41467-025-57354-z

Ma, J., Rabin, S. S., Anthoni, P., Bayer, A. D., Nyawira, S. S., Olin, S., et al. (2022). Assessing the impacts of agricultural managements on soil carbon stocks, nitrogen loss, and crop production – A modelling study in eastern Africa. *Biogeosciences*, *19*(8), 2145–2169. https://doi.org/10.5194/bg-19-2145-2022

Molnar, C. (2025). *Interpretable machine learning: A guide for making black box models explainable* (3rd ed.). Christoph Molnar.

Molnar, C., Bischl, B., & Casalicchio, G. (2018). Iml: An R package for Interpretable Machine Learning. *Journal of Open Source Software*, *3*(26), 786. https://doi.org/10.21105/joss.00786

Müller, K. (2020). Here: A simpler way to find your Files (Version 1.0.1) [Computer software]. https://CRAN.R-project.org/package=here

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nyawira, S. S., Hartman, M. D., Nguyen, T. H., Margenot, A. J., Kihara, J., Paul, B. K., et al. (2021). Simulating soil organic carbon in maize-based systems under improved agronomic management in Western Kenya. *Soil and Tillage Research*, *211*, 105000. https://doi.org/10.1016/j.still.2021.105000

Ojima, D. S., Dirks, B. O. M., Glenn, E. P., Owensby, C. E., & Scurlock, J. O. (1993). Assessment of C budget for grasslands and drylands of the world. *Water, Air, and Soil Pollution*, *70*(1), 95–109. https://doi.org/10.1007/BF01104990

Parton, W. J., Schimel, D. S., Cole, C. V., & Ojima, D. S. (1987). Analysis of factors controlling soil organic matter levels in Great Plains grasslands. *Soil Science Society of America Journal*, *51*(5), 1173–1179. https://doi.org/10.2136/sssaj1987.03615995005100050015x

Pierson, D., Lohse, K. A., Wieder, W. R., Patton, N. R., Facer, J., de Graaff, M.-A., et al. (2022). Optimizing process-based models to predict current and future soil organic carbon stocks at high-resolution. *Scientific Reports*, *12*(1), 10824. https://doi.org/10.1038/s41598-022-14224-8

Rasmussen, C., Heckman, K., Wieder, W. R., Keiluweit, M., Lawrence, C. R., Berhe, A. A., et al. (2018). Beyond clay: Towards an improved set of variables for predicting soil organic matter content. *Biogeochemistry*, *137*(3), 297–306. https://doi.org/10.1007/s10533-018-0424-3

R Core Team. (2024). R: A Language and Environment for Statistical Computing (Version R version 4.4.1) [Computer software]. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rocci, K. S., Cotrufo, M. F., Ernakovich, J., Foster, E., Frey, S., Georgiou, K., et al. (2024). Bridging 20 years of soil organic matter frameworks: Empirical support, model representation, and next steps. *Journal of Geophysical Research: Biogeosciences*, *129*(6), e2023JG007964. https://doi.org/10.1029/2023JG007964

Rocci, K. S., Pierson, D., Jevon, F. V., Polussa, A., Oliverio, A. M., Bradford, M. A., et al. (2025). Integrating microbial community data into an ecosystem-scale model to predict litter decomposition in the face of climate change. *Global Change Biology*, *31*(7), e70352. https://doi.org/10.1111/gcb.70352

Running, S., & Zhao, M. (2021). MODIS/Terra net primary production gap-filled yearly L4 Global 500m SIN grid V061 [Dataset]. *NASA Land Processes Distributed Active Archive Center*. https://doi.org/10.5067/MODIS/MOD17A3HGF.061

Saiz, G., Wynn, J. G., Wurster, C. M., Goodrick, I., Nelson, P. N., & Bird, M. I. (2015). Pyrogenic carbon from tropical savanna burning: Production and stable isotope composition. *Biogeosciences*, *12*(6), 1849–1863. https://doi.org/10.5194/bg-12-1849-2015

