PERSPECTIVE

Check for updates

Soil organic carbon models need independent time-series validation for reliable prediction

Julia Le Noë[1,2]✉, Stefano Manzoni [3,4], Rose Abramoff [5,6], Tobias Bölscher[7],
Elisa Bruni [1,5], Rémi Cardinael [8,9,10], Philippe Ciais [5], Claire Chenu[7],
Hugues Clivot[11], Delphine Derrien[12], Fabien Ferchaud [13], Patricia Garnier[7],
Daniel Goll [5], Gwenaëlle Lashermes [11], Manuel Martin[14], Daniel Rasse[15],
Frédéric Rees[7], Julien Sainte-Marie[16], Elodie Salmon [5], Marcus Schiedung [17],
Josh Schimel [18], William Wieder [19,20], Samuel Abiven[1], Pierre Barré[1],
Lauric Cécillon [1] & Bertrand Guenet[1]

Numerical models are crucial to understand and/or predict past and future soil organic carbon dynamics. For those models aiming at prediction, validation is a critical step to gain confidence in projections. With a comprehensive review of ~250 models, we assess how models are validated depending on their objectives and features, discuss how validation of predictive models can be improved. We find a critical lack of independent validation using observed time series. Conducting such validations should be a priority to improve the model reliability. Approximately 60% of the models we analysed are not designed for predictions, but rather for conceptual understanding of soil processes. These models provide important insights by identifying key processes and alternative formalisms that can be relevant for predictive models. We argue that combining independent validation based on observed time series and improved information flow between predictive and conceptual models will increase reliability in predictions.

S oils have a potential to mitigate climate change[1,2] by sequestering and storing carbon[3,4] (C). However, warming and land use changes are expected to increase decomposition and lead to soil organic carbon (SOC) losses, causing an amplifying feedback on climate change[5]. Whether soils mitigate or exacerbate climatic changes depends on the balance between (i) C inputs from vegetation and exogenous organic matter, and (ii) C outputs from microbial respiration, dissolved organic C leaching, soil erosion, combustion by wildfires and volatile

[1] Laboratoire de Géologie, École normale supérieure, CNRS, PSL Univ., IPSL, Paris, France. [2] Institut des Sciences de l'Ecologie et de l'Environnement de vParis (CNRS, Sorbonne Université, IRD, INRAE, UPEC, Université Paris-Cité), Sorbonne Université, 4 place Jussieu, 75252 Paris, Cedex 05, France. [3] Department of Physical Geography, Stockholm University, Stockholm SE-106 91, Sweden. [4] Bolin Centre for Climate Research, Stockholm University, Stockholm SE-106 91, Sweden. [5] Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, Gif-sur-Yvette, France. [6] Lawrence Berkley National Laboratory, University of California, Berkeley, CA, USA. [7] University of Paris-Saclay, INRAE, AgroParisTech, UMR EcoSys, 91120 Palaiseau, France. [8] AIDA, Univ Montpellier, CIRAD, Montpellier, France. [9] CIRAD, UPR AIDA, Harare, Zimbabwe. [10] Department of Plant Production Sciences and Technologies, University of Zimbabwe, Harare, Zimbabwe. [11] Université de Reims Champagne-Ardenne, INRAE, FARE, UMR A 614, Reims, France. [12] INRAE, Biogéochimie des Ecosystèmes Forestiers, 54000 Nancy, France. [13] BioEcoAgro Joint Research Unit, INRAE, Université de Liège, Université de Lille, Université de Picardie Jules Verne, 02000 Barenton-Bugny, France. [14] INRAE, InfoSol, Orléans, France. [15] Norwegian Institute of Bioeconomy Research (NIBIO), Division of Environment and Natural Resources, Høgskoleveien 7, 1432 Ås, Norway. [16] Université de Lorraine, AgroParisTech, INRAE, SILVA, F-54000 Nancy, France. [17] Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. [18] Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA. [19] Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA. [20] Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA. ✉email: julia.lenoe@ird.fr

organic compounds emissions[6]. Mathematical models are crucial to understand and predict this balance, as they condense complex processes into a mathematical formalism suitable for quantitative analyses.

Soil organic carbon models translate theoretical hypotheses into a simplified overview of the ecological system described by schematic representations and mathematical equations (Fig. 1a). By confronting model simulations against empirical observations, the underlying hypotheses of a model can be tested (Fig. 1b, c). After testing, models can be used to infer the effects of environmental drivers, such as climate change or land use, on SOC dynamics (Fig. 1c). As a tool to connect data and theory, a model is neither true nor false[7]. Rather, the value of a model comes from its ability to explain and account for a set of phenomena[8], so that the validation process of a model depends on the purpose and context in which it is applied (Box 1).

Since the early 1930s, a wide variety of models to represent the dynamics of SOC have been developed for various spatial and temporal scales, climatic conditions, land-uses and land-covers[9,10]. At least two families of model structures can be identified: (i) approaches relying on conceptual SOC pools decaying according to first order kinetics[11,12], and (ii) approaches that resolve microbial and physical processes controlling SOC decay and stabilisation by describing extra-cellular enzymatic reactions, diffusion and/or sorption kinetics[13,14] (Fig. 1b). In the

former category, first-order kinetics imply that decay rates are proportional to the SOC stocks of the various pools considered, with rate modifiers to implicitly represent the effects of key factors, namely soil temperature, soil moisture, and clay content on microbial and physical processes[15,16]. In the latter category, nonlinear kinetics consider the feedbacks between microbial activity and SOC substrates by representing the decay rate as a function of SOC and/or microbial C stocks[17,18]. Yet, no consensus has emerged on any approach for understanding and predicting SOC dynamics[19]. As a consequence of the wide range of approaches, predicted SOC values exhibit large discrepancies across models, irrespective of model category or temporal and spatial scales[20–22], indicating that more robust validation procedures of SOC models need to be developed[23]. Comparison of model performance is also complicated by lacking standardisation of model validation criteria[24] (Fig. 1c). While validation of models is a critical step to improve confidence in SOC model predictions, so far, no comprehensive review of the different approaches to model validation has been undertaken.

