LETTERS
https://doi.org/10.1038/s41558-019-0545-2

nature climate change

There are amendments to this paper

Nitrogen and phosphorus constrain the CO_2 fertilization of global plant biomass

César Terrer[1,2,3*], Robert B. Jackson[1,4], I. Colin Prentice[5,6,7], Trevor F. Keenan[8,9], Christina Kaiser[10,11], Sara Vicca[12], Joshua B. Fisher[13,14], Peter B. Reich[15,16], Benjamin D. Stocker[17], Bruce A. Hungate[18,19], Josep Peñuelas[17,20], Ian McCallum[3], Nadejda A. Soudzilovskaia[21], Lucas A. Cernusak[22], Alan F. Talhelm[23], Kevin Van Sundert[12], Shilong Piao[24,25], Paul C. D. Newton[26], Mark J. Hovenden[27], Dana M. Blumenthal[28], Yi Y. Liu[29], Christoph Müller[30,31], Klaus Winter[32], Christopher B. Field[4], Wolfgang Viechtbauer[33], Caspar J. Van Lissa[34], Marcel R. Hoosbeek[35], Makoto Watanabe[36], Takayoshi Koike[37], Victor O. Leshyk[18,19], H. Wayne Polley[38] and Oskar Franklin[3]

Elevated CO_2 (eCO_2) experiments provide critical information to quantify the effects of rising CO_2 on vegetation[1–6]. Many eCO_2 experiments suggest that nutrient limitations modulate the local magnitude of the eCO_2 effect on plant biomass[1,3,5], but the global extent of these limitations has not been empirically quantified, complicating projections of the capacity of plants to take up CO_2[7,8]. Here, we present a data-driven global quantification of the eCO_2 effect on biomass based on 138 eCO_2 experiments. The strength of CO_2 fertilization is primarily driven by nitrogen (N) in ~65% of global vegetation and by phosphorus (P) in ~25% of global vegetation, with N- or P-limitation modulated by mycorrhizal association. Our approach suggests that CO_2 levels expected by 2100 can potentially enhance plant biomass by 12 ± 3% above current values, equivalent to 59 ± 13 PgC. The future effect of eCO_2 we derive from experiments is geographically consistent with past changes in greenness[9], but is considerably lower than the past effect derived from models[10]. If borne out, our results suggest that the stimulatory effect of CO_2 on carbon storage could slow considerably this century. Our research provides an empirical estimate of the biomass sensitivity to eCO_2 that may help to constrain climate projections.

Levels of eCO_2 affect the functioning and structure of terrestrial ecosystems and create a negative feedback that reduces the rate of global warming[8,9,11–14]. However, this feedback remains poorly quantified, introducing substantial uncertainty in climate change projections[7,8]. Experiments with eCO_2 simulate the response of plants to eCO_2 and thereby provide important empirical and mechanistic

[1]Department of Earth System Science, Stanford University, Stanford, CA, USA. [2]Institut de Ciència i Tecnologia Ambientals, Universitat Autònoma de Barcelona, Barcelona, Spain. [3]Ecosystems Services and Management Program, International Institute for Applied Systems Analysis, Laxenburg, Austria. [4]Woods Institute for the Environment and Precourt Institute for Energy, Stanford University, Stanford, CA, USA. [5]AXA Chair Programme in Biosphere and Climate Impacts, Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, UK. [6]Department of Biological Sciences, Macquarie University, North Ryde, New South Wales, Australia. [7]Department of Earth System Science, Tsinghua University, Beijing, China. [8]Department of Environmental Science, Policy and Management, UC Berkeley, Berkeley, CA, USA. [9]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [10]Department of Microbiology and Ecosystem Science, Division of Terrestrial Ecosystem Research, Faculty of Life Sciences, University of Vienna, Vienna, Austria. [11]Evolution and Ecology Program, International Institute for Applied Systems Analysis, Laxenburg, Austria. [12]Centre of Excellence PLECO (Plants and Ecosystems), Biology Department, University of Antwerp, Wilrijk, Belgium. [13]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. [14]Joint Institute for Regional Earth System Science and Engineering, University of California at Los Angeles, Los Angeles, CA, USA. [15]Department of Forest Resources, University of Minnesota, St. Paul, MN, USA. [16]Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales, Australia. [17]CREAF, Cerdanyola del Vallès, Spain. [18]Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, USA. [19]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA. [20]CSIC, Global Ecology Unit CREAF-CEAB-UAB, Bellaterra, Spain. [21]Environmental Biology Department, Institute of Environmental Sciences, Leiden University, Leiden, the Netherlands. [22]College of Marine and Environmental Sciences, James Cook University, Cairns, Queensland, Australia. [23]Department of Forest, Rangeland and Fire Sciences, College of Natural Resources, University of Idaho, Moscow, ID, USA. [24]Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing, China. [25]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China. [26]Land & Environmental Management, AgResearch, Palmerston North, New Zealand. [27]School of Biological Sciences, University of Tasmania, Hobart, Tasmania, Australia. [28]Rangeland Resources & Systems Research Unit, Agricultural Research Service, United States Department of Agriculture, Fort Collins, CO, USA. [29]School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing, China. [30]Department of Plant Ecology, Justus Liebig University of Giessen, Giessen, Germany. [31]School of Biology and Environmental Science, University College Dublin, Belfield, Ireland. [32]Smithsonian Tropical Research Institute, Balboa, Republic of Panama. [33]Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, the Netherlands. [34]Department of Methodology and Statistics, Utrecht University, Utrecht, the Netherlands. [35]Soil Chemistry, Wageningen University, Wageningen, the Netherlands. [36]Institute of Agriculture, Tokyo University of Agriculture and Technology, Fuchu, Japan. [37]Graduate School of Agriculture, Hokkaido University, Sapporo, Japan. [38]USDA, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX, USA. *e-mail: cesar.terrer@me.com

constraints for climate projections. Numerous eCO_2 experiments have been conducted over the last three decades and they collectively provide strong evidence for a fertilizing effect of eCO_2 on leaf-level photosynthesis[6]. At the ecosystem level, however, individual CO_2 experiments show contrasting results for the magnitude of the growth and biomass response to eCO_2, ranging from strongly positive in some studies[2] to little or no response with N[1], P[5] or water[3] limitations in other studies. Despite this conflicting evidence at the ecosystem scale, a global-scale carbon (C) sink in terrestrial ecosystems is robustly inferred[12].

Here, we synthesize 1,432 observations from 138 eCO_2 studies in grassland, shrubland, cropland and forest systems (Supplementary Figs. 1 and 2 and Supplementary Table 1), encompassing free-air CO_2 enrichment (FACE) and chamber experiments. We train a random-forest meta-analysis model with this dataset and identify the underlying factors that explain variability within it. We use these relationships to estimate the global-scale change in biomass in response to an increase in atmospheric CO_2 from 375 ppm to 625 ppm, which is the increase in CO_2 expected by 2100 in an intermediate emission scenario.

