Supporting Information for "Radiocarbon analysis reveals underestimation of soil organic carbon persistence in new-generation soil models"

Alexander S. Brunmayr[1], Frank Hagedorn[2], Margaux Moreno Duborgel[2,3], Luisa I. Minich[2,3], Heather D. Graven[1]

[1]Imperial College London, Department of Physics

[2]Eidgenössische Forschungsanstalt WSL

[3]ETH Zurich, Department of Earth Sciences

Contents of this file

Introduction

This document provides further information on the specific model versions and implementations used in this study (section S1, Figures S1-S5), and specifies which simulated pools were associated to which measured soil fraction (section S2, Table S1). It also explains how we re-implemented non-isotopic models with ^{14}C (section S3), and why the ^{14}C implementation of SOMic and the newest version of MIMICS are incorrect (section S4, Figures S6-S7). Section S5 gives some more details on Millennial's turnover times. Finally, Figures S8-S12 at the end of this document show plots of model predictions against observations for each model.

S1. Further information on model versions and implementations

The source codes of all the selected model versions are openly available. By having direct access to the code with which the model developers produced their results, we can be more confident that we test an implementation of the models as intended by their respective authors.

Our final implementations of the Millennial, CORPSE, and MIMICS models are available as python modules on our GitHub repository `https://github.com/asb219/evaluate-SOC-models`. Our slightly modified implementation of the MEND model in `https://github.com/asb219/MEND` is added to our repository as a git submodule. Finally, we installed the SOMic model's R package directly from our forked `https://github.com/asb219/somic1` GitHub repository.

S1.1. Millennial

We use Millennial V2 with Michaelis-Menten kinetics as described in Abramoff et al. (2022). We re-implemented the model with ^{14}C in Python based on the original R code in the `https://github.com/rabramoff/Millennial` repository under the tag "v2" (commit `e95bca9` from September 2021). We used the model equations from file `R/models/derivs_V2_MM.R` in the repository and ran the model with the fitted parameter values from the file `Fortran/MillennialV2/simulationv2/soilpara_in_fit.txt` in the repository. The initial condition for both carbon and ^{14}C stocks is found by first solving for a pre-industrial steady state, similarly to the model tutorial `R/simulation/model_tutorial.Rmd` in the repository, and then running the model from steady state for 200 years using time-varying pre-industrial forcing data featuring a seasonal cycle. The final state of that spinup is

then used as the initial condition for the final run of the model over the 1850-2014 period. The model runs with daily time steps, and though the model tutorial uses the 4th order Runge-Kutta integration method, we integrate the equations simply with the forward Euler method, which is stable and precise enough with daily time steps.

S1.2. CORPSE

The CORPSE model was originally described in Sulman, Phillips, Oishi, Shevliakova, and Pacala (2014). There are currently at least six publicly available versions of CORPSE. Since we are mostly interested in carbon dynamics, the lead developer Benjamin Sulman recommended we use the most up-to-date carbon-only implementation in `https://github.com/bsulman/CORPSE-fire-response` (latest commit at time of writing: `19ee2c7` from February 2021). We reimplemented CORPSE with ^{14}C based on the equations in file `CORPSE_array.py` and using the parameter values from file `Whitman_sims.py` in that repository. However, the equation for the clay-related rate modifying factor is taken from file `code/CORPSE_integrate.py` in repository `https://github.com/bsulman/CORPSE-N`, since the model seems to be working more reliably with that version of the equation. Like in Millennial, the initial condition is found by solving for a pre-industrial steady state and spinning up for 200 years from that steady state. If the solver is unable to find a steady state, the model is spun up for 4000 years. The model runs with daily time steps and uses the forward Euler integration method.

S1.3. SOMic

We use version 1.0 of the SOMic model as described in (Woolf & Lehmann, 2019). The original code is available on the GitHub repository `https://github.com/domwoolf/`

somic1 (hash of latest commit at time of writing: be34e56 from June 2019). However, we forked the repository to https://github.com/asb219/somic1 in order to fix a minor issue in its ^{14}C implementation (see reason in section S4.1), and used the version released under the tag "v1.1-asb219" to produce our results. We spin up the model for 5000 years to get the initial carbon and ^{14}C stocks. The model runs with monthly time steps and uses the forward Euler integration method.

S1.4. MEND

We use the latest version of the default MEND model with carbon-nitrogen coupling as described in G. Wang et al. (2022). Our ^{14}C re-implementation is based on the code from commit 92323c7 (from February 2022) of the GitHub repository https://github.com/wanggangsheng/MEND. We forked the repository from that commit to https://github.com/asb219/MEND so we could adapt the model input and output to our purposes. We use all the default model settings and the optimized parameter values provided in the Fortran namelist file MEND_namelist.nml in the repository. The pre-industrial soil carbon and nitrogen stocks are found by initializing the model with the default initial state from file userio/inp/SOIL_ini.dat and spinning up for 400 year with pre-industrial forcing data. The pre-industrial soil ^{14}C levels are found by running the spun-up model for another 1000 years with pre-industrial forcing data. The model runs with hourly time steps and uses the forward Euler integration method.

S1.5. MIMICS

We use the MIMICS-CN v1.0 model, as published in (Kyker-Snowman et al., 2020), because the latest version of MIMICS (Y. Wang et al., 2021) did not correctly imple-

ment ^{14}C (see section SS4.2). The original R code of MIMICS-CN v1.0 is available on `https://zenodo.org/records/3534562`. It already implements stable isotope tracers, but no radioactive isotopes such as ^{14}C, so we re-implemented the model with ^{14}C in python. Like for Millennial and CORPSE, we spin up for 200 years from the pre-industrial steady-state solution. If no steady state can be found, we spin up for 4000 years. The model runs with hourly time steps and uses the forward Euler integration method.

S2. Correspondences between pools and measurable fractions

This section explains how we associate the simulated pools of each model with either the *POM* (particulate organic matter) fraction or the *MAOM* (mineral-associated organic matter) fraction. See Table S1 for a summary of the correspondences between the modeled pools and the *POM* and *MAOM* fractions.

