Key Points:

- Earth system models (ESMs) predicted global soil C gains under the high emission scenario during the 21st century
- Dominant processes underlying soil C change at global or biome scales differed among models
- Model structure accounted for more than 60% variation of the global soil C change predicted by Coupled Model Intercomparison Project phase 5 ESMs, but only 36% for Coupled Model Intercomparison Project phase 6 models

Global-Scale Convergence Obscures Inconsistencies in Soil Carbon Change Predicted by Earth System Models

Zheng Shi[1,2] , Forrest M. Hoffman[1,3] , Min Xu[1], Umakant Mishra[4,5], Steven D. Allison[6,7], Jizhong Zhou[2,8,9,10], and James T. Randerson[6]

[1]Computational Sciences and Engineering Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [2]Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA, [3]Department of Civil & Environmental Engineering, University of Tennessee, Knoxville, TN, USA, [4]Bioscience Division, Sandia National Laboratories, Livermore, CA, USA, [5]Joint Bioenergy Institute, Emeryville, CA, USA, [6]Department of Earth System Science, University of California-Irvine, Irvine, CA, USA, [7]Department of Ecology and Evolutionary Biology, University of California-Irvine, Irvine, CA, USA, [8]School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK, USA, [9]School of Computer Science, University of Oklahoma, Norman, OK, USA, [10]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Abstract Soil carbon (C) responses to environmental change represent a major source of uncertainty in the global C cycle. Feedbacks between soil C stocks and climate drivers could impact atmospheric CO_2 levels, further altering the climate. Here, we assessed the reliability of Earth system model (ESM) predictions of soil C change using the Coupled Model Intercomparison Project phases 5 and 6 (CMIP5 and CMIP6). ESMs predicted global soil C gains under the high emission scenario, with soils taking up 43.9 Pg (95% CI: 9.2–78.5 Pg) C on average during the 21st century. The variation in global soil C change declined significantly from CMIP5 (with average of 48.4 Pg [95% CI: 2.0–94.9 Pg] C) to CMIP6 models (with average of 39.3 Pg [95% CI: 23.9–54.7 Pg] C). For some models, a small C increase in all biomes contributed to this convergence. For other models, offsetting responses between cold and warm biomes contributed to convergence. Although soil C predictions appeared to converge in CMIP6, the dominant processes driving soil C change at global or biome scales differed among models and in many cases between earlier and later versions of the same model. Random Forest models, for soil carbon dynamics, accounted for more than 63% variation of the global soil C change predicted by CMIP5 ESMs, but only 36% for CMIP6 models. Although most CMIP6 models apparently agree on increased soil C storage during the 21st century, this consensus obscures substantial model disagreement on the mechanisms underlying soil C response, calling into question the reliability of model predictions.

Plain Language Summary In this study, we analyzed and compared the changes in soil C predicted by multiple Earth system models (ESMs). The ESMs projected that under a high emission scenario, there would be a global increase in soil C during the 21st century, with an average uptake of 43.9 Pg C. The variation in global soil C change decreased among ESMs suggesting greater agreement among the models in the newer phase. However, the processes driving soil C change at the global or biome level varied among models, and in some cases, even within the same models between different versions. The differences in model structures explained a significant portion of the variation in soil C change predicted the ESMs. Although there was a consensus among the ESMs that soil C storage would increase during the 21st century, there were substantial disagreements regarding the specific mechanisms and spatial patterns of soil C response to climate change. This indicates that while models were converging on the overall trend, there were significant differences in how they represented the processes influencing soil C dynamics.

1. Introduction

Future carbon (C) storage in soil and its potential for climate mitigation remain uncertain in all global change scenarios, especially at a global scale. On one hand, soils could gain soil organic C (SOC) due to CO_2 fertilization via plant photosynthesis (Terrer et al., 2021) and land management (Bossio et al., 2020). On the other hand, empirical evidence points to large C losses from warming (Nottingham et al., 2020; Soong et al., 2021) and associated disturbance (Pellegrini et al., 2018; Walker et al., 2019).

Earth system models (ESMs) are the main tool for predicting global SOC change under future climate scenarios. Although society relies on ESMs for designing climate policy and mitigation strategies, ESM predictions remain uncertain. Global SOC changes predicted by ESMs in the fifth Coupled Model Intercomparison Project (CMIP5) range from a 270 Pg C gain to a 50 Pg C loss (Todd-Brown et al., 2014). Consequently, researchers have sought to improve the reliability of SOC models in ESMs. Reliable models are both accurate—their encoded mechanisms and outputs match the real world—and certain, meaning that their predictions are consistent. Accurate models are not very useful in decision-making if their predictions are highly uncertain. Models that are certain but inaccurate can be misleading because they make biased, incorrect predictions with apparently high confidence. Current ESM projections in soil carbon have low confidence due to outdated knowledge about soil C turnover (Todd-Brown et al., 2014). Addressing structural uncertainties, rather than refining parameters, is crucial for improving confidence in these projections. Initially, this approach may widen the range of soil carbon projections, but it will ultimately enhance confidence in managing climate change.

Given the complexity of mechanisms represented by ESMs, uncertainty in SOC predictions should be expected (Bradford et al., 2016). For example, whether tropical forests will gain or lose C is an open question. Productivity in these forests can respond positively to elevated CO_2 (Schimel et al., 2015), while SOC loss may be less sensitive to warming temperatures such as Q_{10} (Koven et al., 2017), potentially allowing tropical forests to store more SOC in the future. Alternatively, tropical forests may experience regional warming-induced drought (Tao et al., 2022), which may cause large areas of forest mortality and consequently SOC loss. In permafrost regions, soils may lose C with climate change due to more intense warming (Rantanen et al., 2022), vulnerable C stocks (Mishra et al., 2021; Natali et al., 2021), and high temperature sensitivity (Davidson & Janssens, 2006; Koven et al., 2017). Permafrost thaw can also create thermokarst landscapes and expose deep SOC to microbial decomposition (Schuur et al., 2015). On the other hand, release of mineral nitrogen due to warming may enhance plant growth (Melillo et al., 2002) and lead to SOC gain in permafrost soils (Koven et al., 2015). Moreover, the dominance of these mechanisms may vary with time. A recently model analysis of the arctic-boreal carbon sink revealed a tipping point in 2050–2080 when the carbon sink strength starts to decrease (Braghiere et al., 2023).