We included 56 potential predictors of the CO_2 effect (Supplementary Table 2) belonging to four main categories: nutrients (N, P, mycorrhizal association; see ref. [4]), climate (for example, precipitation and temperature), vegetation (age and type) and experimental methodology (for example, the increase in CO_2 concentration (ΔCO_2) and the type of CO_2 fumigation technology). More details on the model selection are available in the Supplementary Discussion.

The random-forest meta-analysis indicated that the most important predictors of the CO_2 fertilization effect on biomass in our dataset were experiment type (FACE or chambers), soil C:N ratio (an indicator of N availability), soil P availability and mycorrhizal type, with different relationships for C:N and P between mycorrhizal types ($y \approx$ Mycorrhizal_type × N + Mycorrhizal_type × P + Fumigation_type, pseudo-$R^2 = 0.94$). A sensitivity test using a larger dataset of 205 studies confirmed the robustness of the relationships described by the statistical model (Supplementary Discussion). Among 56 potential predictors, mycorrhizal type was the primary modulator of above-ground biomass responses to eCO_2 ($P < 0.001$) (Supplementary Fig. 3).

The eCO_2 effect in arbuscular mycorrhizal (AM) plants was best predicted by soil C:N (Fig. 1a, $P < 0.001$), but not significantly by P (Supplementary Fig. 4a, $P = 0.2830$). The C:N ratio of soil organic matter is a proxy for plant N availability because it is associated with stoichiometric limitations of microbial processes in the soil[15]. Although the constraining role of N on CO_2 fertilization has been reported in many eCO_2 studies[1,3,6], here we find that soil C:N is a powerful indicator to quantify the N-limitation on CO_2 fertilization across experiments.

In contrast, the eCO_2 effect in ectomycorrhizal (ECM) plants was best predicted by soil P (Fig. 1b, $P < 0.001$), but not significantly by soil C:N (Supplementary Fig. 4b, $P = 0.1141$). The critical role of P on CO_2 fertilization across a large number of studies was unexpected, but consistent with an increasing body of research[5,16].

Once the effects of mycorrhizal type, C:N, P and fumigation type were accounted for, other predictors such as climate, biome type (for example, temperate tree versus grass) or the age of the vegetation did not explain an important fraction of the variability in the effect (Supplementary Fig. 5). Previous studies have variously attributed differences in the magnitude of the CO_2 effect to either average temperature (MAT) or precipitation (MAP), or to both[17] (see Supplementary Discussion). Using the model $y \approx$ MAT + MAP + Fumigation_type instead of our final model reduced explained variability (R^2) from 0.94 to 0.05. These results suggest that the CO_2 fertilization can only be reliably predicted when nutrient availability is considered.

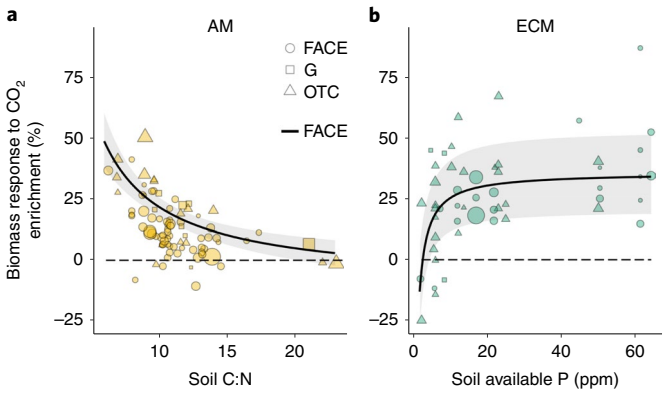

Fig. 1 | Soil C:N and soil P are key plant resources driving the CO_2 fertilization effect on above-ground biomass. Model selection identified the most important drivers of the effect in the dataset of CO_2 experiments ($n = 138$), indicating responses to CO_2 were modulated by mycorrhizal type. **a**,**b**, Meta-analytic scatterplots showing the relationship between the CO_2 effect and soil C:N (an indicator of nitrogen availability) in AM studies ($n = 86$) at 0–10 cm (**a**), and soil available phosphorus in ECM studies ($n = 52$) measured by the Bray method at 0–10 cm (**b**). The type of fumigation technology used (FACE, growth chamber and open top chamber) significantly influenced ($P < 0.001$) the magnitude of the CO_2 effect. Regression lines represent the response found in FACE studies, based on a mixed-effects meta-regression model (pseudo-$R^2 = 0.94$) and their 95% confidence intervals. Dot sizes are drawn proportional to the weights in the model and represent, on average, an increase in atmospheric CO_2 of 250 ppm. G, growth chamber; OTC, open top chamber.

We used the quantitative relationships derived from the meta-analysis to predict the global distribution of the eCO_2 effect based on maps for soil C:N, P and mycorrhizal type. Plant responses to eCO_2 were significantly higher in open top chamber and growth chamber experiments than in FACE (Supplementary Fig. 5, $P < 0.001$) (see Supplementary Discussion), so we included Fumigation_type as a predictor in the scaling model to produce projections that are consistent with the response found in FACE experiments, as they allow CO_2 to be fumigated with as little disturbance as possible.

Our global projections from FACE experiments show a relative increase in biomass of 12 ± 3% (Fig. 2a and Table 1) for the average 250 ppm ΔCO_2 across experiments. The magnitude of the global effect is less than the overall effect of ~20% found previously in meta-analyses[4,6] and the ~30% effect found in several FACE experiments[2,4]. This reduction arises in part because many CO_2 experiments were conducted in relatively fertile soils or under nutrient fertilization regimes. Thus, extrapolating nutrient relationships to areas with naturally poor soils results in a lower global effect. In absolute terms, we estimated a global increase in total biomass of 59 ± 13 PgC for a 250 ppm ΔCO_2 (Fig. 2b and Table 1), scaled from satellite observations of current above-ground biomass[18] and region-specific total to above-ground biomass ratios from the literature (Supplementary Table 4). Global anthropogenic emissions are currently around 10 PgC annually[12], hence the additional C-sequestration in biomass is equivalent to 5–6 years of intermediate CO_2 emissions.