We fill this gap by systematically comparing the validation procedures of ~250 SOC models spanning 90 years of model development history in relation to their scope and main features. We raise three interrelated questions: (i) What are the different SOC model scopes, features and validation procedures in the scientific literature and how have they changed since the first
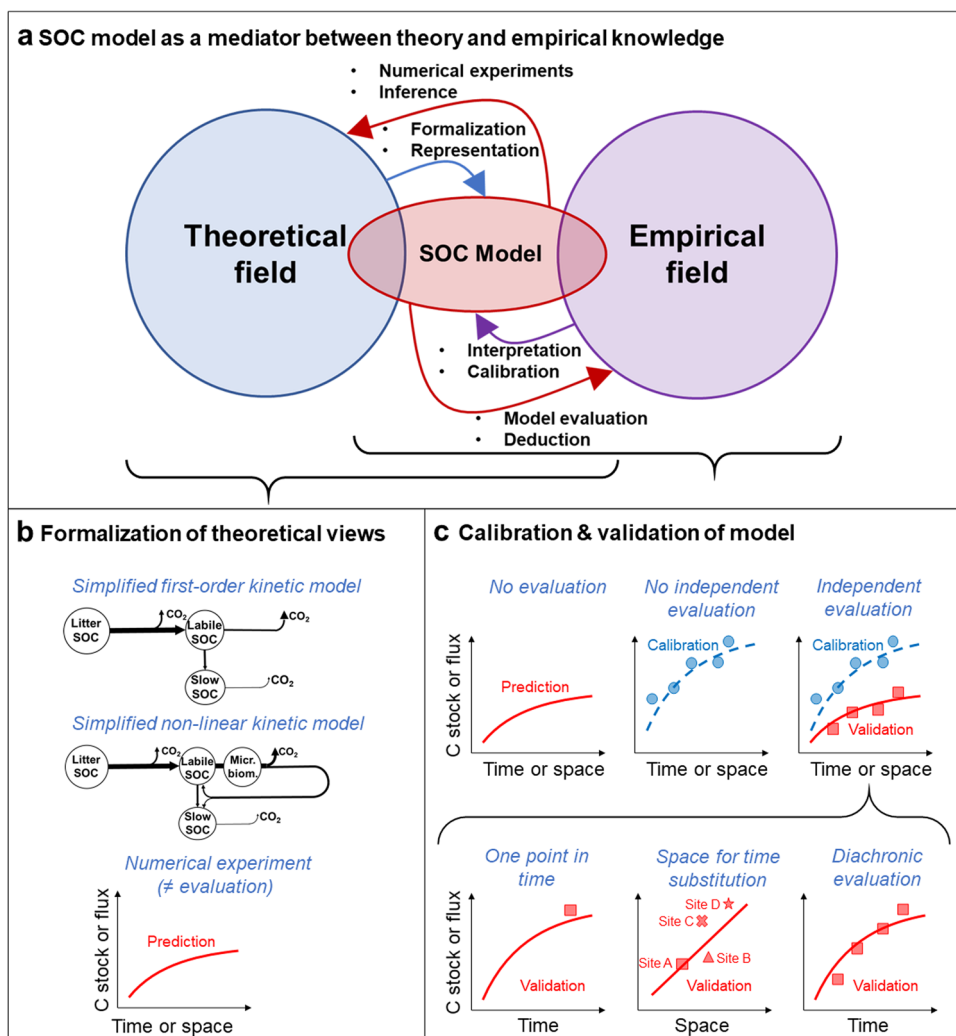
Fig. 1 What is a SOC model? A schematic representation of (**a**) SOC model as a mediator between a theoretical and an empirical field, (**b**) links between models and theory (**c**) as well as between SOC models and empirical field. In **b**, Micr. biom. Stands for microbial biomass.

Box 1 | What is model validation?

Following the seminal definition by Rykiel[24], validation of a model, in the broadest sense, *'is a demonstration that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model'* before specifying that in the case of predictive models, the validation *'involves a comparison of simulated data with data obtained by observation and measurement of the real system'*. In the case of SOC models, the validation most often consists of testing the agreement between simulations (or analytical solutions) and field or laboratory measurements of one or several SOC pools or fluxes[9]. This allows model developers and users to estimate model reliability and accuracy (Fig. 1b).

Independent validation is a necessary condition to assess model robustness in making reliable predictions[24,82]. While calibration involves fitting model parameters to best reproduce some empirical data, independent validation of SOC models tests the model ability to reproduce observations for which it was not calibrated. Therefore, even though several sets of parameters can enable reasonable fit observations, independent validation is needed to identify the most robust set of parameters, i.e., the one that allows reproducing observations in a wide range of contexts. To ensure reliable predictions of SOC dynamics through time, independent validation should compare simulations with diachronic observations (time-series) of SOC stocks or C fluxes. We refer to this type of validation here as 'independent diachronic validation', and describe it more thoroughly hereafter. Another approach to test the ability of a model to capture temporal changes in SOC stocks or C fluxes is the space for time validation, in which simulations are compared to observations along spatial gradients that are caused by processes occurring over different time spans. The assumption that space could be substituted for time can be critical if the land-use history and pedoclimatic contexts differ along the chosen spatial gradient. If that is the case, such contexts act as confounding factors that control, together with time, SOC dynamics[6,83]. As the ability of a model to reproduce temporal shifts in C dynamics is critical for making future projections, examining whether predictive SOC models proceed to independent diachronic validation should be a strong focus of the model development.

SOC models in the 1930's? (ii) How, and to what extent, are different model categories evaluated against observations? (iii) How can the diversity of model's scope and features help enhance the reliability of models aimed at predictions? Answering these questions is crucial for applying and developing adequate SOC models for prediction while leveraging the complementarity of different model scopes and features.

To answer these questions, we summarize model validation methods based on existing definitions in the literature (Box 1) and then conducted a systematic literature review to assess the types of model validation used. The review consisted in four steps: (i) literature screening based on a previous SOC model review from 2009[10] and expert elicitation following a workshop on "Diversity and complementarity in SOC modelling approaches" in October 2021, (ii) literature selection aiming at excluding publications out of the scope of the present study, (iii) systematic literature analysis, and (iv) refinement of literature selection (Fig. S1). Our systematic quantitative review is also complemented by a qualitative analysis of the history of model validation of both first-order and nonlinear SOC models, illustrating different pathways of model validations for three first-order kinetic models (Century, AMG, Q-model) and two nonlinear kinetic models (MIMICS, Millennial) (Figs. S3–5 and Box S1 & S2).