Sakala, D., & Santos, M. J. (2025). Projected land-use change emissions surpass climate change-induced carbon sinks in Sub-Saharan African biomes. *Global Environmental Change*, *94*, 103039. https://doi.org/10.1016/j.gloenvcha.2025.103039

Shi, Z., Allison, S. D., He, Y., Levine, P. A., Hoyt, A. M., Beem-Miller, J., et al. (2020). The age distribution of global soil carbon inferred from radiocarbon measurements. *Nature Geoscience*, *13*(8), 555–559. https://doi.org/10.1038/s41561-020-0596-z

Six, J., Conant, R. T., Paul, E. A., & Paustian, K. (2002). Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils. *Plant and Soil*, *241*(2), 155–176. https://doi.org/10.1023/A:1016125726789

Smith, P., Smith, J. U., Franko, U., Kuka, K., Romanenkov, V. A., Shevtsova, L. K., et al. (2007). Changes in mineral soil organic carbon stocks in the croplands of European Russia and the Ukraine, 1990–2070; comparison of three models and implications for climate mitigation. *Regional Environmental Change*, *7*(2), 105–119. https://doi.org/10.1007/s10113-007-0028-2

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Soetaert, K., & Petzoldt, T. (2010). Inverse modelling, sensitivity and Monte Carlo analysis in R using package FME. *Journal of Statistical Software*, *33*(3), 1–28. https://doi.org/10.18637/jss.v033.i03

Tan, Z., Tieszen, L. L., Liu, S., & Tachie-Obeng, E. (2010). Modeling to evaluate the response of savanna-derived cropland to warming–drying stress and nitrogen fertilizers. *Climatic Change*, *100*(3), 703–715. https://doi.org/10.1007/s10584-009-9688-x

Tao, F., Houlton, B. Z., Huang, Y., Wang, Y.-P., Manzoni, S., Ahrens, B., et al. (2024). Convergence in simulating global soil organic carbon by structurally different models after data assimilation. *Global Change Biology*, *30*(5), e17297. https://doi.org/10.1111/gcb.17297

Torn, M. S., Trumbore, S. E., Chadwick, O. A., Vitousek, P. M., & Hendricks, D. M. (1997). Mineral control of soil organic carbon storage and turnover. *Nature*, *389*(6647), 170–173. https://doi.org/10.1038/38260

Vågen, T.-G., Winowiecki, L. A., Desta, L., Tondoh, J., Weullow, E., Shepherd, K., et al. (2021). Wet chemistry data for a subset of AfSIS: Phase I archived soil samples (version 1.2) [Dataset]. *World Agroforestry (ICRAF)*. https://doi.org/10.34725/DVN/66BFOB

Vågen, T.-G., Winowiecki, L. A., Tondoh, J. E., Desta, L. T., & Gumbricht, T. (2016). Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. *Geoderma*, *263*, 216–225. https://doi.org/10.1016/j.geoderma.2015.06.023

Valbuena, D., Tui, S. H.-K., Erenstein, O., Teufel, N., Duncan, A., Abdoulaye, T., et al. (2015). Identifying determinants, pressures and trade-offs of crop residue use in mixed smallholder farms in Sub-Saharan Africa and South Asia. *Biomass Use Trade-Offs in Cereal Cropping Systems: Lessons and Implications from the Developing World*, *134*, 107–118. https://doi.org/10.1016/j.agsy.2014.05.013