Forests show the largest relative increases in biomass (Table 1 and Fig. 2a). Tropical forests are characterized by low P (Supplementary Fig. 6). However, their association with AM fungi, together with relatively high N (Supplementary Fig. 6), support a widespread, though moderate, biomass enhancement. Our approach does not explicitly include symbiotic acquisition of atmospheric N (N_2-fixation), which is relatively common in tropical forests[19]. Indeed, tropical N_2-fixing

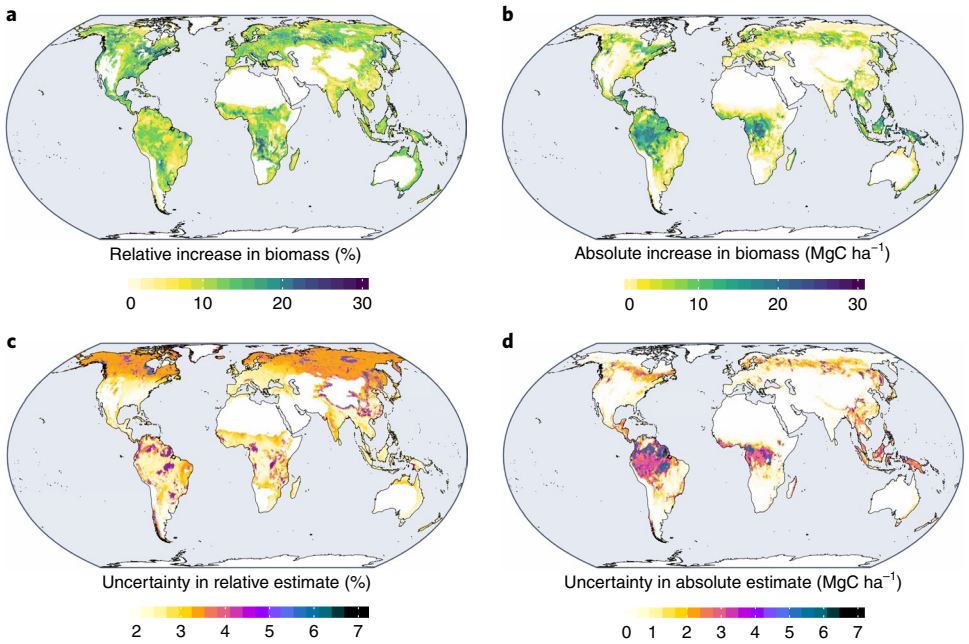

Fig. 2 | Potential above-ground biomass enhancement in terrestrial ecosystems under elevated CO_2. These global estimates were upscaled using the empirical relationships from the key drivers of the effect across experiments synthesized through meta-analysis. **a**–**d**, The effects (**a**,**b**) and uncertainties in the estimations (**c**,**d**) shown in this figure represent an increase in atmospheric CO_2 of 250 ppm in relative (**a**,**c**) and absolute (**b**,**d**) terms. Uncertainties are based on the standard error of the effect adjusted for climate and nutrient sampling coverage, that is, increasing in areas with values of nutrient availability, temperature and precipitation poorly represented by CO_2 experiments.

Table 1 | Summary of upscaled changes in plant above ground and total biomass to elevated CO_2 across biomes

Biome	AM/ECM	C:N	P	AGB relative effect (%)	AGB absolute effect (PgC)	TB absolute effect (PgC)
Boreal forest	20/80	16.5	9	13.5 ± 4	8.1 ± 2.2	10 ± 2.9
Cropland	90/10	11.5	14.5	10 ± 1	2.5 ± 0.4	3.1 ± 0.5
Grassland	80/20	13	14.5	8 ± 1	1.2 ± 0.2	3.7 ± 0.8
Mixed	65/35	14	9.5	10.5 ± 2	2.2 ± 0.5	2.3 ± 0.6
Shrubland	80/20	13	11.5	11.5 ± 2	1.7 ± 0.3	6.8 ± 1.3
Temperate forest	40/60	13	11	14 ± 3	4.2 ± 1.1	4.8 ± 1.4
Tropical forest	80/20	12	7.5	12.5 ± 3	22.7 ± 6.5	31.4 ± 8.9
Global	65/25	13	11	12 ± 3	41.1 ± 9.5	58.7 ± 13.4
Rate					17.3 ± 4.0	24.8 ± 5.7

AM/ECM is the dominance of AM or ECM plants per biome type. C:N and P are the average soil C:N and available phosphorus (ppm) by the Bray method at 0–10 cm. AGB, plant above ground; TB, total biomass. Absolute changes are given for an increase in atmospheric CO_2 of 250 ppm. The final row shows the absolute rate of increase in PgC for a standardized increase in CO_2 of 100 ppm. Uncertainties represent the standard error of the effect.

species can show larger CO_2 effects than non N_2-fixing species[20], and thus the response in tropical forests in our model may be underestimated. Nevertheless, our dataset contains tropical N_2-fixing species[21], indirectly including this effect. Temperate grasslands, which are also dominated by AM plants, show the lowest relative biomass increment as a result of N-limitations. In temperate forests, some of the largest relative increases (~30%) occur in ECM-forests when P is high, but AM-forests show low relative biomass increases due to moderately high C:N (Supplementary Fig. 6).

The absolute eCO_2 effect is dominated by tropical forests (Table 1 and Fig. 2b), consistent with ground-based measurements showing increases in above-ground biomass in recent decades in intact tropical forests[22], with CO_2 identified as the main driver[22,23]. To account for uncertainties, and to highlight the environmental conditions not well represented in eCO_2 experiments, we computed the standard error of the projections (Methods). Wet-tropical and boreal ecosystems show the largest uncertainties in absolute and relative terms, respectively (Fig. 2c,d), reflecting the limited number of studies in ecosystems with extreme values of climate and nutrient availability.

To assess the magnitude of the global eCO_2 effect we derive from FACE, we compared it with the increase in biomass attributed to rising CO_2 concentration (β) from 1980 to 2010 by the TRENDY ensemble of dynamic global vegetation models (DGVMs), standardized to 100 ppm ΔCO_2. Our estimated rate of increase in total biomass is 25 ± 4 PgC 100 ppm^{-1} (5 ± 1% 100 ppm^{-1}), which is substantially below that of most DGVMs (Fig. 3a).

For comparing the geographical distribution of our global eCO_2 effect, we used satellite-based observations of changes in leaf area (greening)[9] attributed to CO_2 rising in the period 1982–2009. Although changes in greenness and above-ground biomass are not

