A hybrid framework for the spin-up and initialization of distributed coupled ecohydrological-biogeochemical models - SUPPLEMENTARY INFORMATION

Taiqi Lian[1], Ziyan Zhang[2], Athanasios Paschalis[3,2], and Sara Bonetti[1]

[1]Laboratory of Catchment Hydrology and Geomorphology, École Polytechnique Fédérale de Lausanne (EPFL), 1951 Sion, Switzerland
[2]Department of Civil and Environmental Engineering, Imperial College London, UK
[3]Department of Civil and Environmental Engineering, University of Cyprus, Nicosia, Cyprus

Correspondence: Sara Bonetti (sara.bonetti@epfl.ch)

Table S1. Examples of spin-up methods and applications used in ecohydrological and biogeochemical models across different spatial scales. The table summarizes the model name, spatial scale, and methodological approach. The category indicates whether the spin-up scheme was newly introduced or applied and eventually compared to other methods.

Source	Model	Scale	Description	Category
Zhan et al. (2003)	Simplified Simple Biosphere Model (SSiB)	Regional scale & Plot scale	Developed semi-analytical solution to stabilize and accelerate spin-up.	Introduced
Thornton & Rosenbloom (2005)	Biome-BGC Model	Plot scale	Introduced multiple spin-up acceleration methods: accelerated decomposition, N addition, General multivariate minimization methods.	Introduced
Basso et al. (2011)	DSSAT-CENTURY Model	Plot scale	Developed SOC pool initialization using site-specific land-use history.	Introduced
Hashimoto et al. (2011)	CENTURY v4 Model	Plot scale	Proposed slow-relaxation scaling of C/N pools during spin-up.	Introduced
Xia et al. (2012)	Community Atmosphere Biosphere Land Exchange Model (CABLE)	Global scale & Plot scale	Developed semi-analytical solution to accelerate coupled C-N model spin-up.	Introduced
Koven et al. (2013)	Community Land Model version 4 (CLM4)	Global scale	Introduced vertically resolved C-N model with pool-specific modified accelerated decomposition spin-up.	Introduced
Ng et al. (2014)	Community Land Model with Carbon-Nitrogen (CLM-CN)	Plot scale	Combined data assimilation with spin-up adjustments.	Introduced
Qu et al. (2018)	Terrestrial Ecosystem Model (TEM)	Plot scale	Developed analytical solution-based fast spin-up algorithm.	Introduced
Bruun & Jensen (2002)	Daisy Model	Plot scale	Compared equilibrium initialization vs. historical simulation.	Tested
Foereid et al. (2012)	DAYCENT Model	Plot scale	Tested steady-state and trend-fitting SOC initialization.	Tested
Nemo et al. (2017)	Rothamsted Carbon Model (RothC)	Plot scale	Evaluated C input adjustment and SOC fractionation methods.	Tested
Dimassi et al. (2018)	CENTURY v4.5 Model	Plot scale	Tested SOC initialization scenarios combining crop history and relaxation.	Tested
Lavin-Gullon et al. (2023)	Regional Climate Model (RCM)	Regional scale	Assessed spin-up time and variability in overlapping time slices.	Tested

Table S2. Comparison between soil organic carbon and nitrogen (SOC, SON) from the steady state condition with the biogeochemistry-only module (i.e., no coupled vegetation-soil biogeochemistry dynamics) and the reference steady state from the most comprehensive plot-scale spin-up (i.e., considering coupled vegetation-soil biogeochemistry dynamics) in the ten cells (see Fig. 2h in the main text). Information on vegetation cover and soil texture in these cells is also provided.

ID	Vege	Sand [%] Clay [%] Organic [%]	SOC [kg C/m^2]			SON [kg N/m^2]		
			Steady	Coevo.	Relative difference (%)	Steady	Coevo.	Relative difference (%)
1	Tree	40.9 27.5 6.8	12.51	13.51	-7.40	0.93	1.04	-10.58
2	Tree	42.0 24.9 9.0	10.46	12.31	-15.03	0.80	0.96	-16.67
3	Tree	42.5 24.3 8.6	10.36	11.94	-13.23	0.81	0.94	-13.83
4	Tree	42.5 25.9 7.2	11.90	13.74	-13.39	0.89	1.06	-16.04
5	Tree	39.5 25.9 8.3	11.53	12.90	-10.62	0.87	1.00	-13.00
6	Grass	41.7 26.5 6.7	6.70	8.42	-20.43	0.84	0.90	-6.67
7	Grass	41.3 26.9 7.1	6.41	8.00	-19.88	0.82	0.86	-4.65
8	Grass	40.4 27.9 8.5	6.32	7.76	-18.56	0.81	0.84	-3.57
9	Grass	40.1 26.7 7.2	6.91	8.29	-16.65	0.87	0.89	-2.25
10	Grass	41.6 26.5 6.3	6.73	8.36	-19.50	0.85	0.90	-5.56

Table S3. Jensen–Shannon divergence ($\times 10^{-4}$) between the probability density functions (PDFs) of soil organic carbon (SOC) and soil organic nitrogen (SON) from different percentages of tracked cells and those from the benchmark simulation across multiple simulation scenarios. Values in parentheses under the Random Soil scenario represent results obtained when clay content was included as an additional predictor in the random forest model.