To assess model reliability within the current generation of ESMs, we evaluated the SOC change under high-emission scenarios in the CMIP Representative Concentration Pathway RCP 8.5 or Shared Socioeconomic Pathways SSP 585. We analyzed 24 available ESMs used in CMIP phase 5 (CMIP5) and phase 6 (CMIP6) to quantify the change in SOC stocks from 1996 to 2100. Based on prior analysis of CMIP5 models, we expected SOC stocks in CMIP6 models to increase over this time (Todd-Brown et al., 2013). We evaluated key mechanisms underlying the SOC changes, including the strength of the CO_2 fertilization effect on ecosystem productivity, the warming effect on soil C turnover, and local conditions such as climate and soil C stocks. We determined the relative importance (RI) of these mechanisms. In addition, we built random forest (RF) surrogate models for SOC dynamics in these ESMs to quantify the uncertainty in SOC change contributed by model structure and model inputs.

2. Methods

2.1. Data Sources

We collected soil C stock outputs from 24 ESMs (Table S1 in Supporting Information S1) based on the availability of the model outputs in CMIP5 (Arora et al., 2013) and CMIP6 (Arora et al., 2020). Results were analyzed from historical experiments with RCP 8.5 in CMIP5 models and SSP585 in CMIP6 models. The historical simulations run from 1850 to 2005 for CMIP5 models and from 1850 to 2014 for CMIP6 models. Eight models reported results in both CMIP5 and CMIP6. Note that GFDL-ESM2G and HadGEM2-ES simulations in CMIP5 start at 1861 and 1860, respectively. RCP8.5 and SSP585 are high radiative forcing scenarios with a prescribed atmospheric CO_2 mole fraction ranging from 378 to 935 ppm. We used model runs with prescribed CO_2 concentrations for consistent comparison of the strength of the CO_2 fertilization effects among models.

2.2. Biome Definition

A biome map of $1° \times 1°$ resolution (Figure S1 in Supporting Information S1) was used for consistent biome-level analysis (Todd-Brown et al., 2013). The original $0.05° \times 0.05°$ land cover map was obtained from the MODIS (Moderate Resolution Imaging Spectroradiometer) Land cover MCD12Q1 product 43. One of 16 land cover types was assigned to each $1° \times 1°$ grid cell by taking the most common land cover from the original underlying

0.05° × 0.05° land cover map. Each 1° × 1° grid cell was assigned to one of 8 biomes: tundra, boreal forest, tropical rainforest, temperate forest, desert and shrubland, grassland and savanna, cropland and urban, or permanent wetland (Todd-Brown et al., 2013). Note that snow and ice are excluded for biome assignment. Briefly, tundra includes grassland and open shrubland north of 55°N. Boreal forest includes evergreen needleleaf forest, deciduous needleleaf forest, woody savanna north of 50°N, mixed forest north of 50°N, and closed shrublands north of 50°N. Tropical rainforest includes evergreen broadleaf forest between 25°N and 25°S. Temperate forest consists of deciduous broadleaf, evergreen broadleaf outside of 25°N–25°S, and mixed forest south of 50°N. Desert and shrubland includes barren or sparsely vegetated, open shrubland south of 55°N, and closed shrubland south of 50°N. Grassland and savanna includes woody savanna south of 50°N, savanna, and grasslands south of 55°N. Cropland and urban includes cropland, urban, and mosaic land cover. Permanent wetlands are unmodified.

All model outputs were regridded to 1° × 1° for biome- and grid-scale comparisons. Our regridding approach assumed conservation of mass. Note that no regridding was done for computations with global total SOC.

2.3. Random Forest Surrogate Models for Soil Carbon Dynamics in ESMs

We used a machine learning RF algorithm to construct a surrogate model for the soil carbon dynamics in each ESM (Figure S2 in Supporting Information S1). The predictors included decadal mean of contemporary (1996–2005) net primary productivity (NPP), surface air temperature (Ta), precipitation (P), soil C content (cSoil), soil C turnover time (τ) and their changes over the 21st century, except for cSoil which is the target variable. In the RFs, we used 300 trees, maximal depth of 20, and minimal sample split of 4. In total, we built 24 RF surrogate models using the regridded 1° × 1° model outputs. By forcing each of the 24 RF-models with all 24 sets of predictors, we obtained 576 (24 × 24) sets of SOC change. The predicted SOC change by the RF-models accounted for uncertainties from the RF model structure and RF predictors.

We computed RI for the predictors in each RF model and across land biomes. Relative importance of a given feature variable was defined as the normalized increase in mean square error when this variable was excluded in the RF algorithm $\mathrm{RI} = \Delta \mathrm{MSE}_i / \left(\sum_{j=1}^{n} \Delta \mathrm{MSE}_j \right)$ where i is a given feature and j are the other features.

We statistically assessed the uncertainties introduced by variations in predictors and RF model structure. The uncertainty analysis is for all land pixels without dividing them into biomes.

$$SS_{\mathrm{model}} = n \times \left(\sum_{i=1}^{n} \left(\overline{X}_{\mathrm{model}} - \overline{X} \right)^2 \right) \tag{1}$$

$$SS_{\mathrm{feature}} = n \times \left(\sum_{i=1}^{n} \left(\overline{X}_{\mathrm{feature}} - \overline{X} \right)^2 \right) \tag{2}$$

$$SS_{\mathrm{interaction}} = \sum_{i=1}^{n \times n} \left(X - \overline{X}_{\mathrm{model}} - \overline{X}_{\mathrm{feature}} + \overline{X} \right)^2, \tag{3}$$

where SS is the sum of square error and n is the number of surrogate models. SS_{model}, SS_{feature}, and $SS_{\mathrm{interaction}}$ are the sum of square errors across models, features and their interactions, respectively. $\overline{X}$, $\overline{X}_{\mathrm{model}}$, and $\overline{X}_{\mathrm{feature}}$ are the overall mean, the mean value across models and features, respectively. The variation explained by RF models is $SS_{\mathrm{model}}/(SS_{\mathrm{model}} + SS_{\mathrm{feature}} + SS_{\mathrm{interaction}})$ and by model features is $SS_{\mathrm{feature}}/(SS_{\mathrm{model}} + SS_{\mathrm{feature}} + SS_{\mathrm{interaction}})$. The remaining variation is attributed to the interaction between RF models and feature variables.