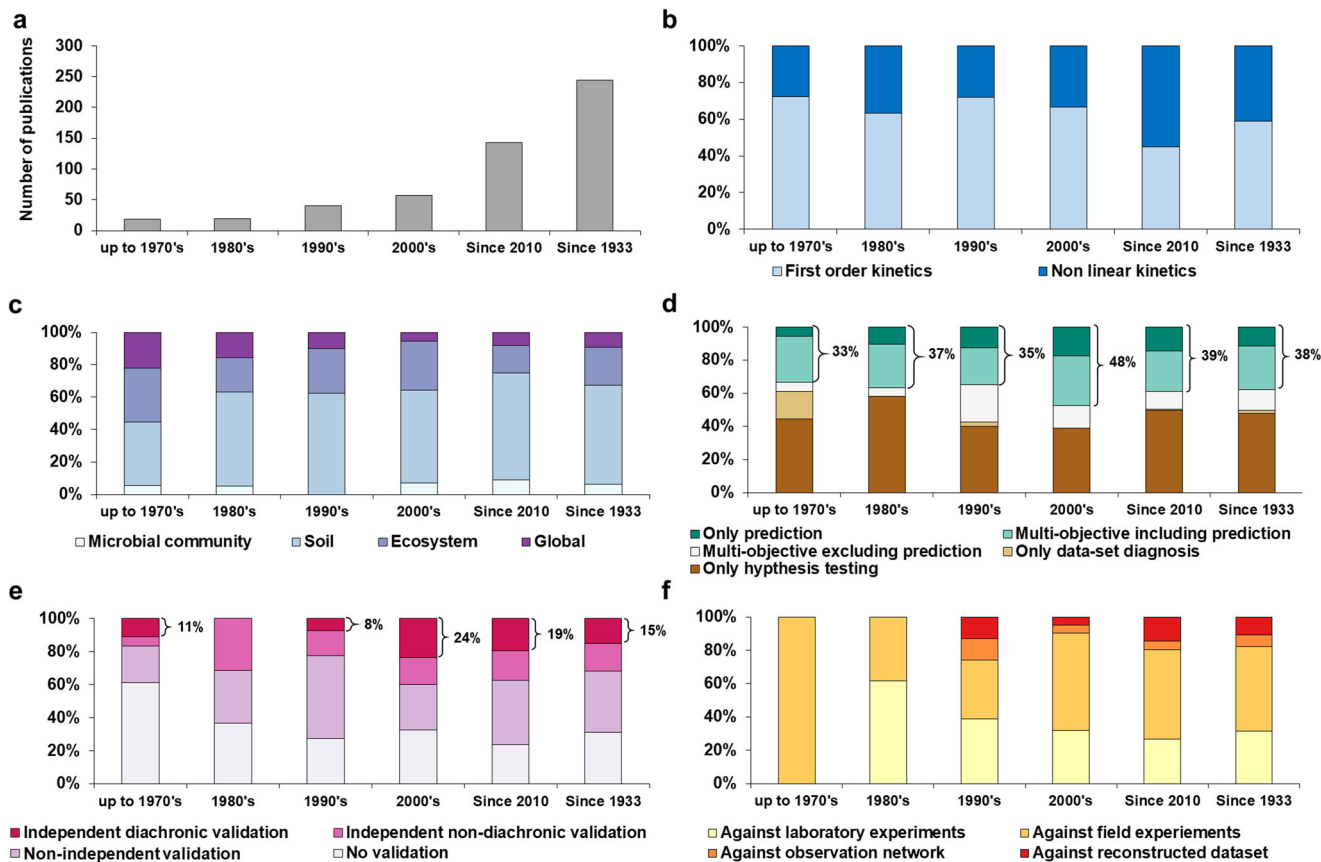**Model features, scopes and validation across time**

Model features, scope, and validation procedures are regarded here as primary data and represented through an historical perspective (Fig. 2), and bar chart analyses (Fig. 3).

Regarding model features, we focused on the representation of decomposition kinetics (Fig. 2b) and the ecological system level of the models (Fig. 2c) (see also SI1). These 'system levels' characterise the context in which models are used, and partly also reflect the spatial scale of application (starting from smallest scales at the 'microbial community' level). We found that 28 to 37% of the 137 models until the 2000's applied non-linear kinetics including both SOC and microbial biomass or enzymes. This number increased to 55% of the 139 models published after 2010 (Fig. 2b). The recent increased number and proportion of SOC models based on nonlinear kinetics can be explained by at least two reasons. First, nonlinear kinetics allow us to account for transient dynamics and feedback responses of soil microorganisms to changing environments and SOC decomposition[25,26]. While linear models rarely capture these effects[27], they were preferred for their simplicity and inherently stable behaviours. In

the last decades the need to include nonlinear effects in the context of climatic changes has motivated a revival of nonlinear models. Second, since the 2010's an increasing number of Earth System Models (ESM) and Dynamic Global Vegetation Models (DGVM) intended to incorporate these feedback effects to predict global C dynamics under climate change[18,28]. This new direction of ESM and DGVM development was probably motivated by their poor SOC prediction performance[29], leading to further publications arguing for inclusion of soil microbial processes in ESM[30]. Yet, most models in the 2010–2021 period were not developed for global scale applications (only 4%, Fig. 2c), with most models describing processes from the soil- to the ecosystem-level (69% and 18%, respectively, of all models in that period, Fig. 2c). A minority of models describe processes at microscopic scales, including models at the microbial community level (9%, Fig. 2c). This dominance of soil- and ecosystem-level models, might reflect the spatial scales at which most data are available (Fig. 2f). Thus, despite the call for including nonlinear kinetics in ESM and DGVM, nonlinear models are still primarily developed for smaller scales. Moreover, models were increasingly needed to provide decision support to enhance SOC sequestration at the plot or farm level, which require local- rather than global-scale predictions[31].

Regarding model scope, we distinguished three objectives: hypothesis-testing and formalisation, data interpretation (diagnostic models), and prediction (prognostic models, see also SI1). Those categories are not exclusive, which means that a given model can fall into more than one category. Because prediction is the scope for which model validation is the most critical, we focus here on this objective (Fig. 2d). Between one third and one half of the models were developed with this objective in mind, with a notable decline between the 2000's and the 2010's, suggesting that most of these newer models are in an early phase of theoretical development and testing before being possibly used for prediction[32,33].