Scenarios/Track cells [%]		10	20	40	60	80
Original	SOC	6.04	2.84	1.05	0.59	0.27
	SON	0.99	0.57	0.28	0.15	0.08
Random Veg	SOC	3.40	2.66	1.13	0.61	0.34
	SON	0.82	0.62	0.29	0.18	0.12
Homog. Veg	SOC	2.49	1.10	0.79	0.31	0.11
	SON	0.98	0.56	0.33	0.18	0.05
Random Soil	SOC	6659 (6.80)	24 (4.05)	21 (2.12)		
	SON	6725 (3.39)	16 (2.26)	2 (1.81)		
Homog. Soil	SOC	4.37	1.36	0.53		
	SON	0.50	0.24	0.12		
0.5 h_{max}	SOC	5.62	2.16	0.88		
	SON	0.98	0.57	0.25		
0.2 h_{max}	SOC	4.08	2.68	0.68		
	SON	0.70	0.37	0.18		

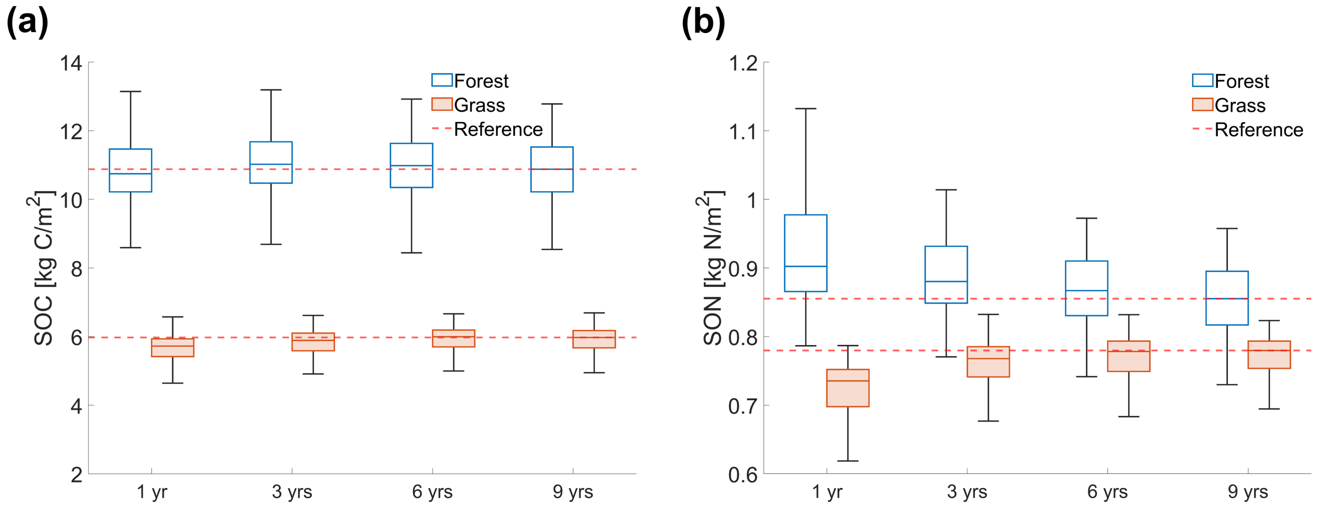

Figure S1. Comparison of steady state of (a) soil organic carbon (SOC) and (b) soil organic nitrogen (SON) using only the biogeochemistry module, with forcing and flux tracking periods of 1 year, 3 years, 6 years, and 9 years (reference). The forcing and flux tracking period is repeated to complete the 1000-year spin-up of the biogeochemistry module.