3. Results

3.1. SOC Change at a Global Scale

We analyzed 24 fully coupled ESMs to investigate the predicted SOC change under the high emission scenarios RCP 8.5 in CMIP5 and SSP585 in CMIP6. Consistent with our expectation that global SOC would increase under global change, the majority (17 out of 24) of these models predicted a net gain in SOC, with a mean of 43.9 Pg [95% CI: 9.2–78.5 Pg] C gain (Figures 1a and 1b) by 2100. Among the 17 models, 11 predicted more than 50 Pg C
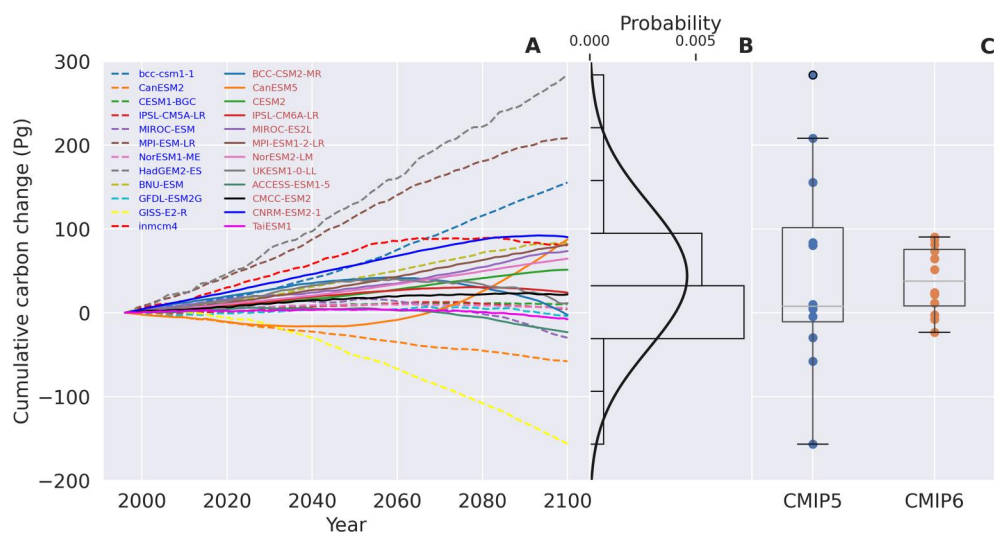
Figure 1. Changes in soil organic carbon under global change during the 21st century. The 12 dashed lines are the Earth system models in Coupled Model Intercomparison Project phase 5 (CMIP5) and the 12 solid lines are those in Coupled Model Intercomparison Project phase 6 (CMIP6). The first eight models reported results in both CMIP5 and CMIP6. The cumulative change of soil organic C (SOC) was tracked from 1996 to the end of the 21st century (a). The probability distribution of the cumulative change showed an average C gain of 43.9 Pg (95% CI: 9.2–78.5 Pg) (b). The variation in SOC change declined significantly from CMIP5 models to CMIP6 models (c).

gain. A gain over 150 Pg C was found in three CMIP5 models (HadGEM2-ES, MPI-ESM-LR, and bcc-csm1-1). In contrast, GISS-E2-R predicted a large SOC loss of about 150 Pg C and CanESM2 predicted a loss over 50 Pg C.

SOC change varied much less in CMIP6 models compared to CMIP5 models (Figure 1c). Most of the CMIP6 models simulated C gain, except for a slight loss in three of the 12 models. The largest gain of nearly 100 Pg C was predicted by CNRM-ESM2-1. Models such as MPI-ESM, MIROC, NorESM, and CNRM in CMIP6 also predicted high C gains between 50 Pg and 100 Pg C. The three models that predicted moderate SOC losses (<30 Pg C) were ACCESS, BCC and TaiESM1.

Moreover, the eight models available in both CMIP5 and CMIP6 simulated rather different SOC changes between the two model versions. For example, the UK model (HadGEM2-ES) predicted a nearly 300 Pg C gain in CMIP5, but little change in CMIP6 (UKESM1-0-LL). Similarly, the MPI model in CMIP5 (MPI-ESM-LR) predicted a gain of more than 200 Pg C, which shrank to less than 100 Pg C in CMIP6 (MPI-ESM1-2-LR). The BCC model predicted a gain of more than 150 Pg C in CMIP5 (bcc-csm1-1), but little net change in the CMIP6 model (BCC-CSM2-MR).

3.2. SOC Change at a Biome Scale

The models predicted diverse biome-level SOC changes (Figure 2, Figure S3 in Supporting Information S1). Due to temperature constraints on decomposition, we expected large SOC losses with global change in high latitude biomes but SOC gains in warm regions. In contrast to this expectation, the biomes in cold regions (i.e., tundra and boreal forest) showed divergent patterns across the models with both C gains and losses simulated. For example, C gains were over 100 Pg C in HadGEM2-ES and over 50 Pg C in bcc-csm1-1 and MPI-ESM-LR. On the contrary, BCC-CSM2-MR lost 100 Pg C and GISS-E2-R lost 70 Pg C from cold regions. In most cases, SOC change in boreal and tundra biomes showed the same direction of response to global change within a given model.

We also found model differences in biome-level SOC change in warm biomes. BCC-CSM2-MR predicted a sink of 60 Pg C in tropical forests, whereas MIROC-ESM predicted a loss of 30 Pg C there. Shrubland and desert gained 60 Pg C in MPI-ESM-LR and grasslands gained 40 Pg C in HadGEM2-ES, but other models showed little change or significant losses in these biomes. SOC change in temperate forest was less variable than for other biomes (Figure 2, Figure S3 in Supporting Information S1).

Large SOC gains predicted in HadGEM2-ES, MPI-ESM-LR, and bcc-csm1-1 were contributed by different biomes. Dry regions were much larger C sinks in MPI-ESM-LR, cold regions sequestered more C in the
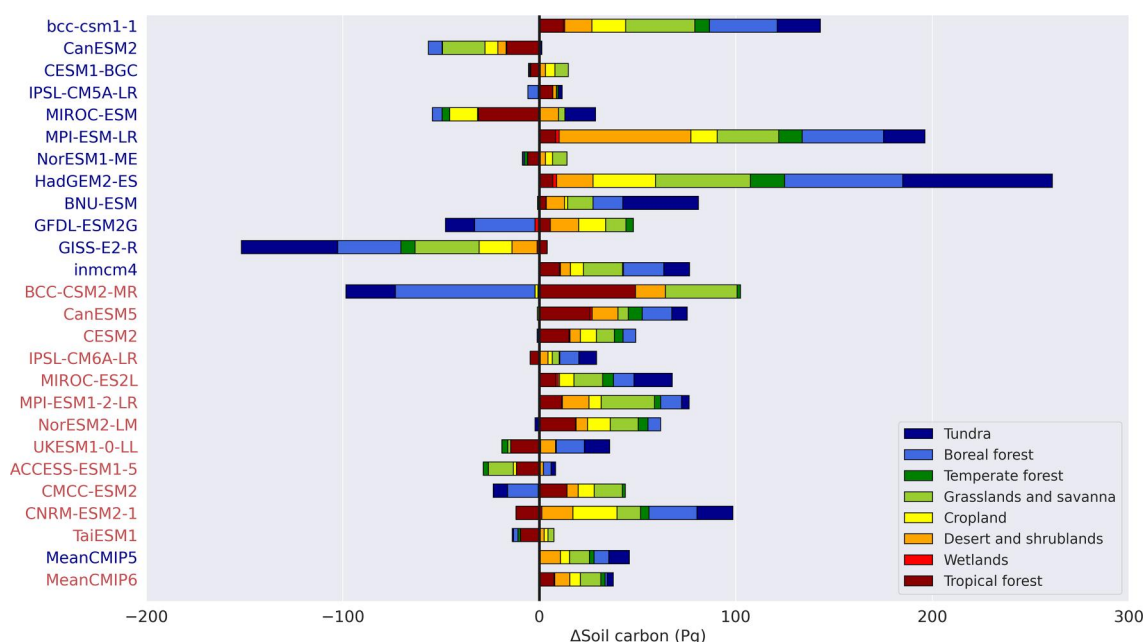
Figure 2. Biome-level change in soil organic carbon. Coupled Model Intercomparison Project phase 5 models are shown in blue text and Coupled Model Intercomparison Project phase 6 models are shown in red text. SOC change varied between biomes in the same model and across models. See Figure S1 in Supporting Information S1 for the biome delineations.