We assume that the *POM* fraction (corresponding to the "light fraction" resulting from density fractionation) is composed of fragmented and partially processed plant litter which is not stabilized in the soil matrix through mineral association. We assume that the *MAOM* fraction (corresponding to the "heavy fraction" resulting from density fractionation) is composed of soil organic carbon which is enclosed in stable aggregates or strongly adsorbed to minerals. Since the live microbial biomass and dissolved organic carbon generally represent a small fraction of soil organic carbon, we can neglect them, so we assume they belong to neither *POM* nor *MAOM*.

S2.1. MEND

We assume that the *POM* fraction is composed of the POM_O and POM_H pools, and that the *MAOM* fraction is composed of the MOM and QOM pools.

List of organic carbon pools in the MEND-new (2022) model by G. Wang et al. (2022) (also see Figure S1):

- POM_O and POM_H (particulate organic matter decomposed by oxidative and hydrolytic enzymes, respectively).
- MOM (mineral-associated organic matter).
- QOM: "active layer of MOM" which can exchange carbon with DOM through adsorption and desorption (G. Wang et al., 2022).

- DOM (dissolved organic matter).
- MB_A and MB_D (active and dormant microbial biomass, respectively).
- EP_O, EP_H, EM: various microbial exo-enzymes.

Note that the "Litter" pool in the MEND model diagram in Figure S1 is not explicitly modeled as a pool, and therefore does not feature in the above list of organic carbon pools.

S2.2. Millennial

We assume that the measured *MAOM* fraction is the sum of the Aggregate C and MAOM pools, and that the *POM* fraction is entirely composed of the POM pool.

List of organic carbon pools in Millennial v2 by Abramoff et al. (2022) (see also Figure S2):

- POM (particulate organic carbon).
- Aggregate C: "stable microaggregates which remain after dispersion in the larger particle size fraction (>50–60 μm)" (Abramoff et al., 2022), so this corresponds to the coarse heavy fraction.
- MAOM (mineral-associated organic carbon): consists of organic matter associated to minerals through sorption (Abramoff et al., 2022).
- Microbial Biomass: live microbial biomass.
- LMWC (low molecular weight carbon): "LMWC could be analogous to dissolved organic C (DOC) if there is enough moisture in the soil matrix, and if we do not consider DOC molecules that are too large to be taken up by microbes" (Abramoff et al., 2022).

S2.3. SOMic

The *MAOM* fraction is composed of the MAC pool, and the *POM* fraction is composed of the SPM and IPM pools.

List of organic carbon pools in SOMic 1.0 by Woolf and Lehmann (2019) (also see Figure S3):

- SPM and IPM (soluble and insoluble plant matter, respectively).
- MAC (mineral-associated carbon): "mineral-sorbed or -occluded SOC" (Woolf & Lehmann, 2019).
- DOC (dissolved organic carbon).
- MB (microbial biomass).

S2.4. CORPSE

We associate the *MAOM* fraction with the SPC_p, CPC_p, and MN_p pools, since they represent mineral-adsorbed and micro-aggregated carbon (Moore et al., 2020). We associate the *POM* fraction with the SPC_u and CPC_u pools, but not the microbial MN_u pool, because POM is mostly composed of unprotected plant-derived carbon.

List of organic carbon pools in the CORPSE-fire-response version of the CORPSE model, first published in Sulman et al. (2014) and last updated in Moore et al. (2020) (see also Figure S4):

- SPC_u, CPC_u, and MN_u (Unprotected simple plant carbon, Unprotected complex plant carbon, and Unprotected microbe necromass, respectively).
- SPC_p, CPC_p, and MN_p (Protected simple plant carbon, Protected complex plant carbon, and Protected microbe necromass): "protected organic matter is inaccessible to microbial

decomposition through chemical sorption to mineral surfaces or occlusion within microaggregates" (Moore et al., 2020).

- LMB (live microbial biomass).

S2.5. MIMICS

According to Kyker-Snowman et al. (2020), the SOM_c pool corresponds to the *POM* fraction, and the SOM_p pool corresponds to the *MAOM* fraction.

List of organic carbon pools in MIMICS-CN v1.0 by Kyker-Snowman et al. (2020) (see also Figure S5):

- LIT_m and LIT_s (metabolic and structural litter, respectively): litter pools which are not considered part of soil organic matter.
- SOM_p (physicochemically protected soil organic matter): "is primarily composed of microbial products that are adsorbed onto mineral surfaces" and is "analogous to heavy fraction or MAOM pools" (Kyker-Snowman et al., 2020).
- SOM_c (chemically recalcitrant soil organic matter): "consists of decomposed or partially decomposed litter" and is "analogous to light fraction or POM pools" (Kyker-Snowman et al., 2020).
- SOM_a (available soil organic matter): "the only SOM pool that is available for microbial decomposition; it contains a mixture of fresh microbial residues, products that are desorbed from the SOMp pool (e.g., Jilling et al., 2018), as well as depolymerized organic matter from the SOMc pool" (Kyker-Snowman et al., 2020). This pool is usually very small and we associate it to neither POM nor MAOM.

- MIC_r and MIC_K ("low-efficiency, r strategist" microbes and "high-efficiency, K strategist" microbes, respectively): live microbial biomass pools.

Note that MIMICS-CN v1.0 also has a Dissolved Inorganic Nitrogen (DIN) pool, which does not contain organic carbon.

S3. Radiocarbon predictions with non-isotopic models

Among the new-generation models selected for this study, SOMic, MIMICS, and MEND have already implemented ^{14}C. However, the most recent and only open-source version of MEND does not include ^{14}C, and SOMic and MIMICS incorrectly implemented their ^{14}C simulations (see section S4). Nevertheless, we can still produce ^{14}C predictions with non-isotopic models by individually tracking the carbon fluxes at every time step and attaching a ^{14}C signal to each flux. Since none of the models define an internal structure for their pools, we will assume by default that the pools are well-mixed, which means that the $\Delta^{14}C$ of a pool's outflux is equal to the pool's $\Delta^{14}C$. This assumption is common practice for ^{14}C modeling in soils (Sierra et al., 2017).