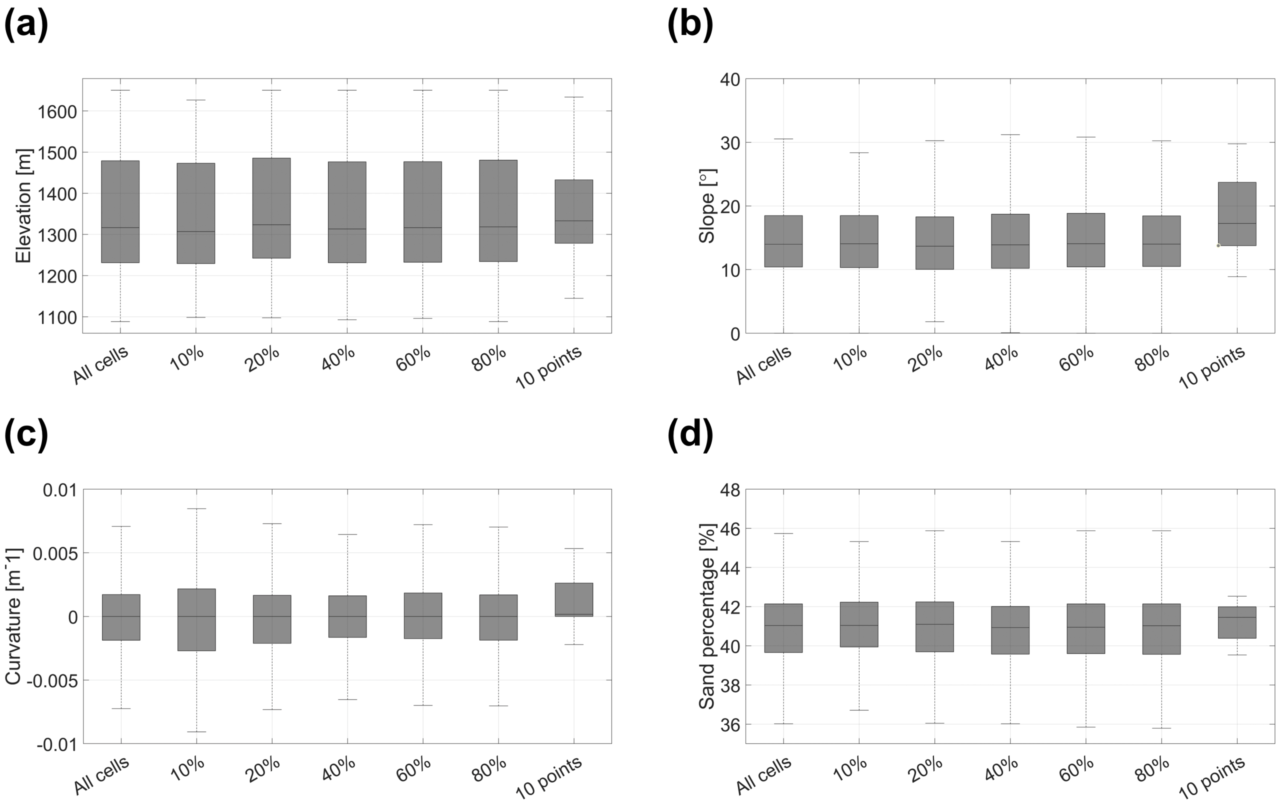

Figure S2. Distribution of topographic and soil attributes for all grid cells and selected subsets used for random forest model training. Panels show the distributions of (a) elevation, (b) slope, (c) curvature, and (d) sand percentage for all cells in the domain ("All cells"), different percentages of tracked training cells (10%–80%), and the 10 selected reference points for additional plot-scale simulation. The distributions indicate that the selected subsets preserve the key topographic and soil characteristics of the full domain.

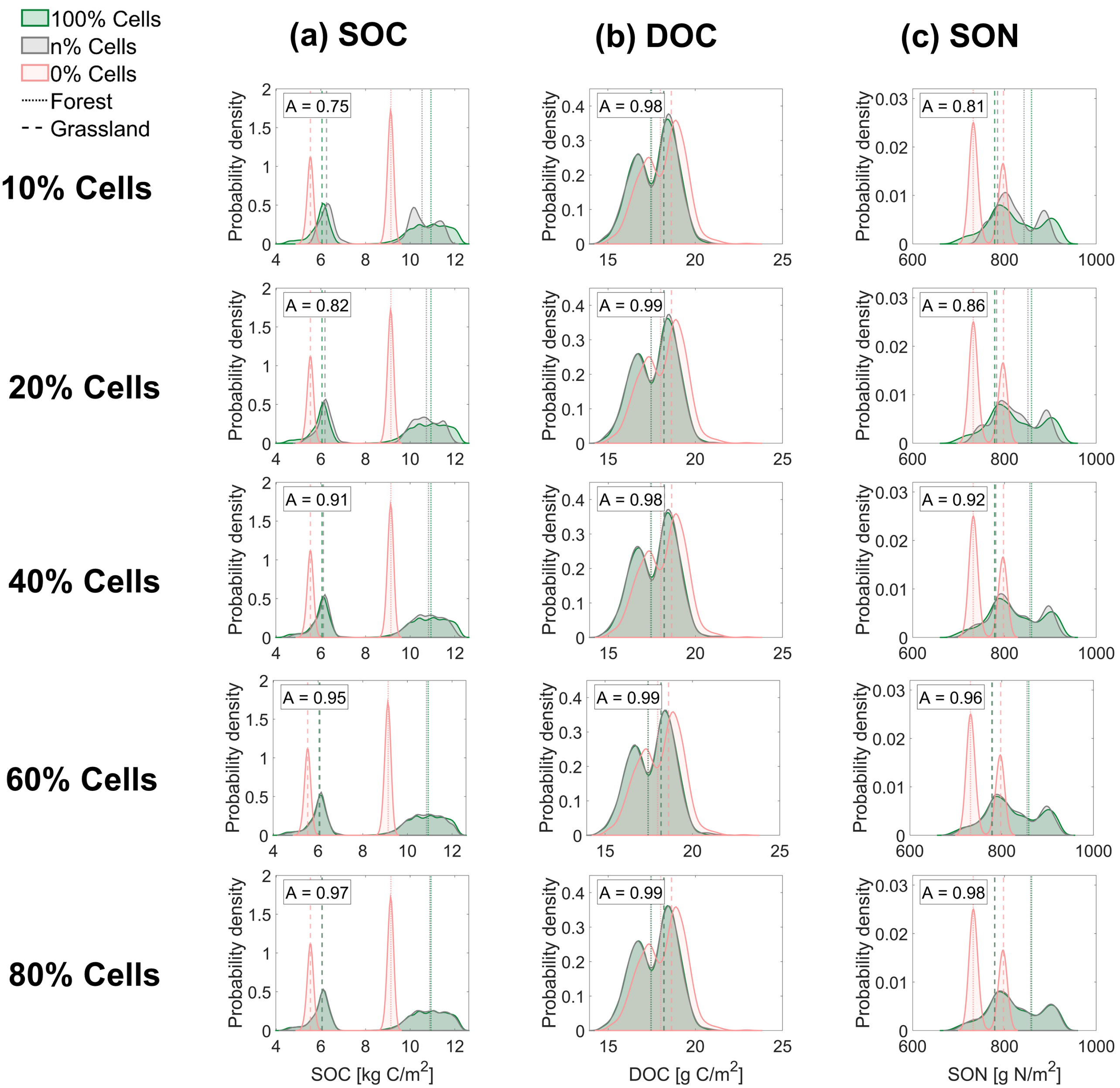

Figure S3. Probability density functions (PDFs) of (a) soil organic carbon (SOC), (b) dissolved organic carbon (DOC), and (c) soil organic nitrogen (SON) under different initialization settings using Random Forest (RF) with $n = 10\%$, 20%, 40%, 60%, and 80%. Each subplot compares the resulting distribution with those from the benchmark simulation using 100% of tracked cells (1000% Cells) and from the simulation without cell tracking or RF (0% Cells). Shaded areas represent the overlapping region with the benchmark distribution. The overlap area between PDFs (A value) in each panel quantifies the similarity between the PDF and the benchmark. Dashed and dotted lines represent median values for grassland and forest areas, respectively.