HadGEM2-ES, and boreal and grassland ecosystems contributed the most C gains in the bcc-csm1-1 model. The large C loss in GISS was attributed to all biomes, with the greatest loss from the cold regions.

Interestingly, several models predicted little SOC change due to opposing responses of SOC in different biomes. For example, a large C loss in permafrost regions was compensated by C gains in other regions in GFDL-ESM2G, BCC-CSM2, and CMCC-ESM2. On the other hand, C gains in the cold regions of UKESM1-0 and the tundra biome of MIROC-ESM were offset by C losses from tropical forest.

At the biome level, models did not predict consistent changes in SOC between CMIP5 and CMIP6 (Figure 2, Figure S3 in Supporting Information S1). Except for MPI, all models present in both CMIPs predicted opposing SOC changes in several biomes. CanESM predicted opposing SOC changes in all biomes between the two model versions. The two MPI models predicted the same direction of SOC change, but with a large difference in magnitude.

3.3. Predictors of SOC Change

At the grid scale, the spatial distribution of SOC change was highly predictable using a RF approach. The predictors included contemporary mean annual temperature (Ta) and precipitation (P), NPP, turnover rate (τ) and their changes at the end of the 21st century, as well as contemporary mean soil C stock (cSoil) (Figure 3, Figures S4 and S5 in Supporting Information S1). Among the predictors, Ta and NPP showed consistent increases during the 21st century (Figures S6–S10 in Supporting Information S1). On average, the nine predictors accounted for more than 80% of the variation in SOC change on a grid level and reached as high as 93% in several models (Figure 3).

The most important predictors varied among models. Change in NPP (ΔNPP) was the most important feature in nine out of the 24 models and was the second most important feature in 11 models (Figures S4 and S5 in Supporting Information S1). Contemporary mean soil C stock and air temperature were the two most dominant variables in five models, and mean annual NPP was the dominant variable in three models. When cSoil was the dominant predictor, the models usually simulated SOC loss in the high latitudes. However, when ΔNPP was the most important variable, there was not a clear trend toward either C loss or gain for a given region. Precipitation and its change were the least important predictors for nearly all models.
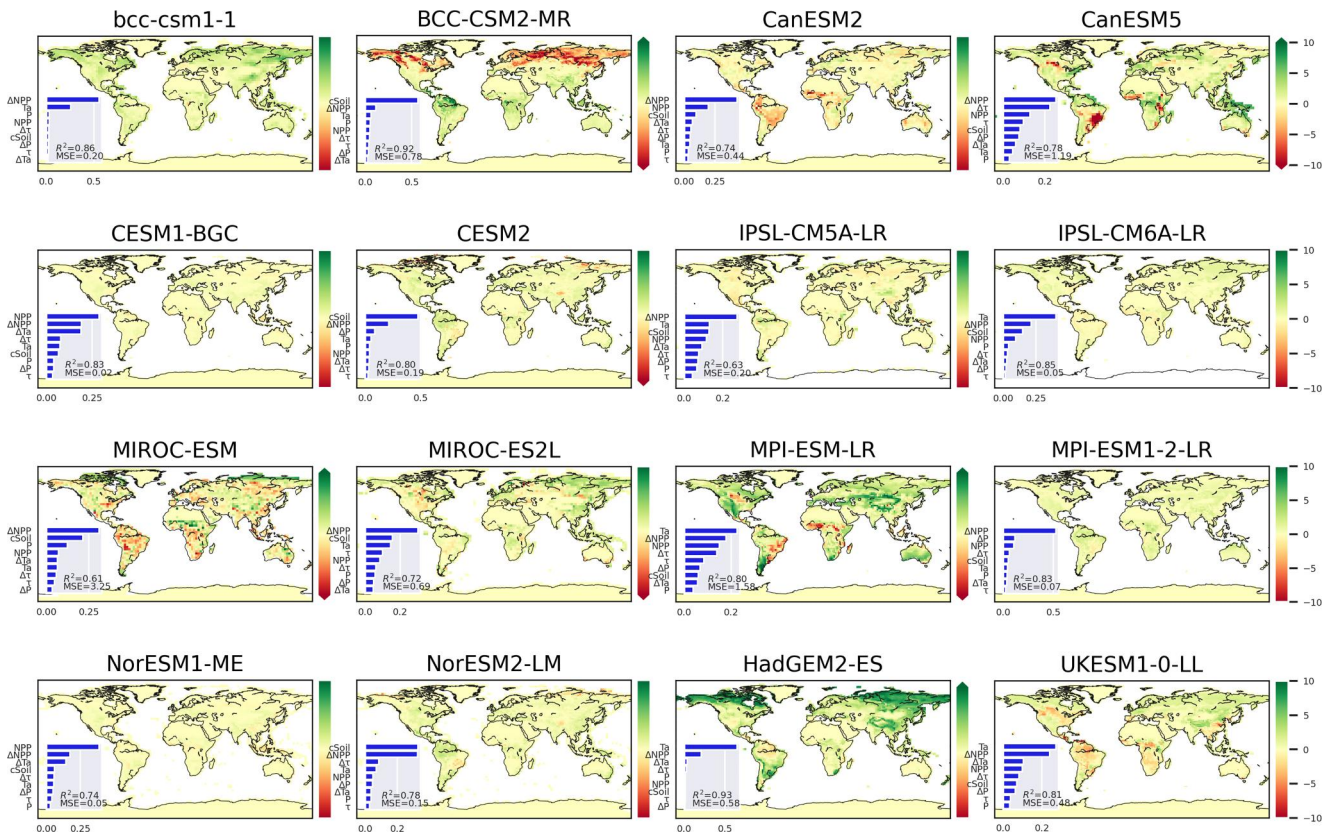
Figure 3. Global distribution of soil organic carbon change for eight models present in both Coupled Model Intercomparison Project and CMIP6. The SOC change was derived as the decadal difference between mean soil C in 2091–2100 and 1996–2005. The inset figures show the relative importance of each predictor for the soil C change generated by the random forest approach. R^2 and mean square error are provided for each model. Note that the color bars share the same scale, with units of kg C m^{-2}. Refer to Supplementary Figures 2 and 3 for the rest of the models.

