A trade-off between plant and soil carbon storage under elevated CO_2

C. Terrer[1,2*], R.P. Phillips[3], B.A. Hungate[4,5], J. Rosende[6], J. Pett-Ridge[1], M. Craig[3,7], K.J. van Groenigen[8], T.F. Keenan[9,10], B.N. Sulman[7], B.D. Stocker[11,12], P.B. Reich[13,14], A.F.A. Pellegrini[2,15], E. Pendall[14], H. Zhang[16], R.D. Evans[17], Y. Carrillo[14], J. B. Fisher[18,19], K. Van Sundert[20], Sara Vicca[20], R.B. Jackson[2].

Affiliations

[1]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550 USA.

[2]Department of Earth System Science, Stanford University, Stanford, CA, USA.

[3]Department of Biology, Indiana University, Bloomington, IN, 47405, USA.

[4]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, Arizona 86011, USA.

[5]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA.

[6]Institut de Ciència i Tecnologia Ambientals, Universitat Autònoma de Barcelona, Barcelona, Spain.

[7]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA.

[8]Department of Geography, College of Life and Environmental Sciences, University of Exeter, Exeter, EX4 4 RJ, UK.

[9]Department of Environmental Science, Policy and Management, UC Berkeley, Berkeley, CA, USA.

[10]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

[11]Department of Environmental Systems Science, ETH, Universitätsstrasse 2, 8092 Zürich, Switzerland

[12]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[13]Department of Forest Resources, University of Minnesota, St. Paul, MN, USA.

[14]Hawkesbury Institute for the Environment, Western Sydney University, Penrith 2571 Australia.

[15]Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom

[16]Environmental Change Institute, School of Geography and the Environment, University of Oxford, OX1 3QY, UK

[17]School of Biological Sciences and the Stable Isotope Core Laboratory, Washington State University, Pullman, WA 99164-4236, USA.

[18]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA, 91109, USA

[19]Joint Institute for Regional Earth System Science and Engineering, University of California at Los Angeles, Los Angeles, CA, 90095, USA

[20]Plants and Ecosystems (PLECO), Biology Department, University of Antwerp, Wilrijk, Belgium

*Correspondence to: terrer@stanford.edu

Abstract

Terrestrial ecosystems remove about 30% of the CO_2 emitted by human activities each year[1], yet the persistence of this carbon sink partly depends on how plant biomass and soil carbon stocks respond to future increases in atmospheric CO_2[2,3]. While plant biomass often increases in elevated CO_2 (eCO_2) experiments[4–6], soil carbon has been observed to increase, remain unchanged, or even decline[7]. The mechanisms driving this variation across experiments remain poorly understood, creating uncertainty in climate projections[8,9]. Here, we synthesized data from 108 eCO_2 experiments and found that the effect of eCO_2 on soil carbon stocks is best explained by a negative relationship with plant biomass: when plant biomass is strongly stimulated by eCO_2, soil carbon accrual declines; conversely, when biomass is weakly stimulated, soil carbon accumulates. This trade-off appears related to plant nutrient acquisition, whereby enhanced biomass requires mining the soil for nutrients, which decreases soil carbon accrual. We found an increase in soil carbon stocks with eCO_2 in grasslands (8±2%) and no increase in forests (0±2%), even though plant biomass in grassland responded less strongly (9±3%) than in forest (23±2%). Ecosystem models do not reproduce this trade-off, which implies that projections of soil carbon may need to be revised.

Introduction

The future of the land sink, especially of soil carbon, is particularly uncertain[9]. Soils can become either sources or sinks of carbon with rising levels of atmospheric CO_2, depending on the prevalence of gains via photosynthesis or losses via respiration[9,10]. This uncertainty in terrestrial ecosystem model projections reflects uncertainty in both the mechanisms and the parameter values controlling soil carbon cycling under eCO_2[11].

Plant growth generally increases in response to eCO_2[4,12], with soil nutrients identified as the dominant factor explaining variability across experiments[12–15]. The effect of eCO_2 on soil carbon stocks (β_{soil}) is more equivocal. Although the expectation is that soil carbon will accrue as eCO_2 increases plant growth[16], a few experiments show increases in β_{soil}, many show no change, and some even show losses[7]. The observed variation in β_{soil} across experiments is puzzling, and there is large disagreement regarding the dominant mechanisms explaining this variation[7,17,18].

A positive relationship between the effects of eCO_2 on plant biomass and soil carbon pools is expected if increased plant production under eCO_2 increases carbon inputs (litter) into the soil. Indeed, a positive relationship between inputs and soil carbon storage is formalized in first order kinetics[16] and is applied in most terrestrial ecosystem models[19,20]. Because the effect of eCO_2 on plant aboveground biomass (β_{plant}) is strongly correlated with the effect of eCO_2 on litter production (Extended Data Fig. 1a, $r=0.81$) and on root production[21], a positive relationship between β_{plant} and β_{soil} can thus be expected based on first order kinetics. This hypothesis, however, ignores soil carbon losses associated with accelerated soil organic matter decomposition sometimes observed under eCO_2[7,18]. Plants acquire limiting resources from soils through carbon investment belowground in root growth, exudates, and symbiotic bacteria and fungi. Accelerated decomposition of soil organic matter fueled by plant carbon inputs can enable plant nutrient uptake (the "priming effect"[22]). The return on this belowground carbon investment is an increase in aboveground biomass production[15]. However, the priming effect can decrease soil carbon[5]. A *negative* relationship between β_{plant} and β_{soil} may thus emerge through the economics of plant resource acquisition.

Here, we evaluate the mechanisms of β_{soil}, including its relationship with β_{plant}, by synthesizing 268 observations of β_{soil} from 108 eCO_2 experiments spanning the globe with coupled β_{plant}-β_{soil} data (Supplementary Table 1) using meta-analysis techniques. We explore how well these mechanisms are represented in ecosystem models, and upscale the geographical distribution of β_{soil} derived from experiments to identify regions where models might be missing important processes.

Results
Predictors of SOC accrual under eCO_2

Overall, eCO_2 increased soil carbon stocks by 4.6% across experiments (Fig. 1, 95%-CI: 1.7% to 7.5%). Given the strong variation in β_{soil} across factors (Fig. 1), we used a random-forest approach in the context of meta-analysis (meta-forest) to quantify the importance of 19 potential predictors (Extended Data Table 1), including climate, soil, plant, and ecosystem variables and their interactions, accounting for covariation across predictors and potential nonlinearities.