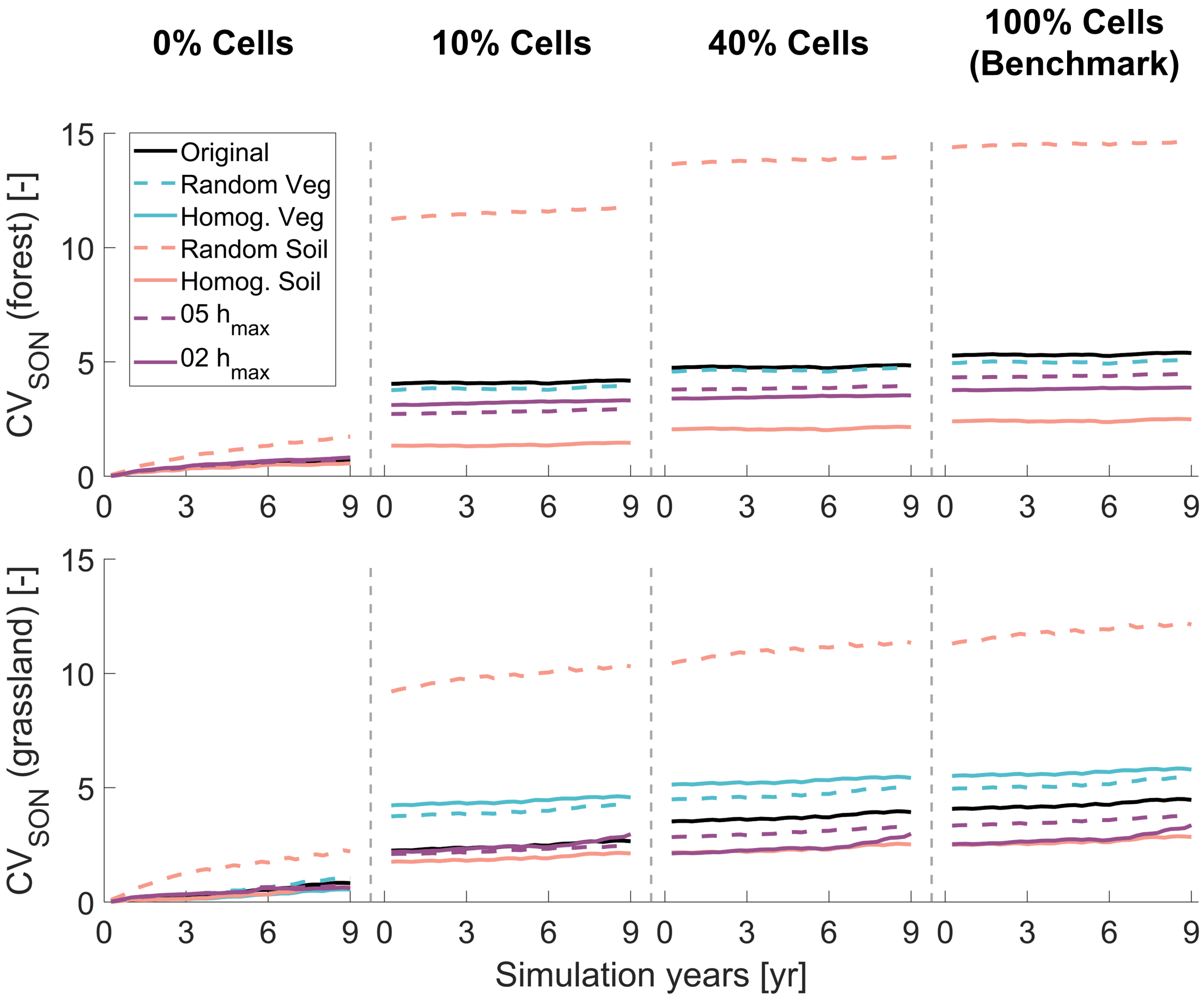

Figure S4. Temporal evolution of the coefficient of variation (CV) of soil organic nitrogen (SON) over a 9-year simulation period in forest (top) and grassland (bottom) under different initialization settings and simulation scenarios. Each column represents one initialization setting: no cells tracking and no random forest (0% Cells), $n = 10\%$ cells tracking and random forest (10% Cells), $n = 40\%$ cells tracking and random forest (40% Cells), and $n = 100\%$ cells tracking (100% Cells). Line colors and styles indicate results from different simulation scenarios, as shown in the legend.