Finally, we characterised validation by considering two relevant axes: the dependence of the validation dataset to the calibration dataset (Fig. 2e) and the source of the validation dataset (Fig. 2f, see also SI1). Four types of model validation dataset are considered: (i) independent diachronic validation, which provides the most confidence in model accuracy[34], (ii) independent non-diachronic validation, which tests model ability to reproduce spatial pattern of SOC stocks, (iii) non-independent validation, which corresponds to model calibration[35,36], and (iv) no validation/calibration (Box 1). Before the 1970's, most models were not

Fig. 2 Temporal evolution of model number, features, scopes and validation procedure. Evolution of (**a**) the number of publications reviewed in the present research and model features (expressed as percentages of all models in each time interval) regarding: (**b**) decomposition kinetics, (**c**) model level of interest, (**d**) model scope, and model validation according to (**e**) the procedures applied and (**f**) validation conditions. Curly brackets in (**d**) group models including 'prediction' among their scopes; curly brackets in (**f**) highlight the percentage of models undergoing independent diachronic validation.
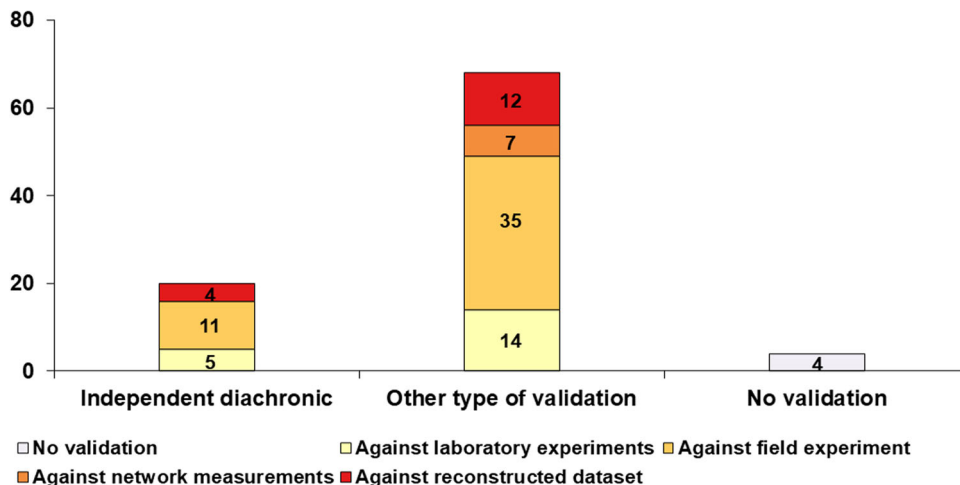
Fig. 3 Validation of models aiming at prediction. Types of validation depending on the source of data used with all SOC models reviewed since 2000—accounting for 92 models in 45 publications. Numbers inside the bar charts indicate the total number of each combination of type and condition of model validation.

evaluated except via qualitative comparisons with observations (Fig. 2e). Since the 1980's, more models have been evaluated, but non-validated models remained an important fraction. However, as the percentage of non-validated models remains fairly constant since the 1990's while the total number of models increases (Fig. 2a), the number of non-validated models has actually increased. Independent diachronic validation, while remaining a

minor fraction, became more widespread in the last two decades (Fig. 2a, e), indicating increased attention to model ability to predict temporal changes in SOC (Fig. 2e).

In addition, we examined four sources of data utilised for validation: (i) against laboratory experiments, (ii) against field experiments, (iii) against observation networks, and (iv) against reconstructed datasets in which observation values from field

measurements are scaled up to construct gridded datasets[37]. Although the validation of SOC models against observation networks and reconstructed data sets increased since the 1990's, we found that the majority of validations are still based on field or laboratory experiments (Fig. 2f), which is expected as the majority of models describe processes at the level of soil or ecosystem (Fig. 2c).

Model features, scopes and validation since 2000
To assess recent trends in SOC model validation, we analysed the relative distributions of model features and scopes depending on the type of validation—either independent diachronic or other types (Fig. S2). We found that 27% of first order kinetics models were validated against diachronic independent data while this was the case for only 15% of nonlinear kinetics models. This indicates that simpler models based on first order kinetics has been validated more thoroughly with independent time-series observations than the 'more mechanistic' microbial models. In contrast, independent validation has been performed in similar proportion to test both models at the soil level or below and models at the ecosystem level or above.

Additionally, we analysed the types of validation used for models proposed since year 2000 and aiming at prediction. To do that, we considered two relevant axes: (i) types and (ii) conditions of model validation. Approximately 40% of model validations were based on field data in predictive models, but were not based on independent and diachronic data (Fig. 3). This proportion has remained almost constant since 1933 (Table S1). Similarly, 15% of validations were based on laboratory data, but they were not independent diachronic (Fig. 3). In several cases, the validation consisted of calibrating the model and assessing its performance against the same calibration data[38,39]. In many cases, independent validation based on space-for-time substitutions was used[17,40–42]. Only 23% of models aiming at prediction were tested using an independent diachronic validation, despite its importance for testing the ability of models to simulate SOC temporal dynamics (Box 1). This group included 14 first order models, namely the LPJ[43], MOMOS[44], CN-SIM[45], CIPS[46], AMG[47], Roth-C[48], PRIM[49], ORCHIDEE-PRIM[50], Yasso[51], N14CP-Agri[52], CASA-CNP[53], DAISY[54] models, as well as three unnamed models[55,56], and six nonlinear models, namely the SOMKO[57], Ecosys[58], BACWAVE-WEB[59], MIMICS[28], CORPSE[53] and one unnamed model[60]. Only 7 models (8%) were evaluated against network measurements at the territorial- or national- level, but not diachronically and independently. However, type of data does not seem to covary with evaluation type, in that a fairly constant proportion of model validations are independent diachronic (26% for lab experiments, 24% for field, 25% for reconstructed).

Recommendations for model validation
Our review identified a lack of independent diachronic validation based on SOC network measurements (Fig. 3). Filling this gap will be crucial to evaluate objectively regional to global SOC predictions in the context of climate and land use changes. Reuse of highly valuable data sets and novel measurement networks will be central to this effort. Both incubation and field experiments provide valuable empirical observations to evaluate SOC models. In particular, existing decadal field experiments[5,61] represent a valuable source of data. They provide long-term time-series data on SOC stocks in different climatic conditions, but remain underutilised. Given the time required to develop high value long-term data sets and measurement networks, much increased access to, valorization and more extensive use of these resources will be key to this effort. At a larger scale, validations against observations from measurement networks allow the predictive

value of SOC models to be assessed under a vast range of land-use and pedoclimatic contexts.