108 We found that β_{plant} is the most important predictor of β_{soil} (Extended Data Fig. 2a,b, n=108),
109 revealing a strong coupling between CO_2-driven changes in plant biomass and soil carbon. In
110 addition, β_{soil} increased with background SOC stocks (Fig. 1), also identified as an important
111 predictor.
112
113 Contrary to expectations from some first order models[19,20], the relationship between β_{soil} and
114 β_{plant} was negative. For the subset (n=73) of field experiments with intact soils (non-potted plants
115 and non-reconstructed soils), we found a significant interaction between β_{plant} and nitrogen (N)-
116 fertilization (Extended Data Fig. 2c, p<0.01). In non-fertilized experiments, the slope between
117 β_{soil} and β_{plant} was significantly negative (Fig. 2a, p<0.0001, R^2=0.67, n=38), whereas in fertilized
118 experiments the slope was less pronounced and nonsignificant (p=0.34, n=35) (Extended Data
119 Fig. 3a). In non-fertilized experiments, increases in plant biomass were associated with
120 decreasing soil carbon stocks (Fig. 2a), consistent with the priming effect. In N-fertilized
121 experiments, eCO_2 generally increased both plant biomass and soil carbon (Extended Data Fig.
122 3b), in line with first order kinetics.
123
124 We propose a framework to explain the negative relationship between β_{soil} and β_{plant} based on
125 plant nutrient acquisition strategies. Symbiotic associations between plants and arbuscular (AM)
126 and ecto (ECM) mycorrhizal fungi mediate β_{plant} (Extended Data Fig. 2d), resulting in much
127 higher β_{plant} in ECM than AM when nutrient availability is low (Fig. 2b). ECM plants efficiently
128 increase N-uptake under eCO_2 (Fig. 2c, n=12), enhancing β_{plant}. However, acquiring N from soil
129 organic matter via priming accelerates soil carbon losses[7], reducing β_{soil} in ECM (Fig. 2b). In
130 contrast, eCO_2 did not significantly affect N-uptake in AM systems (Fig. 2c, n=12). This
131 outcome limits β_{plant} in AM systems but stimulates β_{soil} (Fig. 2b), likely due to increased carbon
132 inputs through fine-root production and rhizodeposition[21,23,24] combined with decreased carbon
133 losses[25]. The composition of the soil organic matter may mediate this effect as well; AM plants
134 produce more easily decomposable litter[26], which enhances mineral-associated soil organic
135 matter (MAOM) formation[27] and results in a greater fraction of soil carbon in MAOM under AM
136 relative to ECM systems[28,29]. Indeed, eCO_2 increases MAOM more strongly in AM systems than
137 in ECM systems (Fig, 2d, n=19). Because MAOM is less accessible to microbial decomposers[30],
138 greater MAOM in AM systems could limit priming-induced losses and promote long-term soil
139 carbon storage.
140
141 We considered three alternative mechanisms that could potentially explain this tradeoff. First,
142 grasses allocate more carbon to roots than trees, which is associated with greater SOC stocks[31,32].
143 Because grassland species associate with AM fungi and the majority of tree species in the dataset
144 associate with ECM, the observed increase in β_{soil} in AM could be driven by ecosystem type
145 rather than mycorrhizal type. However, we found that eCO_2 effects on root biomass and fine-root
146 production were generally lower in grasses than trees, and also in AM than ECM trees (Extended
147 Data Fig. 4). Second, in non-fertilized experiments with available data (n=16), eCO_2 increased
148 litter C:N by 8%, which could reduce the decomposability of litter and the stabilization of carbon
149 in the soil[27]. If litter quality is reduced more in ECM than AM, this could help explain why eCO_2
150 increased SOC in AM, but not in ECM. However, the effect of eCO_2 on litter quality was similar
151 between mycorrhizal types (Extended Data Fig. 4). Finally, contrasting β_{soil} in AM vs. ECM
152 could be driven by larger background SOC in grasslands vs. forests, given that higher SOC is
153 associated with higher β_{soil} (Fig. 1). We found, however, that background SOC was similar

154 between mycorrhizal types and ecosystem types (Extended Data Fig. 4). Thus, differences in root
155 allocation, litter quality and background SOC in grasses vs. trees cannot explain the tradeoff
156 between β_{soil} and β_{plant}. Instead, losses in SOC associated with plant nutrient uptake (priming
157 effect) in ECM, and gains associated with rhizodeposition in AM, are likely key. Experiments
158 including both AM and ECM tree species should be targeted to better understand the impacts of
159 nutrient-acquisition strategies under eCO_2.
160
161 **Upscaling**
162 To explore the potential geographical distribution of β_{soil}, we simulated a global FACE
163 experiment (Fig. 3a). Unlike Fig. 1 where predictors are analyzed individually, our meta-forest
164 model can upscale β_{soil} from experiments while accounting for all important predictors
165 simultaneously on a grid (Extended Data Figs. 5-6, 10-fold cross-validated $R^2 = 0.51$).
166 Grasslands, croplands, and shrublands showed a stronger potential to accumulate soil carbon in
167 response to experimental eCO_2 than forests (Fig. 3a,b). Soils in semi-arid herbaceous ecosystems
168 were particularly responsive to eCO_2, consistent with the results from the Mojave desert FACE
169 experiment that showed eCO_2-driven increases in soil carbon, but not biomass[33]. We identified
170 large areas not currently sampled with eCO_2 experiments, particularly in the tropics and high
171 latitudes (Fig. 3c,d, Extended Data Fig. 6), where new experiments would help reduce
172 uncertainties.
173
174 **Data-model comparison**
175 In addition to the negative relationship between β_{soil} and β_{plant}, we also found a significantly
176 negative relationship between β_{soil} and the effect of eCO_2 on aboveground biomass *production*
177 (Extended Data Fig. 1b, $R^2=0.55$, *p<0.001*), which is strongly correlated with litter production
178 (Extended Data Fig. 1a, $R^2=0.63$, *r=0.81*, *p<0.01*). This result questions the positive relationship
179 between litter inputs and soil carbon stocks encoded in most ecosystem models. Thus, we
180 investigated the relationship between β_{soil} and β_{plant} in models from three different model
181 ensembles (description in Extended Data Table 2). First, models from the FACE model-data
182 synthesis project (FACE-MDS)[34] mimic the experimental treatment in six eCO_2 experiments and
183 allow for a direct comparison with respective observations. While observations from the six
184 experiments included in FACE-MDS showed a *negative* relationship between β_{soil} and β_{plant} (Fig.
185 4a, blue line, $R^2=0.99$, p<0.001), the twelve models simulated a *positive* relationship when
186 pooled by experiment (Fig. 4a, red line, $R^2=0.91$, p<0.01). The relationship across all models
187 individually was positive as well (Extended Data Fig. 7a, dashed line, $R^2=0.37$, p<0.0001), and
188 none of the individual models was able to reproduce the observations. Second, to investigate
189 whether the same relationships emerge across the globe and in simulations where CO_2 increases
190 gradually, we evaluated global century-scale relationships between β_{soil} and β_{plant} from the
191 TRENDY and CMIP5 model ensembles (Fig. 4b,c). Overall, TRENDY and CMIP5 models did
192 not simulate a negative relationship either (Fig. 4b,c). Instead, most models simulated a positive
193 relationship and the vast majority of model simulations fell into the upper-right quadrant
194 (Extended Data Fig. 7b,c), reflecting that inputs drive SOC accumulation in the first-order
195 decomposition soil model structure common to the models.
196
197 In TRENDY and CMIP5 model simulations, β_{soil} was estimated over a much longer time period
198 than in experiments (Extended Data Table 2). Given the relatively slow turnover times of SOC
199 pools, and the slow pace of changes in species composition and evolutionary pressures on both