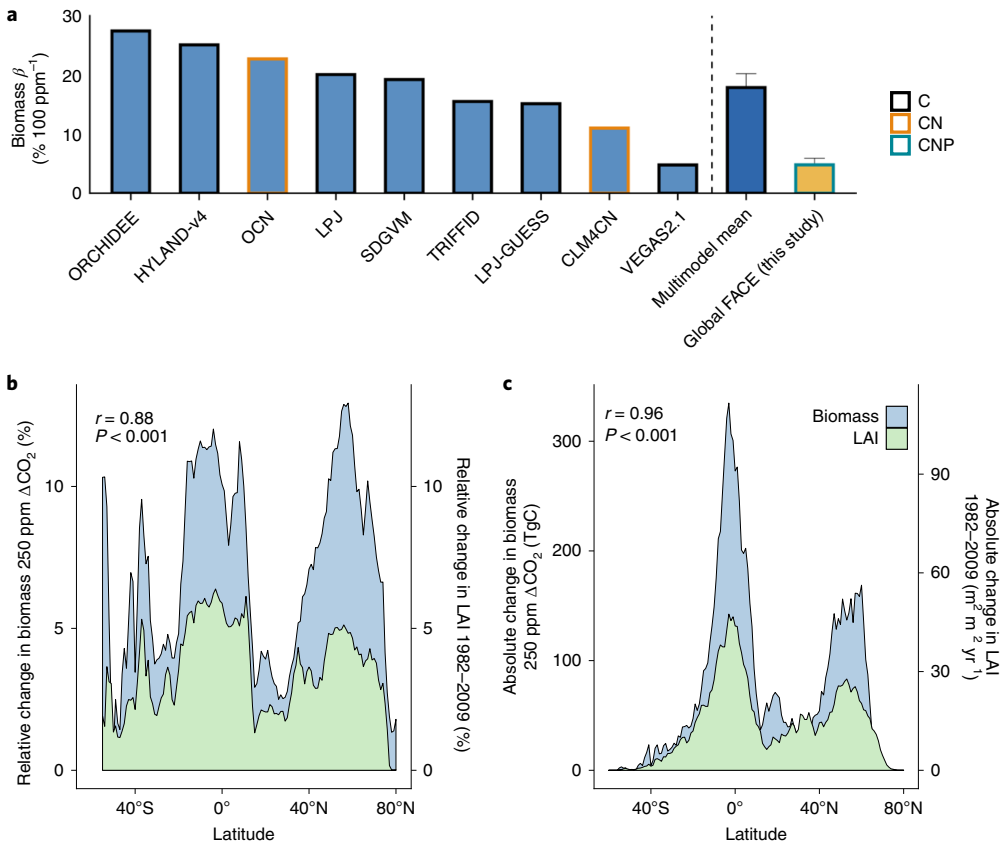

Fig. 3 | Comparison of the global effect of elevated CO_2 with existing independent approaches. a, Comparison of the magnitude of the effect of elevated CO_2 on total biomass and the sensitivity of total biomass to the historical increase in atmospheric CO_2 (β) in the period 1980–2010 as estimated by nine vegetation models. Results were standardized to 100 ppm ΔCO_2. C, carbon-only models; CN, carbon models with coupled nitrogen cycle; CNP, carbon, nitrogen and phosphorus limitations. **b**,**c**, Comparison of the latitudinal distribution of the relative (**b**) and absolute (**c**) effect of elevated CO_2 on above-ground biomass (this study) and past changes in leaf area index (LAI) attributed to the increase in atmospheric CO_2 in the period 1982–2009 (LAI data from ref. [9]). Relative latitudinal changes were computed as the average relative effect of all pixels contained within 1° latitude. Absolute latitudinal changes were computed as the sum of the absolute effect in all pixels contained within 1° latitude.

necessarily correlated, we found an intriguingly strong correlation between the contemporary CO_2-driven increase in greenness and our independently estimated biomass projections (Fig. 3b,c).

In summary, our results suggest that plant biomass responses to eCO_2 are driven primarily by interactions with N and P modulated by mycorrhizal status. N constrains the strength of CO_2 fertilization in most AM plants (Fig. 1a), which currently store ~65% of terrestrial vegetation C[24], probably because the ability of AM fungi to supply plants with N is relatively small[25,26]. In contrast, we observed that P availability alters the biomass response to eCO_2 in ECM plants, which store ~25% of terrestrial vegetation C[24]. The sensitivity of ECM plants to P availability may be driven by the positive effect of eCO_2 on N uptake in ECM plants[26], which, together with widespread N deposition, might reinforce the limiting role of P[27] in the ecosystem.

Although our analysis uses the most comprehensive dataset of eCO_2 observations currently available, it has several limitations. First, our data-driven approach, unlike DGVMs, is not intended to capture the complex interactions that drive long-term changes in the C cycle, such as warming, disturbance, changes in water availability or N deposition. Instead, it is aimed at the empirical quantification of net CO_2 effects, providing constraints on the attribution of modelled biomass responses to CO_2 and a better mechanistic understanding of the underlying drivers of the effect. Second, tropical and boreal ecosystems are under-represented in global eCO_2 experiments (Supplementary Fig. 1). We have accounted for this uncertainty in our estimates, which we also use to highlight the specific regions where eCO_2 experiments are urgently needed. Furthermore, it is critical that comprehensive soil data in eCO_2 experiments are reported, ideally in more long-term studies.

The comparison between the global-level responses to eCO_2 found in FACE and past changes in biomass and greening attributed to CO_2 has important implications. First, the convergence with satellite-based observations of greening supports our projections, indicating that empirical relationships with soil nutrients can be powerful for explaining large-scale patterns of eCO_2 responses, despite ecosystem-level uncertainties. Thus, all else equal, the same ecosystems that are currently responsible for most of the greening[9] and C uptake[11,14] may remain important for future increases in biomass under eCO_2 (see Fig. 3b,c). Second, the effect attributed to rising CO_2 in past decades by DGVMs is larger in magnitude than our predicted effect of increasing CO_2 expected in the future (Fig. 3a). Different hypotheses could help explain this disagreement, including turnover rates of biomass in relatively short-term scale in experiments versus century-scale model simulations. This disagreement could be also explained based on the declining sensitivity of photosynthesis and growth towards higher CO_2 levels[28], or due to an increasingly constraining role of soil nutrients under elevated CO_2 concentration in FACE than under lower CO_2 concentration in models[29]. More targeted model simulations

mimicking experiments are needed to understand the causes of this discrepancy[30].

A key strength of our upscaling approach is that it synthesizes observational evidence at local scales and captures a global view of the eCO_2 effect on plant biomass and its drivers. DGVMs differ at the process level (including the current effects of CO_2 on biomass, see Fig. 3a), and consequently vary when projecting the future. Our data-based approach, along with new data from ongoing experiments, can be updated continuously and used to calibrate DGVMs, providing an empirical constraint for model simulations of the biomass sensitivity to CO_2.

Our results highlight the key role of terrestrial ecosystems, in particular forests, in mitigating the increase in atmospheric CO_2 resulting from anthropogenic emissions. Thus, if deforestation and land use changes continue decreasing the extent of forests, or if warming and other global changes diminish or reverse the land carbon sink, we will lose an important contribution towards limiting global warming.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41558-019-0545-2.

Received: 1 November 2018; Accepted: 4 July 2019;
Published online: 12 August 2019