We run all of the selected models using the forward Euler method to advance from one time step to the next. The models either implicitly or explicitly produce the internal flux matrix Φ^i at each time step i, where $\Phi^i_{jk} \geq 0$ is the flux of carbon from pool k into pool j (with $j \neq k$), and $\Phi^i_{jj} \leq 0$ is the total outflux of carbon out of pool j at time step i. They also define the external influx vector I^i such that $I^i_j \geq 0$ is the influx of carbon entering the modeled system through pool j at time step i. Matrix Φ contains all the fluxes between the pools and out of the system, and vector I contains all the influxes of carbon from outside the system into the modeled pools. We can therefore find the carbon stocks C^{i+1}_j of pool j at time step $i+1$ based on the Φ^i, I^i, and C^i_j of the previous time step i:

$$C^{i+1}_j = C^i_j + I^i_j + \sum_k \Phi^i_{jk}\,, \tag{S1}$$

where the summation of internal fluxes Φ^i_{jk} is performed over all donor pools k to get the total internal carbon flux into pool j (when $k \neq j$), subtracted by the flux out of pool j (when $k = j$).

Assuming the pools are well-mixed, we can now produce ^{14}C predictions by tagging each flux Φ_{jk} with the ^{14}C signal of pool k. We measure the ^{14}C signal in terms of the unitless "absolute Fraction Modern" ($\mathrm{FM_{abs}}$) as defined in Trumbore, Sierra, and Hicks Pries (2016), such that $\mathrm{FM_{abs}} = 1 + (\Delta^{14}\mathrm{C}/1000‰)$. The $\mathrm{FM_{abs}}$ is proportional to the $^{14}\mathrm{C}/^{12}\mathrm{C}$ ratio normalized to a δ^{13}C of $-25‰$ (Trumbore et al., 2016), and is thus proportional to the normalized ratio of ^{14}C to total carbon ($^{14}\mathrm{C}/\mathrm{C}$), considering the negligible abundance of ^{14}C compared to ^{12}C and ^{13}C. Therefore, if we know F^i_j, the $\mathrm{FM_{abs}}$ of pool j at time step i, we can find F^{i+1}_j at time step $i+1$ with the following equation (provided all the pools and the influx have comparable δ^{13}C signals):

$$F^{i+1}_j C^{i+1}_j = (1-\lambda) F^i_j C^i_j + I^i_j F^i_{\mathrm{influx}} + \sum_k \Phi^i_{jk} F^i_k \,, \tag{S2}$$

where C^{i+1}_j is given by equation (S1), λ is the radioactive decay rate of ^{14}C in units of inverse time step size, and F^i_{influx} is the $\mathrm{FM_{abs}}$ of the external carbon influx at time step i given by the forcing data. We can then recover the Δ^{14}C at each time step i and for each pool j with $(F^i_j - 1) \times 1000‰$.

S4. Incorrect or inaccurate ^{14}C implementations

S4.1. SOMic

The SOMic model (Woolf & Lehmann, 2019), as implemented on the GitHub repository `domwoolf/somic1` (commit `be34e56` from June 2019), does not produce accurate ^{14}C predictions. Instead of working with the more typical Δ^{14}C or absolute Fraction Modern ($\mathrm{FM_{abs}}$) units, this implementation tracks the ^{14}C age, which we summarily define as $\mathrm{Age} = -\log(\mathrm{FM_{abs}})\,\lambda^{-1}$, where λ is the radioactive decay rate of ^{14}C. This causes complications when updating the ^{14}C ages of the pools at each time step and when computing the total ^{14}C age of the soil from the ^{14}C ages of the individual pools. Indeed, to find the combined age $\mathrm{Age_{A+B}}$ of pools A and B, the implementation of SOMic takes a weighted average over the ages, which is not entirely accurate:

$$\mathrm{Age_{A+B}} = \frac{C_\mathrm{A}\mathrm{Age_A} + C_\mathrm{B}\mathrm{Age_B}}{C_\mathrm{A} + C_\mathrm{B}}\,, \tag{S3}$$

where Age_i and C_i are the ^{14}C age and the carbon stocks, respectively, of pool i. This weighted average formula is used to integrate the ^{14}C ages of carbon fluxes into the pools at each time step on lines 154-160, and to compute the ^{14}C age of the total soil on line 210 of file `src/SOMIC.cpp` in the `domwoolf/somic1` GitHub repository (commit `be34e56`).

In order to prove that equation (S3) is inaccurate, let us derive how to correctly add the ^{14}C ages of pools A and B. Let $^{14}C_i$ denote the ^{14}C stocks and C_i the total carbon stocks of pool i. Then, by conservation of mass, we have

$$^{14}C_\mathrm{A+B} = {}^{14}C_\mathrm{A} + {}^{14}C_\mathrm{B} \quad \text{and} \quad C_\mathrm{A+B} = C_\mathrm{A} + C_\mathrm{B} \quad \Rightarrow \quad \frac{{}^{14}C_\mathrm{A+B}}{C_\mathrm{A+B}} = \frac{{}^{14}C_\mathrm{A} + {}^{14}C_\mathrm{B}}{C_\mathrm{A} + C_\mathrm{B}}\,. \tag{S4}$$

Since the FM_{abs} is proportional to the $^{14}C/C$ ratio (assuming pools A and B have a similar ^{13}C content), the above is equivalent to

$$F_{A+B} = \frac{C_A F_A + C_B F_B}{C_A + C_B}\,, \tag{S5}$$

where F_i and C_i are the FM_{abs} and carbon stocks, respectively, of pool i. It follows that the combined ^{14}C age of pools A and B is given by

$$\mathrm{Age}_{A+B} = -\lambda^{-1} \cdot \log\left(\frac{C_A \exp\left(-\lambda \cdot \mathrm{Age}_A\right) + C_B \exp\left(-\lambda \cdot \mathrm{Age}_B\right)}{C_A + C_B}\right)\,. \tag{S6}$$

Notice that equation (S3) is the first non-zero term of the above result's Taylor expansion around $\mathrm{Age}_A = 0$, $\mathrm{Age}_B = 0$. This means that equation (S3) works well for ages that are close to zero, i.e. when the $\Delta^{14}C$ is close to zero. However, it fails to accurately predict the propagation of the bomb spike into the soil ecosystem in the latter half of the 20th century, as shown in Figure S6. While the error induced by the incorrect implementation exceeds 25‰ for the total soil $\Delta^{14}C$ in the 1970s, the error in the 2000s and 2010s is only around 10‰, which is relatively minor.