200 plants and soil microbes, long-term effects likely differ to those found in experiments. However,
201 first order models simulate a positive relationship β_{plant}:β_{soil} when they are forced to simulate the
202 temporal scale of experiments (Fig. 4a), suggesting important processes are missing in models.
203 By including explicit links between plant growth, belowground carbon allocation and SOC
204 decomposition rates, models may more effectively reproduce the observed negative relationship
205 between β_{soil} and β_{plant} and improve long-term projections.
206
207 To estimate the error in terrestrial ecosystem model projections of β_{soil} caused by ignoring the
208 tradeoff between β_{soil} and β_{plant}, we calculated "expected"-β_{soil} as a function of our upscaled β_{plant}
209 and the ratio β_{soil}/β_{plant} simulated by CMIP5 models. CMIP5 models overestimated β_{soil} for
210 forests (Fig. 3e,f, red color). In contrast, CMIP5 models underestimated β_{soil} in large areas
211 dominated by grasses (Fig. 3e,f, blue color), likely because they do not account for the effects of
212 rhizodeposition on β_{soil}[21]. Results with TRENDY models were similar (Extended Data Fig. 8).
213
Discussion
215 In summary, our synthesis of experiments shows that soil carbon stocks can increase by ~5% in
216 response to a 65% step increase in CO_2 concentrations, with a strong coupling between CO_2-
217 driven changes in plant aboveground biomass and soil carbon. However, the coupling between
218 plant biomass and soils is an inverse relationship (Fig. 2a, Extended Data Fig. 1b), opposite to
219 that simulated by many ecosystem models (Fig. 4). The effect of eCO_2 on soil carbon storage is
220 dependent on a fine balance between changes in inputs and changes in turnover[18], where the
221 latter is dependent on root-microbe-mineral interactions in the rhizosphere. Our results suggest
222 that rhizosphere responses, and especially priming, explain much of the variation in β_{soil} across
223 experiments (Fig. 2). Most models focus on carbon inputs and underestimate rhizosphere
224 effects[11,20,35], likely explaining the disagreement in β_{soil} between observations and models (Figs.
225 3,4). We propose a framework to explain β_{soil} based on nutrient acquisition strategies[15,36,37]. On
226 one end of the spectrum, substantial acquisition of soil N is possible via priming[5] in ECM plants,
227 causing a stronger plant biomass sink at the expense of soil carbon accrual. On the other end, low
228 nutrient availability strongly constrains the plant biomass sink[38] in AM plants. However, the
229 ecosystem-level sink is not necessarily eliminated; instead, eCO_2 can trigger soil carbon accrual
230 through plant carbon-allocation belowground[21,23,24]. When plant growth is severely limited by N
231 or other nutrients, eCO_2 may only cause a transient priming effect in ECM, with high soil
232 decomposition and insufficient nutrient uptake rendering no ecosystem-level sink[39].
233
234 Our results underline the potential of grassland soils to store carbon as atmospheric CO_2 levels
235 continue to rise. The results also suggest that current state of the art models may overestimate the
236 soil carbon sequestration potential of forests in large parts of the world. Previous studies suggest
237 that the potential of vegetation to take up CO_2 will slow later in this century due to nutrient
238 constraints[12–14,38,39]. Our synthesis indicates that these nutrient constraints extend to carbon
239 storage in ecosystems as a whole — through a partial tradeoff between increased plant growth
240 and soil carbon storage whereby ecosystems where plant growth is more nutrient limited
241 accumulate more carbon belowground. The apparent mismatch between observations and how
242 most models represent the biomass-to-soil link suggests that many terrestrial ecosystem models
243 do not adequately represent the critical processes driving soil carbon accumulation. Models are
244 evolving to include more sophisticated representations of soil nutrient cycling, and some now
245 include microbial activity explicitly[36,40]. This change towards coupled carbon-nutrient cycling

246 mediated by plant-soil interactions is important for more realistically and accurately modeling
247 the carbon cycle today and for projecting the land sink in the future.
248

References (main)

250 1. Friedlingstein, P. *et al.* Global Carbon Budget 2020. *Earth Syst. Sci. Data* **12**, 3269–3340 (2020).
251 2. Schimel, D., Stephens, B. B. & Fisher, J. B. Effect of increasing CO2 on the terrestrial carbon cycle. *Proc.*
252 *Natl. Acad. Sci.* **112**, 436–441 (2015).
253 3. Keenan, T. *et al.* Recent pause in the growth rate of atmospheric CO2 due to enhanced terrestrial carbon
254 uptake. *Nat. Commun.* **7**, 13428 (2016).
255 4. Baig, S., Medlyn, B. E., Mercado, L. M. & Zaehle, S. Does the growth response of woody plants to elevated
256 CO2 increase with temperature? A model-oriented meta-analysis. *Glob. Change Biol.* **21**, 4303–4319 (2015).
257 5. Drake, J. E. *et al.* Increases in the flux of carbon belowground stimulate nitrogen uptake and sustain the long
258 term enhancement of forest productivity under elevated CO2. *Ecol. Lett.* **14**, 349–357 (2011).
259 6. Norby, R. J. *et al.* Forest response to elevated CO2 is conserved across a broad range of productivity. *Proc.*
260 *Natl. Acad. Sci.* **102**, 18052–18056 (2005).
261 7. van Groenigen, K. J., Qi, X., Osenberg, C. W., Luo, Y. & Hungate, B. A. Faster Decomposition Under
262 Increased Atmospheric CO2 Limits Soil Carbon Storage. *Science* **344**, 508 (2014).
263 8. Friedlingstein, P. *et al.* Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. *J. Clim.*
264 **27**, 511–526 (2014).
265 9. Todd-Brown, K. E. O. *et al.* Changes in soil organic carbon storage predicted by Earth system models during
266 the 21st century. *Biogeosciences* **11**, 2341–2356 (2014).
267 10. Heimann, M. & Reichstein, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* **451**,
268 289–292 (2008).
269 11. Bradford, M. A. *et al.* Managing uncertainty in soil carbon feedbacks to climate change. *Nat. Clim. Change* **6**,
270 751–758 (2016).
271 12. Terrer, C. *et al.* Nitrogen and phosphorus constrain the CO 2 fertilization of global plant biomass. *Nat. Clim.*
272 *Change* **9**, 684–689 (2019).
273 13. Reich, P. B., Hungate, B. A. & Luo, Y. Carbon-Nitrogen Interactions in Terrestrial Ecosystems in Response to
274 Rising Atmospheric Carbon Dioxide. *Annu. Rev. Ecol. Evol. Syst.* **37**, 611–636 (2006).
275 14. Norby, R. J. & Zak, D. R. Ecological lessons from free-air CO2 enrichment (FACE) experiments. *Annu. Rev.*
276 *Ecol.* **42**, 181–203 (2011).
277 15. Terrer, C. *et al.* Ecosystem responses to elevated CO2 governed by plant–soil interactions and the cost of
278 nitrogen acquisition. *New Phytol.* **217**, 507–522 (2018).
279 16. Olson, J. S. Energy Storage and the Balance of Producers and Decomposers in Ecological Systems. *Ecology*
280 **44**, 322–331 (1963).
281 17. Hungate, B. A. *et al.* Assessing the effect of elevated carbon dioxide on soil carbon: a comparison of four
282 meta analyses. *Glob. Change Biol.* **15**, 2020–2034 (2009).
283 18. Kuzyakov, Y., Horwath, W. R., Dorodnikov, M. & Blagodatskaya, E. Review and synthesis of the effects of
284 elevated atmospheric CO2 on soil processes: No changes in pools, but increased fluxes and accelerated cycles.
285 *Soil Biol. Biochem.* **128**, 66–78 (2019).
286 19. Tian, H. *et al.* Global patterns and controls of soil organic carbon dynamics as simulated by multiple
287 terrestrial biosphere models: Current status and future directions. *Glob. Biogeochem. Cycles* **29**, 775–792
288 (2015).
289 20. Todd-Brown, K. E. O. *et al.* Causes of variation in soil carbon simulations from CMIP5 Earth system models
290 and comparison with observations. *Biogeosciences* **10**, 1717–1736 (2013).
291 21. Nie, M., Lu, M., Bell, J., Raut, S. & Pendall, E. Altered root traits due to elevated CO2: a meta analysis.
292 *Glob. Ecol. Biogeogr.* **22**, 1095–1105 (2013).
293 22. Kuzyakov, Y. Priming effects: Interactions between living and dead organic matter. *Soil Biol. Biochem.* **42**,
294 1363–1371 (2010).
295 23. Treseder, K. K. A meta analysis of mycorrhizal responses to nitrogen, phosphorus, and atmospheric CO2 in
296 field studies. *New Phytol.* **164**, 347–355 (2004).
297 24. Jastrow, J. D. *et al.* Elevated atmospheric carbon dioxide increases soil carbon. *Glob. Change Biol.* **11**, 2057–
298 2064 (2005).
299 25. Carrillo, Y., Dijkstra, F. A., LeCain, D. & Pendall, E. Mediation of soil C decomposition by arbuscular
300 mycorrizhal fungi in grass rhizospheres under elevated CO2. *Biogeochemistry* **127**, 45–55 (2016).