National and macro-regional soil monitoring networks are routinely resampled and data are becoming more available (e.g., with the second campaign of the French Soil Network Measurement[44] and fourth resampled of the European LUCAS topsoil network[62] and with the regional or national soil monitoring systems existing in 18 European countries[63]), and will therefore enable validating models in a large set of contexts and at large spatial and temporal scales. To this end, some existing limitations still need to be overcome. In particular, the different soil databases already in use[64–66] require harmonisation in terms of spatial resolution, reported variables and measurement methods. Harmonized datasets will provide reliable model input variables[30], such as C inputs from plants or amendments[67] and initialisation of SOC pools[68], allowing meaningful model inter-comparisons. These limitations have been identified as a key research priority to provide compatible benchmark data sets for a consistent validation of SOC models. In the case of SOC model implementation within ESMs, an additional barrier is that ESMs often lack the adequate forcing data at the site resolution. Because data scarcity is still largely limiting large-scale diachronic validation, global datasets, such as the Soil Respiration Database[69], remain invaluable to validate the ability of models to predict the spatial variability of C fluxes from soils—for which time-series are not a prerequisite. However, these datasets are insufficient to validate the predictive accuracy of the temporal dynamics of SOC stocks.

Models tend to become increasingly complex, including more and more biogeochemical processes[70] (Box S2). Yet, adding processes or compartments to SOC models might result in hidden compensating biases arising from the overfitting of model parameters[21,71]. Therefore, evaluating model ability to reproduce newly explicitly represented processes and/or soil compartments is required to address this bias. To this end, validation of new models including e.g., microbial processes should be conducted on field experiments that allow controlling the effects of environmental or management changes on these processes. More generally, it is worth emphasising that the more observations are collected—not only on SOC stocks and SOC compartments, but also on bulk density, vertical SOC profile, clay content, moisture content, C inputs from plants, microbial biomass, etc.—the better it is for stringently validating all model outputs.

To summarize, the next steps in model validation to improve reliability and accuracy should be, in order of priority: (i) To gather and homogenize time series of long-term field trials and measurement networks that could be used as benchmark datasets to systematically validate independently and diachronically all models aimed at predicting SOC dynamics. Such a benchmark dataset would allow stringent comparisons of SOC models, helping to choose the best-suited models depending on the spatio-temporal scale of interest, land-use and pedoclimatic contexts. (ii) To maintain and develop SOC monitoring networks to build spatially-explicit maps of the temporal dynamics of SOC stocks, which in turn will allow validating SOC models from regional to global scales (this type of validation is currently sorely underrepresented, Fig. 3). (iii) To avoid compensating biases, to validate models incorporating new representations of soil processes and/or soil compartment against independent time-series data specifically related to those processes and compartments.

Model diversity can help improve prediction ability
Each model considers a number of biogeochemical processes governing the dynamics of SOC, which is translated into their mathematical formalism (Fig. 1). Can we take advantage of the

diversity of SOC models to improve those aimed at predictions or projections and evaluate their relative performance? To answer this question, let us consider the common object of interest for all these models – the dynamics of SOC – and the role of models as a mediator between empirical and theoretical knowledge[72] (Fig. 1).

Different SOC modelling approaches, reflecting distinct theoretical assumptions, can use the same empirical observations for calibration or validation[53,73]. Comparing the accuracies of SOC models describing different processes with respect to the same set of observations allows us to identify the mathematical formalism achieving the best performance. For instance, Lawrence et al.[25] developed four SOC models of varying mechanistic complexity and compared their ability to simulate soil respiration observed in a laboratory incubation experiment to test which mechanisms improve model performance. However, the use of the models that showed the best performance might be hindered by the low-quality or lack of available input data and parameters when the intended scale is larger or longer than the one for which models were tested and parametrized, for instance when moving from the laboratory incubation to the field level. To overcome this barrier, new indicators based on widely available observational data could be developed to transfer processes identified as important by complex models toward simpler models. One example is the integration of the priming effect (enhanced or inhibited decomposition of native SOC by added fresh organic matter) in SOC models. Priming can be described by modelling microbial biomass and enzymatic reactions (Michaelis-Menten kinetics)[25] (Fig. 1c), but including these processes could be hampered by the lack of reliable input data and parameter estimates at the intended scale of application. To circumvent this issue, an empirical relationship between C input and mineralization rate can be used to surrogate explicit microbial processes and thus represent the priming effect in a simple way[49,74]. Moreover, predictive models can benefit from the comparison with more complex models designed for other goals, as such comparisons provide a safeguard against being locked into one type of model and help to identify processes that may have been neglected or incorrectly represented at the scale at which predictions are made[25,49].

Different SOC models can also be used to run equivalent numerical experiments to compare how their mechanistic representations affect the simulations. Bridging empirical knowledge, condensed into model calibrations, and theoretical knowledge gained by analysing the simulation results, has two virtues in terms of model complementarities. First, it enables researchers to explore how feedback mechanisms can result in distinct behaviours in different SOC models[26,75,76]. For instance, Sainte-Marie et al.[75] explored how different conceptualisations and mathematical formalisations of SOC depolymerisation by different decomposer groups impact the chemistry and amount of organic matter during decomposition and at steady state. This analysis has a heuristic value in itself, independent of the subsequent validation of the model[24]. Second, bridging empirical and theoretical knowledge allows for testing the sensitivity of different models to shifts in driving variables[20,76]. For instance, Ito et al.[20] highlighted the high sensitivities and uncertainties on the impact of land-use on future SOC dynamics at the global level in 15 DGVMs.

Model diversity can also result in synergies when models are implemented into ensemble modelling approaches[77–79]. Considering multiple structurally diverse models for prediction allows to estimate the uncertainty of the simulated variables due to the different processes represented and their sensitivities to driving factors[80]. Weighted averages and model selection based on performance criteria reduce the prediction error and provide more robust predictions, relative to single model simulations[77]. However, standardizing modelling methods remains a fundamental

issue in multi-model inter-comparison exercises[77–79]. For example, the choice of the parameters to calibrate, the initialization method to use for conceptually different soil compartments, and the estimation of forcing variables make model inter-comparisons difficult. Protocols are needed that also account for data availability[81]. Multi-model ensembles thus represent promising tools to account for uncertainties in the simulations and provide greater reliability than individual models.