Interestingly, ΔNPP and Ta were the dominant predictors in the three models with the largest C gains (i.e., HadGEM2-ES, MPI-ESM-LR, and bcc-csm1-1). The large C loss in GISS-E2-R was associated with cSoil and Δτ as dominant predictors. In addition, the dominant predictors changed in five out of the eight models that were compared across CMIP5 and CMIP6. For example, the dominant predictor was ΔNPP in bcc-csm1-1 but cSoil in BCC-CSM2-MR. Similarly, the dominant predictor changed from mean annual NPP to cSoil across the two versions of CESM. In the MIROC and UK ESMs, the dominant predictors remained the same with ΔNPP and cSoil for MIROC, and Ta and ΔNPP for UK ESMs.

At a biome scale, the dominant predictors for SOC change varied greatly among biomes and across models (Figure 4, Figures S11 and S12 in Supporting Information S1). The dominant factors always differed between the two versions of the same model. For example, cSoil determined soil C change in the tundra biome in MIROC-ESM, but not in MIROC-ESM2. Across the two cold biomes (i.e., tundra and boreal forest), the dominant predictors were not always consistent. For example, in both versions of CESM, cSoil determined soil C change in tundra regions, whereas mean annual temperature was more important than cSoil in the boreal forest. On the other hand, in several models, ΔNPP was the most frequent dominant factor across biomes. For example, ΔNPP was overwhelmingly important for almost all biomes in models such as bcc-csm1-1, HadGEM2-ES and MPI-ESM1-2-LR.

3.4. Uncertainty in RF Surrogate Models for Soil Carbon Dynamics in ESMs

We used RF surrogate models for soil carbon dynamics to quantify the uncertainties in SOC change from RF model structure and predictors. Including the uncertainties from surrogate model structures and predictors, soils gained 43.7 Pg C (95% CI: 37.0–50.4 Pg C) during the 21st century (Figure S13 in Supporting Information S1). The variation was mainly determined by model structure (57.7%). RF predictors only accounted for 20.2% of the variation. The remaining variation was explained by the model-predictor interaction. There was a clear difference
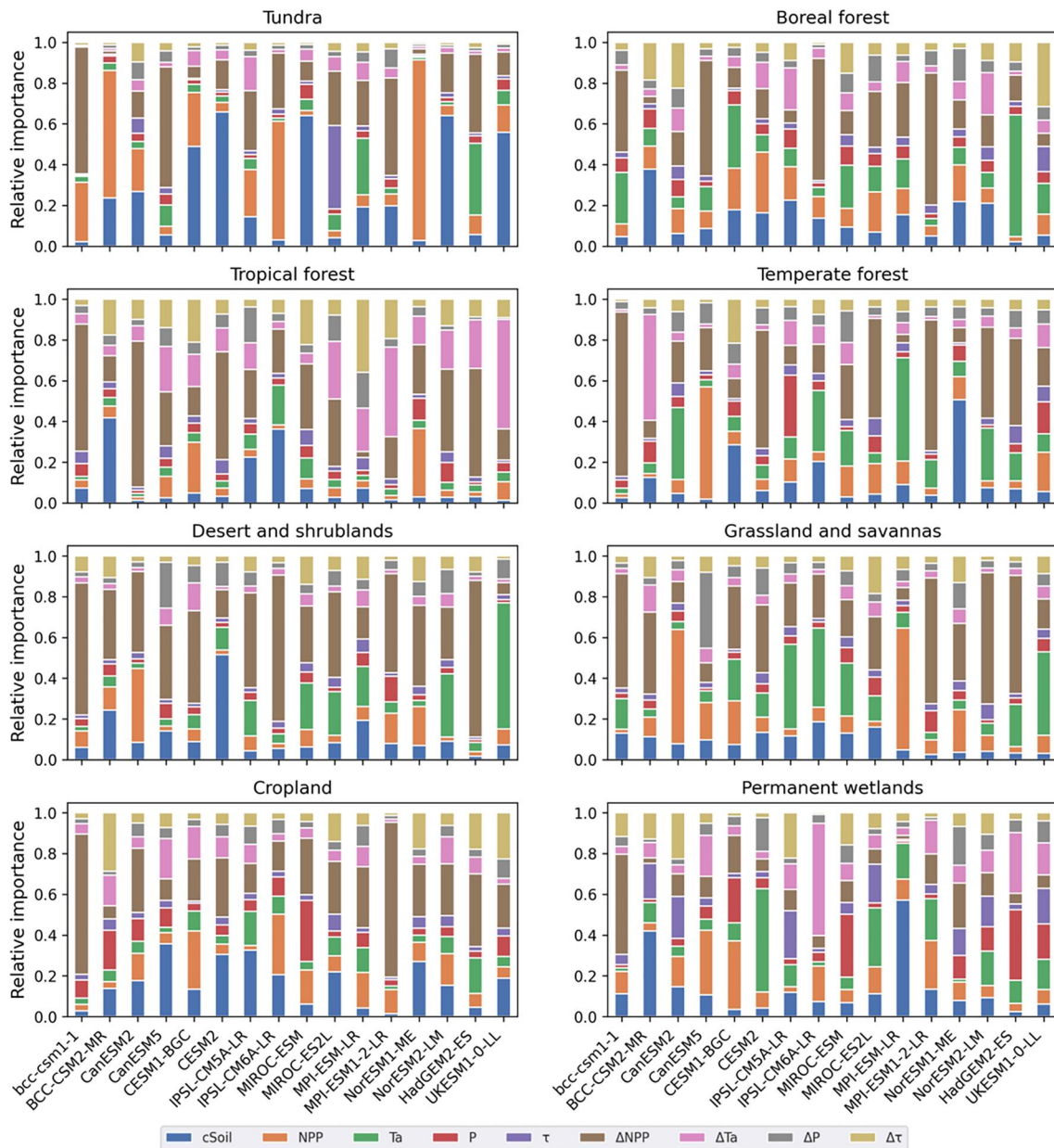
Figure 4. Relative importance (RI) in biome-level predictors for soil organic carbon change. The figure shows the eight models present in both Coupled Model Intercomparison Project and CMIP6. A random forest approach was applied in each of the eight biomes to determine the RI of the nine predictors. Refer to Figures S9 and S10 in Supporting Information S1 for the rest of the models.

in explained variation between CMIP5 and CMIP6 RF models (Figure 5). In CMIP5-RF results, the model structure accounted for more than 60% of the variation whereas it only accounted for 36% in CMIP6-RF results, even lower than the RF predictors (39%). Furthermore, there was a significant statistical difference in SOC change between CMIP5-RF models and CMIP6-RF models ($df = 286$, $F = 2.06$, $p = 0.04$) (Figure 5). CMIP5-RF predicted a soil C gain of 54.3 Pg (95% CI: [37.6–71.1 Pg C]), and CMIP6-RF predicted a soil C gain of 34.2 Pg (95% CI: [24.6–43.8 Pg C]).

