nature ecology & evolution

Article

https://doi.org/10.1038/s41559-023-02093-x

Biome-scale temperature sensitivity of ecosystem respiration revealed by atmospheric CO_2 observations

Received: 5 March 2022

Accepted: 10 May 2023

Published online: 15 June 2023

Check for updates

Wu Sun[1] ✉, Xiangzhong Luo[2,3,7], Yuanyuan Fang[4], Yoichi P. Shiga[5,8], Yao Zhang[2,3,9], Joshua B. Fisher[6], Trevor F. Keenan[2,3] & Anna M. Michalak[1] ✉

The temperature sensitivity of ecosystem respiration regulates how the terrestrial carbon sink responds to a warming climate but has been difficult to constrain observationally beyond the plot scale. Here we use observations of atmospheric CO_2 concentrations from a network of towers together with carbon flux estimates from state-of-the-art terrestrial biosphere models to characterize the temperature sensitivity of ecosystem respiration, as represented by the Arrhenius activation energy, over various North American biomes. We infer activation energies of 0.43 eV for North America and 0.38 eV to 0.53 eV for major biomes therein, which are substantially below those reported for plot-scale studies (approximately 0.65 eV). This discrepancy suggests that sparse plot-scale observations do not capture the spatial-scale dependence and biome specificity of the temperature sensitivity. We further show that adjusting the apparent temperature sensitivity in model estimates markedly improves their ability to represent observed atmospheric CO_2 variability. This study provides observationally constrained estimates of the temperature sensitivity of ecosystem respiration directly at the biome scale and reveals that temperature sensitivities at this scale are lower than those based on earlier plot-scale studies. These findings call for additional work to assess the resilience of large-scale carbon sinks to warming.

The terrestrial carbon–climate feedback, resulting from the sensitivity of the terrestrial carbon sink to the physical climate[1], dominates the uncertainty in climate projections[2–4]. This feedback depends on the difference between the responses of photosynthesis and respiration to a changing climate[5]. Estimates of photosynthesis[6,7] and respiration[8–10] vary substantially across terrestrial biosphere models (TBMs)[11], however. Furthermore, whereas there has recently been a proliferation of novel measurement techniques to better constrain photosynthesis from regional to global scales, such as solar-induced chlorophyll fluorescence[12], near-infrared reflectance of vegetation[13] and

[1]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA. [2]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA. [3]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [4]Bay Area Air Quality Management District, San Francisco, CA, USA. [5]Universities Space Research Association, Mountain View, CA, USA. [6]Schmid College of Science and Technology, Chapman University, Orange, CA, USA. [7]Present address: Department of Geography, National University of Singapore, Singapore, Singapore. [8]Present address: San Francisco, CA, USA. [9]Present address: Sino–French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing, China. ✉e-mail: wsun@carnegiescience.edu; michalak@carnegiescience.edu

https://doi.org/10.1038/s41559-023-02093-x

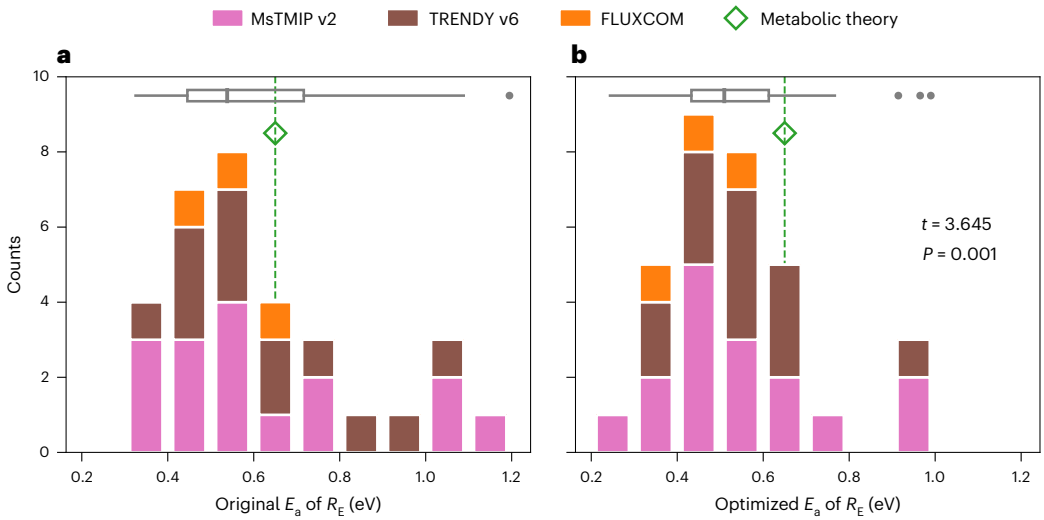


Fig. 1 | Using atmospheric observations to constrain the temperature sensitivity of respiration for North America reduces the spread in estimates across models and suggests that the large-scale sensitivity is lower than that implied by the metabolic theory of ecology and by plot-scale studies. **a**,**b**, Histograms of the original (**a**) and optimized (**b**) aggregate temperature sensitivity of ecosystem respiration for North America, as represented by $N = 32$ independent estimates of E_a, for TBMs in the MsTMIP v2 (pink) and TRENDY v6 (brown) ensembles and data-driven models in the FLUXCOM ensemble (orange). Grey boxplots summarize the estimates across models, with the centre line, bounds of box, whiskers and dots representing the median, first and third quartiles, smallest and largest estimates falling within 1.5× of the interquartile range from the nearest quartiles and outliers beyond that range, respectively. Green diamonds and vertical dashed lines indicate the reference value of 0.65 eV based on the metabolic theory of ecology[29,59,114] and plot-scale estimates[52]. A two-tailed, paired two-sample *t* test confirms a statistically significant difference between model-represented temperature sensitivities before (**a**) and after (**b**) optimization against atmospheric CO_2 observations, as indicated by the *t* statistic and *p* value shown in **b**.

carbonyl sulfide[14,15], respiration remains difficult to constrain at large scales due to the absence of a unique spectral signature or atmospheric tracer.

Accurate climate projections thus hinge on improving our understanding of the magnitude, space–time distribution and climatic sensitivity of respiration. More specifically, constraining the temperature sensitivity of respiration is key to estimating total ecosystem respiration, to assessing the extent to which ecosystem respiration will be amplified by a warming climate and to assessing climate-related risks for regions that can trigger a substantial positive carbon–climate feedback[16–19].

At regional to global scales, ecosystem respiration (R_E)—the sum of autotrophic and heterotrophic respiratory fluxes—can be estimated from prognostic TBMs that parameterize respiration based on responses to environmental and biotic drivers[20,21], from data-driven models that rely on site-level measurements[22], from remotely sensed covariates of respiration[23] or from a combination thereof[24–26]. TBMs represent respiratory processes in diverse ways[27] but ultimately rely on generalizing locally derived relationships based on sparsely distributed observations to continental and global scales[28]. This has led to a large spread in estimates of respiration. In addition, estimates of global ecosystem respiration based on TBMs (76–180 Pg C yr^{-1}) vary more widely than those based on data-driven models (94–109 Pg C yr^{-1}; Extended Data Fig. 1).

Within naturally occurring temperature ranges, ecosystem respiration typically shows an exponential response to temperature[29,30] as described by the Arrhenius equation[31] (Methods). While components of ecosystem respiration are also influenced by moisture[32], phenology[33], photosynthate input[34,35], biomass[35], nutrients[36], litter quality[37] and soil microbial responses[38,39], and especially so at the plot scale and for diurnal to monthly timescales, temperature remains the first-order control of ecosystem respiration (and its primary components) at aggregated spatial (biome to continental) and temporal (monthly to decadal) scales[17,40–43]. This aggregate temperature sensitivity of ecosystem respiration, which incorporates both the direct response of ecosystem respiration to temperature and indirect influences from other climatic and physiological variables, thereby represents the overall response of biome-scale ecosystem respiration to temperature.

Currently, the only available estimates of the biome-scale temperature sensitivity of respiration are from upscaling of plot-scale estimates, which are largely based on observations from eddy covariance towers (for example, Mahecha et al.[44]). Such observations are unevenly distributed across ecosystems[45] and regions[46,47] and represent fluxes for areas only up to several km^2 (refs. 48–50). Therefore, even a network of hundreds of sites may not adequately represent fluxes at biome or continental scales[51]. This sampling limitation has led to a large discrepancy between carbon fluxes upscaled from plot-scale observations and those derived from regional-scale observational constraints[26], which casts doubt on the robustness of using plot-scale estimates to inform certain biome-scale responses. Moreover, cross-site analyses show that temperature sensitivity is similar across ecosystem types[30,44,52], which contradicts anticipated responses from thermal acclimation of autotrophic and heterotrophic respiration; that is, the warmer the climate, the lower the temperature sensitivity[53–55]. Independent empirical estimates of biome-scale temperature sensitivity of ecosystem respiration, such as those based on regional-scale observational constraints, are thus critically needed.

Here we leverage 39,000 atmospheric CO_2 concentration measurements from a network of monitoring stations across North America[56] during the period 2007–2010 to infer temperature sensitivities of ecosystem respiration for North America and major biomes therein (Extended Data Fig. 2a). To do so, we first use model estimates of ecosystem respiration at monthly temporal and 1° × 1° spatial resolution, obtained from the sum of gross primary productivity (GPP) and net ecosystem exchange (NEE, negative for uptake), to evaluate the biome-scale temperature sensitivity of ecosystem respiration as represented by TBM simulations from the Multi-scale Synthesis and Terrestrial Model Intercomparison Project version 2 (MsTMIP v2) (ref. 57) and Trends in Net Land–Atmosphere Exchange version 6 (TRENDY v6) (ref. 58) ensembles and FLUXCOM machine-learning models[26]. We then assess the degree to which space–time variability in observed

atmospheric CO_2 concentrations is captured by carbon flux estimates from this same ensemble of models. Finally, we use the constraint provided by the atmospheric observations to optimize continental- and biome-scale temperature sensitivities of respiration both for individual models and across the model ensemble.

Results

Uncertainty in the temperature sensitivity of respiration

We find that the temperature sensitivity of ecosystem respiration, as represented by the Arrhenius activation energy (E_a, in eV; Methods), ranges widely for the 29 model simulations examined here both for North America (Fig. 1a) and for individual biomes (Extended Data Fig. 3). For North America, for example, E_a ranges from 0.33 to 1.19 eV (Fig. 1a, Extended Data Fig. 4 and Supplementary Table 1), equivalent to Q_{10} values (that is, the factor by which respiration increases for a 10 °C increase in temperature) of 1.5 to 4.9 at 10 °C (Methods and Extended Data Fig. 5). Inferred temperature sensitivities are highly correlated across biomes within the model ensemble; models with a high temperature sensitivity in one biome also tend to have high sensitivities in other biomes (Supplementary Table 2).

Approximately two-thirds of the examined TBM simulations (19 of 29) show an E_a estimate for North America that is lower than site-level E_a values derived from flux tower observations[30,52], which converge around a universal empirical value of E_a = 0.65 eV for organism- and community-level respiration, according to the metabolic theory of ecology[29,59] (Fig. 1a). Note that the model-represented, regional-scale estimates derived here are methodologically comparable with previously reported plot-scale temperature sensitivity (0.65 eV) because both are bottom-up estimates and are derived using temperature as the driver, implicitly accounting for the influence of other covariates[30,52]. This discrepancy in temperature sensitivity between regional and plot scales is also present for individual biomes (Extended Data Fig. 3). Interestingly, all three FLUXCOM models, which are trained on site-level flux observations, also exhibit E_a estimates for North America (0.48 to 0.61 eV; Fig. 1a and Supplementary Table 1) that are lower than those derived directly from half hourly eddy covariance flux tower observations (approximately 0.65 eV (ref. 52)), indicating that the upscaling of site-level flux observations in space and time impacts the observed aggregate temperature sensitivity of ecosystem respiration.

These biome-scale aggregate temperature sensitivity estimates may differ from those prescribed in models for individual respiratory components because the former characterize large-scale phenomenological properties resulting from a wide array of underlying microscopic respiratory reactions and encompass mediating effects from GPP and other covariates. For example, the Community Land Model version 4.5 (CLM4.5) prescribes a Q_{10} value of 1.5 for autotrophic and heterotrophic respiration[60], but the activation energy estimated from its model output is 0.62 eV for North America (Supplementary Table 1), equivalent to an apparent Q_{10} value of 2.4 at 10 °C. On the other hand, it is precisely the fact that these temperature sensitivity estimates reported here characterize the aggregate properties of biomes, which cannot be readily deduced from model parameterizations, that makes them useful for informing the climatic responses of ecosystem respiration and diagnosing implications of model parameterizations.

Empirical constraint on temperature sensitivities

Next we adjust each model's activation energy and baseline respiration to maximize consistency with observations of atmospheric CO_2 concentrations (Methods) and find that the temperature sensitivity of ecosystem respiration is reduced for most (19 of 29) TBMs and all three FLUXCOM models (Fig. 1a,b and Supplementary Table 1). This indicates that the temperature sensitivity across existing prognostic TBMs and data-driven models is higher than what atmospheric observations suggest. The decrease in the optimized temperature sensitivity of ecosystem respiration is especially large for TBMs with high original sensitivity estimates. As a result of the optimization, model spread of the temperature sensitivity of ecosystem respiration also decreases (Fig. 1b), leading to an ensemble mean (± 1 standard error) of 0.54 (± 0.03) eV (and an ensemble median of 0.51 eV). This ensemble mean temperature sensitivity is statistically significantly lower than the previously reported plot-scale temperature sensitivity of 0.65 eV (p = 0.00019, one-tailed Z test).