S4.2. MIMICS

The only MIMICS version already implemented with ^{14}C is published in Y. Wang et al. (2021), and the code is available at `https://data.csiro.au/collection/csiro:47942v1`. However, this ^{14}C implementation is incorrect (see Figure S7).

The time evolution of the carbon stocks in MIMICS is given by function $f(C,t)$, which depends on the carbon stocks vector C and time t. Function f is implemented as subroutine `modelx` in the source file `vsoilmic05f_ms25.f90`. We can write function f in terms of internal carbon transfer matrix A and carbon influx vector I:

$$\mathrm{d}C/\mathrm{d}t = f(C,t) = A(C,t)C + I(t)\,, \tag{S7}$$

where matrix $A(C,t)$ is a function of carbon stocks C and time t, and vector $I(t)$ is time-dependent.

Then, following the same procedure which yielded equation (S2), we can derive the equation governing the evolution of the ^{14}C stocks (^{14}C):

$$\mathrm{d}^{14}C/\mathrm{d}t = -\lambda^{14}C + A(C,t)^{14}C + {}^{14}I(t)\,, \tag{S8}$$

where λ is the radioactive decay rate of ^{14}C, and ^{14}I is the external influx of ^{14}C.

However, in the ^{14}C-implementation of MIMICS, the evolution of the ^{14}C stocks is predicted with

$$\mathrm{d}^{14}C/\mathrm{d}t = -\lambda^{14}C + f(^{14}C,t) = -\lambda^{14}C + A(^{14}C,t)^{14}C + {}^{14}I(t)\,. \tag{S9}$$

The above equation is incorrect because $A(^{14}C,t) \neq A(C,t)$.

S5. Turnover times in the Millennial model

In Millennial version 2 (Abramoff et al., 2022), the POM, MAOM, and Aggregate C pools exchange carbon with each other on the scale of a few months. The aggregate formation rate of the POM pool is between 0.012/day and 0.026/day (k_{pa} in Table A1 of Abramoff et al., 2022), which translates to an average aggregation time of 1-3 months. Meanwhile, the optimized rate of aggregate formation for the MAOM pool is between 0.0038/day and 0.0052/day (k_{ma} in Table A1 of Abramoff et al., 2022), giving MAOM an average aggregation time of 6-8 months. The Aggregate C pool has a breakdown rate of around 0.02/day (k_b in Table A1 of Abramoff et al., 2022), so aggregates have a turnover time of just 50 days. POM and MAOM exchange their carbon rapidly with the Aggregate C pool, which then redistributes the carbon back to the POM and MAOM pools in less than 2 months, on average. This means that the ^{14}C signals of the POM, MAOM, and Aggregate C pools get homogenized within a couple years.

References

Abramoff, R. Z., Guenet, B., Zhang, H., Georgiou, K., Xu, X., Viscarra Rossel, R. A., ... Ciais, P. (2022, January). Improved global-scale predictions of soil carbon stocks with Millennial Version 2. *Soil Biology and Biochemistry*, *164*, 108466. Retrieved 2022-01-19, from https://linkinghub.elsevier.com/retrieve/pii/S0038071721003400 doi: 10.1016/j.soilbio.2021.108466

Graven, H., Allison, C. E., Etheridge, D. M., Hammer, S., Keeling, R. F., Levin, I., ... White, J. W. C. (2017, December). Compiled records of carbon isotopes in atmospheric CO_2 for historical simulations in CMIP6. *Geoscientific Model Development*, *10*(12), 4405–4417. Retrieved 2021-03-10, from https://gmd.copernicus.org/articles/10/4405/2017/ doi: 10.5194/gmd-10-4405-2017

Kyker-Snowman, E., Wieder, W. R., Frey, S. D., & Grandy, A. S. (2020, September). Stoichiometrically coupled carbon and nitrogen cycling in the MIcrobial-MIneral Carbon Stabilization model version 1.0 (MIMICS-CN v1.0). *Geoscientific Model Development*, *13*(9), 4413–4434. Retrieved 2023-06-29, from https://gmd.copernicus.org/articles/13/4413/2020/ doi: 10.5194/gmd-13-4413-2020

Moore, J. A. M., Sulman, B. N., Mayes, M. A., Patterson, C. M., & Classen, A. T. (2020, April). Plant roots stimulate the decomposition of complex, but not simple, soil carbon. *Functional Ecology*, *34*(4), 899–910. Retrieved 2022-08-29, from https://onlinelibrary.wiley.com/doi/10.1111/1365-2435.13510 doi: 10.1111/1365-2435.13510

Sierra, C. A., Müller, M., Metzler, H., Manzoni, S., & Trumbore, S. E. (2017, May). The muddle of ages, turnover, transit, and residence times in the carbon cycle. *Global*

Change Biology, *23*(5), 1763–1773. Retrieved 2021-04-09, from http://doi.wiley.com/10.1111/gcb.13556 doi: 10.1111/gcb.13556

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E., & Pacala, S. W. (2014, December). Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO2. *Nature Climate Change*, *4*(12), 1099–1102. Retrieved 2022-05-29, from http://www.nature.com/articles/nclimate2436 doi: 10.1038/nclimate2436

Trumbore, S. E., Sierra, C. A., & Hicks Pries, C. E. (2016). Radiocarbon Nomenclature, Theory, Models, and Interpretation: Measuring Age, Determining Cycling Rates, and Tracing Source Pools. In E. A. Schuur, E. Druffel, & S. E. Trumbore (Eds.), *Radiocarbon and Climate Change* (pp. 45–82). Cham: Springer International Publishing. Retrieved 2021-04-05, from http://link.springer.com/10.1007/978-3-319-25643-6_3 doi: 10.1007/978-3-319-25643-6_3