Outlook
Soil organic carbon models are expected to predict SOC stock changes and to provide a sound biogeochemical context for simulating coupled soil-vegetation interactions. Yet, reliable model validation lags behind increasing model complexity. Based on a systematic review of ~250 models over 90 years, we advocate for stringent independent diachronic validation of prediction-aimed models at all scales based on sites and networks that provide time series of SOC stocks and/or C fluxes at plot, national and regional scales. Continued efforts to maintain these datasets are thus imperative to increase reliability and accuracy of SOC projections and predictions. In parallel to observation-based validation, predictive models can also be conceptually evaluated by comparing them to models designed for other goals, to identify processes that might have been neglected or incorrectly represented at the scale at which predictions are made.

References
1. IPCC. Summary for Policymakers. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, et al., (eds.)]. (2019).
2. Roe, S. et al. Contribution of the land sector to a 1.5 °C world. Nat. Clim. Chang. 9, 817–828 (2019).
3. Ciais, P. et al. Carbon and other biogeochemical cycles. in Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change 465–570 (Cambridge University Press, 2014).
4. Bossio, D. A. et al. The role of soil carbon in natural climate solutions. Nat. Sustain 3, 391–398 (2020).
5. Walker, T. W. N. et al. Microbial temperature sensitivity and biomass change explain soil carbon loss with warming. Nat. Clim. Change 8, 885–889 (2018).
6. Chenu, C. et al. Increasing organic stocks in agricultural soils: knowledge gaps and potential innovations. Soil Tillage Res. 188, 41–52 (2019).
7. Popper, K. Conjectures and refutations: The growth of scientific knowledge. (Routledge, 1963).
8. Dahan, A. Chapitre 10. Épistémologie de la modélisation, le cas des modèles de climat. In Modélisation de l'environnement: entre natures et sociétés 193–208 https://doi.org/10.3917/quae.herve.2009.01.0193 (Éditions Quæ, 2009).
9. Campbell, E. E. & Paustian, K. Current developments in soil organic matter modeling and the expansion of model applications: a review. Environ. Res. Lett. 10, 123004 (2015).
10. Manzoni, S. & Porporato, A. Soil carbon and nitrogen mineralization: theory and models across scales. Soil Biol. Biochem. 41, 1355–1379 (2009).
11. Falloon, P. & Smith, P. Simulating SOC changes in long-term experiments with RothC and CENTURY: model evaluation for a regional scale application. Soil Use Manag. 18, 101–111 (2006).

12. Parton, W. J. & Rasmussen, P. E. Long-term effects of crop management in wheat-fallow: II. CENTURY model simulations. *Soil Sci. Soc. Am. J.* **58**, 530–536 (1994).

13. Abramoff, R. et al. The Millennial model: in search of measurable pools and transformations for modeling soil carbon in the new century. *Biogeochemistry* **137**, 51–71 (2018).

14. Zhang, H. et al. Microbial dynamics and soil physicochemical properties explain large-scale variations in soil organic carbon. *Glob. Change Biol.* **26**, 2668–2685 (2020).

15. Luo, Z., Feng, W., Luo, Y., Baldock, J. & Wang, E. Soil organic carbon dynamics jointly controlled by climate, carbon inputs, soil properties and soil carbon fractions. *Glob. Change Biol.* **23**, 4430–4439 (2017).

16. Xu, X. et al. Soil properties control decomposition of soil organic carbon: Results from data-assimilation analysis. *Geoderma* **262**, 235–242 (2016).

17. Georgiou, K., Abramoff, R. Z., Harte, J., Riley, W. J. & Torn, M. S. Microbial community-level regulation explains soil carbon responses to long-term litter manipulations. *Nat. Commun.* **8**, 1223 (2017).

18. Wieder, W. R., Bonan, G. B. & Allison, S. D. Global soil carbon projections are improved by modelling microbial processes. *Nat. Clim. Change* **3**, 909–912 (2013).

19. Menichetti, L., Ågren, G. I., Barré, P., Moyano, F. & Kätterer, T. Generic parameters of first-order kinetics accurately describe soil organic matter decay in bare fallow soils over a wide edaphic and climatic range. *Sci. Rep.* **9**, 20319 (2019).

20. Ito, A. et al. Soil carbon sequestration simulated in CMIP6-LUMIP models: implications for climatic mitigation. *Environ. Res. Lett.* **15**, 124061 (2020).

21. Shi, Z., Crowell, S., Luo, Y. & Moore, B. Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nat. Commun.* **9**, 2171 (2018).

22. Wieder, W. R. et al. Carbon cycle confidence and uncertainty: exploring variation among soil biogeochemical models. *Glob. Change Biol.* **24**, 1563–1579 (2018).

23. He, Y. et al. Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century. *Science* **353**, 1419–1424 (2016).

24. Rykiel, E. J. Testing ecological models: the meaning of validation. *Ecol. Modelling* **90**, 229–244 (1996).

25. Lawrence, C. R., Neff, J. C. & Schimel, J. P. Does adding microbial mechanisms of decomposition improve soil organic matter models? A comparison of four models using data from a pulsed rewetting experiment. *Soil Biol. Biochem.* **41**, 1923–1934 (2009).

26. Wutzler, T. & Reichstein, M. Colimitation of decomposition by substrate and decomposers – a comparison of model formulations. *Biogeosciences* **5**, 749–759 (2008).

27. Diele, F., Luiso, I., Marangi, C. & Martiradonna, A. SOC-reactivity analysis for a newly defined class of two-dimensional soil organic carbon dynamics. *Appl. Math. Model.* **118**, 1–21 (2023).

28. Wieder, W. R. et al. Explicitly representing soil microbial processes in Earth system models: Soil microbes in earth system models. *Glob. Biogeochem. Cycles* **29**, 1782–1800 (2015).

29. Todd-Brown, K. E. O., Hopkins, F. M., Kivlin, S. N., Talbot, J. M. & Allison, S. D. A framework for representing microbial decomposition in coupled climate models. *Biogeochemistry* **109**, 19–33 (2012).

30. Luo, Y. et al. Toward more realistic projections of soil carbon dynamics by Earth system models. *Glob. Biogeochem. Cycles* **30**, 40–56 (2016).

31. Van de Broek, M. et al. Assessing the climate regulation potential of agricultural soils using a decision support tool adapted to stakeholders' needs and possibilities. *Front. Environ. Sci.* **7**, 131 (2019).