Furthermore, we find that the adjustments to E_a needed to maximize the consistency with observed atmospheric CO_2 variability, ΔE_a, are linearly and negatively correlated with the original E_a estimates of the models (Fig. 2; dashed lines), both for the North American domain (Fig. 2a) and for individual biomes (Fig. 2b–d). This means that atmospheric observations can not only be used to adjust model-specific temperature sensitivities but also to infer an optimal sensitivity ($\hat{E}_{a,opt}$) across models. This ensemble optimal sensitivity is represented by the activation energy corresponding to zero adjustment (that is, $\Delta E_a = 0$) on the linear relationship between E_a and ΔE_a.

We find that the best estimate of the temperature sensitivity of ecosystem respiration ($\hat{E}_{a,opt}$) across models is 0.43 ± 0.06 eV (1σ uncertainty; equivalent to Q_{10} = 1.9 ± 0.2 at 10 °C) for North America. This sensitivity is substantially and significantly lower than the previous global E_a estimate of 0.65 eV (ref. 52) (p = 0.00022, one-tailed Z test) derived from half hourly, site-level flux observations but is remarkably similar to the E_a estimate that represents the temperature sensitivity on an annual timescale (0.42 eV (ref. 52)). Moreover, the optimal sensitivity (0.43 ± 0.06 eV; Fig. 2a) is also lower than the corrected ensemble mean (0.54 ± 0.03 eV, mean ± 1 s.e.) (p = 0.060, one-tailed Z test) or median sensitivity (0.51 eV; Fig. 1b) across the models examined here. This difference also shows that absent an observational constraint, the mean or median responses across a model ensemble may not be a good estimate of actual sensitivity.

We also find substantial variability in temperature sensitivity across biomes, namely, 0.38 ± 0.03 eV (Q_{10} = 1.7 ± 0.1 at 10 °C) for croplands, 0.50 ± 0.03 eV (Q_{10} = 2.1 ± 0.1 at 10 °C) for evergreen needleleaf forests and 0.53 ± 0.05 eV (Q_{10} = 2.2 ± 0.1 at 10 °C) for deciduous broadleaf and mixed forests, which contrasts with earlier site-level studies that had suggested that E_a is uniform across a range of ecosystem types[30,52]. Tropical and Arctic biomes are not considered individually here due to the low sensitivities of available atmospheric CO_2 observations to surface fluxes in these regions (Extended Data Fig. 2b).

Additional analyses confirm that the temperature sensitivities inferred here are not substantially impacted by potential confounding effects of soil moisture or radiation, by co-variability with GPP, by thermal acclimation or by lateral fluxes (Supplementary Notes 1–5 and Supplementary Figs. 1–7). These inferred temperature sensitivities therefore do represent the overall biome-scale response of ecosystem respiration to climatic temperature gradients.

Correcting temperature sensitivity improves model NEE skill

Having obtained observationally constrained estimates of temperature sensitivities of ecosystem respiration (Fig. 2), we then examine how these estimates impact the ability of NEE estimates from TBMs and data-driven models to explain observed atmospheric CO_2 variability.

Surprisingly, although space–time variability in atmospheric CO_2 concentrations results from space–time patterns in NEE, we find that GPP estimates explain the variability in observed atmospheric CO_2 concentrations better than corresponding NEE estimates for 12 of the 29 original TBM simulations in the MsTMIP and TRENDY ensembles (Fig. 3; circles, squares or diamonds; Extended Data Fig. 6). The difference between the explanatory power of monthly averaged NEE estimates and that of GPP estimates ($\Delta R^2 = R^2_{NEE} - R^2_{GPP}$) ranges from ΔR^2 = 0.10 for the simulation for which the improvement from GPP to NEE is greatest to ΔR^2 = −0.26 for the simulation where the deterioration relative to the explanatory power of GPP is greatest, highlighting the misrepresentation of ecosystem respiration in the latter set of models.

https://doi.org/10.1038/s41559-023-02093-x

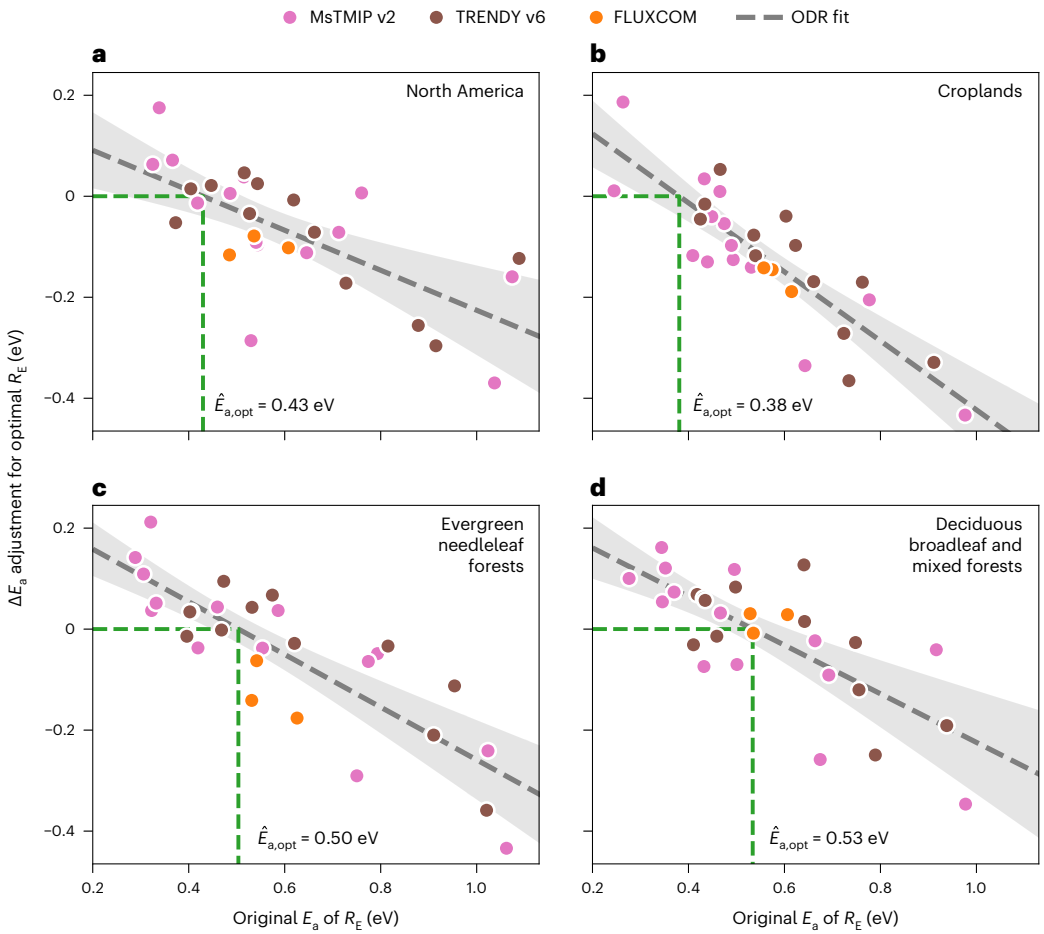


Fig. 2 | Optimal North American and biome-specific estimates of the temperature sensitivity of ecosystem respiration for individual models and the model ensemble. Adjustments to model-specific estimates of the activation energy are determined by maximizing consistency with observed atmospheric CO_2 variability (Methods). **a**–**d**, Relationships between adjustments to model-specific estimates of the activation energy needed to maximize consistency with observed atmospheric CO_2 variability (ΔE_a; vertical axis) and the original estimates of activation energy (E_a; horizontal axis) for models in the MsTMIP v2 (pink), TRENDY v6 (brown) and FLUXCOM (orange) model ensembles in the North American domain (**a**), croplands (**b**), evergreen needleleaf forests (**c**) and deciduous broadleaf and mixed forests (**d**). Note that only models for which the explanatory power of simulated GPP (R^2_{GPP}) exceeds that of shortwave radiation (R^2_{SW}) are included (Extended Data Fig. 6), leaving two models in the MsTMIP ensemble and one model in the TRENDY ensemble excluded (Methods). The grey dashed lines represent the best orthogonal distance regression (ODR) fit between ΔE_a and E_a estimates across the three ensembles of models ($N = 29$), with light grey shading indicating the 95% prediction interval. The optimal E_a ($\hat{E}_{a,opt}$) corresponds to the point where the ODR fit line crosses $\Delta E_a = 0$ eV (that is, no adjustment to E_a is needed) and is indicated using a green dashed line and listed for each biome in the corresponding panel.

For the FLUXCOM models (not considered TBMs because of their data-driven nature; Methods), NEE estimates neither markedly improve nor degrade the degree to which flux estimates reproduce observed atmospheric CO_2 variability relative to GPP estimates ($\Delta R^2 \leq 0.02$; Fig. 3). Note that three of the 29 TBM simulations have GPP estimates that explain an even smaller portion of the observed atmospheric CO_2 variability than does shortwave radiation ($R^2_{SW} = 0.23$; symbols with empty left portion in Fig. 3; Extended Data Fig. 6), a first-order climatic driver of GPP, and these simulations are therefore excluded from further analysis (consistent with Fig. 2; Methods).

Correcting ecosystem respiration by adjusting temperature sensitivity and baseline respiration (Extended Data Fig. 7 and Methods) substantially improves the degree to which NEE estimates reproduce observed atmospheric CO_2 variability for those models for which the performance of NEE trailed that of GPP (Fig. 3a, red shaded area). For four of the ten models in this group (SDGVM, CLASS-CTEM-N, CABLE and TRIPLEX-GHG; Supplementary Table 3), NEE estimates in fact outperform GPP estimates once ecosystem respiration is corrected. Only one model, BIOME-BGC (Supplementary Table 3), which has the lowest temperature sensitivity among all investigated models ($E_a = 0.33$ eV; Supplementary Table 1), shows a minor degradation after respiration is corrected.

By contrast, correcting ecosystem respiration does not consistently improve the performance of NEE estimates for models for which the explanatory power of NEE already exceeded that of GPP (Fig. 3a, blue shaded area). For these models, bias in ecosystem respiration may have been offset by corresponding bias in GPP.

Interestingly, correcting ecosystem respiration using a single value of the ensemble optimal temperature sensitivity (0.43 eV for North America) across all models and biomes yields an improvement in the performance of models' estimates of NEE that is almost as large (mean increase in R^2_{NEE} of 0.10 for models for which $R^2_{NEE} < R^2_{GPP}$; Fig. 3b) as adjusting both the respiration sensitivity and baseline respiration to model-specific optimized values (mean increase in R^2_{NEE} of 0.12 for the same set of models; Fig. 3a), further supporting the robustness of the overall estimate of temperature sensitivity for North America.

The improvement in model explanatory power of NEE after correcting ecosystem respiration indicates a dominant role of the temperature sensitivity bias in causing the underperformance in carbon flux estimates, although other sources of bias also exist. Attribution of