301   26.   Averill, C., Bhatnagar, J. M., Dietze, M. C., Pearse, W. D. & Kivlin, S. N. Global imprint of mycorrhizal
302   fungi on whole-plant nutrient economics. *Proc. Natl. Acad. Sci.* **116**, 23163–23168 (2019).

303   27.   Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K. & Paul, E. The Microbial Efficiency-Matrix
304   Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization:
305   do labile plant inputs form stable soil organic matter? *Glob. Change Biol.* **19**, 988–995 (2013).

306   28.   Cotrufo, M. F., Ranalli, M. G., Haddix, M. L., Six, J. & Lugato, E. Soil carbon storage informed by particulate
307   and mineral-associated organic matter. *Nat. Geosci.* **12**, 989–994 (2019).

308   29.   Craig, M. E. *et al.* Tree mycorrhizal type predicts within-site variability in the storage and distribution of soil
309   organic matter. *Glob. Change Biol.* **24**, 3317–3330 (2018).

310   30.   Schmidt, M. W. I. *et al.* Persistence of soil organic matter as an ecosystem property. *Nature* **478**, 49–56
311   (2011).

312   31.   Jobbágy, E. G. & Jackson, R. B. The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate
313   and Vegetation. *Ecol. Appl.* **10**, 423–436 (2000).

314   32.   Sokol, N. W., Kuebbing, S. E., Karlsen Ayala, E. & Bradford, M. A. Evidence for the primacy of living root
315   inputs, not root or shoot litter, in forming soil organic carbon. *New Phytol.* **221**, 233–246 (2019).

316   33.   Evans, R. D. *et al.* Greater ecosystem carbon in the Mojave Desert after ten years exposure to elevated CO_2.
317   *Nat. Clim. Change* **4**, 394–397 (2014).

318   34.   Walker, A. P. *et al.* *FACE-MDS Phase 2: Model Output.*
319   https://www.osti.gov/dataexplorer/biblio/dataset/1480327 (2018) doi:10.15485/1480327.

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

322   36.   Sulman, B. N. *et al.* Diverse Mycorrhizal Associations Enhance Terrestrial C Storage in a Global Model.
323   *Glob. Biogeochem. Cycles* **33**, 501–523 (2019).

324   37.   Shi, M., Fisher, J. B., Brzostek, E. R. & Phillips, R. P. Carbon cost of plant nitrogen acquisition: global carbon
325   cycle impact from an improved plant nitrogen cycle in the Community Land Model. *Glob. Change Biol.* **22**,
326   1299–1314 (2016).

327   38.   Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E. & McMurtrie, R. E. CO2 enhancement of forest
328   productivity constrained by limited nitrogen availability. *Proc. Natl. Acad. Sci.* **107**, 19368–19373 (2010).

329   39.   Jiang, M. *et al.* The fate of carbon in a mature forest under carbon dioxide enrichment. *Nature* **580**, 227–231
330   (2020).

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

333

334

335

FIGURE LEGENDS

Fig. 1. Meta-analysis of the effect of elevated CO_2 on soil carbon (%) across different factors. n=108. Overall means and 95% confidence intervals are given; we interpret CO_2 effects when the zero line is not crossed by the confidence intervals. Arrows represent 95% confidence intervals that extend beyond the limits of the plot. Soil carbon stocks represent values in ambient CO_2 plots as a continuous variable, here expressed as intervals of equal sample size for illustration purposes. Values in parenthesis are sample sizes.
CO_2 effects represent, on average, an increase in CO_2 from 372 ppm to 616 ppm. FACE: Free Air CO_2 Enrichment, OTC: Open Top Chamber, AM: arbuscular mycorrhizal, AM-ER: mix of AM and ericoid mycorrhizal, ECM: ectomycorrhizal, N-fixer: fixation of atmospheric nitrogen.

Fig. 2. Elevated CO_2 experiments show an inverse relationship between the effects of elevated CO_2 on plant biomass and soil carbon (a). This inverse relationship can be explained by the different efficiencies in plant nutrient uptake (c) between arbuscular mycorrhizal (AM) and ectomycorrhizal (ECM) nutrient-acquisition strategies driving opposite effects on plant biomass and soil carbon pools (b), including mineral-associated soil organic matter (MAOM) stocks (d). Regression line (a) is based on a quadratic mixed-effects meta-regression model and 95% confidence interval (R^2=0.67, p<0.0001, n=38). Dots in (a) represent the individual experiments in the meta-analysis, with dot sizes proportional to model weights. Dots in (b-d) represent overall effect sizes from a meta-analysis and 95% confidence intervals. Data shown here are for non-fertilized experiments (see Extended Data Fig. 3 for nutrient-fertilized experiments).

Fig. 3. Effect of elevated CO_2 (~240 ppm) on soil carbon stocks upscaled from 108 CO_2 experiments. (a,b) Relative effect of elevated CO_2 on soil carbon upscaled based on a meta-forest approach with data from CO_2 experiments. (c, d) Standard error in (a). Green dots represent the location of the CO_2 experiments included in the analysis. (e,f) Difference between expected CO_2 effects on soil carbon stocks based on CMIP5 models and upscaled based on experiments (shown in a). Expected values result from the relationship between β_{soil} and β_{plant} coded in models. Positive values (red color) indicate an overestimation by models; negative values (blue color) indicate an underestimation by models. Shaded areas between -15 to 15 and from 60 to 90 degrees in latitude represent ecosystems not well sampled by experiments and are excluded from the analysis. Boxplots show the median, the first to third quartile, the 1.5x interquartile ranges, and outliers.

Figure 4. Comparison of modeled and measured relationships between aboveground biomass and soil carbon responses to CO_2. (a) Relationship observed (blue) and modeled (red) across 6 eCO_2 experiments. Model results are based on 12 models applied to the same 6 experiments with a common forcing and initialization protocol. Experiments included are Duke FACE (DUKE), Kennedy Space Center (KSCO), Nevada Desert FACE (NDFF), Oak Ridge FACE (ORNL), Prairie PHACE (PHAC), and Rhinelander (RHIN). Regression line across observations in (a) is based on a quadratic meta-regression model. Modeled simulations averaged in (a) for each experiment are from the FACE Model-Data-Synthesis project phase 2. (b,c) Global-scale relationship simulated by ecosystem models (b) from the TRENDY ensemble for the historical increase in CO_2 since the year 1700 and (c) from the CMIP5 ensemble for an increase in CO_2 from 372 ppm to 616 ppm as in eCO_2 experiments. Dotted lines are the 1:1 line.

Methods

Overview. Here, we collect data on the effects of elevated CO_2 on soil carbon stocks (β_{soil}) in both relative and absolute terms and synthesize them through meta-analysis. We also collect data on climatic, experimental, and vegetation characteristics that could potentially explain variability in β_{soil} ("predictors"). In Fig. 1, we show a descriptive meta-analysis of overall β_{soil} across different predictor factors. We next combine the strengths of meta-analysis (e.g. accounting for within-study variability, weights) with random-forest (e.g. computational efficiency, nonlinearities, interactions) – meta-forest – to quantify the relative importance of 19 predictors in explaining variation in β_{soil} in the dataset. In Fig. 2, we describe the regression between β_{soil} and its most important predictor (β_{plant}), and explore the possible mechanisms underlying this relationship. In Fig. 3, we apply the data-trained meta-forest model to upscale β_{soil}. Finally, we investigate whether the emerging relationship between β_{soil} and β_{plant} found in experiments is represented in models (Fig. 4).

Data collection. We have compiled the openly available Report of Mutualistic Associations, Nutrients, and Carbon under eCO_2 (ROMANCE) v1.0 dataset[41] with data on soil organic carbon (SOC) and plant biomass from eCO_2 experiments. Expanding van Groenigen et al's 2014 meta-analysis[7] of 53 experiments reporting soil organic carbon (SOC) data, we used Google Scholar to gather a total of 166 studies related to eCO_2 experiments, published from 1st January 2013–1st May 2019. Search terms were either "elevated CO_2", "increased CO_2" or "CO_2 enrichment" and either "soil carbon" or "plant biomass." To account for experiments that could have been omitted by van Groenigen et al. prior to 2013, we consulted the list of CO_2 experiments from INTERFACE (https://www.bio.purdue.edu/ INTERFACE/experiments.php), the Global List of FACE Experiments from the Oak Ridge National Laboratory (http://facedata.ornl.gov/global_face.html), the ClimMani database on manipulation experiments (www.climmani.org) and the database described by Dieleman et al.[42]. We recorded the structure of each eCO_2 experiment from the papers, taking into consideration the start date and total duration of the experiment (years), and the location of the experiment (coordinates). When the data were presented in figures, mean values and standard error were extracted using GraphClick.

For this meta-analysis, only one datum per experiment was considered to avoid pseudoreplication. The effects of eCO_2 on soil C pools are modulated by increases in soil C inputs from plant litter as well as feedbacks between plants and soils altering soil biogeochemical cycles that can take several years to occur. Thus, we used the most recent measurements in each experiment as the most representative data of the effect of eCO_2 on SOC.

For plant biomass, measurements across different time-points were combined so that only one effect size was analyzed per study. The combined effect size and variance that account for the correlation among the different time-point measurements was calculated following the method described in Borenstein et al.[43], using a conservative approach by assuming non-independency of multiple outcomes ($r = 1$) and performed using the MAd package in R[44]. We collected data on both aboveground biomass stocks and production. When aboveground biomass production data were unavailable, we collected plant data in the following order or preference: NPP, aboveground biomass increment, foliage production and yield. When biomass or soil data were not reported, studies were excluded. We also included the data on litter production reported by Song and Wan[45] to study the interactions with aboveground biomass and production data.

Soil carbon measurements in the dataset were reported at different depths, varying from 5 to 30 cm maximum depth, with an average depth of ~20 cm. When upscaling eCO_2 effects on

430   SOC through meta-forest, we included a fixed value of 0-30 cm in depth as a covariate to control
431   for the influence of soil depth, interpolating predictions for the same soil depth of models.
432   SOC data reported as concentrations were transformed to stocks ($g\ m^{-2}$) using soil bulk
433   density. When bulk density was not reported, we used data reported for similar experiments
434   within the same site or assumed a bulk density of $1\ g\ cm^{-3}$. Assumptions are indicated in the
435   dataset.
436   Studies from ROMANCE v1.0 were not included in the meta-analysis if they met any of
437   the following exclusion criteria: i) studies with no SOC data; ii) papers with no plant biomass
438   data; iii) studies where the duration of the eCO_2 experiment lasted less than 0.5 years. A total of
439   138 independent experiments were collected, of which, 108 were included in the final analysis
440   based on these exclusion criteria.
441
442   **Meta-analysis.** Two types of effect size were calculated:1) the log response ratio (mean response
443   in elevated-to-ambient CO_2 plots), to measure effect sizes in relative terms (%) for each
444   experiment; and 2) the raw mean difference, to compute effect sizes in absolute terms ($g\ m^{-2}$).
445   For each experiment, we collected data on SOC stocks, standard deviation and sample size under
446   elevated and ambient (control) CO_2 plots. Effect sizes were calculated using the *escalc* function
447   from the R package *metafor*[46]. We calculated overall effects in a weighted, mixed-effects model
448   using the rma.mv function in *metafor*. The potential non-independency of studies within the
449   same site (e.g. different species, different treatments) was accounted for by including "site" as a
450   random effect. Effect size measurements from individual studies in the meta-analysis were
451   weighted by the inverse of the variance[47]. 13% of studies did not report standard deviations,
452   which were thus imputed using Rubin and Schenker's[48] resampling approach from studies with
453   similar means. These calculations were performed using the R package *metagear*[49].
454
455   **Variable Importance and upscaling approach.** 19 potential moderators were coded (Extended
456   Data Table 1). Including all 19 moderators in a meta-regression risks overfitting the model.
457   Therefore, we applied the R package '*metaforest*'[50] to identify potentially relevant moderators in
458   predicting β_{soil} across the complete dataset of 108 studies. The approach is based on the machine-
459   learning 'random forest' algorithm, which is robust to overfitting, and is integrated in a meta-
460   analytic context by incorporating the variance and weight of each experiment as in classic meta-
461   analysis (see above).
462   As an initial step, we conducted variable pre-selection by including the 19 predictors in
463   *metaforest* with 10,000 iterations and replicated 100 times with a recursive algorithm in the
464   *preselect* function from *metafor*[46]. Moderators that consistently displayed negative variable
465   importance (i.e., that showed a reduction in predictive performance) were dropped using the
466   *preselect_vars* function. Moderators that improved predictive performance were then carried
467   forward to optimize the model. Parameters of the meta-forest model were optimized using the
468   *train* function from the *caret* package[51], and calculated 10-fold cross validated R^2 with 75% of
469   the data used as training data and 25% for validation. Unlike maximum likelihood model-
470   selection approaches, this method can handle many potential predictors and their interactions and
471   considers non-linear relationships. Partial dependence plots were produced which visualize the
472   association of each moderator with the effect size, while accounting for the average effect of all
473   other moderators.
474   As a sensitivity test, and to identify important interactions between predictors, we ran an
475   alternative model-selection procedure using maximum likelihood estimation. For this purpose,

476 we used the rma.mv() function from the *metafor* R package[46] and the glmulti() function from the
477 *glmulti* R package[52] to automate fitting of all possible models containing the 5 most important
478 predictors and their interactions (level=2). Model selection was based on Akaike Information
479 Criterion corrected for small samples (AICc), with the relative importance value for a particular
480 predictor equal to the sum of the Akaike weights (probability that a model is the most plausible
481 model) for the models in which the predictor appears.
482   Finally, the data-trained meta-forest model was applied to global gridded data of pre-
483 selected predictors (see Extended Data Table 1 for gridded data sources) to estimate the effect of
484 elevated CO_2 on SOC. The resulting global maps are geographically constrained to ecosystems
485 best represented by experiments. We remove the estimates for latitudes comprised between -15
486 to 15 degrees, corresponding to tropical ecosystems not sampled by experiments (Fig. 3c, green
487 dots), and from 60 to 90 degrees.
488
489 **Nitrogen fertilization and soil disturbance.** We used the information reported in the papers to
490 assess whether the soils were exposed to external inputs of N fertilization ("yes") or not ("no").
491 Experiments were also classified as either having "disturbed" or "intact" soils as noted in the
492 papers. If not, experiments which used pots or reconstructed soils were categorized as disturbed.
493 We used the same approach and classification as in ref[53].
494   To upscale the effect of nitrogen fertilization and disturbance on β_{soil}, we reclassified the
495 ESA CCI land cover map https://www.esa-landcover-cci.org/?q=node/164. Reclassification files
496 are accessible online https://figshare.com/account/projects/74721/articles/11710155. For
497 example, we classify "Cropland, rainfed"-"Herbaceous cover" (class 11) and "Cropland,
498 irrigated or post-flooding" (class 20) as fertilized.
499
500 **Nutrient-acquisition strategy classification.** We considered the importance of the type of
501 symbiotic association as a driver of eCO_2 effects on soil C. Mycorrhizal status includes AM,
502 ECM and a mix of AM and ericoid (ER) mycorrhizal fungal associations. Here, we also
503 considered some plant species known to associate with N-fixing microorganisms. We refer to
504 this classification as "Symbiotic", because it includes both mycorrhizal status and N-fixation.
505 Together, these four symbiosis types represent different mechanisms plants use to acquire
506 nutrients[15].
507   We assessed the impact of the dominant symbiotic association type by classifying all
508 studies as ECM, AM, AM-ER, and N-fixers, using the check-lists by Wang et al.[54] and Maherali
509 et al.[55], with additional classifications derived from the literature. Species that associate with
510 both ECM and AM (e.g. Populus spp.) were classified as ECM because these species can
511 potentially benefit from increased N-availability due to the presence of ECM fungi[56]. Most of the
512 N-fixers in the dataset were associated with both N-fixing symbionts as well as AM fungi, but
513 we classified them as N-fixers because these species can potentially benefit from N acquired
514 through N-fixation.
515
516 **MAOM data.** We retrieved data on mineral-associated organic matter (MAOM) and particulate
517 organic matter (POM) for the subset of studies employing size or density fractionation of soil
518 organic matter (n = 19). Because of methodological differences, POM is loosely defined as
519 organic matter recovered in the total coarse (typically > 53µm) or light (typically < 1.6 g cm-3)
520 soil fraction. Where MAOM was not reported, it was estimated based on mass balance by
521 subtracting the POM fraction from total C.