32. la Cecilia, D., Riley, W. J. & Maggi, F. Biochemical modeling of microbial memory effects and catabolite repression on soil organic carbon compounds. *Soil Biol. Biochem.* **128**, 1–12 (2019).

33. Wutzler, T., Zaehle, S., Schrumpf, M., Ahrens, B. & Reichstein, M. Adaptation of microbial resource allocation affects modelled long term soil organic matter and nutrient cycling. *Soil Biol. Biochem.* **115**, 322–336 (2017).

34. Barré, P. et al. La séquestration de carbone dans les sols agricoles, forestiers et urbains: état des lieux des méthodes d'évaluation et de quantification. *Etude et Gestion des Sols* **27**, 305–320 (2020).

35. Ghezzehei, T. A., Sulman, B., Arnold, C. L., Bogie, N. A. & Berhe, A. A. On the role of soil water retention characteristic on aerobic microbial respiration. *Biogeosciences* **16**, 1187–1209 (2019).

36. Ginovart, M., López, D. & Gras, A. Individual-based modelling of microbial activity to study mineralization of C and N and nitrification process in soil. *Nonlinear Anal. Real World Appl.* **6**, 773–795 (2005).

37. Hashimoto, S. et al. Global spatiotemporal distribution of soil respiration modeled using a global database. *Biogeosciences* **12**, 4121–4132 (2015).

38. Ahrens, B. et al. Bayesian calibration of a soil organic carbon model using measurements of soil organic carbon and heterotrophic respiration as joint constraints. *Biogeosciences* **11**, 2147–2168 (2014).

39. Flores, O. et al. KEYLINK: towards a more integrative soil representation for inclusion in ecosystem scale models—II: model description, implementation and testing. *PeerJ* **9**, e10707 (2021).

40. Huang, Y., Guenet, B., Wang, Y. L. & Ciais, P. Global simulation and evaluation of soil organic matter and microbial carbon and nitrogen stocks using the microbial decomposition model ORCHIMIC v2.0. *Glob. Biogeochem. Cycles* **35**, 1–20 (2021).

41. Keyvanshokouhi, S. et al. Effects of soil process formalisms and forcing factors on simulated organic carbon depth-distributions in soils. *Sci. Total Environ.* **652**, 523–537 (2019).

42. Saito, K., Machiya, H., Iwahana, G., Yokohata, T. & Ohno, H. Numerical model to simulate long-term soil organic carbon and ground ice budget with permafrost and ice sheets (SOC-ICE-v1.0). *Geosci. Model Dev.* **14**, 521–542 (2021).

43. Sitch, S. et al. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model: LPJ dynamic global vegetation model. *Glob. Change Biol.* **9**, 161–185 (2003).

44. Pansu, M., Bottner, P., Sarmiento, L. & Metselaar, K. Comparison of five soil organic matter decomposition models using data from a ^{14}C and ^{15}N labeling field experiment: Comparison of five soil organic matter models. *Glob. Biogeochem. Cycles* **18**, n/a–n/a (2004).

45. Petersen, B. M. et al. CN-SIM: a model for the turnover of soil organic matter. II. Short-term carbon and nitrogen development. *Soil Biol. Biochem.* **37**, 375–393 (2005).

46. Kuka, K., Franko, U. & Rühlmann, J. Modelling the impact of pore space distribution on carbon turnover. *Ecol. Model.* **208**, 295–306 (2007).

47. Saffih-Hdadi, K. & Mary, B. Modeling consequences of straw residues export on soil organic carbon. *Soil Biol. Biochem.* **40**, 594–607 (2008).

48. Farina, R., Coleman, K. & Whitmore, A. P. Modification of the RothC model for simulations of soil organic C dynamics in dryland regions. *Geoderma* **200–201**, 18–30 (2013).

49. Guenet, B., Moyano, F. E., Peylin, P., Ciais, P. & Janssens, I. A. Towards a representation of priming on soil carbon decomposition in the global land biosphere model ORCHIDEE (version 1.9.5.2). *Geosci. Model Dev.* **9**, 841–855 (2016).

50. Guenet, B. et al. Impact of priming on global soil carbon stocks. *Glob. Change Biol.* **24**, 1873–1883 (2018).

51. Mao, Z. et al. Modeling soil organic carbon dynamics in temperate forests with Yasso07. *Biogeosciences* **16**, 1955–1973 (2019).

52. Janes-Bassett, V., Davies, J., Rowe, E. C. & Tipping, E. Simulating long-term carbon nitrogen and phosphorus biogeochemical cycling in agricultural environments. *Sci. Total Environ.* **714**, 136599 (2020).

53. Jian, J. et al. Leveraging observed soil heterotrophic respiration fluxes as a novel constraint on global-scale models. *Glob. Change Biol.* **27**, 5392–5403 (2021).

54. Laub, M. et al. DRIFTS band areas as measured pool size proxy to reduce parameter uncertainty in soil organic matter models. *Biogeosciences* **17**, 1393–1413 (2020).

55. Nicolardot, B., Recous, S. & Mary, B. Simulation of C and N mineralisation during crop residue decomposition: A simple dynamic model based on the C:N ratio of the residues. *Plant Soil* **228**, 83–103 (2001).

56. Meurer, K. H. E. et al. Modelling dynamic interactions between soil structure and the storage and turnover of soil organic matter. *Biogeosciences* **17**, 5025–5042 (2020).

57. Gignoux, J. et al. Design and test of a generic cohort model of soil organic matter decomposition: the SOMKO model: SOMKO: a generic decomposition model. *Glob. Ecol. Biogeogr.* **10**, 639–660 (2001).

58. Grant, R. Modeling transformations of soil organic carbon and nitrogen at differing scales of complexity. Modeling carbon and nitrogen dynamics for soil management (2001).

59. Zelenev, V. V., van Bruggen, A. H. C., Leffelaar, P. A., Bloem, J. & Semenov, A. M. Oscillating dynamics of bacterial populations and their predators in response to fresh organic matter added to soil: The simulation model 'BACWAVE-WEB'. *Soil Biol. Biochem.* **38**, 1690–1711 (2006).

60. Chakrawal, A., Herrmann, A. M. & Manzoni, S. Leveraging energy flows to quantify microbial traits in soils. *Soil Biol. Biochem.* **155**, 108169 (2021).