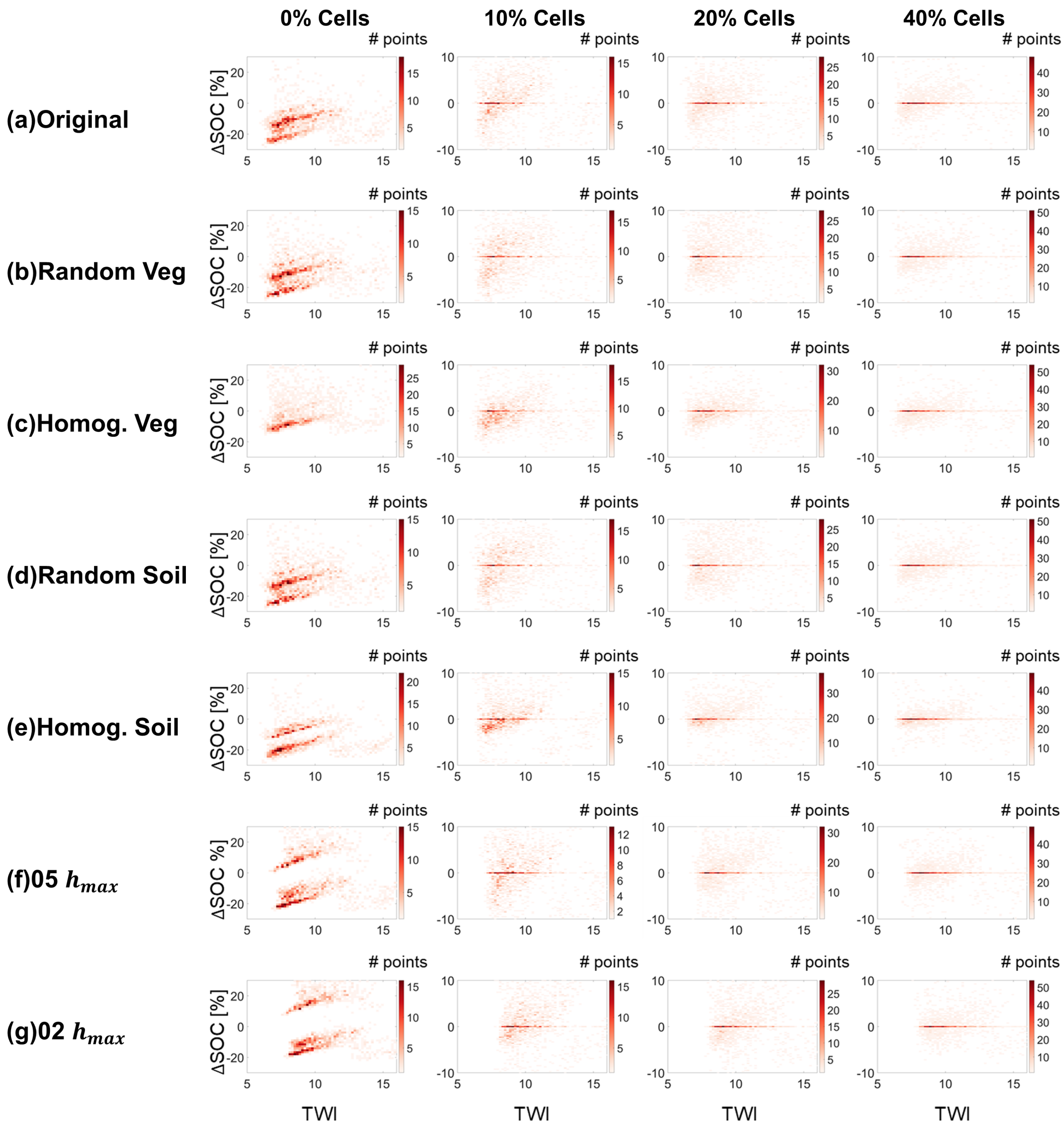

Figure S5. Density plots showing the relationship between relative SOC bias ($\Delta SOC = (SOC_{\text{simulation}} - SOC_{\text{benchmark}})/SOC_{\text{benchmark}}$), where $SOC_{\text{simulation}}$ represents results from different initialization settings, and the topographic wetness index (TWI) across all simulation scenarios. Results are shown for (a) Original, (b) Random Veg, (c) Homog. Veg, (d) Randon Soil, (e) Homog. Soil, (f) 0.5 h_{max}, and (g) 0.2 h_{max} simulation scenarios. Columns represent different initialization settings: no cell tracking or RF ($n = 0\%$, 0% Cells), and RF-based initialization using $n = 10\%$, 20%, and 40% of tracked cells. The red color scale indicates the number of grid cells per bin.