Vanlauwe, B., Descheemaeker, K., Giller, K. E., Huising, J., Merckx, R., Nziguheba, G., et al. (2015). Integrated soil fertility management in Sub-Saharan Africa: Unravelling local adaptation. *SOIL*, *1*(1), 491–508. https://doi.org/10.5194/soil-1-491-2015

von Fromm, S. F., Doetterl, S., Butler, B. M., Aynekulu, E., Berhe, A. A., Haefele, S. M., et al. (2023). Controls on timescales of soil organic carbon persistence across Sub-Saharan Africa. *Global Change Biology*, *30*(1), e17089. https://doi.org/10.1111/gcb.17089

von Fromm, S., Rocci, K., & Abramoff, R. (2026). Data and code: Evaluating soil carbon models for Sub-Saharan Africa: Revealing knowledge gaps in subtropical and tropical soil biogeochemistry (Version v1.0.0) [Computer software]. *Zenodo*. https://doi.org/10.5281/ZENODO.18249449

von Fromm, S. F., Hoyt, A. M., Lange, M., Acquah, G. E., Aynekulu, E., Berhe, A. A., et al. (2021). Continental-scale controls on soil organic carbon across Sub-Saharan Africa. *SOIL*, *7*(1), 305–332. https://doi.org/10.5194/soil-7-305-2021

von Fromm, S. F., Hoyt, A. M., Sierra, C. A., Georgiou, K., Doetterl, S., & Trumbore, S. E. (2024). Controls and relationships of soil organic carbon abundance and persistence vary across pedo-climatic regions. *Global Change Biology*, *30*(5), e17320. https://doi.org/10.1111/gcb.17320

von Fromm, S. F., Jungkunst, H. F., Amenkhienan, B., Hall, S. J., Georgiou, K., Pries, C. H., et al. (2025). Moisture and soil depth govern relationships between soil organic carbon and oxalate-extractable metals at the global scale. *Biogeochemistry*, *168*(1), 20. https://doi.org/10.1007/s10533-025-01208-9

von Fromm, S. F., Olson, C. I., Monroe, M. D., Sierra, C. A., Driscoll, C. T., Groffman, P. M., et al. (2025). Temporal and spatial dynamics of soil carbon cycling and its response to environmental change in a Northern hardwood Forest. *Global Change Biology*, *31*(5), e70250. https://doi.org/10.1111/gcb.70250

Wang, L., Abramowitz, G., Wang, Y.-P., Pitman, A., Ciais, P., & Goll, D. S. (2025). Using explainable AI to diagnose the representation of environmental drivers in process-based soil organic carbon models. *Biogeosciences*, *22*(23), 7845–7863. https://doi.org/10.5194/bg-22-7845-2025

Wei, X., Hayes, D. J., Fraver, S., & Chen, G. (2018). Global pyrogenic carbon production during recent decades has created the potential for a large, long-term sink of atmospheric CO2. *Journal of Geophysical Research: Biogeosciences*, *123*(12), 3682–3696. https://doi.org/10.1029/2018JG004490

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., & Bonan, G. B. (2014). Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model. *Biogeosciences*, *11*(14), 3899–3917. https://doi.org/10.5194/bg-11-3899-2014

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., Taylor, P. G., & Bonan, G. B. (2015). Representing life in the Earth system with soil microbial functional traits in the MIMICS model. *Geoscientific Model Development*, *8*(6), 1789–1808. https://doi.org/10.5194/gmd-8-1789-2015

Zhang, L., Yang, L., Crowther, T. W., Zohner, C. M., Doetterl, S., Heuvelink, G. B. M., et al. (2025). Mapping global distributions, environmental controls, and uncertainties of apparent topsoil and subsoil organic carbon turnover times. *Earth System Science Data*, *17*(6), 2605–2623. https://doi.org/10.5194/essd-17-2605-2025

Zhang, Y., Lavallee, J. M., Robertson, A. D., Even, R., Ogle, S. M., Paustian, K., & Cotrufo, M. F. (2021). Simulating measurable ecosystem carbon and nitrogen dynamics with the mechanistically defined MEMS 2.0 model. *Biogeosciences*, *18*(10), 3147–3171. https://doi.org/10.5194/bg-18-3147-2021

21698961, 2026, 5, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026JG009726 by EBMG ACCESS - ZIMBABWE, Wiley Online Library on [18/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License