4. Discussion

Earth system models in CMIP5 and CMIP6 predicted an average global SOC gain under high emission scenarios by 2100. The variation in SOC change was significantly lower in CMIP6 compared with CMIP5 results.
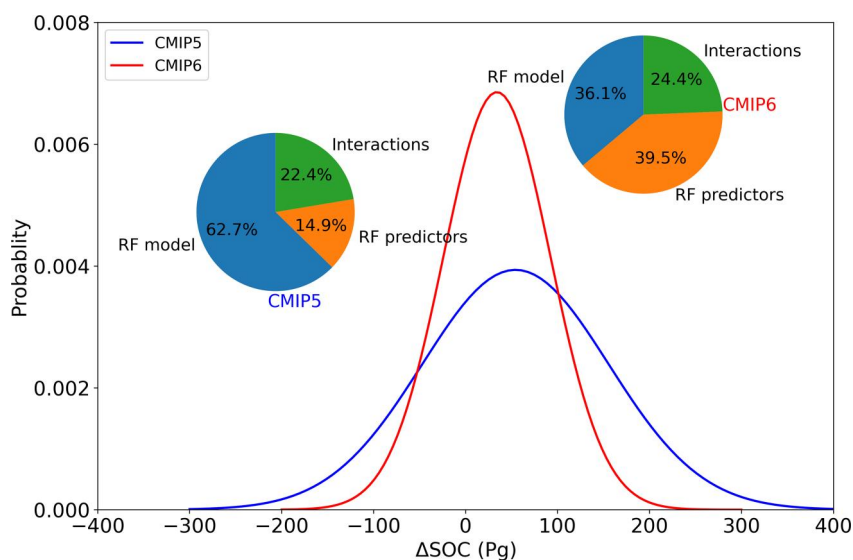
Figure 5. Uncertainty of soil carbon change in random forest surrogate models. CMIP5-RF predicted a soil C gain of 54.3 Pg (95% CI: [37.6–71.1 Pg C]), and CMIP6-RF predicted a soil C gain of 34.2 Pg (95% CI: [24.6–43.8 Pg C]). The importance of model structure declined from Coupled Model Intercomparison Project phase 5 to Coupled Model Intercomparison Project phase 6.

Similarly, previous studies showed increasing convergence of predicted contemporary SOC in CMIP6 relative to CMIP5 (Braghiere et al., 2023; Varney et al., 2022). However, we conclude that the apparent convergence was a result of inherent divergence in dominant mechanisms and processes in the models. In other words, the low variation (convergence) in SOC change in CMIP6 was achieved by revising different processes (divergence) among models.

The tighter clustering of modeled SOC change in CMIP6 likely reflects a "regression toward the mean" phenomenon rather than a real increase in model performance. When we examined the key drivers of variation in SOC change, it was clear that different models arrived at similar answers for very different reasons. Model developers may have adjusted model parameters and structures to reduce large deviations from data-driven estimates of global soil C stocks following CMIP5 analyses (Todd-Brown et al., 2013, 2014). Still, a high risk of model bias remains due to missing process representations and limited diversity in model structure. Additionally, more SOC observations are also beneficial for model development. If the mean SOC itself contains large uncertainty, "regression toward the mean" would be less meaningful.

Increasing confidence in ESM predictions requires reducing spread among models, but only after fully representing key Earth system processes and their associated uncertainties in the model structure (Bradford et al., 2016). Notably, none of the ESMs except GFDL represents explicit microbial processes (e.g., direct uptake of soil carbon by microbes such as through Michaelis-Menten kinetics), yet models including microbial processes predict rather different C dynamics (Allison et al., 2010; Wieder et al., 2013) and sink capacity (Hararuk et al., 2015; Shi et al., 2018). Many CMIP6 models omit or underrepresent other fundamental mechanisms critical for long-term C dynamics, such as nutrient constraints on vegetation production under global change; large-scale vegetation mortality risk under warming-associated long-term drought and fire or insect outbreak in forest ecosystems; permafrost thaw leading to large and fast carbon release from tundra regions; land-use changes and management practices such as deforestation, biomass burning and forest degradation. Soil carbon saturation is another mechanism which may prevent continuing soil carbon gain.

4.1. Negative Feedback of Soil Carbon to Global Change

Our analysis calls into question current ESM predictions of SOC balance. In most CMIP6 system models, the net balance of SOC change was positive, resulting in a negative soil carbon-climate feedback during the 21st century. In these models, higher C inputs due to elevated CO_2 and warming more than compensated for C losses triggered

by climate change. Moreover, the rate of SOC increase was not predicted to decline, meaning that these models are projecting soils to remain C sinks beyond the 21st century (Koven et al., 2021).

At the same time, many empirical studies conflict with these projections. A recent meta-analysis of manipulative field experiments shows that soil did not gain C under the combined treatment of warming and CO_2 fertilization (Bai et al., 2023). Further evidence shows that elevated CO_2 stimulates the accumulation of soil C in the short term, but the SOC gains do not persist in the longer term (van Groenigen et al., 2017), potentially due to warming-induced SOC loss (Bai et al., 2023). Synthesizing >100 CO_2 enrichment experiments, Terrer et al. (2021) reported a trade-off as CO_2 concentrations rise, with larger increases in plant biomass associated with lower soil carbon gains due to priming effects. Additionally, at larger scales, empirical studies have shown declining ecosystem productivity in tropical forest due to land use and climate change (Gatti et al., 2023), indicating that soil carbon loss may follow.

The large differences in soil carbon-climate feedback between the two CMIP phases are striking. None of the eight models included in both CMIPs showed similar predictions of SOC change across the versions. Several models even predicted opposite responses of global total soil C. The C sinks significantly shrank in CMIP6 relative to CMIP5 for the three models that predicted the largest SOC gains. Particularly, the two largest SOC gains with more than 200 Pg in CMIP5 models (HadGEM2-ES and MPI-ESM-LR) shrank to less than 100 Pg due likely to the newly added nitrogen cycle (Jones et al., 2016). Indeed, we found that the increase of NPP under global change is much less in the two models of CMIP6 than those of CMIP5. However, the shrunken SOC gain can be caused by other mechanisms. For example, neither bcc-csm1-1 nor BCC-CSM2-MR had nitrogen cycles in the land components. However, it managed to reduce SOC gain from more than 150 Pg to nearly 0. It was realized by greater sensitivity of high-latitude SOC to temperature change. These findings indicate significant changes in ESM structure and parameterization, which were likely intended to improve model realism. Still, disparate predictions of SOC change at the biome scale suggest that attempts at improving the realism of CMIP6 models may not yet have improved accuracy.