https://doi.org/10.1038/s41559-023-02093-x

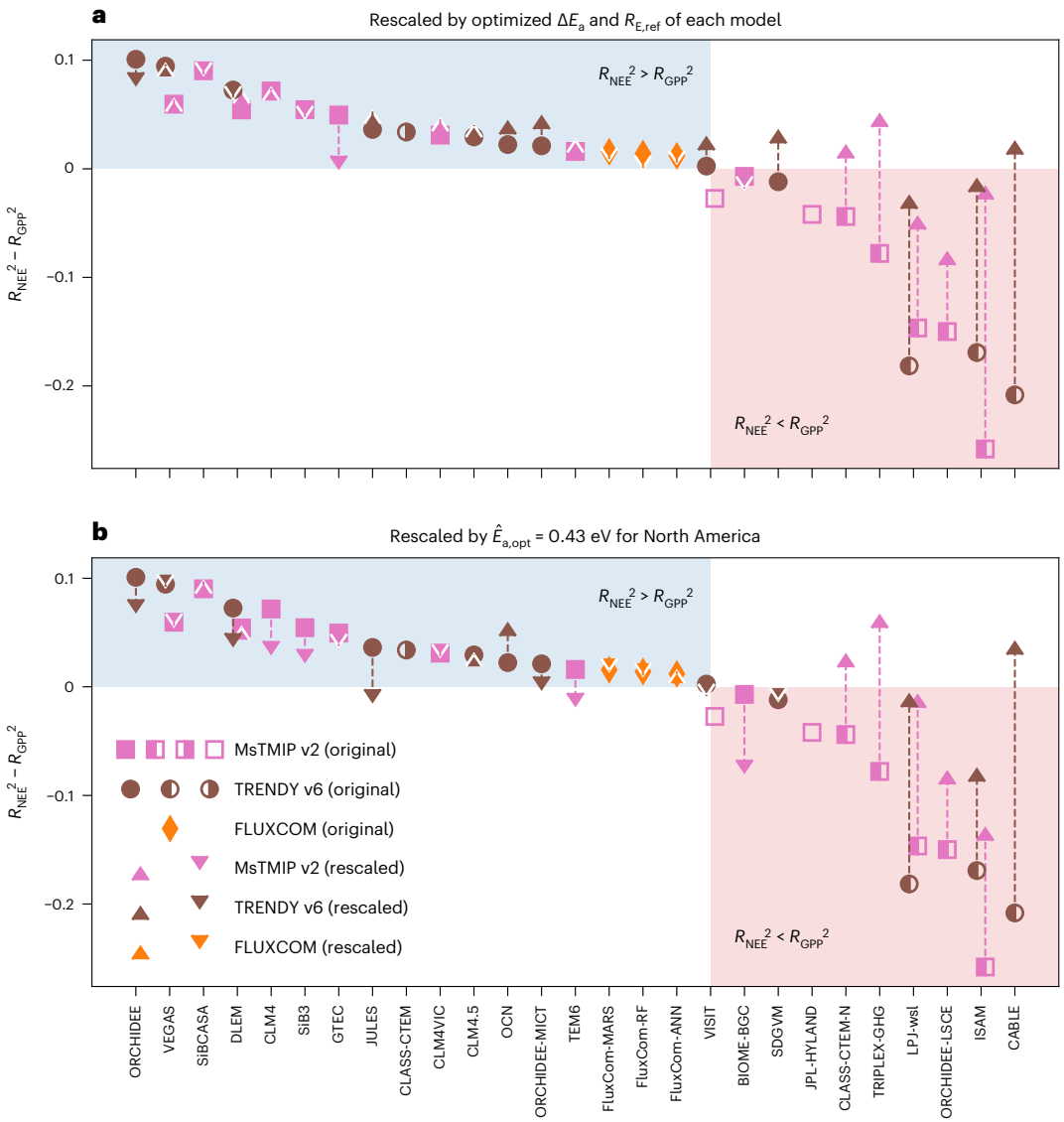


Fig. 3 | Rescaling ecosystem respiration by adjusting E_a to optimized values leads to substantial improvement in NEE explanatory power for models for which the explanatory power of NEE originally trailed that of GPP. a, The difference between the explanatory power of NEE (R^2_{NEE}) and GPP (R^2_{GPP}) before and after rescaling based on model-specific optimal adjustments to the temperature sensitivity (ΔE_a in Fig. 2a) and the baseline respiration rate at 10 °C for North America (Extended Data Fig. 7a). **b**, The same information from **a** after rescaling based on an overall optimal temperature sensitivity derived for North America ($\hat{E}_{a,opt}$ = 0.43 eV; Fig. 2a). The R^2 differences ($R^2_{NEE} - R^2_{GPP}$) of the original MsTMIP, TRENDY and FLUXCOM models are represented by squares, circles and diamonds, respectively. A symbol that is filled in the left (right) half indicates that the corresponding model's GPP (NEE) estimates explain a higher fraction of CO_2 variability than does incoming shortwave radiation (R^2_{SW} = 0.23). Solid symbols indicate models for which both GPP and NEE estimates outperform shortwave radiation in explaining atmospheric CO_2 variability, whereas empty symbols indicate models for which neither GPP nor NEE estimates outperform shortwave radiation. Models for which the explanatory power of GPP lags behind that of shortwave radiation (CLASS-CTEM, VISIT and JPL-HYLAND) are excluded from the rescaling. The blue shaded area highlights models for which the original estimates of NEE outperform GPP in terms of explanatory power, while the red shaded area includes models for which GPP outperforms NEE. The R^2 differences after rescaling respiration are indicated by triangles, with upward orientation indicating an improvement in the explanatory power and downward orientation indicating the opposite. Most models for which the explanatory of NEE originally lagged that of GPP (red shaded regions of both panels) show a marked improvement in explanatory power once the temperature sensitivity of ecosystem respiration is adjusted to optimized values. Conversely, adjusting the temperature sensitivity of respiration does not have a large or consistent impact for those models where the explanatory power of NEE was already superior to that of GPP (blue shaded regions). See Supplementary Table 3 for a list of all model names, abbreviations and references.

model NEE bias to respiration parameters has been difficult up to now because of equifinality—that is, in the presence of respiration bias, the bias in NEE estimates may still be alleviated by compensation from bias in GPP estimates. Here because rescaling ecosystem respiration based on an overall optimal temperature sensitivity for North America alone (Fig. 3b; Methods) leads to substantial improvement in NEE explanatory power for models for which the performance of NEE trails that of GPP ($\Delta R^2 < 0$; Fig. 3b, red shaded area), respiration bias is probably a more important contributor to the NEE bias than is GPP bias for these models. This notion is also corroborated by changes in the seasonal cycles of NEE after imposing the optimal temperature sensitivity on the model ensemble (Supplementary Figs. 8 and 9) and by additional tests on the influences of GPP uncertainty on the inferred temperature sensitivity (Supplementary Figs. 1 and 2).

Discussion

Our findings highlight the spatial-scale dependence of the temperature sensitivity of respiration. Several factors probably contribute to the temperature sensitivity at the biome scale and seasonal to interannual timescales being lower than that reported from plot-scale studies. Temporal aggregation may smooth out short-term (for example, half hourly, typical for plot-scale observations) responses, leading to markedly reduced temperature sensitivity on annual timescales[30,52]. For soil heterotrophic respiration—a major component of ecosystem respiration—climatological temperature sensitivity has also been shown to differ from instantaneous temperature sensitivity[17]. Indeed, we find that the relationship between biome-scale temperature sensitivity estimates and mean air temperatures of the studied biomes (Supplementary Fig. 10) is broadly consistent with the previously observed response of climatological Q_{10} for soil carbon decomposition to the mean air temperature in the temperate domain[17]. Despite the difference in respiratory components, such qualitative consistency lends support to the robustness of the inferred climatologically relevant temperature sensitivity for temperate North American biomes.

Moreover, biome-scale temperature sensitivities may further differ from plot-scale temperature sensitivities because the responses observed here incorporate indirect sensitivities to temperature via drivers such as soil moisture[32], nutrients[36] and phenology[33]. Indeed, some plot-scale studies show a lower temperature sensitivity than the biome-scale temperature sensitivity inferred here after removing influences from low-frequency variability ($Q_{10} = 1.4 \pm 0.1$ (ref. 44)) or hydrometeorological drivers ($Q_{10} = 1.6 \pm 0.1$ (ref. 61)). In addition, as ecosystems in cooler (warmer) climates tend to have higher (lower) baseline respiration rates[29], accounting for this spatial gradient in baseline respiration potentially dampens the inferred biome-scale temperature sensitivity. Ecosystem state variables such as GPP and phenology may also regulate the baseline respiration[62]. The relative contribution of these factors at the biome scale merits further research. Given that the sensitivities inferred here encompass these various additional factors, they are likely to more aptly reflect the bulk response of respiration to future warming on aggregate space–time scales and therefore inform climatic responses.

In addition to differences in sensitivity across scales, several known challenges with flux measurements and partitioning may have also led to an overestimate of the temperature sensitivity of ecosystem respiration in earlier studies even at the plot scale. These known challenges can cause ecosystem respiration to be undercounted when temperatures are low (at night or in the dormant season) and overestimated when temperatures are high (during the day or in the growing season). Nighttime respiration can be underestimated in conditions of weak turbulence and canopy CO_2 storage build-up[63]. Moreover, for sites in uneven terrain, there is no accepted way to reliably account for advective CO_2 fluxes[64], which may lead to a further underestimation of respiratory carbon loss at night. Partitioning daytime respiration from NEE measurements also remains challenging because the relationship between nighttime respiration and temperature often does not hold during the daytime due to light inhibition of leaf respiration[65–67] and non-temperature controls of nighttime autotrophic respiration[68]. As a result, using the relationship between nighttime respiration and temperature to partition daytime fluxes often leads to overestimation of daytime respiration[69]. Furthermore, this overestimation of daytime respiration is particularly strong in the growing season (up to 23%) (ref. 69), further enhancing the bias of inferred temperature sensitivity on an annual timescale. These problems all contribute to a potential positive bias in the plot-scale temperature sensitivity and, in turn, the size of the difference in temperature sensitivity across scales reported here.

The difference in the temperature sensitivity of ecosystem respiration between croplands and forests observed here (Fig. 2b-d) and the absence of such biome specificity in plot-scale studies may arise from several factors. Croplands typically show a higher carbon use efficiency (ratio between net primary productivity and GPP) than unmanaged forests[70], that is, a lower autotrophic fraction in ecosystem respiration. This, in turn, means that the fraction of soil heterotrophic respiration, which is less sensitive to air temperature than is aboveground respiration, is higher in croplands than forests, thereby leading to the overall lower temperature sensitivity of ecosystem respiration in croplands. Management practices such as harvest, irrigation and tillage can cause dramatic changes in ecosystem respiration at weekly to monthly timescales[71], though their impact on the temperature sensitivity at longer timescales remains poorly understood. The dearth of cropland sites in FLUXNET observations[45] may explain the lack of biome specificity in the temperature sensitivity inferred in plot-scale studies and the fact that the FLUXCOM models[26], trained on FLUXNET observations, capture the temperature sensitivity best in deciduous broadleaf and mixed forests (Fig. 2d) but least well in croplands (Fig. 2b). To reach a robust understanding of cropland carbon cycling, future investigations may need to examine the partitioning among different respiratory components, quantify the impact of management practices on the temperature sensitivity of ecosystem respiration and expand the eddy covariance network in cropland areas.

Given that the temperature sensitivity of ecosystem respiration constitutes a leading-order positive climate feedback, the findings presented here also open several avenues for advancing the understanding of terrestrial carbon–climate feedbacks.

First, although observed atmospheric CO_2 variability incorporates influences of both photosynthesis and ecosystem respiration, the inference of an ensemble optimal temperature sensitivity for respiration for North America (Fig. 2a) and major biomes (Fig. 2b–d) indicates that the current level of uncertainty in GPP[7,72,73] does not obscure information about ecosystem respiration (Supplementary Fig. 1). Given this finding, partitioning of net carbon fluxes into photosynthesis and respiration at regional scales is, in principle, achievable in a manner akin to the respiration-based partitioning widely used at eddy covariance sites[63,74]; such partitioning would overcome a key observational challenge in constraining regional-scale responses of carbon fluxes to climate. Pursuing this goal may require expanding and optimizing atmospheric observational networks to better resolve space–time variability in photosynthesis and respiration[75], especially in sparsely sampled tropical and Arctic biomes.

Second, our findings establish a link between the explanatory power of modelled NEE and bias in the temperature sensitivity of ecosystem respiration, which is useful for refining TBMs. Although the temperature sensitivity of either GPP or ecosystem respiration may be adjusted to improve the consistency between NEE and observed atmospheric CO_2 concentrations, in reality, GPP and ecosystem respiration are intricately tied by carbon allocation and decomposition processes. Such coupling does not always allow errors in ecosystem respiration to be cancelled by corresponding errors in GPP. Here the substantial improvements in model explanatory power resulting from correcting the temperature sensitivity of ecosystem respiration (Fig. 3) provide robust evidence that current model representation of respiratory processes contribute substantially to uncertainty in NEE estimates, and in turn, uncertainty in the response of the terrestrial carbon sink to future warming. We suggest that calibration of the biome-scale temperature sensitivity of ecosystem respiration against multi-scale constraints be prioritized in model development to reduce compensating errors in ecosystem respiration and GPP, which could otherwise meet model benchmarks but for incorrect reasons.

Finally, if bias in the temperature sensitivity of ecosystem respiration represented by standalone TBM simulations is indicative of that in coupled Earth system model simulations, we may expect the response of ecosystem respiration to warming to be stronger in those models than what current atmospheric CO_2 observations suggest as well. This difference also highlights a gap in the predictive understanding of how large-scale carbon sinks respond to future warming, because much of

current understanding derives from Earth system model simulations. Meanwhile, in the long run, the response of ecosystem respiration to climatic warming may be additionally influenced by acclimation to a warmer climate[53–55,76,77], changes in soil microbial community composition[78], limitation of labile carbon pools[79] and soil warming[80], among other factors. As early warning signs of potential saturation and destabilization of regional carbon sinks appear[81], there is an urgent need to assess the resilience of large-scale carbon sinks to climatic warming by synthesizing multi-scale observational constraints with Earth system models that embed state-of-the-art mechanistic understanding of respiration. Looking forward, we anticipate that the expanding network of ground-based, airborne and satellite remote-sensing observations of atmospheric CO_2 concentrations will further elucidate the climatic sensitivities of ecosystem respiration across regions and spatiotemporal scales and thereby inform respiration-driven terrestrial carbon–climate feedbacks.

Methods

Model estimates of terrestrial carbon fluxes

Monthly model estimates of GPP and NEE were obtained from the MsTMIP (version 2, spanning 1901–2010)[57,82–86] and the TRENDY (version 6, spanning 1960–2016)[58,87,88] ensembles for the period 2007–2010. We chose this study period because of the overlapping temporal coverage of carbon flux estimates (MsTMIP ends in 2010) and high-resolution atmospheric transport (starting in 2007). Each model ensemble used a set of standardized climate drivers to drive individual model runs, thereby minimizing model divergence contributed from climate drivers. There were 29 simulations from 24 independent TBMs. In addition, we used three machine-learning models from the FLUXCOM ensemble for comparison[26]. This yielded $N = 32$ pairs of simulated GPP and NEE estimates for the analysis of ecosystem respiration. The model output was regridded to 1° × 1° resolution. A list of all models is provided in Supplementary Table 3.

Estimating the temperature sensitivity of respiration

Air temperature data were obtained from the North American Regional Reanalysis data[89], regridded to the same monthly, 1° × 1° resolution as the carbon flux estimates.

Ecosystem respiration (R_E) was calculated from the sum of model estimates of GPP and NEE per definition (note that NEE is negative when there is net ecosystem uptake).

We used the Arrhenius equation to describe the relationship between ecosystem respiration (R_E, μmol m^{-2} s^{-1}) and air temperature (T, K):

$$R_E(T) = R_{E,ref} \exp\left[-\frac{E_a}{k_B}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right], \quad (1)$$

where E_a (eV) is the activation energy of ecosystem respiration, $k_B = 8.617333262 \times 10^{-5}$ eV K^{-1} is the Boltzmann constant, $T_{ref} = 10$ °C or 283.15 K is the reference temperature and $R_{E,ref}$ (μmol m^{-2} s^{-1}) is the baseline ecosystem respiration rate at T_{ref}.

We estimated E_a and $R_{E,ref}$ for each model for North America (Fig. 1a) and then also separately for several major biomes (croplands, evergreen needleleaf forests and deciduous broadleaf and mixed forests; Extended Data Fig. 2a) through a linear regression:

$$\ln R_E(T) = \ln R_{E,ref} - \frac{E_a}{k_B}\left(\frac{1}{T} - \frac{1}{T_{ref}}\right). \quad (2)$$

R_E values that were negative or close to zero were filtered out. We used all valid monthly, grid-cell-level R_E and corresponding temperature values during the period 2007–2010 to estimate E_a and $R_{E,ref}$ for each study domain (North America or individual biomes), accounting for both spatial and temporal variabilities in temperature and R_E. This treatment took the assumption of 'trading space for time'[54,90], given the short span of the study period. Note that we did not consider temperature thresholds in the response of ecosystem respiration[52] because monthly mean temperature rarely dips below the previously identified low temperature threshold (−24.8 °C) and most models do not show a discontinuity above the high temperature threshold (15.1 °C) in the study domain.

The Q_{10} formulation expresses respiration as an exponential function of temperature, that is, a linear relationship between $\ln R_E$ and T. Hence we linearized $\ln R_E$ against T to obtain the apparent Q_{10} as a function of E_a and T to aid interpretation of estimated E_a values in the context of reported Q_{10} values in the literature:

$$Q_{10}(E_a, T) = \exp\left(\frac{\partial \ln R_E}{\partial T}\Delta T_{10}\right) = \exp\left(\frac{E_a}{k_B T^2}\Delta T_{10}\right), \quad (3)$$

where $\Delta T_{10} = 10$ K. For example, at $T = 283.15$ K (10 °C), $E_a = 0.65$ eV is equivalent to an apparent Q_{10} of 2.6 (Extended Data Fig. 5).

Optimizing the temperature sensitivity of respiration

Observations of atmospheric CO_2 concentrations are sensitive to net CO_2 fluxes from large regions (10^6 km^2) (ref. 91) and thus provide top-down constraints on carbon flux estimates for large biomes and continental regions[73,92–96], provided that atmospheric transport is adequately resolved to link fluxes to observed concentrations. While atmospheric CO_2 observations can inform the mass balance and space–time variability of net carbon fluxes, they lack direct component-level constraints on photosynthesis and respiration. Consequently, TBM-based estimates[97] or photosynthetic proxies[9,98] have been needed to separately inform estimates of photosynthesis and respiration from atmospheric CO_2 observations.

Here we optimized the continental- and biome-scale temperature sensitivity of ecosystem respiration for individual models by maximizing the consistency between transported signals of ecosystem respiration from bottom-up models and the components of space–time variability in CO_2 concentrations caused by ecosystem respiration (Extended Data Fig. 2).

To do so, atmospheric measurements of CO_2 mixing ratio during 2007–2010 were obtained from 44 tall towers across North America, sourced from the ObsPack CO_2 GlobalViewPlus v3.2 data product[56,99]. We used three hourly averaged CO_2 measurements centred in the afternoon (15:00 local time for most sites and 16:00 or 17:00 for a few remaining sites). From these measurements, we subtracted background CO_2 values and fossil fuel influences (FFDAS v2 (Fossil Fuel Data Assimilation System version 2)) (ref. 100) to derive the net influence of terrestrial biospheric carbon fluxes on atmospheric CO_2 mixing ratio. The resulting observations of CO_2 depletion or enhancement were then filtered to remove influences from non-terrestrial fluxes or data points that showed a large model–data mismatch. In sensitivity analyses, we also subtracted the influence of lateral fluxes obtained from the Global Stocktake data product[101–104] (Supplementary Notes 2). We then removed the influence of GPP on atmospheric CO_2 observations in the same way. This yielded the respiratory component of the space–time variability in atmospheric CO_2 observations. Because biomass-burning emissions (~0.06 Pg C yr^{-1}) (ref. 105) were minor compared with NEE over North America (−0.7 Pg C yr^{-1}) (ref. 95) during the study period (2007–2010)—and certainly dwarfed by model spread in ecosystem respiration—influences from biomass-burning emissions on the temperature sensitivity of ecosystem respiration were not considered.

Sensitivities of CO_2 mixing ratio measurements to surface fluxes, also known as transport footprints, were produced from high-resolution (10 km for temperate North America and 40 km for tropical and high-latitude North America) WRF–STILT (Weather Research and Forecasting – Stochastic Time Inverted Lagrangian Transport) model runs[106] and post aggregated to a 1° × 1° resolution as part

of the CarbonTracker–Lagrange project (https://gml.noaa.gov/ccgg/carbontracker-lagrange/). The potential transport error is demonstrably minor relative to the spread in TBM flux estimates[107,108].

We then calculated the adjustments to the temperature sensitivity (E_a) and the baseline respiration rate ($R_{E,ref}$) by minimizing the mismatch between the transported signal of model estimates of ecosystem respiration and the respiratory component of atmospheric CO_2 variability. The original estimates of ecosystem respiration (R_E, equation (1)) were adjusted as follows:

$$R_E^* = R_E \alpha \exp\left[-\frac{\Delta E_a}{k_B}\left(\frac{1}{T}-\frac{1}{T_{ref}}\right)\right], \quad (4)$$

where R_E^* (μmol m^{-2} s^{-1}) is the adjusted estimate of ecosystem respiration, ΔE_a (eV) is the adjustment to E_a needed to minimize the mismatch and $\alpha > 0$ is a dimensionless scaling factor for the baseline respiration rate.

Like the original estimates of E_a, the adjustments ΔE_a were estimated over the North American domain and for major biomes. The parameter α accounted for the adjustment to the magnitude of ecosystem respiration (Extended Data Fig. 7), thereby allowing the bias from temperature sensitivity and that from the seasonal cycle amplitude to be separately constrained.

To obtain an estimate of the overall optimal E_a for North America (and for individual biomes) across models, we fitted linear relationships between ΔE_a adjustments and the original E_a estimates for individual models using the orthogonal distance regression. Orthogonal distance regression was used because it accounts for errors in both the regressor and the response[109]. The fitting was performed using the SciPy ODR function[110,111]. An optimal value of E_a, that is, $\hat{E}_{a,opt}$ was derived at the intersection of $\Delta E_a = 0$ and the best fit line for each domain (North America or individual biomes; Fig. 2).

Evaluating carbon flux estimates using observations

We used in situ measurements of atmospheric CO_2 mixing ratio and the transport footprints to determine the extent to which regional-scale estimates of GPP and NEE are consistent with observations. As in Sun et al.[73], we calculated the transported signal of GPP or NEE estimates for each individual model and assessed how well the transported signal explains the biospheric component of the observed CO_2 space–time variability using the coefficient of determination (R^2).

We used the explanatory power of shortwave radiation, $R^2_{SW} = 0.23$, as a minimum benchmark to exclude a small subset of models with very low explanatory power ($N = 3$) from the subsequent analysis on ecosystem respiration. Shortwave radiation data were obtained from the North American Regional Reanalysis data[89]. Because shortwave radiation is a first-order climatic driver of GPP[112] and influential in determining NEE space–time variability[113], we would expect carbon flux estimates to explain at least as much portion of observed atmospheric CO_2 variability as the shortwave radiation.

Evaluating temperature sensitivity-corrected NEE estimates

We assessed how updating the temperature sensitivity of respiration ('Optimizing the temperature sensitivity of respiration') impacted the explanatory power of NEE estimates.

To do so, we adjusted ecosystem respiration estimates for each model according to the optimal ΔE_a and α derived for that model in the optimization procedure (equation (4)). Note that correction to the baseline respiration rate through α changes the mean ecosystem respiration magnitude (equation (4)), which avoids erroneously adjusting the temperature sensitivity in response to a bias in overall magnitude.

In a second analysis, we instead adjusted E_a of all models to the optimal value over North America ($\hat{E}_{a,opt} = 0.43$ eV) estimated from the model ensemble. The adjustment of R_E was conducted by multiplying R_E by the temperature sensitivity adjustment, $\exp(-\frac{\Delta E_a}{k_B T})$, followed by rescaling the magnitude to conserve the mean R_E. We also capped the adjusted R_E at its original maximum value $\max(R_E)$ within a relative tolerance level (≤1%) to prevent unrealistically high values being produced by the exponential function.

The performance of the two sets of adjusted NEE estimates in explaining the space–time variability in atmospheric CO_2 measurements was evaluated using the same approach as described in the previous section.

Reporting summary

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

Data availability

Results presented in this study are available at https://doi.org/10.5281/zenodo.7874439. The ObsPack GLOBALVIEWplus CO_2 data product is available at https://www.esrl.noaa.gov/gmd/ccgg/obspack. The CarbonTracker–Lagrange WRF–STILT footprints are available at https://gml.noaa.gov/ccgg/carbontracker-lagrange/. The FFDAS v2 data product is available at https://ffdas.rc.nau.edu/. The North American Regional Reanalysis data can be obtained at https://psl.noaa.gov/data/gridded/data.narr.html. The MsTMIP v2 model ensemble is available at https://nacp.ornl.gov/. The TRENDY v6 model ensemble is available at https://blogs.exeter.ac.uk/trendy/. The FLUXCOM model ensemble is available at http://www.fluxcom.org/.

Code availability

Code used to generate figures and results is available at https://doi.org/10.5281/zenodo.7874439.

References

1. Friedlingstein, P., Dufresne, J.-L., Cox, P. M. & Rayner, P. How positive is the feedback between climate change and the carbon cycle? *Tellus B* **55**, 692–700 (2003).
2. Friedlingstein, P. et al. Climate-carbon cycle feedback analysis: results from the C^4MIP model intercomparison. *J. Clim.* **19**, 3337–3353 (2006).
3. Friedlingstein, P. et al. Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Clim.* **27**, 511–526 (2014).
4. Arora, V. K. et al. Carbon-concentration and carbon-climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences* **17**, 4173–4222 (2020).
5. Heimann, M. & Reichstein, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* **451**, 289–292 (2008).
6. Beer, C. et al. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329**, 834–838 (2010).
7. Anav, A. et al. Spatiotemporal patterns of terrestrial gross primary production: a review. *Rev. Geophys.* **53**, 785–818 (2015).
8. Shao, P., Zeng, X., Moore, D. J. P. & Zeng, X. Soil microbial respiration from observations and earth system models. *Environ. Res. Lett.* **8**, 034034 (2013).
9. Byrne, B. et al. Evaluating GPP and respiration estimates over northern midlatitude ecosystems using solar-induced fluorescence and atmospheric CO_2 measurements. *J. Geophys. Res. Biogeosci.* **123**, 2976–2997 (2018).
10. Jian, J. et al. Leveraging observed soil heterotrophic respiration fluxes as a novel constraint on global-scale models. *Glob. Change Biol.* **27**, 5392–5403 (2021).
11. Fisher, J. B., Huntzinger, D. N., Schwalm, C. R. & Sitch, S. Modeling the terrestrial biosphere. *Annu. Rev. Environ. Resour.* **39**, 91–123 (2014).
12. Frankenberg, C. et al. New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophys. Res. Lett.* **38** (2011).

13. Badgley, G., Field, C. B. & Berry, J. A. Canopy near-infrared reflectance and terrestrial photosynthesis. *Sci. Adv.* **3**, e1602244 (2017).
14. Campbell, J. E. et al. Photosynthetic control of atmospheric carbonyl sulfide during the growing season. *Science* **322**, 1085–1088 (2008).
15. Whelan, M. E. et al. Reviews and syntheses: carbonyl sulfide as a multi-scale tracer for carbon and water cycles. *Biogeosciences* **15**, 3625–3657 (2018).
16. Koven, C. D. et al. Permafrost carbon-climate feedbacks accelerate global warming. *Proc. Natl Acad. Sci.* **108**, 14769–14774 (2011).
17. Koven, C. D., Hugelius, G., Lawrence, D. M. & Wieder, W. R. Higher climatological temperature sensitivity of soil carbon in cold than warm climates. *Nat. Clim. Change* **7**, 817–822 (2017).
18. Schuur, E. A. G. et al. Climate change and the permafrost carbon feedback. *Nature* **520**, 171–179 (2015).
19. Bouskill, N. J., Riley, W. J., Zhu, Q., Mekonnen, Z. A. & Grant, R. F. Alaskan carbon-climate feedbacks will be weaker than inferred from short-term experiments. *Nat. Commun.* **11**, 5798 (2020).
20. Koven, C. D. et al. The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences* **10**, 7109–7131 (2013).
21. Atkin, O. K. et al. In *Plant Respiration: Metabolic Fluxes and Carbon Balance* (eds Tcherkez, G. & Ghashghaie, J.) 107–142 (Springer International Publishing, 2017).
22. Luyssaert, S. et al. CO_2 balance of boreal, temperate, and tropical forests derived from a global database. *Glob. Change Biol.* **13**, 2509–2537 (2007).
23. Konings, A. G. et al. Global satellite-driven estimates of heterotrophic respiration. *Biogeosciences* **16**, 2269–2284 (2019).
24. Yuan, W. et al. Redefinition and global estimation of basal ecosystem respiration rate. *Glob. Biogeochem. Cycles* **25** (2011).
25. Ai, J. et al. MODIS-based estimates of global terrestrial ecosystem respiration. *J. Geophys. Res. Biogeosci.* **123**, 326–352 (2018).
26. Jung, M. et al. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences* **17**, 1343–1365 (2020).
27. Schwalm, C. R. et al. A model-data intercomparison of CO_2 exchange across North America: results from the North American Carbon Program site synthesis. *J. Geophys. Res. Biogeosci.* **115**, G00H05 (2010).
28. Bonan, G. B. In *Climate Change and Terrestrial Ecosystem Modeling* 1–24 (Cambridge Univ. Press, 2019).
29. Enquist, B. J. et al. Scaling metabolism from organisms to ecosystems. *Nature* **423**, 639–642 (2003).
30. Yvon-Durocher, G. et al. Reconciling the temperature dependence of respiration across timescales and ecosystem types. *Nature* **487**, 472–476 (2012).
31. Arrhenius, S. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Z. Phys. Chem.* **4U**, 226–248 (1889).
32. Reichstein, M. et al. Ecosystem respiration in two Mediterranean evergreen holm oak forests: drought effects and decomposition dynamics. *Funct. Ecol.* **16**, 27–39 (2002).
33. Lindroth, A. et al. Leaf area index is the principal scaling parameter for both gross photosynthesis and ecosystem respiration of Northern deciduous and coniferous forests. *Tellus B* **60**, 129–142 (2008).
34. Migliavacca, M. et al. Semiempirical modeling of abiotic and biotic factors controlling ecosystem respiration across eddy covariance sites. *Glob. Change Biol.* **17**, 390–409 (2011).
35. Collalti, A. et al. Plant respiration: controlled by photosynthesis or biomass? *Glob. Change Biol.* **26**, 1739–1753 (2020).
36. Reich, P. B. et al. Scaling of respiration to nitrogen in leaves, stems and roots of higher land plants. *Ecol. Lett.* **11**, 793–801 (2008).
37. Fierer, N., Craine, J. M., McLauchlan, K. & Schimel, J. P. Litter quality and the temperature sensitivity of decomposition. *Ecology* **86**, 320–326 (2005).
38. Schipper, L. A., Hobbs, J. K., Rutledge, S. & Arcus, V. L. Thermodynamic theory explains the temperature optima of soil microbial processes and high Q_{10} values at low temperatures. *Glob. Change Biol.* **20**, 3578–3586 (2014).
39. Alster, C. J., Koyama, A., Johnson, N. G., Wallenstein, M. D. & von Fischer, J. C. Temperature sensitivity of soil microbial communities: an application of macromolecular rate theory to microbial respiration. *J. Geophys. Res. Biogeosci.* **121**, 1420–1433 (2016).
40. Phillips, S. C. et al. Interannual, seasonal, and diel variation in soil respiration relative to ecosystem respiration at a wetland to upland slope at Harvard Forest. *J. Geophys. Res. Biogeosci.* **115**, G02019 (2010).
41. Piao, S. et al. Forest annual carbon cost: a global-scale analysis of autotrophic respiration. *Ecology* **91**, 652–661 (2010).
42. Bond-Lamberty, B. & Thomson, A. Temperature-associated increases in the global soil respiration record. *Nature* **464**, 579–582 (2010).
43. Hursh, A. et al. The sensitivity of soil respiration to soil temperature, moisture, and carbon supply at the global scale. *Glob. Change Biol.* **23**, 2090–2103 (2017).
44. Mahecha, M. D. et al. Global convergence in the temperature sensitivity of respiration at ecosystem level. *Science* **329**, 838–840 (2010).
45. Pastorello, G. et al. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci. Data* **7**, 225 (2020).
46. Schimel, D. et al. Observing terrestrial ecosystems and the carbon cycle from space. *Glob. Change Biol.* **21**, 1762–1776 (2015).
47. Villarreal, S. & Vargas, R. Representativeness of FLUXNET sites across Latin America. *J. Geophys. Res. Biogeosci.* **126**, e2020JG006090 (2021).
48. Schuepp, P. H., Leclerc, M. Y., MacPherson, J. I. & Desjardins, R. L. Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation. *Boundary Layer Meteorol.* **50**, 355–373 (1990).
49. Chen, B. et al. Assessing eddy-covariance flux tower location bias across the Fluxnet-Canada Research Network based on remote sensing and footprint modelling. *Agric. For. Meteorol.* **151**, 87–100 (2011).
50. Chu, H. et al. Representativeness of eddy-covariance flux footprints for areas surrounding AmeriFlux sites. *Agric. For. Meteorol.* **301-302**, 108350 (2021).
51. Villarreal, S. et al. Ecosystem functional diversity and the representativeness of environmental networks across the conterminous United States. *Agric. For. Meteorol.* **262**, 423–433 (2018).
52. Johnston, A. S. A. et al. Temperature thresholds of ecosystem respiration at a global scale. *Nat. Ecol. Evol.* **5**, 487–494 (2021).
53. Tjoelker, M. G., Oleksyn, J. & Reich, P. B. Modelling respiration of vegetation: evidence for a general temperature-dependent Q_{10}. *Glob. Change Biol.* **7**, 223–230 (2001).
54. Niu, B. et al. Warming homogenizes apparent temperature sensitivity of ecosystem respiration. *Sci. Adv.* **7**, eabc7358 (2021).
55. Crous, K. Y., Uddling, J. & De Kauwe, M. G. Temperature responses of photosynthesis and respiration in evergreen trees from boreal to tropical latitudes. *New Phytol.* **234**, 353–374 (2022).
56. Masarie, K. A., Peters, W., Jacobson, A. R. & Tans, P. P. ObsPack: a framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements. *Earth Syst. Sci. Data* **6**, 375–384 (2014).

57. Huntzinger, D. N. et al. Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Sci. Rep.* **7**, 4765 (2017).
58. Le Quéré, C. et al. Global carbon budget 2017. *Earth Syst. Sci. Data* **10**, 405–448 (2018).
59. Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M. & Charnov, E. L. Effects of size and temperature on metabolic rate. *Science* **293**, 2248–2251 (2001).
60. Oleson, K. W. et al. *Technical Description of Version 4.5 of the Community Land Model (CLM).* Report No. NCAR/TN-503+STR (National Center for Atmospheric Research, 2013); https://doi.org/10.5065/D6RR1W7M
61. Fan, N. et al. Global apparent temperature sensitivity of terrestrial carbon turnover modulated by hydrometeorological factors. *Nat. Geosci.* **15**, 989–994 (2022).
62. Migliavacca, M. et al. Influence of physiological phenology on the seasonal pattern of ecosystem respiration in deciduous forests. *Glob. Change Biol.* **21**, 363–376 (2015).
63. Goulden, M. L., Munger, J. W., Fan, S.-M., Daube, B. C. & Wofsy, S. C. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. *Glob. Change Biol.* **2**, 169–182 (1996).
64. Aubinet, M. et al. Direct advection measurements do not help to solve the night-time CO_2 closure problem: evidence from three different forests. *Agric. For. Meteorol.* **150**, 655–664 (2010).
65. Kok, B. On the interrelation of respiration and photosynthesis in green plants. *Biochim. Biophys. Acta* **3**, 625–631 (1949).
66. Wehr, R. et al. Seasonality of temperate forest photosynthesis and daytime respiration. *Nature* **534**, 680–683 (2016).
67. Yin, X., Niu, Y., van der Putten, P. E. L. & Struik, P. C. The Kok effect revisited. *New Phytol.* **227**, 1764–1775 (2020).
68. Bruhn, D. et al. Nocturnal plant respiration is under strong non-temperature control. *Nat. Commun.* **13**, 5650 (2022).
69. Keenan, T. F. et al. Widespread inhibition of daytime ecosystem respiration. *Nat. Ecol. Evol.* **3**, 407–415 (2019).
70. Campioli, M. et al. Biomass production efficiency controlled by management in temperate and boreal ecosystems. *Nat. Geosci.* **8**, 843–846 (2015).
71. Eugster, W. et al. Management effects on European cropland respiration. *Agric. Ecosyst. Environ.* **139**, 346–362 (2010).
72. Ryu, Y., Berry, J. A. & Baldocchi, D. D. What is global photosynthesis? History, uncertainties and opportunities. *Remote Sens. Environ.* **223**, 95–114 (2019).
73. Sun, W. et al. Midwest US croplands determine model divergence in North American carbon fluxes. *AGU Adv.* **2**, e2020AV000310 (2021).
74. Reichstein, M. et al. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Glob. Change Biol.* **11**, 1424–1439 (2005).
75. Shiga, Y. P., Michalak, A. M., RandolphKawa, S. & Engelen, R. J. In-situ CO_2 monitoring network evaluation and design: a criterion based on atmospheric CO_2 variability. *J. Geophys. Res. Atmos.* **118**, 2007–2018 (2013).
76. Burton, A. J., Melillo, J. M. & Frey, S. D. Adjustment of forest ecosystem root respiration as temperature warms. *J. Integr. Plant Biol.* **50**, 1467–1483 (2008).
77. Bradford, M. A. et al. Thermal adaptation of soil microbial respiration to elevated temperature. *Ecol. Lett.* **11**, 1316–1327 (2008).
78. Crowther, T. W. & Bradford, M. A. Thermal acclimation in widespread heterotrophic soil microbes. *Ecol. Lett.* **16**, 469–477 (2013).
79. Sullivan, P. F., Stokes, M. C., McMillan, C. K. & Weintraub, M. N. Labile carbon limits late winter microbial activity near Arctic treeline. *Nat. Commun.* **11**, 4024 (2020).
80. Natali, S. M. et al. Large loss of CO_2 in winter observed across the northern permafrost region. *Nat. Clim. Change* **9**, 852–857 (2019).
81. Fernández-Martínez, M. et al. Diagnosing destabilization risk in global land carbon sinks. *Nature* **615**, 848–853 (2023).
82. Huntzinger, D. N. et al. The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project—part 1: overview and experimental design. *Geosci. Model Dev.* **6**, 2121–2133 (2013).
83. Huntzinger, D. N. et al. NACP MsTMIP: global 0.5-degree model outputs in standard vormat, Version 1.0. *NASA Earth Data* https://doi.org/10.3334/ORNLDAAC/1225 (2018).
84. Huntzinger, D. N. et al. NACP MsTMIP: global 0.5-degree model outputs in standard format, Version 2.0. *NASA Earth Data* https://doi.org/10.3334/ORNLDAAC/1599 (2020).
85. Wei, Y. et al. NACP MsTMIP: global and North American driver data for multi-model intercomparison. *NASA Earth Data* https://doi.org/10.3334/ORNLDAAC/1220 (2014).
86. Wei, Y. et al. The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project—part 2: environmental driver data. *Geosci. Model Dev.* **7**, 2875–2893 (2014).
87. Sitch, S. et al. Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Glob. Change Biol.* **14**, 2015–2039 (2008).
88. Sitch, S. et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653–679 (2015).
89. Mesinger, F. et al. North American regional reanalysis. *Bull. Am. Meteorol. Soc.* **87**, 343–360 (2006).
90. Pickett, S. T. A. In *Long-Term Studies in Ecology: Approaches and Alternatives* (ed Likens, G. E.) 110–135 (Springer, 1989).
91. Gloor, M. et al. What is the concentration footprint of a tall tower? *J. Geophys. Res. Atmos.* **106**, 17831–17840 (2001).
92. Kaminski, T., Knorr, W., Rayner, P. J. & Heimann, M. Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle. *Glob. Biogeochem. Cycles* **16**, 1066 (2002).
93. Schuh, A. E. et al. Evaluating atmospheric CO_2 inversions at multiple scales over a highly inventoried agricultural landscape. *Glob. Change Biol.* **19**, 1424–1439 (2013).
94. Fang, Y., Michalak, A. M., Shiga, Y. P. & Yadav, V. Using atmospheric observations to evaluate the spatiotemporal variability of CO_2 fluxes simulated by terrestrial biospheric models. *Biogeosciences* **11**, 6985–6997 (2014).
95. Hu, L. et al. Enhanced North American carbon uptake associated with El Niño. *Sci. Adv.* **5**, eaaw0076 (2019).
96. Byrne, B. et al. Improved constraints on northern extratropical CO_2 fluxes obtained by combining surface-based and space-based atmospheric CO_2 measurements. *J. Geophys. Res. Atmos.* **125**, e2019JD032029 (2020).
97. Schuh, A. E. et al. A regional high-resolution carbon flux inversion of North America for 2004. *Biogeosciences* **7**, 1625–1644 (2010).
98. Hu, L. et al. COS-derived GPP relationships with temperature and light help explain high-latitude atmospheric CO_2 seasonal cycle amplification. *Proc. Natl Acad. Sci.* **118**, e2103423118 (2021).
99. Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957–2016; obspack_co2_1_GLOBALVIEWplus_v3.2_2017-11-02. *Cooperative Global Atmospheric Data Integration Project* https://doi.org/10.15138/G3704H (2017).
100. Asefi-Najafabady, S. et al. A multiyear, global gridded fossil fuel CO_2 emission data product: evaluation and analysis of results. *J. Geophys. Res. Atmos.* **119**, 10213–10231 (2014).
101. Byrne, B. et al. National CO_2 budgets (2015-2020) inferred from atmospheric CO_2 observations in support of the global stocktake. *Earth Syst. Sci. Data* **15**, 963–1004 (2023).

102. Byrne, B. et al. Pilot top-down CO_2 budget constrained by the v10 OCO-2 MIP Version 1. *Committee on Earth Observation Satellites* https://doi.org/10.48588/npf6-sw92 (2022).
103. Ciais, P. et al. Definitions and methods to estimate regional land carbon fluxes for the second phase of the REgional Carbon Cycle Assessment and Processes Project (RECCAP-2). *Geosci. Model Dev.* **15**, 1289–1316 (2022).
104. Deng, Z. et al. Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions. *Earth Syst. Sci. Data* **14**, 1639–1675 (2022).
105. Randerson, J. T., Van Der Werf, G. R., Giglio, L., Collatz, G. J. & Kasibhatla, P. S. Global fire emissions database, Version 4.1 (GFEDv4). *NASA Earth Data* https://doi.org/10.3334/ORNLDAAC/1293 (2018)
106. Nehrkorn, T. et al. Coupled weather research and forecasting-stochastic time-inverted lagrangian transport (WRF-STILT) model. *Meteorol. Atmos. Phys.* **107**, 51–64 (2010).
107. Bagley, J. E. et al. Assessment of an atmospheric transport model for annual inverse estimates of California greenhouse gas emissions. *J. Geophys. Res. Atmos.* **122**, 1901–1918 (2017).
108. Rastogi, B. et al. Evaluating consistency between total column CO_2 retrievals from OCO-2 and the in situ network over North America: implications for carbon flux estimation. *Atmos. Chem. Phys.* **21**, 14385–14401 (2021).
109. Boggs, P. T., Spiegelman, C. H., Donaldson, J. R. & Schnabel, R. B. A computational examination of orthogonal distance regression. *J. Econom.* **38**, 169–201 (1988).
110. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
111. Boggs, P. T., Byrd, R. H., Rogers, J. E. & Schnabel, R. B. *User's Reference Guide for ODRPACK Version 2.01 Software for Weighted Orthogonal Distance Regression*. Report No. NISTIR 4834 (Applied Computational Mathematics Division, National Institute of Standards and Technology, 1992).
112. Monteith, J. L. Solar radiation and productivity in tropical ecosystems. *J. Appl. Ecol.* **9**, 747–766 (1972).
113. Fang, Y. & Michalak, A. M. Atmospheric observations inform CO_2 flux responses to enviroclimatic drivers. *Glob. Biogeochem. Cycles* **29**, 555–566 (2015).
114. Allen, A. P. & Gillooly, J. F. The mechanistic basis of the metabolic theory of ecology. *Oikos* **116**, 1073–1077 (2007).

Acknowledgements

This study was funded by the National Aeronautics and Space Administration (NASA) through the Terrestrial Ecology Interdisciplinary Science award number NNH17AE86I and the Carbon Monitoring System award number 80NSSC18K0165. W.S. acknowledges additional support from the Carnegie Institution for Science's endowment fund. X.L. acknowledges support from Singapore's Ministry of Education. T.F.K. acknowledges additional support from the Reducing Uncertainty in Biogeochemical Interactions through Synthesis and Computation Scientific Focus Area (RUBISCO SFA), which is sponsored by the Regional and Global Model Analysis Program in the Climate and Environmental Sciences Division of the Office of Biological and Environmental Research in the US Department of Energy Office of Science, and a NASA award number 80NSSC21K1705.
We acknowledge the National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Laboratory and site principal investigators for providing atmospheric CO_2 observations. We thank the following individuals for collecting and providing the atmospheric CO_2 data: A. Andrews, S. Biraud, K. Davis, S. De Wekker, M. L. Fischer, T. Griffis, B. Law, N. Miles, J. W. Munger, M. J. Parker, S. Richardson, B. Stephens, C. Sweeney, P. Tans, K. Thoning, M. Torn, S. Wofsy, and D. Worthy. See Supplementary Table 4 for detailed site information. Measurements at Walnut Grove, CA, were partially supported by grants from the California Energy Commission Public Interest Environmental Research Program to the Lawrence Berkeley National Laboratory, which operates under US Department of Energy under contract DE-AC02-05CH11231. WRF–STILT atmospheric transport simulations were supported by the NASA grant 'Regional Inverse Modeling in North and South America for the NASA Carbon Monitoring System.' We thank T. Nehrkorn, J. Henderson and J. Eluszkiewicz at Atmospheric and Environmental Research Inc. for conducting WRF–STILT simulations and the CarbonTracker–Lagrange project team for providing WRF–STILT transport footprints.
We thank all data contributors. We thank S. Sitch, P. Friedlingstein and all modellers who have contributed to the Trends in Net Land–Atmosphere Exchange project (TRENDY; http://sites.exeter.ac.uk/trendy). Funding for the Multi-scale synthesis and Terrestrial Model Intercomparison Project (MsTMIP; https://nacp.ornl.gov/MsTMIP.shtml) activity was provided through NASA Research Opportunities in Space and Earth Science (ROSES) grant number NNX10AG01A. Data management support for preparing, documenting and distributing model driver and output data was performed by the Modeling and Synthesis Thematic Data Center at Oak Ridge National Laboratory (ORNL; https://nacp.ornl.gov), with funding through NASA ROSES grant number NNH10AN681. Finalized MsTMIP data products are archived at the ORNL DAAC (https://daac.ornl.gov). We thank the FLUXCOM Initiative (https://www.fluxcom.org) for FLUXCOM data products. We thank K. Gurney for the FFDAS v2 dataset. We thank B. Byrne and the Committee on Earth Observation Satellites for providing gridded lateral carbon flux estimates, accessed at https://doi.org/10.48588/npf6-sw92. We thank F. Chevallier, P. Ciais and the European Union's CoCO2 project (https://coco2-project.eu/) for generating the gridded lateral carbon flux dataset. We acknowledge the National Centers for Environmental Prediction North American Regional Reanalysis data provided by the NOAA Physical Sciences Laboratory, obtained from https://psl.noaa.gov/. The computations presented in this study were conducted through the Carnegie Institution for Science's partnership in the Resnick High Performance Computing Center (https://www.hpc.caltech.edu), a facility supported by Resnick Sustainability Institute at the California Institute of Technology.
This study is a contribution to the North American Carbon Program (https://www.nacarbon.org).

Author contributions

W.S. and A.M.M. envisioned and designed the study and developed the methods. W.S. performed the analyses and visualization with input from A.M.M. W.S. and A.M.M. led the interpretation of the results with input from X.L., Y.F., Y.P.S., J.B.F., T.F.K. and Y.Z. W.S., X.L., Y.F., Y.P.S. and Y.Z. curated and processed the data. T.F.K., A.M.M. and J.B.F. acquired funding that supported the study and supervised the work. W.S. wrote the original draft of the manuscript with input from A.M.M. All authors provided feedback on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-023-02093-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-023-02093-x.

Correspondence and requests for materials should be addressed to Wu Sun or Anna M. Michalak.

Peer review information *Nature Ecology & Evolution* thanks Ashley Ballantyne, Kristine Crous and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

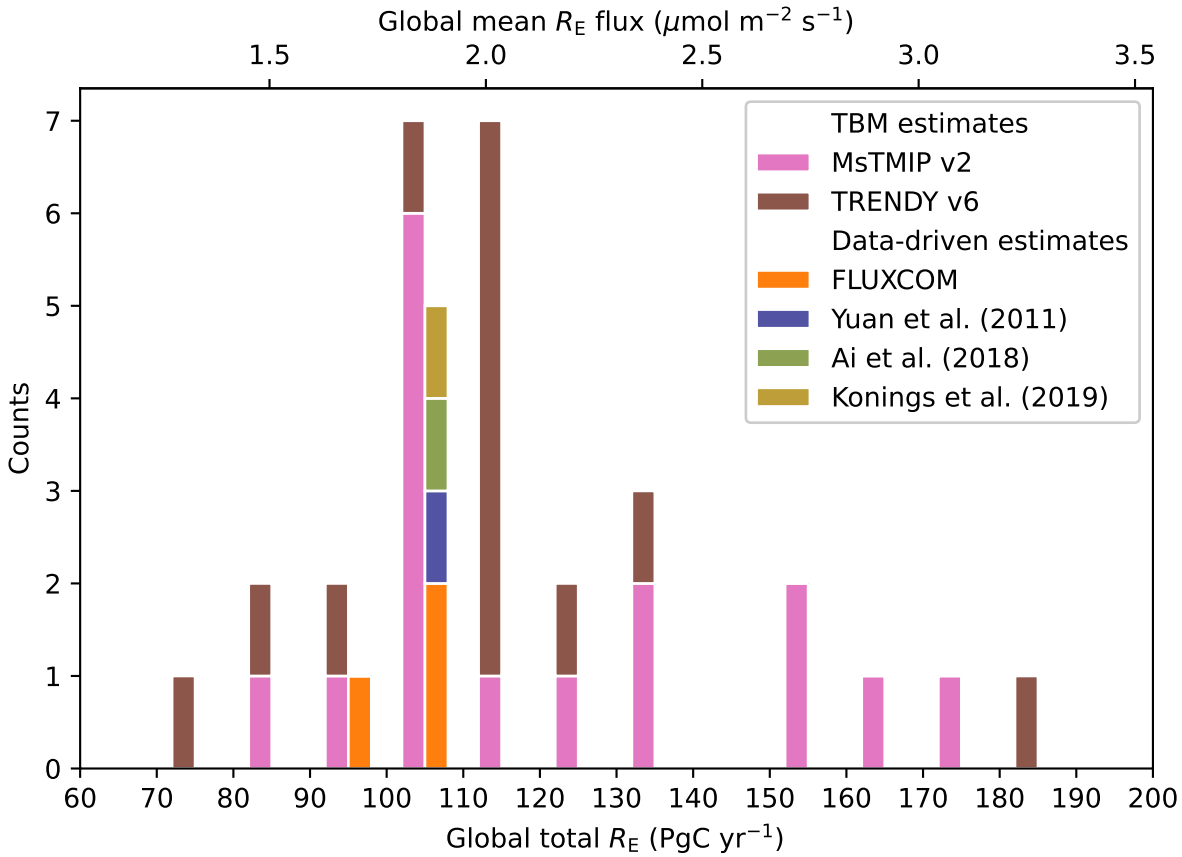
Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

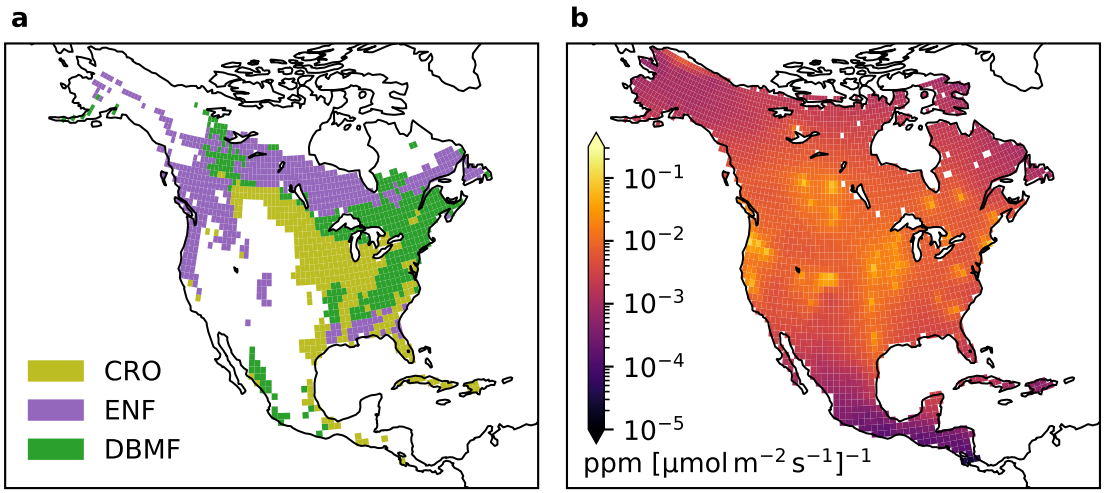
© The Author(s) 2023

https://doi.org/10.1038/s41559-023-02093-x



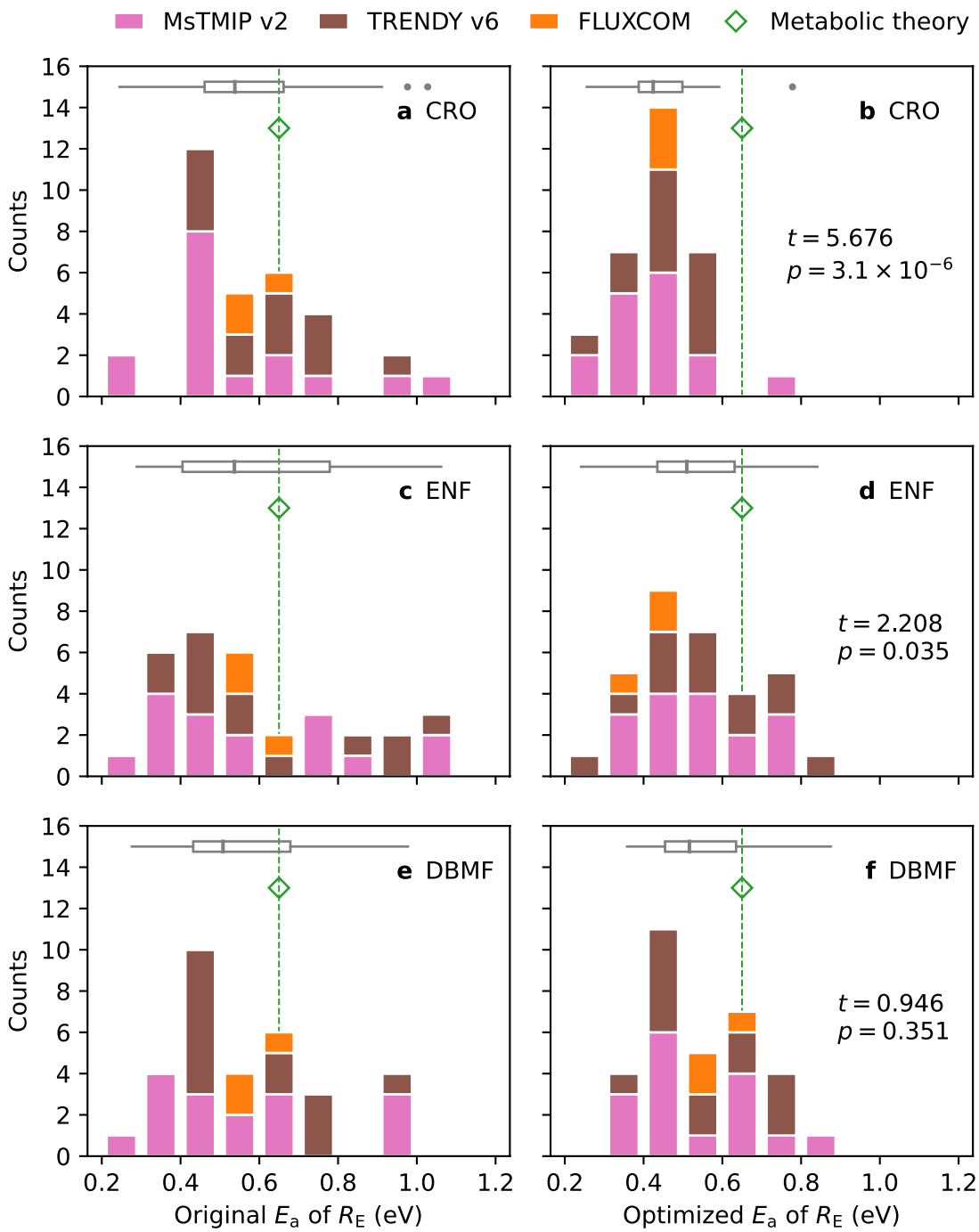
Extended Data Fig. 1 | Estimates of global terrestrial ecosystem respiration vary more widely among terrestrial biosphere models than among data-driven models. Terrestrial biosphere model (TBM) estimates of the global total ecosystem respiration (R_E) are shown from the MsTMIP v2 (pink) and TRENDY v6 (brown) model intercomparison projects for the period 2007–2010. Data-driven estimates of global total R_E include those from three FLUXCOM models (orange) for the same period (2007–2010). In addition, three remote-sensing based estimates from different periods are also included as data-driven estimates, namely, estimates from Yuan et al.[24] for 2000–-2003 (blue), Ai et al.[25] for 2001–2010 (green), and Konings et al.[23] for 2010–2012 (beige). Note that the global estimate from Ai et al.[25] is rescaled to be consistent with the global domain in Yuan et al.[24], assuming the same ratio between global R_E and the total R_E for land south of 65°N and excluding mosaic crops and natural vegetation. The estimate of total ecosystem respiration from Konings et al.[23] is calculated as the sum of heterotrophic and autotrophic respiration estimates presented in their data product. The top axis shows R_E in equivalent flux units if averaged over the global land area (1.49×10^8 km^2).

https://doi.org/10.1038/s41559-023-02093-x



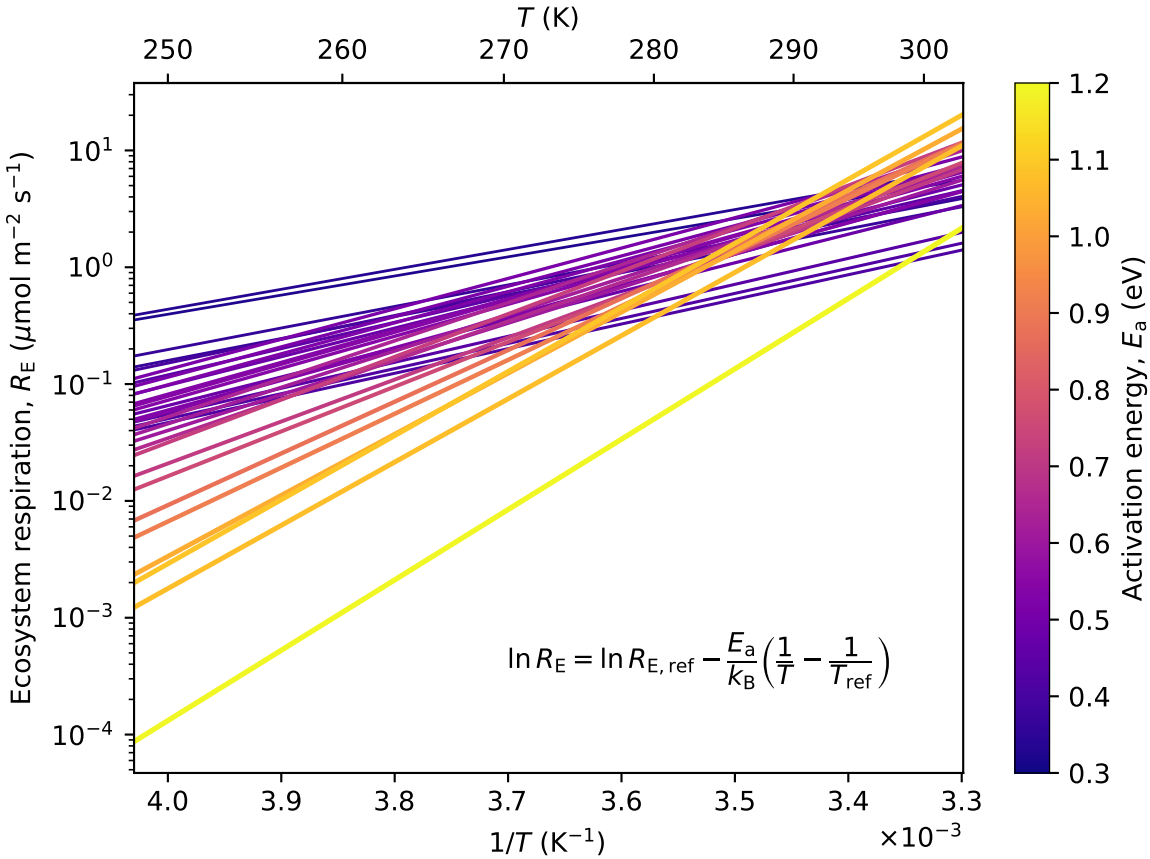
Extended Data Fig. 2 | Atmospheric CO_2 measurements from existing observational networks are most sensitive to fluxes from mesic biomes in temperate and boreal North America. (a) The distribution of major biomes in North America: croplands (CRO; olive), evergreen needleleaf forests (ENF; purple), and deciduous broadleaf and mixed forests (DBMF; green). **(b)** Average sensitivity of atmospheric CO_2 measurements to surface fluxes (ppm [µmol m^{-2} s^{-1}]$^{-1}$) over North America during 2007–2010. Hotspots indicate sites of the CO_2 continuous-monitoring network (Supplementary Table 4). The base map was generated using matplotlib v3.7.1 (https://matplotlib.org) and Cartopy v0.21.1 (https://scitools.org.uk/cartopy).

https://doi.org/10.1038/s41559-023-02093-x



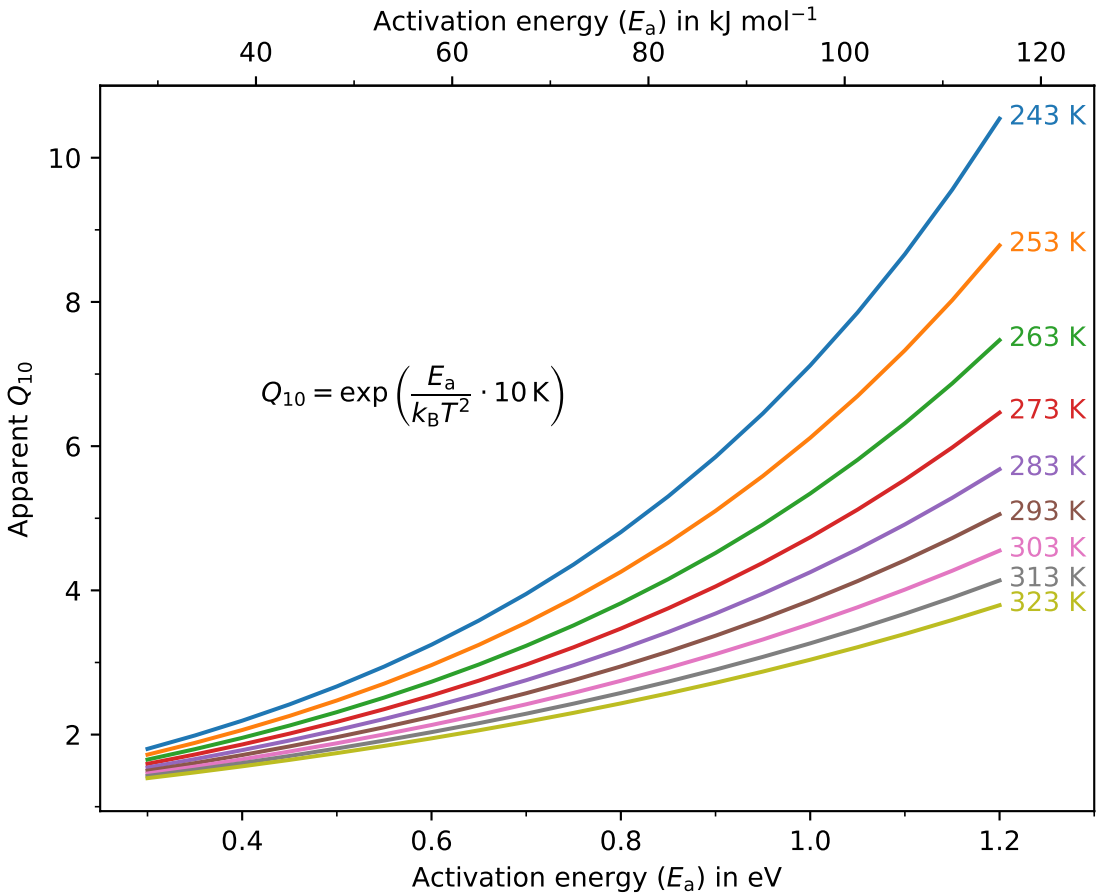
Extended Data Fig. 3 | Using atmospheric observations to constrain the temperature sensitivity of respiration for North American biomes reduces the spread in estimates across models and suggests that the large-scale sensitivity is lower than that implied by the metabolic theory of ecology and by plot-scale studies. Similar to Fig. 1, but for individual biomes. Histograms of the original (left) and optimized (right) aggregate temperature sensitivity of ecosystem respiration for croplands (CRO; **a, b**), evergreen needleleaf forests (ENF; **c, d**), and deciduous broadleaf and mixed forests (DBMF; **e, f**), as represented by $N = 32$ independent estimates of the activation energy (E_a), for TBMs in the MsTMIP v2 (pink) and TRENDY v6 (brown) ensembles and data-driven models in the FLUXCOM ensemble (orange). Grey boxplots summarize the estimates across models, with the centre line, bounds of box, whiskers, and dots representing the median, first and third quartiles, smallest and largest estimates falling within 1.5× of the interquartile range from the nearest quartiles, and outliers beyond that range, respectively. Green diamonds and vertical dashed lines indicate the reference value of 0.65 eV based on the metabolic theory of ecology[29,59,114] and plot-scale estimates[52]. Similar to Fig. 1, *t* statistics and *p* values in (**b, d, f**) indicate differences between model-represented temperature sensitivities before and after optimization against atmospheric CO_2 observations for CRO, ENF, and DBMF, respectively, according to two-tailed, paired two-sample *t* tests.

https://doi.org/10.1038/s41559-023-02093-x



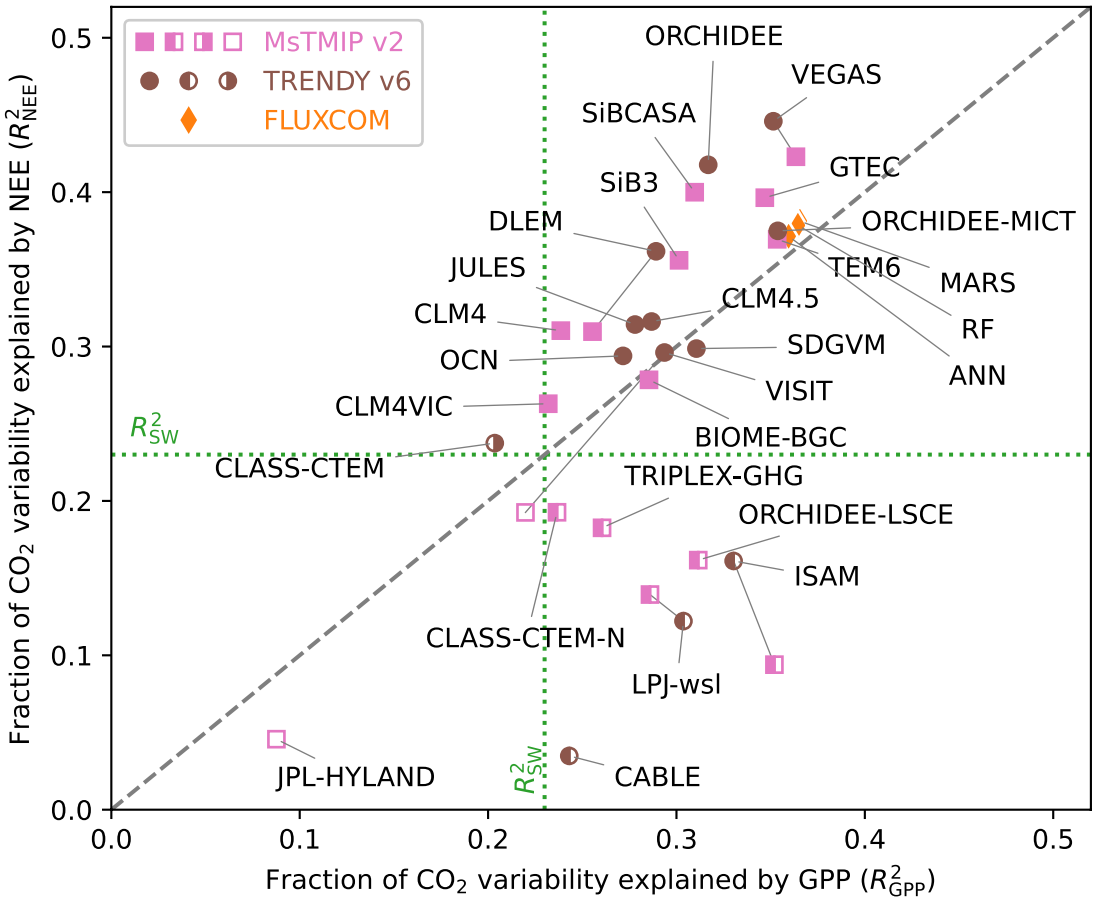
Extended Data Fig. 4 | Fitted temperature responses of model estimates of ecosystem respiration. The natural log of ecosystem respiration (R_E) varies linearly with the reciprocal of temperature ($1/T$), according to the Arrhenius equation. Each line is derived from ecosystem respiration estimates from a model simulation in the ensemble and is colour-coded according to the fitted activation energy (E_a, colourbar).

https://doi.org/10.1038/s41559-023-02093-x



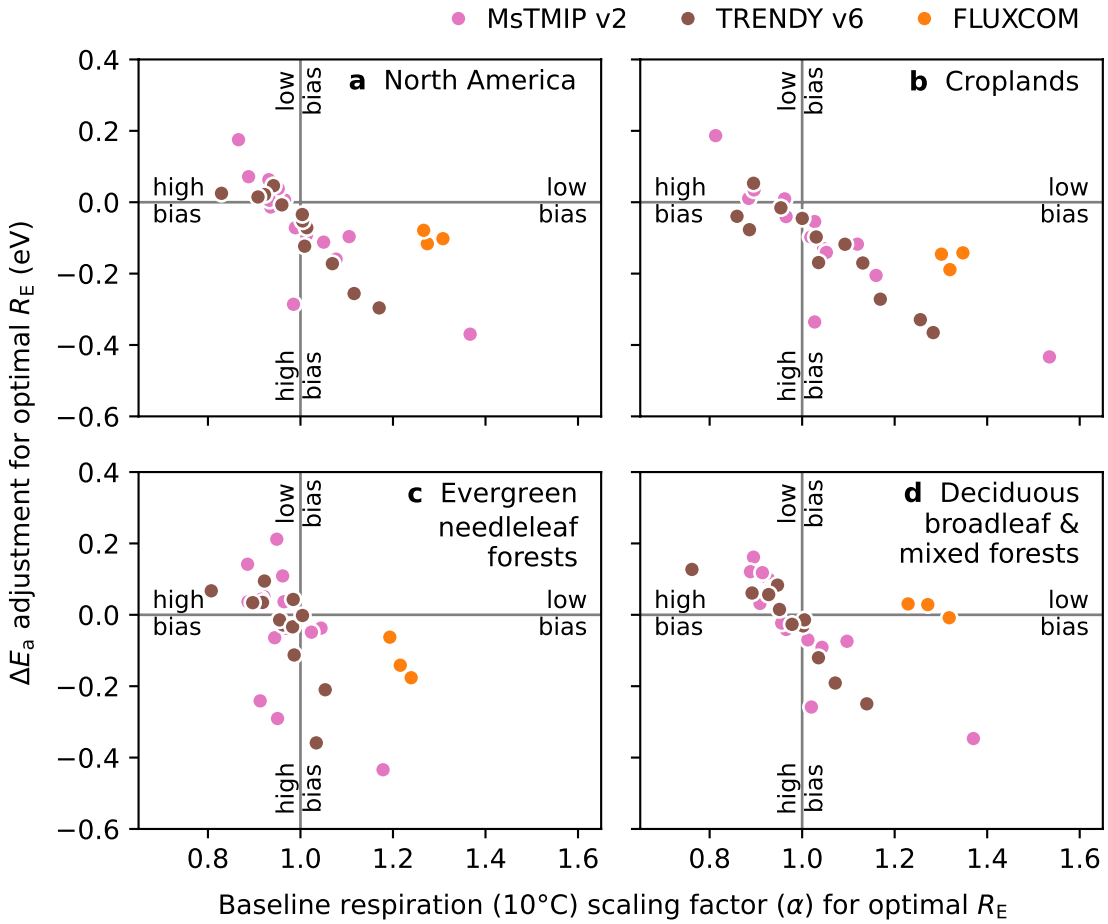
Extended Data Fig. 5 | The relationship between the apparent Q_{10} and the activation energy (E_a) of a reaction is temperature dependent. For a chemical reaction with a certain activation energy (E_a), Q_{10} becomes lower as the temperature increases. See Methods for a derivation of this relationship.

https://doi.org/10.1038/s41559-023-02093-x



Extended Data Fig. 6 | For some models, NEE estimates do not outperform the same models' GPP estimates in explaining observed atmospheric CO_2 variability. Fractions of observed CO_2 variability explained by NEE estimates (R^2_{NEE}) vs. those explained by GPP estimates (R^2_{GPP}) from simulations from the MsTMIP v2 (pink squares), TRENDY v6 (brown circles), and FLUXCOM (orange diamonds) model ensembles for 2007–2010. A model that falls below the one-to-one dashed line has GPP estimates that explain observed atmospheric CO_2 variability better than that same model's NEE estimates. Similar to Fig. 3, symbols that are half filled to the left (right) indicate models whose GPP (NEE) estimates explain a higher fraction of CO_2 variability than does incoming shortwave radiation ($R^2_{SW} = 0.23$; green dotted line). Data points that represent the same model participating in different ensembles (for example, LPJ-wsl in MsTMIP and TRENDY) are linked with thin lines.

https://doi.org/10.1038/s41559-023-02093-x



Extended Data Fig. 7 | Attribution of respiration bias to the temperature sensitivity and the baseline respiration rate based on maximal consistency with atmospheric CO_2 observations for (a) the North American domain, (b) croplands, (c) evergreen needleleaf forests, and (d) deciduous broadleaf and mixed forests. TBMs and the FLUXCOM models are colour-coded circles, as in the legend. Note that a negative ΔE_a adjustment means that the original E_a parameter of the model has a high bias, and a positive ΔE_a adjustment means a low bias in E_a. Similarly, for the baseline respiration adjustment, a ratio greater than one indicates a low bias in the original baseline respiration of the model, and a ratio smaller than one indicates a high bias in baseline respiration. For most TBMs, a low bias in E_a is compensated by a high bias in the baseline respiration (the 2nd quadrant), or a high bias in E_a is compensated by a low bias in the baseline respiration (the 4th quadrant). However, FLUXCOM models tend to show a low bias in the baseline respiration and a small to no bias in temperature sensitivity.

nature portfolio

Corresponding author(s): Wu Sun
Anna M. Michalak

Last updated by author(s): Apr 28, 2023

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the 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
☒	☐ A statement on 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 statistical parameters 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

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We did not use software in data collection. Publicly available data were provided by individual PIs.
Data analysis	Code used to generate figures and results is archived at https://doi.org/10.5281/zenodo.7874439. The following software environment is required to run the code: Python >= 3.8 (Python 2 not supported) jupyterlab >= 3.2.1 numpy >= 1.21.2 scipy >= 1.7.2 pandas >= 1.3.4 xarray >= 0.19.0 statsmodels >= 0.13.0 matplotlib >= 3.5.0 cartopy >= 0.18.0 seaborn >= 0.11.2

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Results presented in this study are available at https://doi.org/10.5281/zenodo.7874439. The ObsPack GLOBALVIEWplus CO2 data product is available at https://www.esrl.noaa.gov/gmd/ccgg/obspack. The CarbonTracker-Lagrange WRF-STILT footprints are available at https://www.esrl.noaa.gov/gmd/ccgg/carbontracker-lagrange/. The FFDAS v2 data product is available at https://ffdas.rc.nau.edu/. The North American Regional Reanalysis data can be obtained at https://psl.noaa.gov/data/gridded/data.narr.html. The MsTMIP v2 model ensemble is available at https://nacp.ornl.gov/. The TRENDY v6 model ensemble is available at https://sites.exeter.ac.uk/trendy/. The FLUXCOM model ensemble is available at http://www.fluxcom.org/.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

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

Study description	We use observations of atmospheric CO2 concentrations from a network of towers together with carbon flux estimates from state-of-the-art terrestrial biosphere models to characterize the large-scale temperature sensitivity of ecosystem respiration for various North American biomes.
Research sample	39,217 observations of atmospheric CO2 concentration from a network of 44 towers over North America during 2007–2010; 29 terrestrial biosphere model simulations from 24 independent models; 3 data-driven models for carbon flux estimates in the FLUXCOM ensemble
Sampling strategy	We used all available data after filtering and quality control.
Data collection	Atmospheric CO2 concentrations were collected by individual site PIs across North America. The WRF-STILT transport footprints were provided by the NOAA CarbonTracker-Lagrange Team. Terrestrial biosphere model and data-driven model simulations were provided by individual modelers.
Timing and spatial scale	The study domain covers the North American continent and individual biomes therein. Atmospheric CO2 concentration data were reported at three-hourly time scale from 44 sites across North America, spanning the period 2007–2010. The WRF-STILT transport footprints were at three-hourly time scale and 1° × 1° spatial resolution. Carbon flux estimates from terrestrial biosphere models were at the monthly time scale. All carbon flux estimates from terrestrial biosphere models were harmonized at 1° × 1° spatial resolution.
Data exclusions	We used all relevant data to derive estimates of temperature sensitivity for individual models. For the derivation of an overall temperature sensitivity across models, three terrestrial biosphere models were excluded due to low explanatory power of their estimates of gross primary productivity. This practice has been documented in the methods and reporting.
Reproducibility	Uncertainty ranges of estimates were reported. We have also released the code used to perform the analysis in a publicly accessible

Reproducibility	repository.
Randomization	Randomization is not relevant as we examined the effect over the entire study domain.
Blinding	Blind is not relevant because the work was not experimental in nature.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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