61. Cardinael, R. et al. Sustaining maize yields and soil carbon following land clearing in the forest–savannah transition zone of West Africa: Results from a 20-year experiment. *Field Crops Res.* **275**, 108335 (2022).

62. Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A. & Fernández-Ugalde, O. LUCAS Soil, the largest expandable soil dataset for Europe: a review. *Eur. J. Soil Sci.* **69**, 140–153 (2018).

63. Bispo, A., Arrouays, D., Saby, N. P. A., Boulonne, L. & Fantappiè, M. Towards climate-smart sustainable management of agricultural soils. EJP SOIL-Deliverable 6.3. Proposal of methodological development for the LUCAS programme in accordance with national monitoring programmes. 1–135 (2021). https://ejpsoil.eu/fileadmin/projects/ejpsoil/WP6/EJP_SOIL_Deliverable_6.3_Dec_2021_final.pdf

64. van Leeuwen, J. P. et al. Gap assessment in current soil monitoring networks across Europe for measuring soil functions. *Environ. Res. Lett.* **12**, 124007 (2017).

65. Chen, S. et al. Model averaging for mapping topsoil organic carbon in France. *Geoderma* **366**, 114237 (2020).

66. Malhotra, A. et al. The landscape of soil carbon data: emerging questions, synergies and databases. *Prog. Phys. Geography: Earth Environ.* **43**, 707–719 (2019).

67. Riggers, C., Poeplau, C., Don, A., Frühauf, C. & Dechow, R. How much carbon input is required to preserve or increase projected soil organic carbon stocks in German croplands under climate change? *Plant Soil* **460**, 417–433 (2021).

68. Dimassi, B. et al. The impacts of CENTURY model initialization scenarios on soil organic carbon dynamics simulation in French long-term experiments. *Geoderma* **311**, 25–36 (2018).

69. Jian, J. et al. Soil CollectionA Global Database of Soil Respiration Data, Version 5.0. 0 MB https://doi.org/10.3334/ORNLDAAC/1827 (2021).

70. Abramoff, R. Z. et al. Improved global-scale predictions of soil carbon stocks with Millennial Version 2. *Soil Biol. Biochem.* **164**, 108466 (2022).

71. Carvalhais, N. et al. Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval: Steady state effects on c flux modeling. *Glob. Biogeochem. Cycles* **22**, n/a–n/a (2008).

72. Blankinship, J. C. et al. Improving understanding of soil organic matter dynamics by triangulating theories, measurements, and models. *Biogeochemistry* **140**, 1–13 (2018).

73. Sulman, B. N. et al. Multiple models and experiments underscore large uncertainty in soil carbon dynamics. *Biogeochemistry* **141**, 109–123 (2018).

74. Mary, B., Clivot, H., Blaszczyk, N., Labreuche, J. & Ferchaud, F. Soil carbon storage and mineralization rates are affected by carbon inputs rather than physical disturbance: evidence from a 47-year tillage experiment. *Agriculture, Ecosystems & Environment* **299**, 106972 (2020).

75. Sainte-Marie, J. et al. C-STABILITY an innovative modeling framework to leverage the continuous representation of organic matter. *Nat Commun* **12**, 810 (2021).

76. Todd-Brown, K. E. O. et al. Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences* **10**, 1717–1736 (2013).

77. Riggers, C. et al. Multi-model ensemble improved the prediction of trends in soil organic carbon stocks in German croplands. *Geoderma* **345**, 17–30 (2019).

78. Bruni, E. et al. Multi-modelling predictions show high uncertainty of required carbon input changes to reach a 4‰ target. *Eur. J Soil Sci.* **73**, 1–26 (2022).

79. Farina, R. et al. Ensemble modelling, uncertainty and robust predictions of organic carbon in long-term bare-fallow soils. *Glob. Change Biol.* **27**, 904–928 (2021).

80. Tebaldi, C. & Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. *Phil. Trans. R. Soc. A* **365**, 2053–2075 (2007).

81. Eyring, V. et al. Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).

82. Araújo, M. B., Pearson, R. G., Thuiller, W. & Erhard, M. Validation of species–climate impact models under climate change. *Glob. Change Biol.* **11**, 1504–1513 (2005).

83. Jonard, M. et al. Forest soils in France are sequestering substantial amounts of carbon. *Sci. Total Environ.* **574**, 616–628 (2017).

Author contributions

J.L.N. organized the workshop on "Diversity and complementarity in SOC modelling approaches" that brought together an international panel together an international panel of senior and early-career scientists, who are all authors of the present research, wrote the first draft of the manuscript, led and contributed to the literature review and revised the manuscript. S.M., took part in the workshop, contributed to writing the first draft, provided to the team database on SOC model he previously established, contributed to the literature review and contributed to revising the manuscript. R.A., took part in the workshop, reviewed the first draft, contributed to the literature review, T.B., took part in the workshop, reviewed the first draft. E.B., took part in the workshop, reviewed the first draft, contributed to the literature review and contributed to revising the manuscript. R.C., took part in the workshop, reviewed the first draft. P.C. took part in the workshop, reviewed the first draft, and contributed to revising the manuscript. C.C., took part in the workshop, reviewed the first draft, and contributed to revising the manuscript. H.C., took part in the workshop, reviewed the first draft, contributed to the literature review. D.D. took part in the workshop, reviewed the first draft. F.F. took part in the workshop, reviewed the first draft, contributed to the literature review. P.G. took part in the workshop, reviewed the first draft. D.G., took part in the workshop, reviewed the first draft, contributed to the revising the manuscript. G.L., took part in the workshop, reviewed the first draft. M.M., took part in the workshop, reviewed the first draft. D.R., took part in the workshop, reviewed the first draft. F.R., took part in the workshop, reviewed the first draft. J.S.-M., took part in the workshop, reviewed the first draft. E.S., took part in the workshop, reviewed the first draft. M.S., took part in the workshop, reviewed the first draft and contributed to revising the manuscript. J.S., took part in the workshop, reviewed the first draft and contributed to revising the manuscript. W.W. took part in the workshop, reviewed the first draft. S.A., co-organized the workshop, contributed to writing the first draft, contributed to the literature review. P.B., co-organized the workshop, contributed to writing the first draft. L.C., co-organized the workshop, reviewed the first draft. B.G., co-organized the workshop, contributed to writing the first draft, contributed to the literature review.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-00830-5.

Correspondence and requests for materials should be addressed to Julia Le Noë.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Clare Davis. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.