Wang, G., Gao, Q., Yang, Y., Hobbie, S. E., Reich, P. B., & Zhou, J. (2022, March). Soil enzymes as indicators of soil function: A step toward greater realism in microbial ecological modeling. *Global Change Biology*, *28*(5), 1935–1950. Retrieved 2023-06-23, from https://onlinelibrary.wiley.com/doi/10.1111/gcb.16036 doi: 10.1111/gcb.16036

Wang, Y., Zhang, H., Ciais, P., Goll, D., Huang, Y., Wood, J. D., ... Prescher, A. (2021, April). Microbial Activity and Root Carbon Inputs Are More Important than Soil Carbon Diffusion in Simulating Soil Carbon Profiles. *Journal of Geophysical Research: Biogeosciences*, *126*(4). Retrieved 2022-07-10, from https://onlinelibrary.wiley.com/doi/10.1029/2020JG006205 doi: 10.1029/

2020JG006205

Woolf, D., & Lehmann, J. (2019, December). Microbial models with minimal mineral protection can explain long-term soil organic carbon persistence. *Scientific Reports*, *9*(1), 6522. Retrieved 2021-10-01, from http://www.nature.com/articles/s41598-019-43026-8 doi: 10.1038/s41598-019-43026-8

Table S1. Correspondences between simulated carbon pools and the POM fraction, MAOM fraction, or other carbon reservoirs.

Model	POM fraction	MAOM fraction	Other SOC pools	Litter pools
MEND	POM_O, POM_H	MOM, QOM	DOM, MB_A, MB_D, EP_O, EP_H, EM	
Millennial	POM	MAOM, Aggregate C	LMWC, Microbial Biomass	
SOMic	SPM, IPM	MAC	DOC, MB	
CORPSE	SPC_u, CPC_u	SPC_p, CPC_p, MN_p	MN_u, LMB	
MIMICS	SOM_c	SOM_p	SOM_a, MIC_r, MIC_K	LIT_m, LIT_s

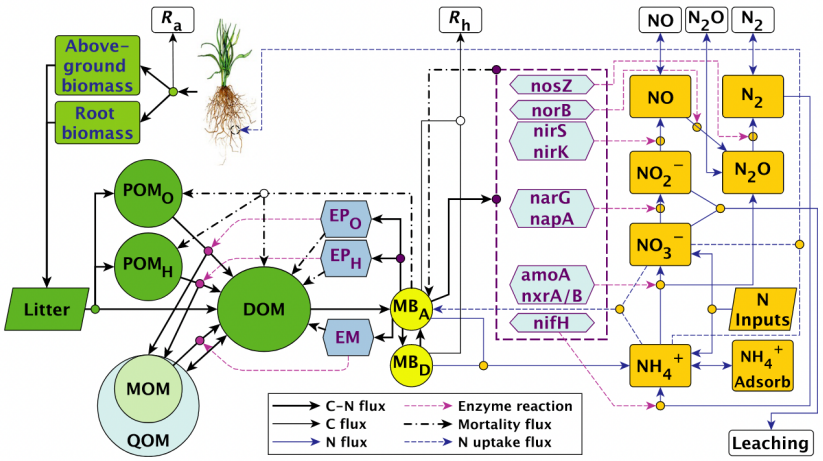

Figure S1. MEND-new (2022) model diagram from G. Wang et al. (2022)

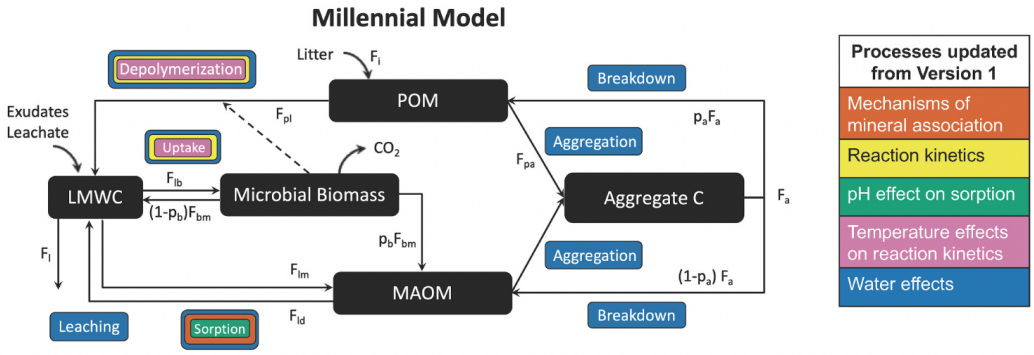

Figure S2. Millennial V2 diagram from Abramoff et al. (2022)

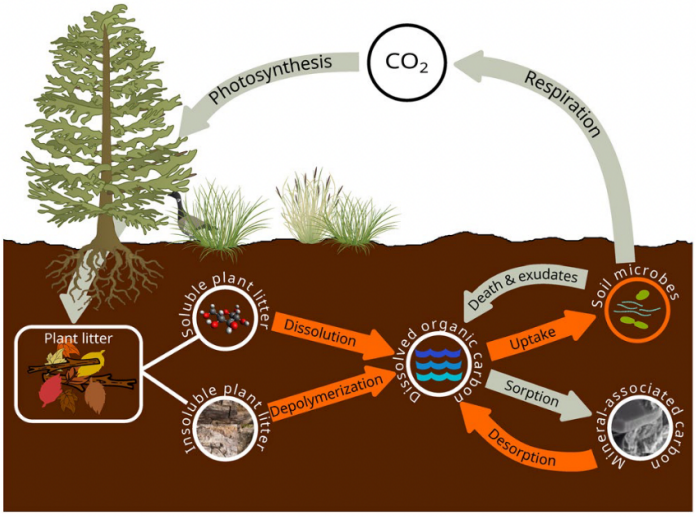

Figure S3. SOMic 1.0 diagram from Woolf and Lehmann (2019)