4.2. Biome-Level SOC Change

In many CMIP6 models, divergent biome-level responses led to global convergence in predicted SOC change. In half of the models, most or all of the biomes responded with small increases in soil C with global change, contributing to convergence. However, these small, positive responses were associated with different processes across biomes and models. Some models (e.g., CanESM5) simulated greater SOC due to increased productivity to global change, whereas other models predicted SOC increase in the areas with high local productivity. Furthermore, the opposite responses between biomes in cold regions and those in other regions in some models also contributed to the global convergence—SOC losses in cold regions were compensated by gains in other regions (e.g., BCC-CSM2-MR).

We expected large SOC losses with global change in high latitude biomes due to a high magnitude of warming and the destabilization of permafrost (Natali et al., 2021; Schuur et al., 2022). Contrary to this expectation, most of the CMIP models predicted SOC gains in tundra and boreal regions. We saw large C losses in only three models (GFDL-ESM2G, GISS-E2-R, and BCC-CSM2-MR) resulting from different mechanisms. Faster turnover in a warmer climate led to the large loss in GISS-E2-R whereas large amounts of soil C in the high latitudes were vulnerable to the increased temperature in BCC-CSM2-MR. A C loss up to 50 Pg was also modeled in GFDL-ESM2G related to high SOC storage. Alternatively, high latitude soils might store more C if climate warming alleviates temperature limitation of plant growth. In the Hadley and MPI models, temperature was indeed the most dominant factor for SOC change, and both models showed greater productivity.

Besides the cold biomes, the biomes in hot or dry regions also showed divergent responses. In the tropical forest, we saw either a large increase (BCC-CSM2-MR) or a decrease (MIROC-ESM) in soil C, driven by changes in NPP. Empirical studies have shown inconsistent responses of productivity to elevated CO_2 and warming in tropical forest ecosystems. These drivers could increase productivity (Schimel et al., 2015), but tropical trees may also be susceptible to mortality caused by climate warming and drought. In dry biomes, elevated CO_2 can increase plant water use efficiency, potentially boosting plant growth and SOC storage (Morgan et al., 2011). Indeed, SOC increased by 80 Pg in the shrubland and desert ecosystems simulated by MPI-ESM-LR. However, we did not see consistently large increases in the dry regions of other models.

4.3. Relative Importance in Model Predictors

Surprisingly, nine predictors could account for most of the spatial variation in SOC change (∼93%) using our RF technique. We expected to see that changes in NPP and turnover rate would be two dominant factors determining SOC change (Todd-Brown et al., 2014). Although we found dominance of ΔNPP in many models, increased NPP did not always translate into an increase in SOC. In some models, an greater NPP meant less soil C loss (see CanESM2 and MIROC-ESM in Figures 3 and 5). Furthermore, the same two factors, air temperature together with ΔNPP, predicted more than 100 Pg SOC gain in three models. These results are consistent with the idea that global change removes temperature constraints on plant growth in cold regions. In contrast to ΔNPP, changes in turnover rate rarely explained spatial variation in SOC change (but see GISS-E2-R in Figure 3). In addition, changes in climate (Ta and P) had minimal direct impact in most models, though climate may have affected NPP and turnover time indirectly. The RF approach represents nonlinear relationships among predictors that likely affected SOC change.

Contemporary SOC was also a good predictor in some of the CMIP6 models. We chose this predictor under the assumption that large SOC pools would release more C under warming, potentially dominating over the CO_2 fertilization effect on productivity. Indeed, several models showed net C losses in cold regions with high C stocks. We caution that incomplete model spin-up could affect current SOC and its future trajectory. For example, in cold regions where models might not reach equilibrium, soils may gain C due to the internal dynamics of the models. However, it is difficult to separate and quantify the spin-up effect with the data currently available in CMIPs.

5. Conclusion

Our analysis refutes the idea that convergence in model-predicted SOC change is due to promising improvements in model performance. Clearly it is possible to reduce the spread in model predictions of global soil C change. However, the underlying discrepancies in biome-scale changes and driving mechanisms both within and across models are a cause for concern. Even though ESM predictions now cluster more tightly around the mean soil C gain, they all changed something different. These underlying issues suggest that the models have not in fact reached a consensus and that predictive confidence is low for soil C feedbacks to climate change. Moreover, uncertainty has not yet peaked because the models are still missing key processes. Some models may be getting the right answers for the wrong reasons, or all the models may be getting the same wrong answer (a bias). There is a risk that the current modeled potential for SOC storage is overestimated. If future models include additional processes such as microbial constraints, thermal and water dynamics in permafrost, and updated fire regimes in tropical and tundra regions, soils could turn from projected sinks into sources of CO_2 with serious consequences for climate change mitigation.

Conflict of Interest

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

Data Availability Statement

CMIP5 and CMIP6 model outputs are downloaded from the Earth System Grid Federation (ESGF) data archive (https://esgf-node.llnl.gov/projects/). All datasets used in the study are publicly available.

All codes used in this study can be accessed by Shi (2024) https://doi.org/10.5281/zenodo.10573693.

References

Allison, S. D., Wallenstein, M. D., & Bradford, M. A. (2010). Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, *3*(5), 336–340. https://doi.org/10.1038/ngeo846

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., et al. (2013). Carbon–Concentration and carbon–climate feedbacks in CMIP5 Earth system models. *Journal of Climate*, *26*(15), 5289–5314. https://doi.org/10.1175/jcli-d-12-00494.1

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., et al. (2020). Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences*, *17*(16), 4173–4222. https://doi.org/10.5194/bg-17-4173-2020

Bai, T., Wang, P., Qiu, Y., Zhang, Y., & Hu, S. (2023). Nitrogen availability mediates soil carbon cycling response to climate warming: A meta-analysis. *Global Change Biology*, *29*(9), 2608–2626. https://doi.org/10.1111/gcb.16627

Bossio, D. A., Cook-Patton, S. C., Ellis, P. W., Fargione, J., Sanderman, J., Smith, P., et al. (2020). The role of soil carbon in natural climate solutions. *Nature Sustainability*, *3*(5), 391–398. https://doi.org/10.1038/s41893-020-0491-z

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

Braghiere, R. K., Fisher, J. B., Miner, K. R., Miller, C. E., Worden, J. R., Schimel, D. S., & Frankenberg, C. (2023). Tipping point in North American Arctic-Boreal carbon sink persists in new generation Earth system models despite reduced uncertainty. *Environmental Research Letters*, *18*(2), 025008. https://doi.org/10.1088/1748-9326/acb226