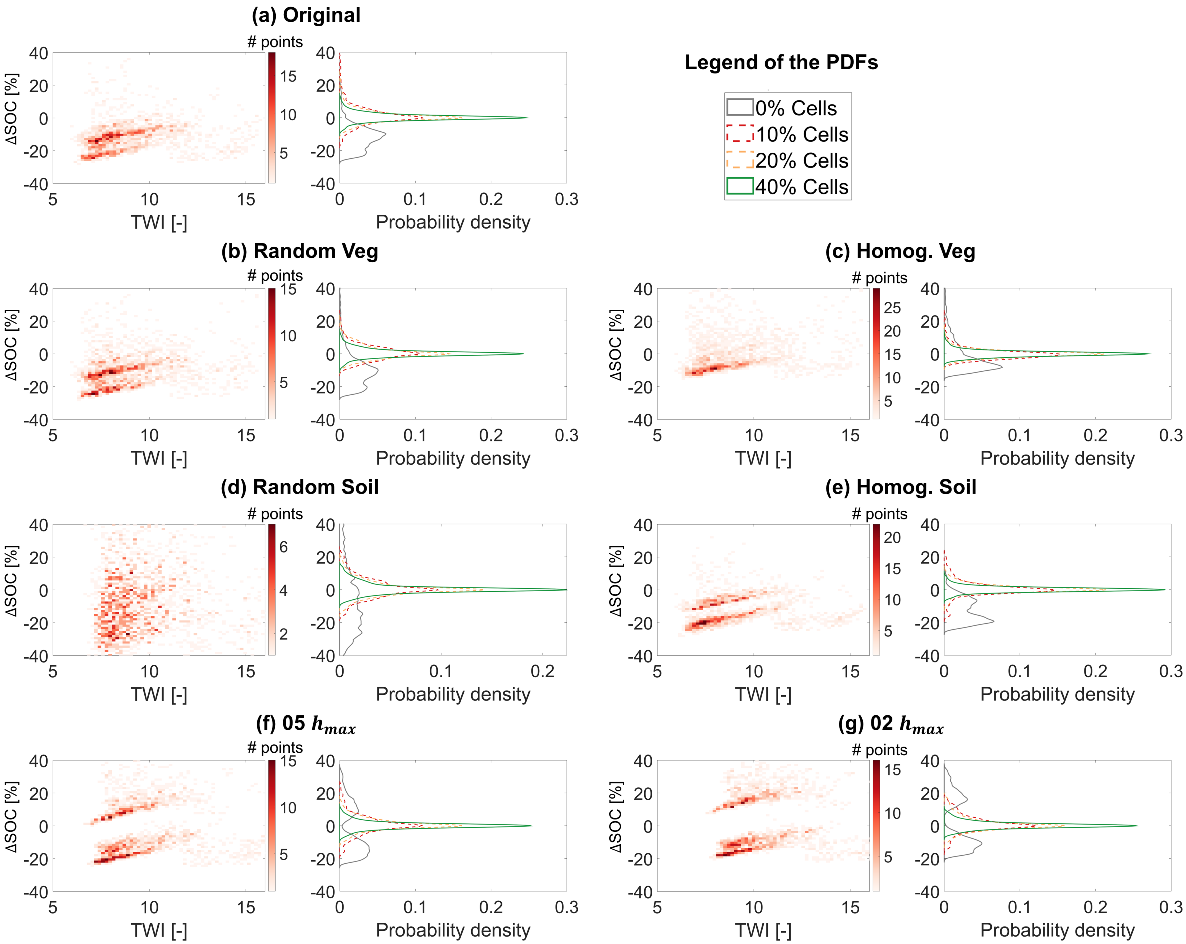

Figure S6. Density maps showing the relationship between SOC bias from the simulations without cell tracking and random forest ($n = 0\%$, 0% Cells) and topography wetness index (TWI) across different simulation scenarios (left panels of each subfigure; $\Delta SOC = SOC_{simulation} - SOC_{benchmark}$). Probability density functions (PDFs) of ΔSOC from $n = 0\%$ simulation (in gray), and with RF-based initialization using $n =$ 10% (red), 20% (yellow), and 40% (green) of tracked cells (right panels). Results are shown for (a) Original, (b) Random Veg, (c) Homog. Veg, (d) Randon Soil, (e) Homog. Soil, (f) 0.5 h_{max}, and (g) 0.2 h_{max} simulation scenarios.

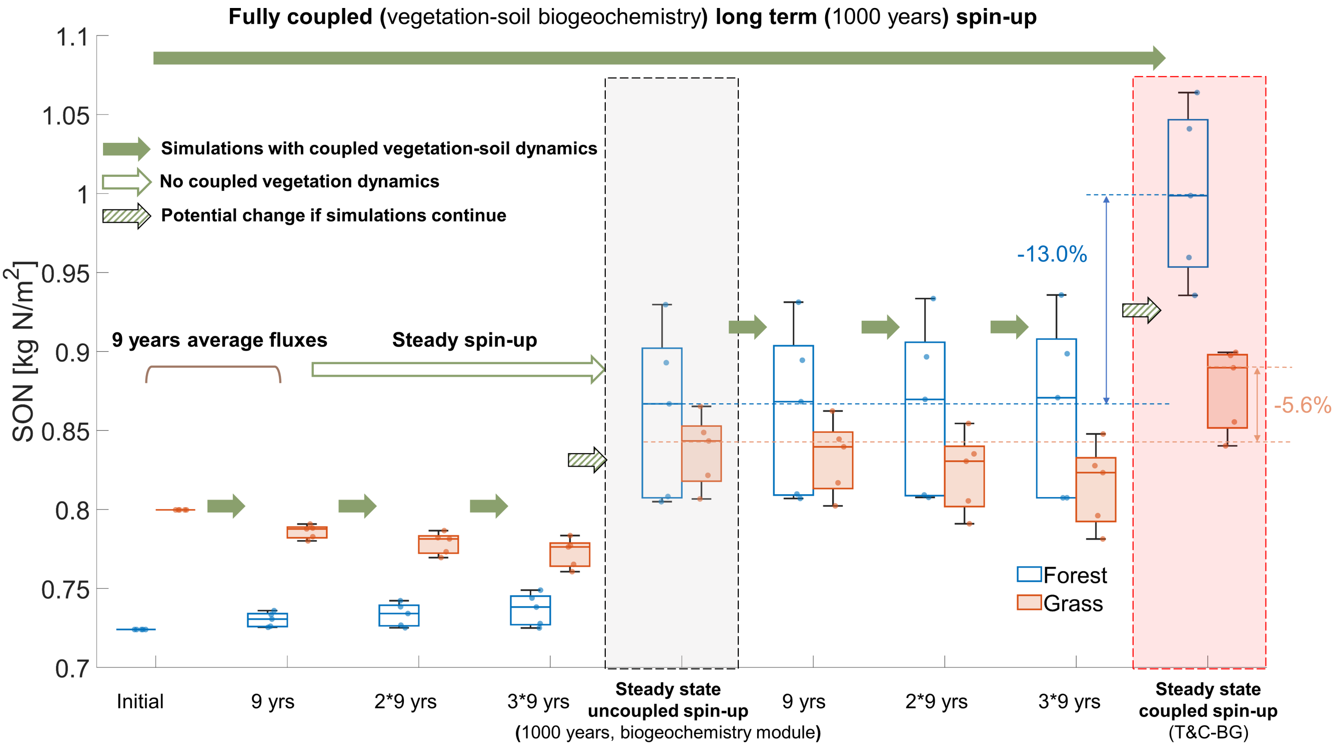

Figure S7. Evolution of soil organic nitrogen (SON) during different initialization schemes and simulation durations at 10 selected sites (Fig. 2h in the main text). Blue/orange boxplots represent forest/grassland sites. Arrows indicate simulation transitions between stages. The initial values (far left) are followed by 3*9-year simulations using the fully coupled T&C-BG model (solid green arrows). The average vegetation fluxes are extracted from the first 9 years to obtain a steady state condition with the biogeochemistry-only module (outlined green arrow, middle gray box). This is followed by three additional 9-year simulations. A comprehensive long-term spin-up using the fully coupled T&C-BG model (i.e., considering coupled vegetation and soil biogeochemical dynamics) is shown in the right red box. Striped arrows indicate the potential direction of SON change if coupled simulations continue beyond the current duration.