References

1. Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E. & McMurtrie, R. E. CO_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proc. Natl Acad. Sci. USA* **107**, 19368–19373 (2010).
2. McCarthy, H. R. et al. Re-assessment of plant carbon dynamics at the Duke free-air CO_2 enrichment site: interactions of atmospheric [CO_2] with nitrogen and water availability over stand development. *New Phytol.* **185**, 514–528 (2010).
3. Reich, P. B., Hobbie, S. E. & Lee, T. D. Plant growth enhancement by elevated CO_2 eliminated by joint water and nitrogen limitation. *Nat. Geosci.* **7**, 920–924 (2014).
4. Terrer, C., Vicca, S., Hungate, B. A., Phillips, R. P. & Prentice, I. C. Mycorrhizal association as a primary control of the CO_2 fertilization effect. *Science* **353**, 72–74 (2016).
5. Ellsworth, D. S. et al. Elevated CO_2 does not increase eucalypt forest productivity on a low-phosphorus soil. *Nat. Clim. Change* **320**, 279–282 (2017).
6. Ainsworth, E. A. & Long, S. P. What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytol.* **165**, 351–372 (2005).
7. Friedlingstein, P. et al. Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Clim.* **27**, 511–526 (2014).
8. Ciais, P. et al. in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. et al.) 465–570 (IPCC, Cambridge Univ. Press, 2013).
9. Zhu, Z. et al. Greening of the Earth and its drivers. *Nat. Clim. Change* **6**, 791–795 (2016).
10. Sitch, S. et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653–679 (2015).
11. Keenan, T. et al. Recent pause in the growth rate of atmospheric CO_2 due to enhanced terrestrial carbon uptake. *Nat. Commun.* **7**, 13428 (2016).
12. Le Quéré, C. et al. Global Carbon Budget 2018. *Earth Syst. Sci. Data* **10**, 2141–2194 (2018).
13. Campbell, J. E. et al. Large historical growth in global terrestrial gross primary production. *Nature* **544**, 84–87 (2017).
14. Schimel, D., Stephens, B. B. & Fisher, J. B. Effect of increasing CO_2 on the terrestrial carbon cycle. *Proc. Natl Acad. Sci. USA* **112**, 436–441 (2015).
15. Manzoni, S., Jackson, R. B., Trofymow, J. A. & Porporato, A. The global stoichiometry of litter nitrogen mineralization. *Science* **321**, 684–686 (2008).
16. Hoosbeek, M. R. Elevated CO_2 increased phosphorous loss from decomposing litter and soil organic matter at two FACE experiments with trees. *Biogeochemistry* **127**, 89–97 (2016).
17. Fernández-Martínez, M. et al. Global trends in carbon sinks and their relationships with CO_2 and temperature. *Nat. Clim. Change* **10**, 1–79 (2018).
18. Liu, Y. Y. et al. Recent reversal in loss of global terrestrial biomass. *Nat. Clim. Change* **5**, 470–474 (2015).
19. Ter Steege, H. et al. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* **443**, 444–447 (2006).
20. Nasto, M. K., Winter, K., Turner, B. L. & Cleveland, C. C. Nutrient acquisition strategies augment growth in tropical N_2 fixing trees in nutrient poor soil and under elevated CO_2. *Ecology* **100**, e02646 (2019).
21. Cernusak, L. A. et al. Responses of legume versus nonlegume tropical tree seedlings to elevated CO_2 concentration. *Plant Physiol.* **157**, 372–385 (2011).
22. Qie, L. et al. Long-term carbon sink in Borneo's forests halted by drought and vulnerable to edge effects. *Nat. Commun.* **8**, 1966 (2017).
23. Almeida Castanho, A. D. et al. Changing Amazon biomass and the role of atmospheric CO_2 concentration, climate, and land use. *Glob. Biogeochem. Cycles* **30**, 18–39 (2016).
24. Soudzilovskaia, N. A. et al. Global mycorrhizal plants distribution linked to terrestrial carbon stocks. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/331884v2 (2018).
25. Hodge, A. & Storer, K. Arbuscular mycorrhiza and nitrogen: implications for individual plants through to ecosystems. *Plant Soil* **386**, 1–19 (2015).
26. Terrer, C. et al. Ecosystem responses to elevated CO_2 governed by plant–soil interactions and the cost of nitrogen acquisition. *New Phytol.* **217**, 507–522 (2018).
27. Peñuelas, J. et al. Human-induced nitrogen-phosphorus imbalances alter natural and managed ecosystems across the globe. *Nat. Commun.* **4**, 2934 (2013).
28. De Kauwe, M. G., Keenan, T., Medlyn, B. E., Prentice, I. C. & Terrer, C. Satellite based estimates underestimate the effect of CO_2 fertilization on net primary productivity. *Nat. Clim. Change* **6**, 892–893 (2016).
29. Wieder, W. R., Cleveland, C. C., Smith, W. K. & Todd-Brown, K. Future productivity and carbon storage limited by terrestrial nutrient availability. *Nat. Geosci.* **8**, 441–444 (2015).
30. Medlyn, B. E. et al. Using ecosystem experiments to improve vegetation models. *Nat. Clim. Change* **5**, 528–534 (2015).

Acknowledgements

We thank C. Körner, R. Norby, M. Schneider, Y. Carrillo, E. Pendall, B. Kimball, M. Watanabe, T. Koike, G. Smith, S.J. Tumber-Davila, T. Hasegawa, B. Sigurdsson, S. Hasegawa, A.L. Abdalla-Filho and L. Fenstermaker for sharing data and advice. This research is a contribution to the AXA Chair Programme in Biosphere and Climate Impacts and the Imperial College initiative Grand Challenges in Ecosystems and the Environment. Part of this research was developed in the Young Scientists Summer Program at the International Institute for Systems Analysis, Laxenburg (Austria) with financial support from the Natural Environment Research Council (UK). C.T. also acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities through the María de Maeztu programme for Units of Excellence (grant no. MDM-2015-0552). I.C.P. acknowledges support from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant no. 787203 REALM). S.V. and K.v.S. acknowledge support from the Fund for Scientific Research, Flanders (Belgium). T.F.K. acknowledges support by the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy under contract DE-AC02-05CH11231 as part of the RuBiSCo SFA. J.P. acknowledges support from the European Research Council through Synergy grant no. ERC-2013-SyG-610028 'IMBALANCE-P'. T.F.K. and J.B.F. were supported in part by NASA IDS Award no. NNH17AE86I. J.B.F. was also supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research. J.B.F. contributed to this research from Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. California Institute of Technology. N.A.S. was supported by Vidi grant no. 016.161.318 by the Netherlands Organization for Scientific Research. This paper is a contribution to the Global Carbon Project.

Author contributions

The study was originally conceived and developed by C.T., with ideas and contributions by R.J., I.C.P., O.F., T.F.K., P.B.R., C.K., S.V, B.S. and J.B.F. Data from DGVMs were analysed by T.F.K. Analysis of drivers was done by C.T and P.B.R. Statistical analysis was carried out by C.T., C.J.v.L. and W.V. Spatial analysis was done by C.T. and I.M. P.B.R., B.A.H., L.A.C., A.F.T., P.C.D.N., M.J.H., D.M.B., C.M., K.W., C.B.F., M.R.H., M.W., T.K., H.W.P. and many others ran the experiments. The initial manuscript was written by C.T. with input from all authors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41558-019-0545-2.

Correspondence and requests for materials should be addressed to C.T.

Peer review information: *Nature Climate Change* thanks Shu Kee Lam, Bassil El Masri and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Methods

Overview. The goal of this paper is to scale the effects of eCO_2 on biomass globally. This scaling requires a quantification of 'current' plant biomass and its distribution worldwide together with a model based on the environmental drivers (predictors) that statistically best explain the observations derived from eCO_2 studies. We collected data on above-ground biomass (Supplementary Figs. 1 and 2 and Supplementary Table 1) because (1) above-ground biomass is the metric most commonly reported in eCO_2 studies and (2) satellites can only detect above-ground biomass; thus, upscaling the effects of eCO_2 on above-ground biomass avoids some of the uncertainties related to modelled products of plant productivity or total (above-ground and below-ground) biomass.