522
523 **FACE Model-Data-Synthesis.** We use data from the FACE Model-Data Synthesis (FACE-
524 MDS) Project Phase 2[57–62], in which 12 models were applied to 6 eCO$_2$ experiments. Each model
525 covered the time periods representative of the FACE experiments, following a standardized
526 protocol including meteorological forcing, CO$_2$ concentration, site history, and vegetation
527 characteristics for each site.
528   Experiments included in the FACE-MDS Phase 2 were Duke FACE[63], Kennedy Space
529 Center[64], Nevada Desert FACE[65], Oak Ridge FACE[38], Prairie PHACE[66,67] and Rhinelander[68].
530 Models included were CLM4.0[69], CLM4.5, DAYCENT, CABLE, JULES[70], LPJ-GUESS, OCN,
531 TECO, ORCHIDEE[71], GDAY, ISAM, and SDGVM. See ref.[61] for an overview of model
532 structures and processes. As in the observational data, we compared relative changes in
533 aboveground biomass and soil carbon stocks of each experiment for eCO$_2$ relative to control
534 treatments.
535
536 **TRENDY models** We use model outputs from the TRENDY v7 S1 simulations, where each
537 model is driven by standardized forcings of observed increasing CO$_2$ for years 1700-2018 CE,
538 and constant preindustrial climate and land use. We selected six models that provided outputs for
539 aboveground vegetation carbon (taken as the sum of wood and leaf carbon), soil carbon, and
540 NPP (CABLE-POP[72], CLM5.0[73], ISAM[74], LPJ-GUESS[75], ORCHIDEE[71], ORCHIDEE-CNP[76]).
541 Wood carbon often includes coarse roots in models. Here, we evaluate relative changes and
542 numbers are not sensitive to the exact definition. Description of models can be found in ref[77].
543 Briefly, ORCHIDEE-CNP includes an interactive N and phosphorus cycle, whereas ORCHIDEE
544 is a C-only model. The rest have coupled C-N cycles. Relative changes were calculated based on
545 means over ten initial years (i, vary depending on the model) and j = 2008-2017 as $(C_j - C_i)/C_i$.
546 To reduce effects of discrepant response timescales of soil C and biomass, we estimated the
547 steady-state soil C storage (C^*) as:

548 $$C^* = \frac{C_j}{1 - \frac{\Delta C_j}{NPP_j}} \ ,$$

549 where ΔC_j is the change in soil C over the years 2008-2017. The relative change in soil C is then
550 taken as $(C^* - C_i) / C_i$. Data shown in Fig. 4 is based on pooled data from all six models. We
551 randomly sampled outputs from N gridcells for each model in order not to bias the visualization
552 towards models with a large number of gridcells (i.e. higher resolution). N is chosen as the
553 number of gridcells in the model with the coarsest resolution.
554
555 **CMIP5 Models - Expected β$_{soil}$.** We used projected SOC (Csoil) and biomass pool (Cveg)
556 responses to rising CO$_2$ as simulated by CMIP5 models as a comparison for the upscaled values
557 we derive from experiments. Specifically, we used data from the experiment "esmFixClim1", in
558 which CO$_2$ is increased by 1% per year from 285 ppm. In the esmFixClim1 experiment, the
559 increase in [CO$_2$] only affects vegetation and not the radiation code of the models, enabling a
560 quantification of the effect of eCO$_2$ in isolation (e.g. excluding warming), and thus a close
561 comparison with eCO$_2$ experiments. At a [CO$_2$] increasing rate of +1% year^{-1}, [CO$_2$] reaches 372
562 ppm (average concentration in ambient CO$_2$ plots in the dataset) in the 28th year and 616 ppm
563 (average concentration in elevated CO$_2$ plots in the dataset) in the 78th year. ΔCveg and ΔCsoil
564 were calculated as the difference between the respective carbon stocks in the 28th and the 78th
565 year.

566   Though plants in both experiments and our CMIP5 dataset see a similar increase in $[CO_2]$,
567   experiments simulate a step increase in CO_2 over half a decade, whereas the increase in CO_2 in
568   CMIP5 models is much slower and over the course of 50 years (Extended Data Table 2). As soil
569   organic matter turns over slowly, the resulting β_{soil} from experiments is lower than $\Delta Csoil$ from
570   models, and the comparison not meaningful. We thus focus on the specific relationship $\beta_{plant}:\beta_{soil}$
571   in experiments vs. models. Here, we calculated the spatially explicit ratio of $\Delta Cveg_{CMIP}$ to
572    $\Delta Csoil_{CMIP}$. This was done for five Earth System Models in the CMIP5 ensemble with
573   esmFixClim1 simulations (CanESM2; GFDL-ESM2M; HadGEM2-ES; IPSL-CM5A-LR; MPI-
574   ESM-LR). Then, we calculate "expected" β_{soil} from CMIP5 applying the same β_{plant} used for
575   experiments with the model-average $\Delta Cveg_{CMIP}$ to $\Delta Csoil_{CMIP}$ ratio:
576
577    $$expected\ (Mg\ C\ ha^{-1})\ = \beta plant\ \times \frac{\Delta Csoil_{CMIP}}{\Delta Cveg_{CMIP}}$$
578
579   With β_{plant} as the effect of elevated CO_2 on plant biomass derived from eCO_2
580   experiments. We then computed the difference between the expected (modeled) and observed
581   (upscaled) effects of elevated CO_2 on β_{soil}. As both expected and upscaled β_{soil} use the same β_{plant},
582   this transformation allows us to directly tackle the consequences of the different $\beta_{soil}/\beta_{plant}$ ratios
583   between experiments and models. We acknowledge, however, that the ratio is likely to change
584   over time, so the comparison needs to be interpreted with caution. We found, however, that first-
585   order models also simulate a positive relationship between β_{soil} and β_{plant} when forced to simulate
586   the same duration as experiments (Fig. 4a), suggesting that the sign of the $\beta_{soil}:\beta_{plant}$ relationship
587   in CMIP5 models would not likely reverse if CMIP5 models would be forced to simulate a step
588   increase in CO_2 over 5 years as in experiments.
589

Methods references

590
591   41.   Terrer, C. Report of Mutualistic Associations, Nutrients, and Carbon under eCO2 (ROMANCE) v1.0 dataset.
592   (2020) doi:10.6084/m9.figshare.11704491.v7.
593   42.   Dieleman, W. I. J. *et al.* Simple additive effects are rare: a quantitative review of plant biomass and soil
594   process responses to combined manipulations of CO2 and temperature. *Glob. Change Biol.* **18**, 2681–2693
595   (2012).
596   43.   Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. Multiple Outcomes or Time-Points within
597   a Study. in 225–238 (John Wiley & Sons, Ltd, 2009).
598   44.   Del Re, A. C. & Hoyt, W. T. MAd: Meta-Analysis with Mean Differences. *R Package Version 08-2* (2014).
599   45.   Song, J. & Wan, S. A global database of plant production and carbon exchange from global change
600   manipulative experiments. (2020) doi:10.6084/m9.figshare.7442915.v9.
601   46.   Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. *J. Stat. Softw.* **36**, (2010).
602   47.   Osenberg, C. W., Sarnelle, O., Cooper, S. D. & Holt, R. D. Resolving ecological questions through meta-
603   analysis: goals, metrics, and models. *Ecology* **80**, 1105–1117 (1999).
604   48.   Rubin, D. B. & Schenker, N. Multiple imputation in health are databases: An overview and some
605   applications. *Stat. Med.* **10**, 585–598 (1991).
606   49.   Lajeunesse, M. J. Facilitating systematic reviews, data extraction and meta analysis with the metagear
607   package for r. *Methods Ecol. Evol.* **7**, 323–330 (2016).
608   50.   Van Lissa, C. J. MetaForest: Exploring heterogeneity in meta-analysis using random forests. (2017)
609   doi:10.31234/osf.io/myg6s.
610   51.   Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **28**, (2008).
611   52.   Calcagno, V. & de Mazancourt, C. glmulti: an R package for easy automated model selection with
612   (generalized) linear models. *J. Stat. Softw.* **34**, (2010).
613   53.   van Groenigen, K. J. *et al.* Element interactions limit soil carbon storage. *Proc. Natl. Acad. Sci.* **103**, 6571–
614   6574 (2006).

615 54. Wang, B. & Qiu, Y. L. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**,
616 299–363 (2006).
617 55. Maherali, H., Oberle, B., Stevens, P. F., Cornwell, W. K. & McGlinn, D. J. Mutualism Persistence and
618 Abandonment during the Evolution of the Mycorrhizal Symbiosis. *Am. Nat.* **188**, E113–E125 (2016).
619 56. Terrer, C., Vicca, S., Hungate, B. A., Phillips, R. P. & Prentice, I. C. Mycorrhizal association as a primary
620 control of the CO fertilization effect. *Science* **353**, 72–74 (2016).
621 57. Walker, A. *et al.* FACE-MDS Phase 2: Model Output. (2018) doi:10.15485/1480327.
622 58. Medlyn, B. E. *et al.* Using ecosystem experiments to improve vegetation models. *Nat. Clim. Change* **5**, 528–
623 534 (2015).
624 59. Zaehle, S. *et al.* Evaluation of 11 terrestrial carbon–nitrogen cycle models against observations from two
625 temperate Free Air CO2 Enrichment studies. *New Phytol.* **202**, 803–822 (2014).
626 60. De Kauwe, M. G. *et al.* Where does the carbon go? A model-data intercomparison of vegetation carbon
627 allocation and turnover processes at two temperate forest free-air CO 2 enrichment sites. *New Phytol.* **203**,
628 883–899 (2014).
629 61. Walker, A. P. *et al.* Comprehensive ecosystem model data synthesis using multiple data sets at two
630 temperate forest free air CO2 enrichment experiments: Model performance at ambient CO2 concentration. *J.*
631 *Geophys. Res. Biogeosciences* **119**, 937–964 (2014).
632 62. Walker, A. P. *et al.* Decadal biomass increment in early secondary succession woody ecosystems is increased
633 by CO 2 enrichment. *Nat. Commun.* **10**, 454 (2019).
634 63. Schlesinger, W. *et al.* The Duke Forest FACE Experiment: CO2 Enrichment of a Loblolly Pine Forest. in
635 197–212 (2006). doi:10.1007/3-540-31237-4_11.
636 64. Hungate, B. A. *et al.* Cumulative response of ecosystem carbon and nitrogen stocks to chronic CO2 exposure
637 in a subtropical oak woodland. *New Phytol.* **200**, 753–766 (2013).
638 65. Jordan, D. N. *et al.* Biotic, abiotic and performance aspects of the Nevada Desert Free-Air CO2 Enrichment
639 (FACE) Facility. *Glob. Change Biol.* **5**, 659–668 (1999).
640 66. Carrillo, Y., Dijkstra, F., LeCain, D., Blumenthal, D. & Pendall, E. Elevated CO2 and warming cause
641 interactive effects on soil carbon and shifts in carbon use by bacteria. *Ecol. Lett.* **21**, 1639–1648 (2018).
642 67. Mueller, K. E. *et al.* Impacts of warming and elevated CO2 on a semi arid grassland are non additive, shift
643 with precipitation, and reverse over time. *Ecol. Lett.* **19**, 956–966 (2016).
644 68. Zak, D. R., Pregitzer, K. S., Kubiske, M. E. & Burton, A. J. Forest productivity under elevated CO2 and O3:
645 positive feedbacks to soil N cycling sustain decade long net primary productivity enhancement by CO2.
646 *Ecol. Lett.* **14**, 1220–1226 (2011).
647 69. Oleson, K. *et al.* Technical description of version 4.5 of the Community Land Model (CLM). (2013)
648 doi:10.5065/D6RR1W7M.
649 70. Clark, D. B. *et al.* The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon
650 fluxes and vegetation dynamics. *Geosci. Model Dev.* **4**, 701–722 (2011).
651 71. Krinner, G. *et al.* A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system.
652 *Glob. Biogeochem. Cycles* **19**, (2005).
653 72. Haverd, V. *et al.* A new version of the CABLE land surface model (Subversion revision r4601) incorporating
654 land use and land cover change, woody vegetation demography, and a novel optimisation-based approach to
655 plant coordination of photosynthesis. *Geosci. Model Dev.* **11**, 2995–3026 (2018).
656 73. Lawrence, D. M. *et al.* The Community Land Model Version 5: Description of New Features, Benchmarking,
657 and Impact of Forcing Uncertainty. *J. Adv. Model. Earth Syst.* **11**, 4245–4287 (2019).
658 74. Meiyappan, P., Jain, A. K. & House, J. I. Increased influence of nitrogen limitation on CO2 emissions from
659 future land use and land use change. *Glob. Biogeochem. Cycles* **29**, 1524–1548 (2015).
660 75. Smith, B. *et al.* Implications of incorporating N cycling and N limitations on primary production in an
661 individual-based dynamic vegetation model. *Biogeosciences* **11**, 2027–2054 (2014).
662 76. Goll, D. S. *et al.* A representation of the phosphorus cycle for ORCHIDEE (revision 4520). *Geosci. Model*
663 *Dev.* **10**, 3745–3770 (2017).
664 77. Friedlingstein, P. *et al.* Global Carbon Budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019).
665 78. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high resolution grids of monthly climatic
666 observations – the CRU TS3.10 Dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
667 79. Soudzilovskaia, N. A. *et al.* Global mycorrhizal plant distribution linked to terrestrial carbon stocks. *Nat.*
668 *Commun.* **10**, 1–10 (2019).
669 80. Hengl, T. *et al.* SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **12**,
670 e0169748 (2017).

81.   Batjes, N. H. Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. *Geoderma* **269**, 61–68 (2016).

82.   Shangguan, W., Dai, Y., Duan, Q., Liu, B. & Yuan, H. A global soil data set for earth system modeling. *J. Adv. Model. Earth Syst.* **6**, 249–263 (2014).

Code availability

The R code used in the analysis presented in this paper is available in GitHub and can be accessed at https://github.com/cesarterrer/SoilC_CO₂

Data availability

All the empirical data that support the main findings of this study have been deposited in Figshare: (https://figshare.com/account/home#/projects/74721) and GitHub (https://github.com/cesarterrer/SoilC_CO₂). FACE-MDS data can be accessed at https://www.osti.gov/dataexplorer/biblio/dataset/1480327. CMIP5 data can be accessed at https://esgf-index1.ceda.ac.uk/search/cmip5-ceda/. TRENDY data can be requested at http://dgvm.ceh.ac.uk/index.html.

Author contributions

C.T and R.P.P. conceived the original idea. C.T. designed the paper, with R.B.J., B.A.H. and K.J.vG contributing to the development of the conceptual framework. J.R. and C.T. collected the biomass and soil carbon data from experiments. M.C. collected MAOM data. K.vS and S.V. collected litter data. C.T. ran the statistical analyses and upscaling. B.D.S ran the analysis with TRENDY models. B.N.S., C.T., and B.A.H. ran the comparison with the FACE MDS data.

718 T.F.K, H.Z., and C.T. analyzed CMIP5 data. P.B.R., B.A.H., E.P., Y.C., R.D.E, R.B.J., and
719 many others ran the experiments. C.T. and B.A.H. wrote the first draft, with input from all
720 authors.
721
722 **Competing interests**
723 The authors declare no competing interests.
724
725 **Additional information**
726 **Supplementary Information** is available for this paper
727 Correspondence and requests for materials should be addressed to C.T: terrer@stanford.edu
728
729 **Extended Data**
730
731 **Extended Data Figure 1 | Relationship of the effects of eCO$_2$ on aboveground biomass**
732 **production with litter production (a) and soil carbon storage (b)**. Results for non-fertilized
733 field eCO$_2$ experiments (n=10, and n=35, respectively). Dots represent individual experiments,
734 with dot size in (b) proportional to the weights in the meta-regression.
735
736 **Extended Data Figure 2 | Variable importance of 19 predictors of the effects of CO$_2$ on soil**
737 **carbon and biomass stocks**. (a-b) Variable importance of the effect of CO$_2$ on soil carbon
738 stocks in (a) relative (%) and (b) absolute terms (g m^{-2}) across the full dataset (n=108). (c)
739 Variable importance of the effect of CO$_2$ on soil carbon stocks (%) across the subset of eCO$_2$
740 experiments in "intact" soils (n=73). (d) Variable importance of the effect of CO$_2$ on plant
741 aboveground biomass (n=138). Variable importance in (a,b,d) is quantified based on a meta-
742 forest model. Variable importance in (c) is quantified based on the sum of AICc weights, which
743 allows for the quantification of the importance of interactions between predictors. As an initial
744 step, moderators that consistently displayed negative variable importance (i.e., that showed a
745 reduction in predictive performance) were automatically dropped.
746
747 **Extended Data Figure 3 | Effects of elevated CO$_2$ on soil carbon and plant biomass in the**
748 **subset of nitrogen-fertilized eCO$_2$ studies (n=35)**, expressed as a regression (a) and overall
749 effects in meta-analysis (b). Dot sizes in (a) represent the individual studies and are drawn
750 proportional to the weights in the model. The regression with the subset of non-fertilized studies
751 is also shown here for comparison. Dots in (b) represent the effect sizes and 95% confidence
752 intervals from the meta-analysis.
753
754 **Extended Data Figure 4 | Analysis of variables potentially explaining the observed effects of**
755 **elevated CO$_2$ on soil carbon.** Effects of elevated CO$_2$ on root biomass (n=45), fine-root
756 production (n=11), litter C:N (n=16) and background soil carbon stocks (n=38), between
757 ecosystem types (grassland vs. forest) and nutrient-acquisition strategies (arbuscular mycorrhizae
758 –AM– vs. ectomycorrhizae –ECM–). Boxplots show the median, the first to third quartile, the
759 1.5x interquartile ranges, and outliers.
760
761 **Extended Data Figure 5 | Partial dependence plots of the six most important predictors of**
762 **the effect of elevated CO$_2$ on soil carbon stocks across 108 experiments.** The figure shows the
763 predicted CO$_2$ effect (yi) in relative (a) and absolute terms (b) across each predictor and the most

important interaction between predictors (right panels) in a random-forest meta-analysis. Error bands represent 95% confidence intervals. Partial regression plots give a graphical depiction of the marginal effect of a variable on the response and the shape and direction of the relationship. Little variation in yi across the values of a predictor generally reflects the low predictive power of the predictor for yi. However, important predictors may show little variation in yi when involved in interactions, so the right panels show the most important interaction in the model. More details about the different predictors in Extended Data Table 1. From a total of 19 predictors, only the six most important predictors and the most important interaction are shown here.

Extended Data Figure 6 | Representativeness of the upscaling predictors of the effect of elevated CO_2 on soil carbon stocks. Histograms showing the distribution of both the predictors in the training dataset of CO_2 experiments and the data used to upscale the global distribution of the effect. Predictions exclude regions between -15 to 15 and from 60 to 90 degrees latitude due to the lack of experiments.

Extended Data Figure 7 | Relationship between the effects of CO_2 on aboveground biomass and soil carbon across individual models from three model ensembles. (a) FACE Model Data Synthesis Phase 2. Individual model results are represented by colored symbols and lines. Each symbol represents one site; lines represent model-specific linear regressions. To ease interpretation of the results and the comparison with Fig. 4 axis limits are set. Dashed line and error band represent the linear regression line and standard error across all experiment-by-model results. (b) TRENDY v7 models. (c) CMIP5 models.

Extended Data Figure 8 | Difference between expected CO_2 effects on soil carbon stocks based on TRENDY models and upscaled based on experiments. Expected values result from the relationship between β_{soil} and β_{plant} coded in models. Positive values (red color) indicate an overestimation by models; negative values (blue color) indicate an underestimation by models.

Extended Data Table 1 | List of predictors used to examine and upscale the effects of elevated CO_2 on soil carbon.

Extended Data Table 2 | Synthetic description of the basic characteristics of three model ensembles in terms of their treatment of CO_2 effects.

Ecosystem type
Agricultural (27)
Grassland (24)
Shrubland (6)
Tree Stand (50)

Nitrogen fertilization
Yes (55)
No (53)

Soil C stocks (g m^{-2})
[193,1512) (36)
[1512,2927) (36)
[2927,8021] (36)

Experiment type
Chamber (20)
FACE (59)
OTC (29)

Disturbance
Disturbed (30)
Intact (78)

Nutrient strategy
AM (47)
ECM (46)
AM-ER (4)
N-fixer (10)

Overall effect

-10   0   10   20

CO_2 effect on soil carbon (%)

a

CO$_2$ effect on soil carbon (%) (y-axis, ranging from -20 to 40)
CO$_2$ effect on plant biomass carbon (%) (x-axis, ranging from 0 to 60)

b

Plants Soils

CO$_2$ effect on carbon pools (%) (y-axis, ranging from -10 to 40)
AM ECM (x-axis)

c

CO$_2$ effect on N-uptake (%) (y-axis, ranging from 0 to 15)
AM ECM (x-axis)

d

CO$_2$ effect on MAOM (%) (y-axis, ranging from -2 to 2)
AM ECM (x-axis)

a

Upscaled eCO$_2$ effect on soil C (%)

-5　0　5　10　15　20　25

b

eCO$_2$ effect on soil C (%)

Agricultural　Forest　Grassland　Shrubland

c

Uncertainty in soil C estimate (%)

0　2　4　6　8

d

Uncertainty (%)

Agricultural　Forest　Grassland　Shrubland

e

(Expected - Upscaled) change in soil C (%)

-15　-10　-5　0　5　10　15

f

Expected - Upscaled (%)

Agricultural　Forest　Grassland　Shrubland