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, *440*(7081), 165–173. https://doi.org/10.1038/nature04514

Gatti, L. V., Cunha, C. L., Marani, L., Cassol, H. L. G., Messias, C. G., Arai, E., et al. (2023). Increased Amazon carbon emissions mainly from decline in law enforcement. *Nature*, *621*(7978), 318–323. https://doi.org/10.1038/s41586-023-06390-0

Hararuk, O., Smith, M. J., & Luo, Y. (2015). Microbial models with data-driven parameters predict stronger soil carbon responses to climate change. *Global Change Biology*, *21*(6), 2439–2453. https://doi.org/10.1111/gcb.12827

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., et al. (2016). C4MIP—The coupled climate–carbon cycle model Intercomparison Project: Experimental protocol for CMIP6. *Geoscientific Model Development*, *9*(8), 2853–2880. https://doi.org/10.5194/gmd-9-2853-2016

Koven, C. D., Arora, V. K., Cadule, P., Fisher, R. A., Jones, C. D., Lawrence, D. M., et al. (2021). 23rd century surprises: Long-term dynamics of the climate and carbon cycle under both high and net negative emissions scenarios. *Earth System Dynamics Discussion*, *2021*, 1–32.

Koven, C. D., Hugelius, G., Lawrence, D. M., & Wieder, W. R. (2017). Higher climatological temperature sensitivity of soil carbon in cold than warm climates. *Nature Climate Change*, *7*(11), 817–822. https://doi.org/10.1038/nclimate3421

Koven, C. D., Lawrence, D. M., & Riley, W. J. (2015). Permafrost carbon−climate feedback is sensitive to deep soil carbon decomposability but not deep soil nitrogen dynamics. *Proceedings of the National Academy of Sciences*, *112*(12), 3752–3757. https://doi.org/10.1073/pnas.1415123112

Melillo, J. M., Steudler, P. A., Aber, J. D., Newkirk, K., Lux, H., Bowles, F. P., et al. (2002). Soil warming and carbon-cycle feedbacks to the climate system. *Science*, *298*(5601), 2173–2176. https://doi.org/10.1126/science.1074153

Mishra, U., Hugelius, G., Shelef, E., Yang, Y., Strauss, J., Lupachev, A., et al. (2021). Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks. *Science Advances*, *7*(9), eaaz5236. https://doi.org/10.1126/sciadv.aaz5236

Morgan, J. A., LeCain, D. R., Pendall, E., Blumenthal, D. M., Kimball, B. A., Carrillo, Y., et al. (2011). C4 grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. *Nature*, *476*(7359), 202–205. https://doi.org/10.1038/nature10274

Natali, S. M., Holdren, J. P., Rogers, B. M., Treharne, R., Duffy, P. B., Pomerance, R., & MacDonald, E. (2021). Permafrost carbon feedbacks threaten global climate goals. *Proceedings of the National Academy of Sciences*, *118*(21), e2100163118. https://doi.org/10.1073/pnas.2100163118

Nottingham, A. T., Meir, P., Velasquez, E., & Turner, B. L. (2020). Soil carbon loss by experimental warming in a tropical forest. *Nature*, *584*(7820), 234–237. https://doi.org/10.1038/s41586-020-2566-4

Pellegrini, A. F. A., Ahlström, A., Hobbie, S. E., Reich, P. B., Nieradzik, L. P., Staver, A. C., et al. (2018). Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity. *Nature*, *553*(7687), 194–198. https://doi.org/10.1038/nature24668

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., et al. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, *3*(1), 168. https://doi.org/10.1038/s43247-022-00498-3

Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO_2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences*, *112*(2), 436–441. https://doi.org/10.1073/pnas.1407302112

Schuur, E. A. G., Abbott, B. W., Commane, R., Ernakovich, J., Euskirchen, E., Hugelius, G., et al. (2022). Permafrost and climate change: Carbon cycle feedbacks from the warming Arctic. *Annual Review of Environment and Resources*, *47*(1), 343–371. https://doi.org/10.1146/annurev-environ-012220-011847

Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., et al. (2015). Climate change and the permafrost carbon feedback. *Nature*, *520*(7546), 171–179. https://doi.org/10.1038/nature14338

Shi, Z. (2024). Global-scale convergence obscures inconsistencies in soil carbon change predicted by Earth system model. *Zenodo*. https://doi.org/10.5281/zenodo.10573693

Shi, Z., Crowell, S., Luo, Y., & Moore, B. (2018). Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nature Communications*, *9*(1), 2171. https://doi.org/10.1038/s41467-018-04526-9

Soong, J. L., Castanha, C., Pries, C. E. H., Ofiti, N., Porras, R. C., Riley, W. J., et al. (2021). Five years of whole-soil warming led to loss of subsoil carbon stocks and increased CO_2 efflux. *Science Advances*, *7*(21), eabd1343. https://doi.org/10.1126/sciadv.abd1343

Tao, S., Chave, J., Frison, P.-L., Le Toan, T., Ciais, P., Fang, J., et al. (2022). Increasing and widespread vulnerability of intact tropical rainforests to repeated droughts. *Proceedings of the National Academy of Sciences*, *119*(37), e2116626119. https://doi.org/10.1073/pnas.2116626119

Terrer, C., Phillips, R. P., Hungate, B. A., Rosende, J., Pett-Ridge, J., Craig, M. E., et al. (2021). A trade-off between plant and soil carbon storage under elevated CO_2. *Nature*, *591*(7851), 599–603. https://doi.org/10.1038/s41586-021-03306-8

Todd-Brown, K. E. O., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., et al. (2014). Changes in soil organic carbon storage predicted by Earth system models during the 21st century. *Biogeosciences*, *11*(8), 2341–2356. https://doi.org/10.5194/bg-11-2341-2014

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

van Groenigen, K. J., Osenberg, C. W., Terrer, C., Carrillo, Y., Dijkstra, F. A., Heath, J., et al. (2017). Faster turnover of new soil carbon inputs under increased atmospheric CO_2. *Global Change Biology*, *23*(10), 4420–4429. https://doi.org/10.1111/gcb.13752

Varney, R. M., Chadburn, S. E., Burke, E. J., & Cox, P. M. (2022). Evaluation of soil carbon simulation in CMIP6 Earth system models. *Biogeosciences*, *19*, 4671–4704. https://doi.org/10.5194/bg-19-4671-2022

Walker, X. J., Baltzer, J. L., Cumming, S. G., Day, N. J., Ebert, C., Goetz, S., et al. (2019). Increasing wildfires threaten historic carbon sink of boreal forest soils. *Nature*, *572*(7770), 520–523. https://doi.org/10.1038/s41586-019-1474-y

Wieder, W. R., Bonan, G. B., & Allison, S. D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, *3*(10), 909–912. https://doi.org/10.1038/nclimate1951