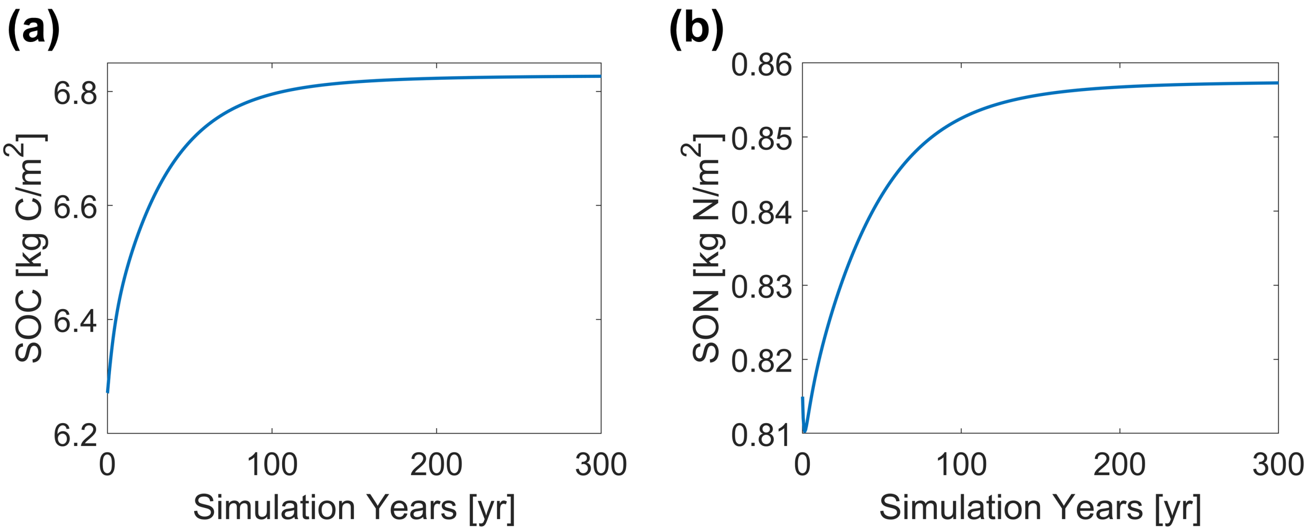

Figure S8. Example of temporal evolution of (a) soil organic carbon (SOC) and (b) soil organic nitrogen (SON) in a one-dimensional spin-up without coupled vegetation-soil biogeochemistry dynamics, based on repeated meteorological forcing at a representative grassland site in Erlenbach. SOC increases gradually over time, while SON initially declines before rising toward a steady state. The figure illustrates the slow convergence of soil carbon and nitrogen pools even under simplified conditions.

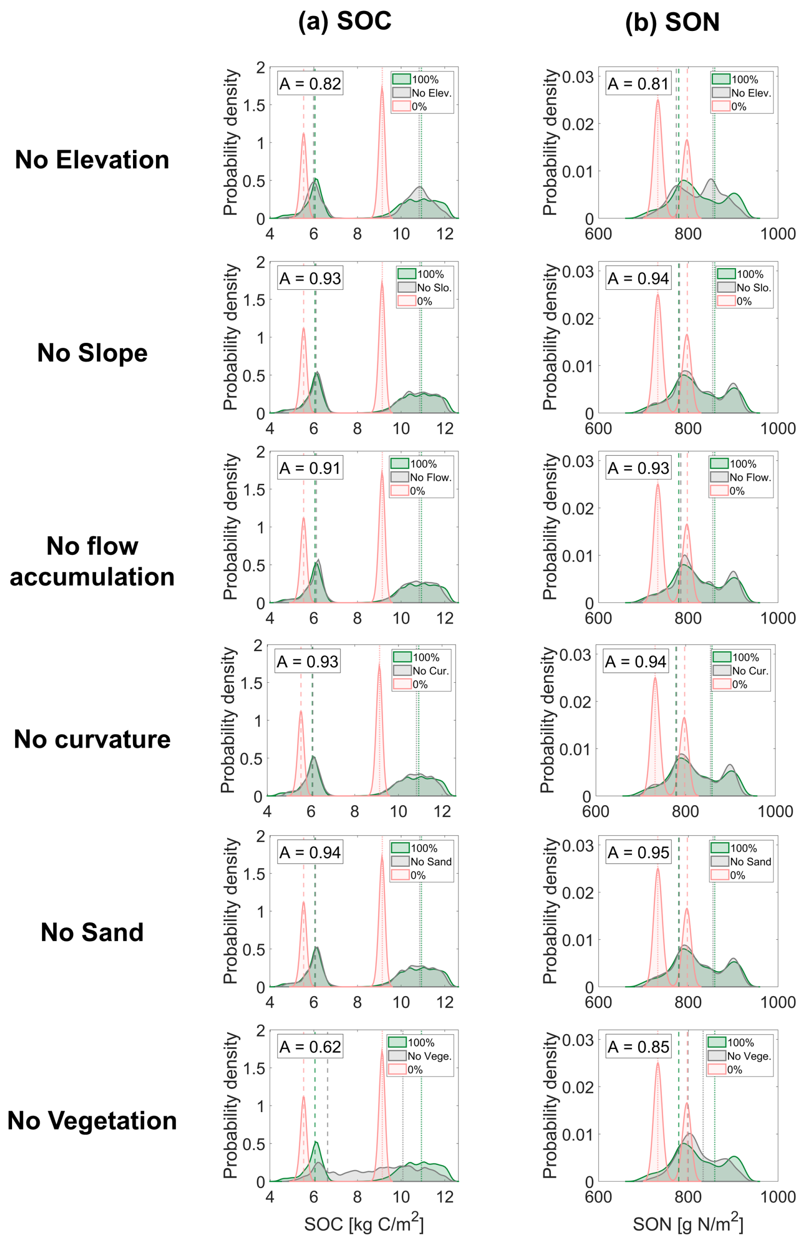

Figure S9. Sensitivity analysis of predictor importance in the Random Forest (RF) model for initializing spatial distributions of (a) soil organic carbon (SOC) and (b) soil organic nitrogen (SON) in the Erlenbach catchment. The analysis is conducted under the original simulation scenario using $n = 40\%$ of tracked cells for RF initialization. Each row shows the probability density functions (PDFs) of model outputs after removing one predictor from the RF training: elevation, slope, flow accumulation area, curvature, sand percentage, and vegetation type. Results are compared with the benchmark simulation using $n = 100\%$ tracked cells and the simulation without RF (No RF). Overlap area between PDFs (A value) quantifies the similarity between the RF-based distributions and the benchmark. Dashed and dotted lines represent median values for grassland and forest areas, respectively.

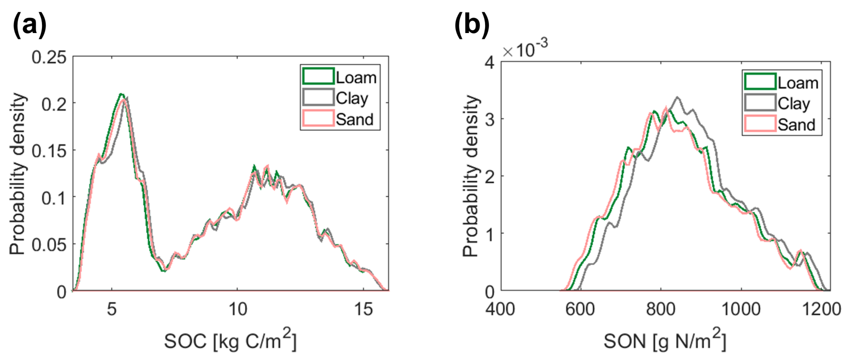

Figure S10. Probability density function of the spatial distribution of soil organic carbon (SOC) and soil organic nitrogen (SON) using three soil texture settings as initial conditions in step (a) in the initialization procedure. The spatial distribution results are the output of step (d) in Fig. 2 in the main text.

References

Basso, B., Gargiulo, O., Paustian, K., Robertson, G. P., Porter, C., Grace, P. R., and Jones, J. W.: Procedures for Initializing Soil Organic Carbon Pools in the DSSAT-CENTURY Model for Agricultural Systems, Soil Science Society of America Journal, 75, 69–78, https://doi.org/10.2136/sssaj2010.0115, 2011.

Bruun, S. and Jensen, L. S.: Initialisation of the soil organic matter pools of the Daisy model, Ecological Modelling, 153, 291–295, https://doi.org/10.1016/S0304-3800(02)00017-0, 2002.

Dimassi, B., Guenet, B., Saby, N. P., Munoz, F., Bardy, M., Millet, F., and Martin, M. P.: The impacts of CENTURY model initialization scenarios on soil organic carbon dynamics simulation in French long-term experiments, Geoderma, 311, 25–36, https://doi.org/10.1016/j.geoderma.2017.09.038, 2018.

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

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

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B., Lawrence, D. M., and Swenson, S. C.: The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10, 7109–7131, https://doi.org/10.5194/bg-10-7109-2013, 2013.

Nemo, Klumpp, K., Coleman, K., Dondini, M., Goulding, K., Hastings, A., Jones, M. B., Leifeld, J., Osborne, B., Saunders, M., Scott, T., Teh, Y. A., and Smith, P.: Soil Organic Carbon (SOC) Equilibrium and Model Initialisation Methods: an Application to the Rothamsted Carbon (RothC) Model, Environmental Modeling & Assessment, 22, 215–229, https://doi.org/10.1007/s10666-016-9536-0, 2017.

Ng, G. C., Bedford, D. R., and Miller, D. M.: A mechanistic modeling and data assimilation framework for Mojave Desert ecohydrology, Water Resources Research, 50, 4662–4685, https://doi.org/10.1002/2014WR015281, 2014.

Qu, Y., Maksyutov, S., and Zhuang, Q.: Technical Note: An efficient method for accelerating the spin-up process for process-based biogeochemistry models, Biogeosciences, 15, 3967–3973, https://doi.org/10.5194/bg-15-3967-2018, 2018.

Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, Ecological Modelling, 189, 25–48, https://doi.org/10.1016/j.ecolmodel.2005.04.008, 2005.

Xia, J. Y., Luo, Y. Q., Wang, Y.-P., Weng, E. S., and Hararuk, O.: A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state, Geoscientific Model Development, 5, 1259–1271, https://doi.org/10.5194/gmd-5-1259-2012, 2012.

Zhan, X., Xue, Y., and Collatz, G.: An analytical approach for estimating CO2 and heat fluxes over the Amazonian region, Ecological Modelling, 162, 97–117, https://doi.org/10.1016/S0304-3800(02)00405-2, 2003.