From an initial pool of 56 potential predictors, we selected the most important predictors based on variable importance metrics from random-forest meta-analysis. We built a mixed-effects meta-regression model with the most important predictors of the effect, and applied this model with global maps to scale the effects of eCO_2 on above-ground biomass.

Finally, our results were evaluated in terms of distribution and magnitude. For the distribution of the effect, we compared the latitudinal distribution of our estimates with the latitudinal effects of CO_2 on changes in greenness (LAI) in the past three decades[9]. For the magnitude of the effect, we compared our sensitivity of biomass changes to eCO_2 with the sensitivity of biomass changes to the historical increase in atmospheric CO_2 (β) derived from the TRENDY ensemble of global vegetation models[10].

Data collection. We collected 1,432 above-ground biomass observations from 205 studies that met our criteria (below), of which 138 had data for all predictors considered and were therefore included in our analysis. Repeated measurements over time within the same plots (that is, annual or seasonal measurements) were considered non-independent, and were thus aggregated so that only one synthetic measurement per study was included in the meta-analysis. Different species or treatments within the same site were considered independent, but we included 'site' as a random effect in the mixed-effects meta-analysis to account for this potential source of non-independency (see Meta-analysis). 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 databases described by Dieleman et al.[31], Baig et al.[32] and Terrer et al.[4,26,33]. We used Google Scholar to locate the most recent publications for each of the previously listed databases.

We included as many observations as possible for our analysis. Criteria for exclusion from the main analysis were: (1) soil C:N and N content data for the specific soils in which the plants were grown were not reported—for example, studies that included a N fertilization treatment were only included when C:N was measured in situ, and not in unfertilized plots; (2) species did not form associations with either AM or ECM—only species in two studies were non-mycorrhizal, insufficient to identify the drivers of the eCO_2 response in this group; and (3) the duration of the experiment was less than 2 months.

We considered the inclusion of factorial $CO_2 \times$ warming or $CO_2 \times$ irrigation studies when specific soil data for those additional treatments were measured and reported. These treatments were treated as independent and were included in the dataset using the specific MAT and MAP for the warming and irrigation treatments, respectively. Approximately a quarter of the studies were irrigated, with irrigation more common in cropland studies. In those cases, and if the total amount of water used in irrigation was not indicated, we assigned the historical maximum value of MAP extracted from the coordinates of the site in the period 1900–2017 from ref. [34]. Although in some studies we found soil data for several soil depth profiles, soil data were most commonly reported for a depth of 0–10 cm. We thus collected soil data at 0–10 cm, and scaled CO_2 effects using global gridded datasets for this depth increment.

Data for MAP, MAT, soil C:N, soil N content, pH, available P and vegetative and experimental predictors were reported in the literature. Data for the rest of the predictors were not commonly reported, so we extracted these data from global gridded datasets (Supplementary Table 2).

We used the check-lists in refs. [35,36], with additional classifications derived from the literature, to classify plant species as ECM, AM or non-mycorrhizal. Species that form associations with both ECM and AM fungi (for example, *Populus spp.*) were classified as ECM because these species can potentially benefit from increased N availability due to the presence of ECM fungi[4,26], as hypothesized. Overall, CO_2 responses from species associated with AM and ECM were similar to strictly ECM species, and their exclusion did not alter the results of the meta-analysis, as found previously[4].

Where possible, data were collected at the species level, and different species from the same site were considered independent when grown in monoculture with sufficient replication (that is, multiple plots of the same species and multiple individuals of the same species in the same plot).

Using these criteria, we found a total of 205 studies with data on above-ground biomass, with 138 of them including data for all the predictors considered, and thus included in the main analysis. Additionally, we ran a sensitivity test including data from our full dataset of 205 studies, estimating missing soil N and P data from proxies, in the following order of preference: (1) from studies that, due to proximity, used similar soils; (2) from gridded datasets (Supplementary Table 2) in the case of non-fertilized soils; and (3) using the mean values in the dataset for fertilized and non-fertilized studies within ecosystem types. For example, if a study comprised of temperate trees in a fertilized soil did not report soil data, and the characteristics of these soils could not be estimated from similar known soils, we assigned missing data as the average values in the dataset for temperate trees in fertilized soils.

An overview of the experiments included in the main analysis is in Supplementary Table 1, data included in the meta-analysis in Supplementary Fig. 2 and location of the studies in Supplementary Fig. 1. An overview of the studies excluded from the main analysis is given in Supplementary Table 3, and included in a sensitivity test.

Model selection and relative importance. We used random-forest model selection in the context of meta-analysis to identify the most important predictors of the CO_2 effect in the dataset. This method has the advantage over maximum likelihood model-selection approaches that can handle many potential predictors and their interactions, and considers nonlinear relationships.

Some of the 56 potential predictors included in the analysis were extracted from global datasets using the coordinates of the experiments (Supplementary Table 2), and thus included missing values. Because random-forest and meta-analysis require complete data, and no methods for multiple imputation are currently available, we applied single imputation using the missForest[37] algorithm. Like any random forests-based technique, the main advantage of this method is that it does not make any distributional assumptions, which means it easily handles (multivariate) non-normal data and complex interactions and nonlinear relations amongst the data.

Some of the potential predictors provided redundant and potentially correlated information (that is, multiple methods to measure soil P and multiple climate predictors) (see Supplementary Table 2). We used principal component analysis (PCA) for dimensional reduction, extracting components from map-based, potentially redundant predictors.

We included all field-based predictors, together with PCA map-based predictors, in a bootstrapped random-forest meta-analysis recursive preselection with the metaforest[38] R package. We trained a random-forest meta-analysis with preselected predictors and calculated variable importance with metaforest[38] (Supplementary Fig. 3). Based on partial dependence plots (Supplementary Fig. 5), we used reciprocal transformations for nonlinear predictors showing ceiling/floor effects. We included the ten most important predictors in a mixed-effects meta-regression model with the metafor[39] R package, including reciprocal transformations for nonlinear predictors and potential interactions. Finally, we pruned the model once, keeping only significant predictors.

As a sensitivity test, we ran an alternative model-selection procedure using maximum likelihood estimation. For this purpose, we used the rma.mv() function from the metafor R package[39] and the glmulti() function from the glmulti R package[40] to automate fitting of all possible models containing the seven most important predictors and their interactions. Model selection was based on Akaike Information Criterion corrected for small samples as criterion, using a genetic algorithm for faster fitting of all potential models. The relative importance value for a particular predictor was equal to the sum of the Akaike weights (probability that a model is the most plausible model) for the models in which the predictor appears. A cut-off of 0.8 was set to differentiate between important and redundant predictors, so that predictors with relative importance near or less than 0.8 are considered unimportant.