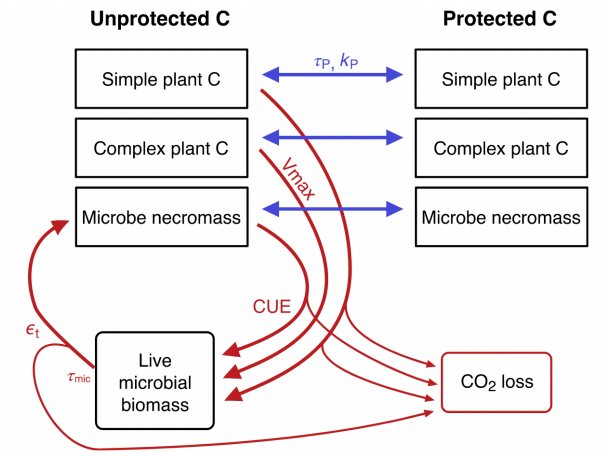

Figure S4. CORPSE diagram from Moore et al. (2020)

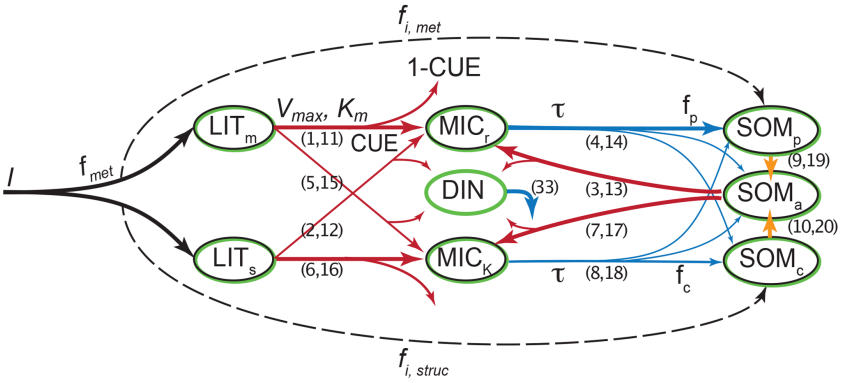

Figure S5. MIMICS-CN v1.0 diagram from Kyker-Snowman et al. (2020)

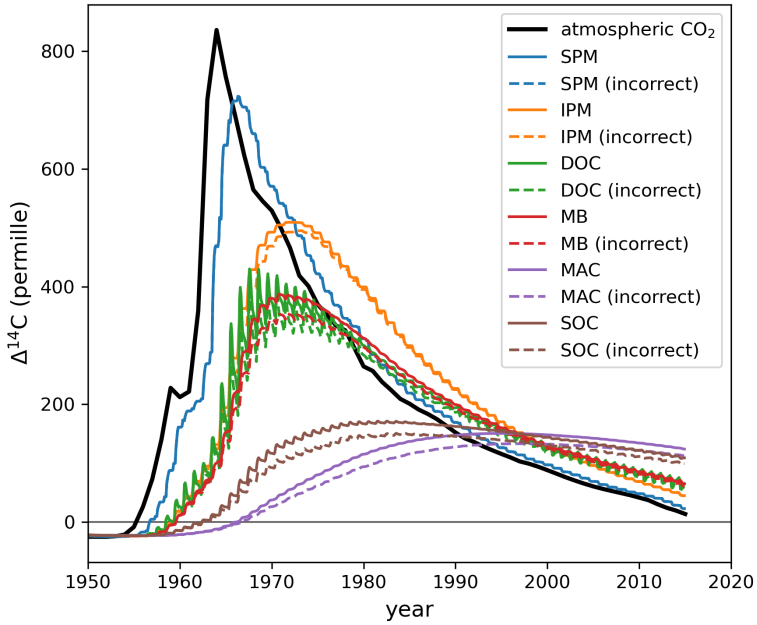

Figure S6. Comparison of Δ^{14}C predicted by SOMic with the correct and incorrect ^{14}C implementations. The atmospheric $\Delta^{14}CO_2$ of the Northern Hemisphere (source: Graven et al., 2017) is plotted for reference. SOMic pool names: SPM, soluble plant matter; IPM, insoluble plant matter; DOC, dissolved organic carbon; MB, microbial biomass; MAC, mineral-associated carbon; SOC, total soil organic carbon.

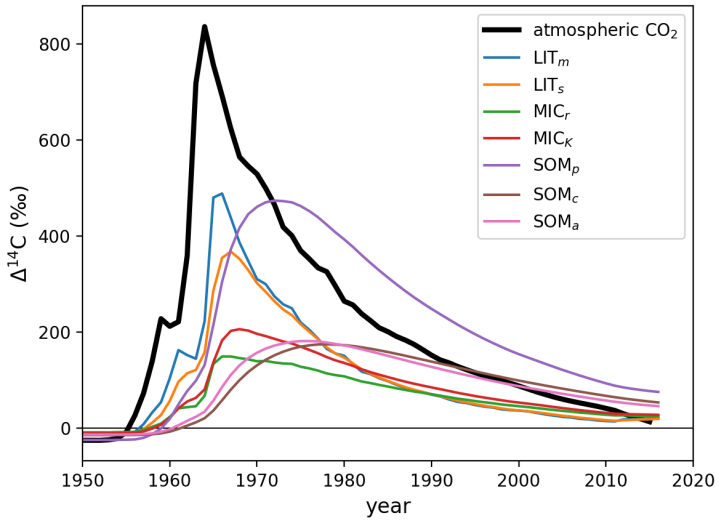

Figure S7. Δ^{14}C output of MIMICS (Y. Wang et al., 2021) with incorrect isotopic implementation. The atmospheric $\Delta^{14}CO_2$ of the Northern Hemisphere (source: Graven et al., 2017) is plotted for reference. MIMICS pool names: LIT_m, metabolic litter; LIT_s, structural litter; MIC_r, r-strategist microbes; MIC_K, K-strategist microbes; SOM_p, physically protected soil organic matter; SOM_c, chemically protected soil organic matter; SOM_a, active soil organic matter.

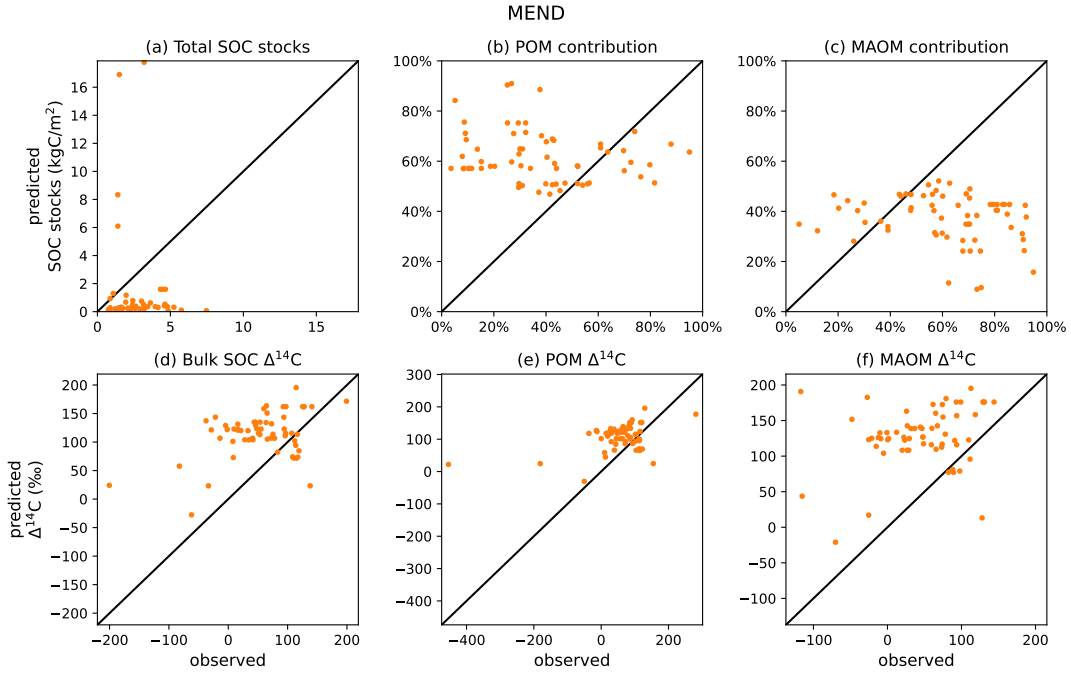

Figure S8. Predictions vs observations plots for the MEND model.

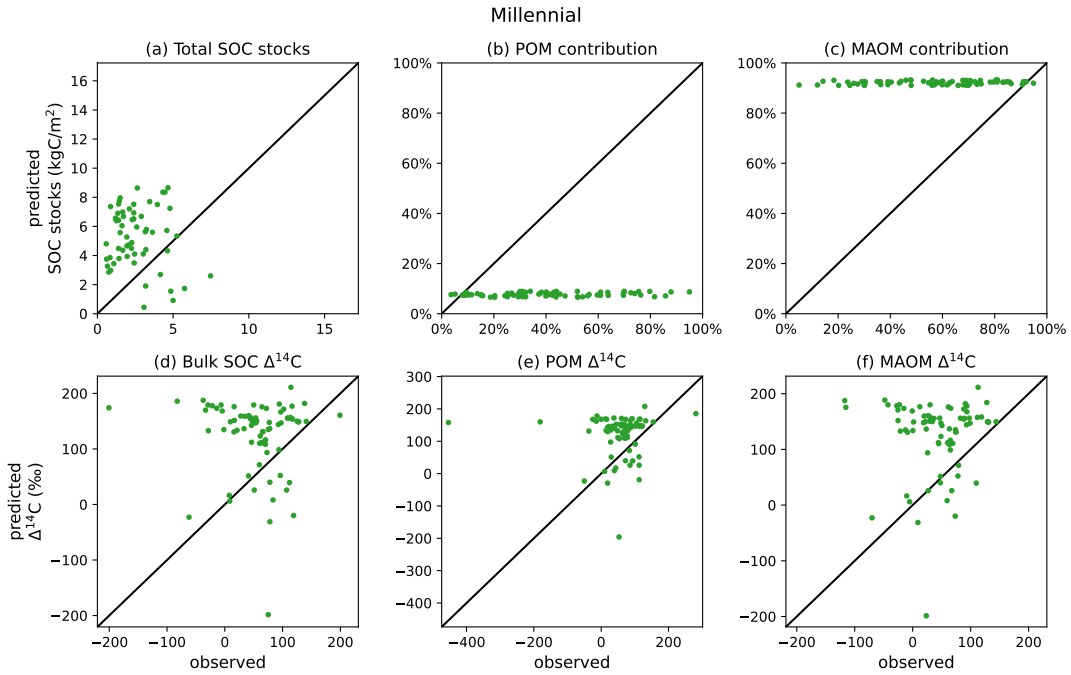

Figure S9. Predictions vs observations plots for the Millennail model.

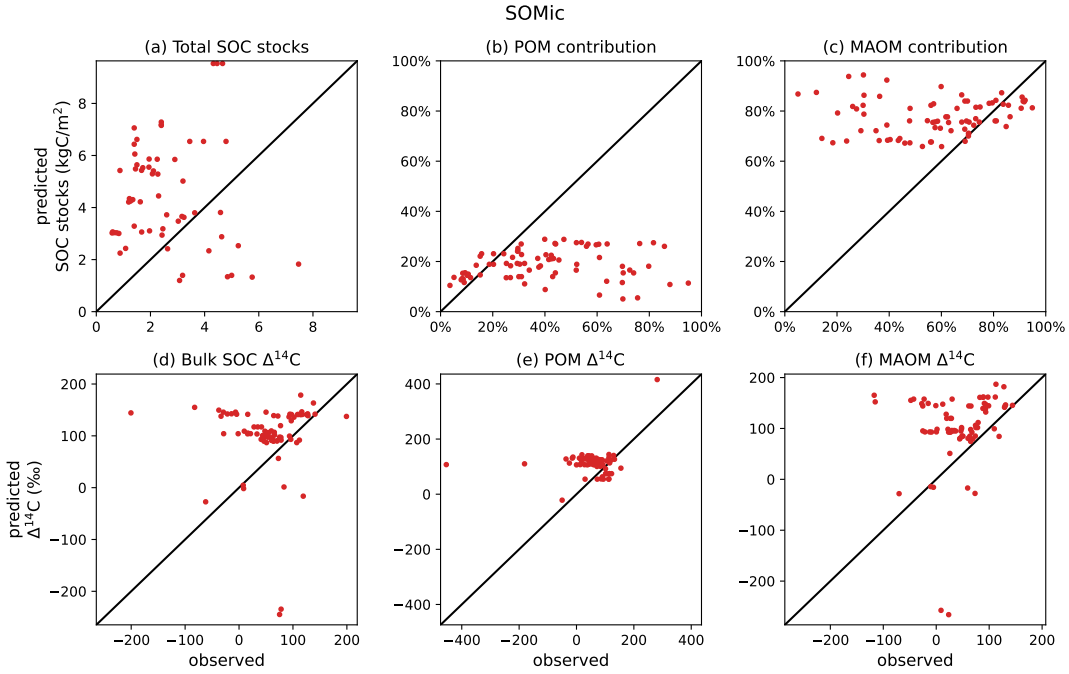

Figure S10. Predictions vs observations plots for the SOMic model.

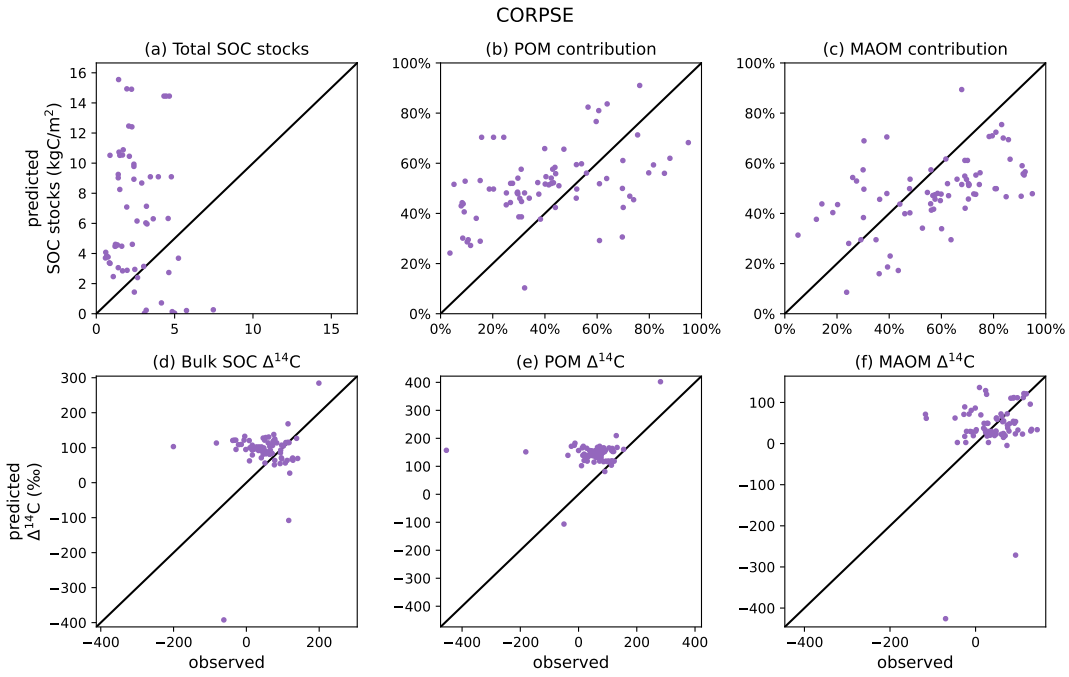

Figure S11. Predictions vs observations plots for the CORPSE model.

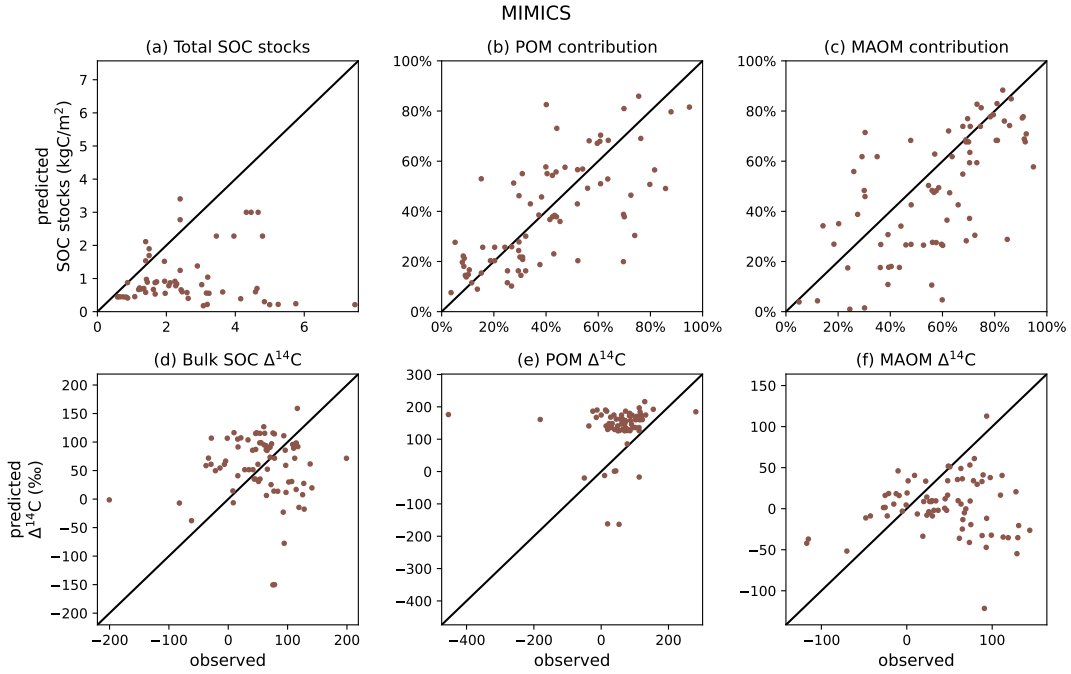

Figure S12. Predictions vs observations plots for the MIMICS model.