Meta-analysis. We used the response ratio (mean response in elevated-to-ambient CO_2 plots) to measure effect sizes[41]. We calculated the natural logarithm of the response ratio (logR) and its variance for each experimental unit to obtain a single response metric in a weighted, mixed-effects model using the rma.mv function in the R package 'metafor'[39]. We included 'site' as a random effect (because several sites contributed more than one effect size and assuming different species or treatments within one site are not fully independent), and weighting effect size measurements from individual studies by the inverse of the variance[42]. Some 5% of studies did not report standard deviations, which were thus imputed using Rubin and Schenker's[43] resampling approach from studies with similar means and performed using the R package metagear[44].

Measurements across different time-points (that is, over several years or harvests) were considered non-independent, and we computed a combined effect across multiple outcomes (for example, time-points) so that only one effect size was analysed per study. The combined variance that accounts for the correlation among the different time-point measurements was calculated following the method described in Borenstein et al.[45], using a conservative approach by assuming non-independency of multiple outcomes ($r = 1$) and performed using the MAd package in R[46].

We considered nonlinear mixed-effects meta-regression models, which were fitted using reciprocal transformations (1/variable).

Quantification of uncertainties. Extrapolating the empirical relationships that drive biomass responses to eCO_2 (for example, $y \approx$ C:N) in the dataset to the globe has an error associated with the mixed-effects meta-regressions. For the case of soil C:N, for example, this error is large for high C:N values, as the representativeness of soils with high C:N values in the dataset is lower, increasing uncertainty in the regression. For the projections of the eCO_2 effect, we limited the maps of C:N and P to be constrained by the minimum and maximum values in the dataset of eCO_2 studies, thus assuming saturating responses to avoid extremely high or low (negative) effects that are not representative of the observed effects. For the projection of uncertainties in Fig. 2, however, we aimed at representing not only the uncertainty associated with the representativeness of the most important predictors (C:N and P), but also the uncertainty associated with the lower sampling effort in areas with extreme climate (for example, very dry and warm—deserts—or cold and dry—boreal—or wet and warm—tropical). We therefore ran an alternative model that included temperature and precipitation in addition to C:N and P. We extrapolated the standard error of this alternative model using unconstrained maps of temperature, precipitation, C:N and P to account for the higher level of uncertainty in areas with climate and soil values that are not well represented by eCO_2 experiments. Thus, uncertainties in our projections represent the unconstrained standard error of the mixed-effects meta-regression, with larger values under soil and climate conditions that are not adequately studied due to low sample size.

Global estimates of N and P availability. N can be limiting for plants (1) if there is little total N content or (2) because N is bound in organic matter with a high C:N ratio. In the latter case, soil microbes that degrade the organic matter become N-limited, resulting in low amounts of free N available for plant uptake. Therefore, soil N content and C:N ratio were included as potential predictors of the CO_2 effect. Other potential predictors, such as nitrate and ammonium contents and N mineralization, were not generally available and were therefore not included in the analysis.

Because soil C:N ratio was an important predictor of the CO_2-driven increase in biomass in our dataset (Fig. 1a and Supplementary Fig. 3), we used a global dataset of soil C:N ratio from ISRIC-WISE on a 30×30 arcsec grid[47] to upscale this effect. The range of C:N values covered by eCO_2 experiments is representative of the range of C:N values represented in the C:N map[47].

Arid regions typically have very low soil C:N ratios as a result of a small organic C pool and also low N content[48,49]. Therefore, soil C:N is not a good indicator of N availability in arid soils, and the model would overestimate the CO_2 effect in these areas because it would assume relatively high N availability. To avoid the overestimation of the CO_2 effect in arid areas with low C:N, yet low N availability, we followed the approach of Wang et al.[50], who found a threshold of 0.32 in aridity index (ratio of precipitation to mean temperature) below which plant N uptake is limited by water availability, and characterized by low soil C:N despite extremely low soil N availability. We converted areas with aridity index <0.32 to null values in the map of soil C:N, thereby treating these areas as missing data for analyses including soil C:N. We used the aridity data from the CGIAR-CSI Global-Aridity Database[51]. In our dataset of CO_2 experiments, the Nevada Desert FACE fell within this category, with low soil C:N, but low total N[52], and no CO_2 effect on biomass[53], supporting this assumption. Running the model strictly in areas with aridity index >0.32 resulted in 0.4 PgC less than by running the model globally. This small difference was the result of the extremely low above-ground biomass in arid regions (Supplementary Fig. 7), rendering small absolute increases in biomass when incorporated in the analysis. Nevertheless, these areas were not included in the final analysis because it is not likely they could increase their biomass under elevated CO_2 due to extremely low N availability. In areas outside this maximum aridity threshold limiting nitrogen uptake, we studied the impact of climatic and water availability predictors in explaining the magnitude of the CO_2 effect.

The amount of P in the soil estimated by the Bray method was one of the important predictors of the biomass responses to eCO_2. We constrained the map of P amount by the minimum and maximum values of P in the dataset of eCO_2 studies, 2–64 ppm, assuming these values are representative of the conditions at <2 and >64 ppm.

Climate data. For the model selection analysis (Fig. 2) we used MAT and MAP data for the individual studies reported in the papers. As an alternative climatic predictors to MAT and MAP to account for the effect of temperature and water availability, we tested additional predictors not commonly reported in the papers, calculated using temperatures and precipitation values from CRU or extracted from other gridded datasets (Supplementary Table 2).

Current above-ground biomass. As global estimates of current above-ground biomass carbon we used passive microwave-based global above-ground biomass carbon from Liu et al.[18] (v.1.0) at 0.25° resolution and available online for the period 1993–2012 (http://www.wenfo.org/wald/global-biomass/).

Land cover types. Calculations of changes in biomass in response to CO_2 across biomes were performed through zonal statistics with the land cover maps from ESA (http://maps.elie.ucl.ac.be/CCI/viewer/download.php) at 300 m resolution (Table 1) and MODIS IGBP (http://glcf.umd.edu/data/lc/) at 5′ resolution (Supplementary Table 4). Both maps were aggregated by dominant classes. The indication of climatic region (that is, temperate, boreal, tropical) within forest land cover types was based on the classification by Pan et al.[54].

Changes in LAI. In order to evaluate the geographical patterns of our predictions, we compared the latitudinal distribution of the effects of elevated CO_2 on above-ground biomass with changes in LAI attributed to CO_2 in the period 1982–2009 (ref. [9]). We used LAI data from three different satellite records and averaged them, as described in ref. [9]. The attribution of the relative and absolute effects of CO_2 on LAI was estimated through vegetation models, as described in Zhu et al.[9].

For the calculation of the effects of elevated CO_2 on biomass, regions where water availability limits N uptake (aridity index < 0.32) were excluded from the analysis (see Global estimates of N and P availability). Thereby, for the comparison of biomass and LAI changes, these arid regions were excluded from both maps.

Global vegetation models. In order to evaluate the magnitude of the sensitivity of plant biomass to eCO_2 derived from our analysis, we analysed biomass β for the historical increase in atmospheric CO_2 derived from the DGVMs considered in the TRENDY intercomparison project (http://dgvm.ceh.ac.uk/node/9). We used TRENDY-v1, which includes nine DGVMs with common input forcing data, varying CO_2 only from 1980 to 2010 (S1) and calculated biomass β as the change in biomass relative to the change in atmospheric CO_2. For more details on the TRENDY model simulations see Sitch et al.[10].

Calculation of total biomass carbon. The TRENDY models considered here output total biomass (above ground + below ground), whereas our results refer to above-ground biomass only. In order to compare the magnitude of the eCO_2 effect derived from models and our approach, we have estimated the potential effect of eCO_2 on total biomass using region-specific ratios of total biomass and above-ground biomass reported in the literature (Supplementary Table 4).

Reporting Summary. Further information on the research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The biomass data from CO_2 experiments summarized in Supplementary Fig. 2 supporting the findings of this study are available in published papers, and soil and climate data required to upscale CO_2 effects are available in published datasets (Supplementary Table 2). Raw data can be obtained from the corresponding author on reasonable request.

Code availability

The R code used in the analysis presented in this paper is available online and can be accessed at https://github.com/cesarterrer/CO2_Upscaling.

References

31. Dieleman, W. I. J. et al. Simple additive effects are rare: a quantitative review of plant biomass and soil process responses to combined manipulations of CO_2 and temperature. *Glob. Change Biol.* **18**, 2681–2693 (2012).
32. Baig, S., Medlyn, B. E., Mercado, L. M. & Zaehle, S. Does the growth response of woody plants to elevated CO_2 increase with temperature? A model-oriented meta-analysis. *Glob. Change Biol.* **21**, 4303–4319 (2015).
33. Terrer, C. et al. Response to comment on 'Mycorrhizal association as a primary control of the CO_2 fertilization effect'. *Science* **355**, 358–358 (2017).
34. Harris, I., Jones, P. D., Osborn, T. J. & Lister, D. H. Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
35. Maherali, H., Oberle, B., Stevens, P. F., Cornwell, W. K. & McGlinn, D. J. Mutualism persistence and abandonment during the evolution of the mycorrhizal symbiosis. *Am. Nat.* **188**, E113–E125 (2016).
36. Wang, B. & Qiu, Y. L. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**, 299–363 (2006).
37. Stekhoven, D. J. & Buhlmann, P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2011).
38. Van Lissa, C. J. MetaForest: exploring heterogeneity in meta-analysis using random forests. Preprint at https://psyarxiv.com/myg6s/ (2017).
39. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 3 (2010).
40. Calcagno, V. & de Mazancourt, C. glmulti: an R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software* **34**, 12 (2010).
41. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).
42. Osenberg, C. W., Sarnelle, O., Cooper, S. D. & Holt, R. D. Resolving ecological questions through meta-analysis: goals, metrics, and models. *Ecology* **80**, 1105–1117 (1999).
43. Rubin, D. B. & Schenker, N. Multiple imputation in health-care databases: an overview and some applications. *Stat. Med.* **10**, 585–598 (1991).
44. Lajeunesse, M. J. Facilitating systematic reviews, data extraction and meta-analysis with the metagear package for R. *Methods Ecol. Evol.* **7**, 323–330 (2016).

45. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. in *Introduction to Meta-Analysis* (eds Borenstein, M. et al.) 225–238 (John Wiley & Sons, Ltd, 2009).
46. Del Re, A. C. & Hoyt, W. T. MAd: meta-analysis with mean differences. R version 0.8-2 (2014).
47. Batjes, N. H. Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. *Geoderma* **269**, 61–68 (2016).
48. Post, W. M., Pastor, J., Zinke, P. J. & Stangenberger, A. G. Global patterns of soil nitrogen storage. *Nature* **317**, 613–616 (1985).
49. Jiao, F., Shi, X.-R., Han, F.-P. & Yuan, Z.-Y. Increasing aridity, temperature and soil pH induce soil C-N-P imbalance in grasslands. *Sci. Rep.* **6**, 19601 (2016).
50. Wang, C. et al. Aridity threshold in controlling ecosystem nitrogen cycling in arid and semi-arid grasslands. *Nat. Commun.* **5**, 4799 (2013).
51. Zomer, R. J., Trabucco, A., Bossio, D. A. & Verchot, L. V. Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agric. Ecosyst. Environ.* **126**, 67–80 (2008).
52. Billings, S. A., Schaeffer, S. M. & Evans, R. D. Trace N gas losses and N mineralization in Mojave Desert soils exposed to elevated CO_2. *Soil Biol. Biochem.* **34**, 1777–1784 (2002).
53. Evans, R. D. et al. Greater ecosystem carbon in the Mojave Desert after ten years exposure to elevated CO_2. *Nat. Clim. Change* **4**, 394–397 (2014).
54. Pan, Y. et al. A large and persistent carbon sink in the world's forests. *Science* **333**, 988–993 (2011).

nature research

Corresponding author(s): César Terrer

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	We used R (version 3.5.0). The R code used in the analysis presented in this paper is available online and can be accessed at https://github.com/cesarterrer/CO2_Upscaling.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

A data availability statement was presented in the manuscript. The data of all the analyses are publicly accessible.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Synthesis study based on a meta-analysis of data from CO2 experiments. We identify the best descriptors ("moderators") of the CO2 effect on biomass in the dataset based on random-forest meta-analysis. We design a nonlinear mixed-effects meta-regression with selected moderators, including "site" as a random factor, and use this statistical model to project the CO2 fertilization effect on biomass globally.
Research sample	Published data from CO2 experiments, excluding pot studies and plants grown in soils of unknown properties.
Sampling strategy	We compiled data from as many studies as possible, as long as they met the criteria specified in methods. We prioritized studies using natural soils.
Data collection	We used previously compiled datasets, such as those by 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 databases described by Dieleman et al. (2012), Baig et al. (2015), and Terrer et al. (2016, 2017, 2018). We used Google Scholar to locate the most recent publications for each of the previously listed experiments, as well as soil data for as many experiments as possible. If no soil data were found, the experiment was excluded from the main analysis, but included in a sensitivity analysis with inferred data from proxies.
Timing and spatial scale	The most recent data were sampled in the field in 2015 and published in 2018; we thus collected data until the date of submission. Experiments were not excluded based on date of publication, as long as they met the criteria specified in methods and data about soil properties were reported.
Data exclusions	We included as many observations as possible for our analysis. Criteria for exclusion from the main analysis were: i) Soil C:N and N content data for the specific soils in which the plants were grown were not reported. For example, studies that included a N fertilization treatment were only included when C:N was measured in situ, and not in unfertilized plots. ii) species did not form associations with either AM or ECM. Only species in two studies were non-mycorrhizal, insufficient to identify the drivers of the eCO2 response in this group; and iii) the duration of the experiment was less than 2 months.
Reproducibility	Our analyses were based on published data (included in Supplementary Fig. 2) and public satellite products and global datasets (Supplementary Table 2), and the results can be reproduced using the R code in https://github.com/cesarterrer/CO2_Upscaling
Randomization	Grouping was based on a machine-learning model selection approach
Blinding	Our study is based on existing data, therefore blinding is not relevant.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging