

Vegetation biogeography is a main source of uncertainty in modelling the land carbon cycle

Received: 20 February 2025

Accepted: 4 December 2025

Published online: 17 December 2025

 Check for updates

Ruiying Zhao¹✉, Xiangzhong Luo^{1,2}✉, Anthony P. Walker³,
Forrest M. Hoffman⁴ & Lian Pin Koh²

The terrestrial biosphere exchanges a large amount of CO₂ with the atmosphere through photosynthesis and respiration, determining the magnitude of land carbon sink and consequently influencing the rate of global warming. The magnitudes of global photosynthesis and respiration, however, vary widely across models (100–200 PgC/year), constituting a key and persistent source of uncertainty in carbon cycle and climate modelling. Here, we argue that the uncertainty in the land carbon cycle modelling is largely attributable to the uncertainty in biogeography – the distribution of plant functional types (PFTs). Using an ensemble of dynamic global vegetation models (DGVMs), we find a strong dependence of total photosynthesis on total area for each PFT. The dependence allows us to reduce the spread of land carbon cycle estimates by ~75% using remote sensing-based PFT maps. We further find that 56 ± 21% of climate-driven changes in global photosynthesis modelled by DGVMs are caused by changes in PFT distribution in the last two decades. Our study identifies vegetation biogeography as a main controlling factor of uncertainty in land carbon cycle modelling and highlights the importance of biogeography-climate interactions in carbon cycle and climate studies.

The land carbon cycle is a critical component of the global carbon budget¹, modulating the warming rate of the global climate by absorbing approximately 25% of anthropogenic CO₂ emissions^{2,3}. Photosynthesis (i.e., gross primary productivity; GPP) and ecosystem respiration (ER) are the two primary carbon fluxes in the land carbon cycle and determine the magnitude of the sink¹. However, global estimates of GPP and ER from models are highly uncertain, ranging from roughly 100 to 200 PgC/year^{4,5}. The considerable uncertainty in the land carbon cycle persists across several generations of dynamic global vegetation models (DGVMs) and Earth system models (ESMs), posing a long-standing challenge in carbon cycle simulations and climate predictions^{6,7}.

Considerable efforts have been made to identify the sources of uncertainty in models for land carbon cycle simulations, such as

assessing climate forcing datasets^{8–10} and improving representations of ecological processes (e.g., thermal acclimation of photosynthesis¹¹, explicit representation of microbial processes¹², plant hydraulics¹³, and competition for resources¹⁴), and parameterizations (e.g., photosynthetic capacity^{15,16} and microbial carbon use efficiency¹⁷) in process-based models. However, these adjustments have yet to significantly reduce the uncertainty in GPP and ER, as evidenced by the similar spread of land carbon simulations from DGVMs over the past decades^{1,18}. Meanwhile, remote sensing-based proxies¹⁹ and data-driven machine learning approaches^{20,21} have also been developed to provide estimates of GPP and ER. While these approaches are reported to estimate global GPP and ER from observations and have less uncertainty, they have limitations in identifying the mechanisms that

¹Department of Geography, National University of Singapore, Singapore, Singapore. ²Center for Nature-based Climate Solutions, National University of Singapore, Singapore, Singapore. ³Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA. ⁴Computational Science and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. ✉e-mail: ruiying@nus.edu.sg; xzluo.remi@nus.edu.sg

are missed or misrepresented in carbon cycle simulations. These limitations include prescribed climate sensitivity²², lack of representation in the tropics²⁰, empirical nature of the scaling factor between GPP and spectral signals²³, and lack of or inconsistent consideration of CO₂ fertilization effect^{24,25}.

Here, we hypothesize that the key source of uncertainty in land carbon cycle simulations is biogeography – the distribution of plant functional types (PFTs), based on multiple lines of evidence. Biogeography is known to affect the simulations of the land carbon cycle in DGVMs^{26–28}, causing up to a 24% difference in global GPP simulations within a single model^{26–28}. The large impact implies that the inter-model differences in estimated land carbon budget are potentially linked to the inter-model differences in biogeography simulations. In particular, a recent study²⁹ shows that the inter-model differences in C₄ vegetation distribution, a key PFT, explained over 80% of the uncertainties in GPP estimates for C₄ vegetation, providing direct evidence of the role of biogeography in land carbon cycle simulations. Additionally, the current framework for improving land carbon cycle simulations in DGVMs implies the importance of biogeography, given that changes in model parameters, structures, and even the emerging climate sensitivities of simulated carbon fluxes can also be driven by PFT change, as DGVMs largely regard these processes as PFT-specific^{26,30–32}.

To test the hypothesis on the dependence of the simulated land carbon cycle on biogeography, we use the outputs from an ensemble of state-of-the-art DGVMs to assess the relationships between total GPP and total area for each PFT. We aim to quantify the dependence of PFT total GPP on PFT distribution and quantify the contribution of vegetation biogeography to the uncertainties in global GPP and ER estimation (see Methods). The dependence will then be applied to narrow inter-model discrepancies in GPP and ER using remote sensing-based (RS) PFT maps. Using the PFT-specific relationships between total GPP and total area, and the simulated changes in PFT distributions from some DGVMs, we further explore the contribution of biogeography changes to the changes in global GPP over the past two decades.

Results

The relationships between PFT GPP and area

We first established the relationships between total GPP (i.e., in PgC/year) and total area (i.e., in km²) for each PFT, using the simulations from an ensemble of DGVMs (see Methods). For each of the twelve PFT types on the land surface, we observed a strong and significant relationship ($p < 0.001$) between GPP and area across models (Fig. 1). Using alternative approaches (i.e., linear regression without forcing through zero or multivariate regression), we consistently observed the strong dependence of total GPP on area for most of the PFTs (Fig. 1; Supplementary Fig. 10; Supplementary Table 5). The strong relationships indicate that the inter-model discrepancy of total GPP of each PFT is highly relevant to the difference in the prescribed or simulated PFT area across DGVMs, and that DGVMs simulate a similar per-area GPP on average for each PFT (i.e., the slopes of the GPP-area relationships in Fig. 1). Specifically, the average per-area GPP is 2.48 ± 0.07 PgC/year per 10⁶ km² for evergreen broadleaf forests (EBF), 1.38 ± 0.08 PgC/year for temperate evergreen needleleaf forests (ENFt), 0.94 ± 0.02 PgC/year for boreal evergreen needleleaf forests (ENFb), 1.22 ± 0.01 PgC/year for deciduous broadleaf forests (DBF), 1.10 ± 0.02 PgC/year for deciduous needleleaf forests (DNF), 0.66 ± 0.02 PgC/year for shrubs (SHR), 0.78 ± 0.06 PgC/year for warm C₃ grasses (C3Gw), 0.67 ± 0.03 PgC/year for cool C₃ grasses (C3Gk), 0.34 ± 0.04 PgC/year for cold C₃ grasses (C3Gc), 0.94 ± 0.10 PgC/year for C₃ crops (C3C), 1.40 ± 0.09 PgC/year for C₄ grasses (C4G), and 1.38 ± 0.04 PgC/year for C₄ crops (C4C).

Updated land carbon cycle estimation using remotely sensed PFT area

Since there is a large spread in PFT area across DGVMs, the uncertainty in PFT area propagates to the substantial uncertainty in global GPP

estimation (Fig. 1). The global GPP estimated by the ensemble of DGVMs ranged from 113.7 to 194.0 PgC/year, with a mean and a standard deviation (SD) of 146.32 ± 26.39 PgC/year (Fig. 2a). The global ER estimated by the DGVMs was 143.39 ± 25.99 PgC/year (Fig. 2b), and had the magnitude similar to GPP since global GPP and ER were strongly coupled across DGVMs (Fig. 2c). The tight relationships between PFT-specific total GPP and area across DGVMs (Fig. 1) allowed us to reduce the uncertainty in global GPP and then ER (Fig. 2c) if we have robust PFT distribution maps from observations.

We, therefore, used the PFT area from five widely-used RS PFT maps – including MODIS v6.1³³, ESA CCI v2.0.8³⁴, CGLS_LC100 v3³⁵, GLC_FCS30D³⁶, and GLASS-GLC³⁷ (see Methods) – in the GPP-area relationships (Fig. 1) to estimate the PFT-specific total GPP, and then summed up to global GPP. We found that the global GPP was updated to 135.66 ± 6.54 PgC/year on average from 2001 to 2019. The uncertainty in the global GPP estimate was reduced by approximately 75% (i.e., from 26.39 PgC/year to 6.54 PgC/year; Fig. 2a). The reduced uncertainty in GPP is mainly due to the reduced uncertainty in RS PFT mapping compared to those used in DGVMs. Specifically, the total GPP of PFTs are 41.35 ± 4.67 PgC/year for EBF, 6.76 ± 1.41 PgC/year for ENFt, 5.20 ± 0.92 PgC/year for ENFb, 14.95 ± 2.59 PgC/year for DBF, 5.27 ± 0.93 PgC/year for DNF, 8.21 ± 1.84 PgC/year for SHR, 7.51 ± 1.03 PgC/year for C3Gw, 5.80 ± 1.08 PgC/year for C3Gk, 1.42 ± 0.20 PgC/year for C3Gc, 10.55 ± 2.36 PgC/year for C3C, 24.07 ± 0.98 PgC/year for C4G, and 4.51 ± 0.15 PgC/year for C4C. The constrained global GPP was within the range of other GPP estimates such as those based on solar-induced fluorescence^{19,38} and soil respiration³⁹, but lower than those inferred from plant carbonyl sulfide⁴⁰ and higher than those upscaled from eddy covariance observations using machine learning^{20,21} (Fig. 2a). The constrained global GPP was further used to estimate global ER (Fig. 2c), with the uncertainty reduced from 25.99 to 6.53 PgC/year (Fig. 2b). Using linear regression not forced through zero, we still achieved a substantial yet smaller reduction (i.e., 48% compared to 75%) in the spread of inter-model GPP and ER (Supplementary Fig. 11).

Contribution of PFT change to global GPP change

We further used the relationships between PFT-specific total GPP and area to estimate the changes in global GPP induced by PFT area changes (Fig. 2d), using the annually adjusted PFT area from four DGVMs that enable dynamic biogeography simulations (Supplementary Table 1). From 2001 to 2019, these DGVMs demonstrated that elevated CO₂ increased global GPP by 23.47 ± 3.25 PgC/year and climate change increased global GPP by 2.05 ± 0.45 PgC/year. Of the global GPP increase induced by elevated CO₂, we found that approximately 20 ± 4% of the increase resulted from elevated CO₂-driven PFT changes, primarily due to a decrease in C₄ grass distribution and an increase in C₃ plant distribution simulated by these DGVMs (Supplementary Fig. 3). Similarly, of the global GPP increase induced by climate change, we found that 56 ± 21% resulted from climate change-driven PFT changes. Among these DGVMs, the climate change-driven PFT changes include the expansion of C₄ grasses, C₄ crops, and deciduous forests in high latitudes, in replacement of C₃ grasses and evergreen forests. The expansion of C₄ was likely due to its anatomical traits, which confer higher water-use efficiency^{41,42}.

Large uncertainty in biogeography simulations

We further examined the multi-year average PFT distribution simulated by each DGVM in the framework of Whittaker bioclimatic schemes⁴³ (Supplementary Fig. 4), where the distribution was presented in a space of mean annual temperature (MAT) and mean annual precipitation (MAP). While these biogeography simulations showed similar PFT types along the gradient of warm to cold (e.g., EBF to DBF, and to ENF), and wet to dry (e.g., EBF to SHR, and to C3G and C4G), there were clear differences in the bioclimatic boundaries of all PFTs between models, as well as between models and RS PFT maps. For

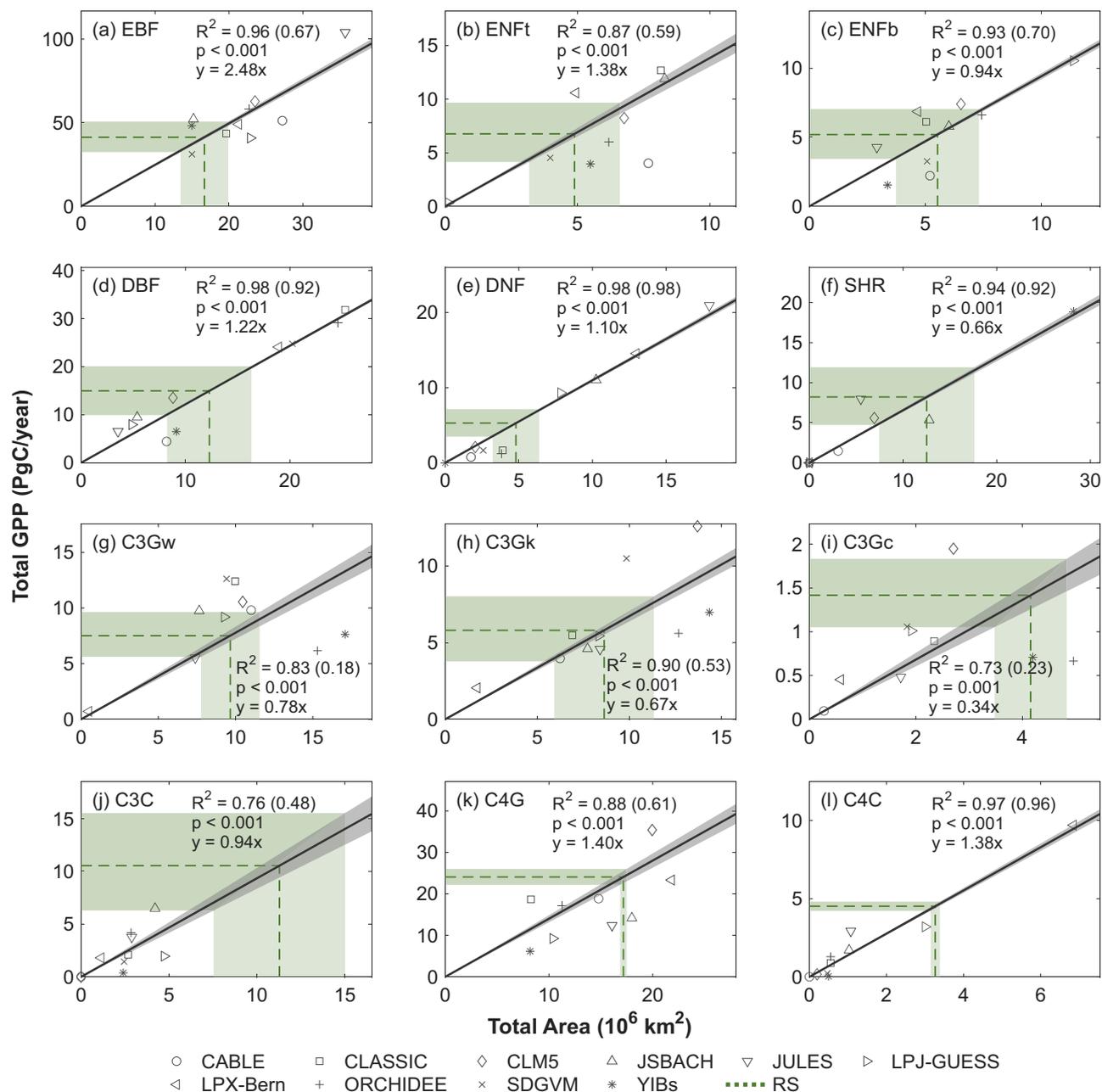


Fig. 1 | The relationships between total gross primary productivity (GPP) and total area for each PFT. The panels represent the relationships between GPP and area for **a** evergreen broadleaf forests (EBF), **b** temperate evergreen needleleaf forests (ENFt), **c** boreal evergreen needleleaf forests (ENFb), **d** deciduous broadleaf forests (DBF), **e** deciduous needleleaf forests (DNF), **f** shrubs (SHR), **g** warm C₃ grasses (C3Gw), **h** cool C₃ grasses (C3Gk), **i** cold C₃ grasses (C3Gc), **j** C₃ crops (C3C), **k** C₄ grasses (C4G), and **l** C₄ crops (C4C). The uncertainty of the relationship was quantified as the upper and lower boundary of the relationship through a bootstrapping method for the regression (see Methods). The regression lines were

forced through (0,0) as in theory there would be no GPP if PFT area was zero. The strength of the relationships was assessed by the coefficient of determination (R^2). Meanwhile, we also provided R^2 for the regression lines without forcing through (0,0) in parentheses. The vertical green lines stand for the average PFT areas from five remote sensing-based (RS) PFT maps, including MODIS v6.1³³, ESA CCI v2.0.8³⁴, CGLS_LC100 v3³⁵, GLC_FCS30D³⁶, and GLASS-GLC³⁷. The vertical shadings indicate the one standard deviation of the PFT area estimated by the five RS PFT maps. The horizontal green lines and shading intervals represent the constrained total GPP for each PFT using RS PFT areas.

example, the MAT range for C4G was narrow (20–30 °C) for CLASSIC and RS PFT maps, but much broader (–20–30 °C) for JULES. The MAP range for DBF was 0–2000 mm/year for JSBACH and 1000–2000 mm/year for RS PFT maps, but could reach 0–4000 mm/year for LPX_Bern and ORCHIDEE (Supplementary Fig. 4). Examining the spatial distribution of PFTs across models (Supplementary Fig. 5; Supplementary Fig. 6), we identified the hotspots for PFT disagreement between DGVMs. Those hotspots of biogeography uncertainty include EBF in the southern U.S. and Mexico, ENF in the boreal-temperate

transition zones, DNF in North America, C3G in the low latitudes, C4G in Central Asia, and DBF and SHR across the globe. Given the difference in the mechanisms each model uses to adjust PFT distribution, it is challenging to generalise the reasons for biases in PFT distribution across models.

Moreover, we found that RS PFT maps also show large differences in their estimates of PFT total area and spatial distributions (Fig. 3; Supplementary Table 3). For example, there were large spreads in the estimates of global EBF area (i.e., from 13.97 to 21.56 × 10⁶ km²) and C3C

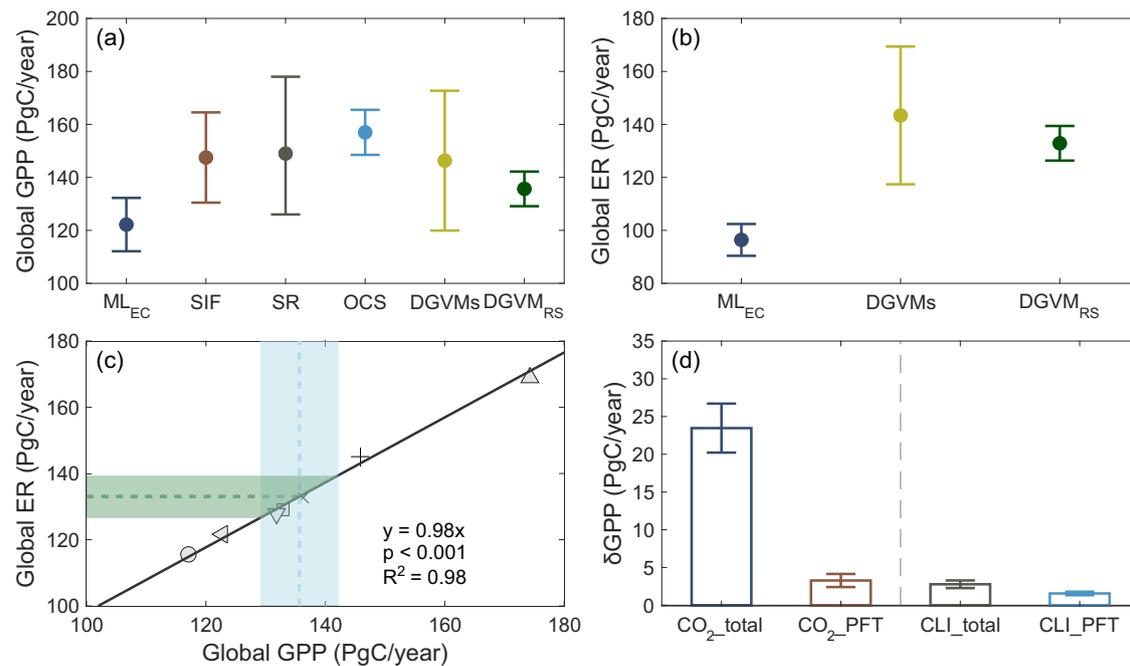


Fig. 2 | The constrained global gross primary productivity (GPP) and ecosystem respiration (ER). **a** The comparison of global GPP estimates from dynamic global vegetation models (DGVMs) from TRENDY v9 project, the constrained GPP using remote sensing-based PFT maps (DGVM_{RS}), and the GPP estimates from other approaches, including upscaled eddy covariance fluxes using machine learning^{20,21} (ML_{EC}), solar-induced fluorescence^{19,50,51} (SIF), soil respiration³⁹ (SR) and plant carbonyl sulfide⁴⁰ (OCS). The ML_{EC} includes the RS_METEO output of FLUXCOM – the ensemble estimates upscaled using different machine learning methods and

different meteorological forcings³⁰, as well as the X-BASE product from the FLUXCOM-X²¹. The solid lines represent the mean and the error bars represent one standard deviation of the GPP estimates. **b** The comparison of global ER estimates from DGVMs, DGVM_{RS}, and ML_{EC}. **c** The relationship between global ER and GPP across DGVMs. **d** The total GPP changes driven by elevated CO₂ (CO₂_total), the total GPP changes driven by CO₂-driven PFT changes (CO₂_PFT), the total GPP changes driven by climate change (CLI_total) and the total GPP changes driven by climate change-driven PFT changes (CLI_PFT).

area (from 7.66 to 16.34 × 10⁶ km²) across the five remote sensing products. The hotspots of biogeography disagreement among RS PFT maps include EBF in the eastern U.S. and western Europe, ENF in northern Canada and eastern Russia, DNF in North America and Europe, and SHR and C3G across the globe. The discrepancy across the RS PFT maps propagated into the uncertainty of our constrained global GPP (Fig. 1; Fig. 2).

Discussion

In this study, we demonstrated that biogeography is a main source of uncertainty in modelling the key fluxes in the land carbon cycle, including global photosynthesis and ecosystem respiration. Using the strong relationships between PFT-specific total GPP and area, and five RS PFT maps, we reduced the uncertainty in global GPP and ER estimates by approximately 75%. The changes in PFT distribution driven by climate change and elevated CO₂ have also contributed substantially to the changes in global GPP over the past decades, accounting for 20 ± 4% of the elevated CO₂-driven increase in global GPP and 56 ± 21% of the climate change-driven increase in global GPP.

Our study found that DGVMs consistently show large differences in biogeography simulations (i.e., PFT distribution), and the issue persists across simulation scenarios (Supplementary Fig. 9). The reasons for these differences are manifold, including the base map of land cover adopted from remote sensing, bioclimatic boundaries, and the considerations of fire, mortality, reproduction, and competition for resources^{44,45}. We noted that some models (e.g., JSBACH and YIBs) did not enable the dynamic vegetation component and adopted prescribed PFT maps in their simulations, i.e., do not simulate changes in biogeography over time. Additionally, when a remote sensing PFT map (e.g., ESA CCI) is used to prescribe PFTs in a model (e.g., ORCHIDEE), the uncertainty in the remote sensing map will inevitably propagate into the uncertainty of model simulations. Recent progress in

implementing demography⁴⁶, traits at the leaf and individual levels⁴⁷, fire⁴⁸, and land-atmosphere feedback⁴⁹ in some DGVMs can likely affect the simulations of biogeography. However, these processes have not been consistently implemented across DGVMs. Using a consistent biogeography baseline in DGVMs can be a first step to reduce PFT-induced uncertainty in land carbon cycle simulations.

Our study demonstrated that using PFT maps from remote sensing reduced the spread in modelled GPP by 75% in our statistical framework. However, this does not imply that biogeography alone accounts for 75% of the total reduced uncertainty in GPP modelling, since the effects of model structures and parameters on biogeography are often intertwined with their effects on biogeochemistry in dynamic vegetation models. For instance, PFT-specific photosynthetic parameters influence carbon sequestration rates (biogeochemistry)⁵⁰, which in turn affect the competition among PFTs (e.g., foliage area), driving shifts in PFT distribution (biogeography) and impacting parameterization again for carbon sequestration²⁶. What we presented is the combined role of biogeographic processes – both directly linked to PFT distribution (e.g., climate boundaries) and indirectly linked to PFT distribution (e.g., photosynthesis parameterizations) – in contributing to uncertainties in land carbon cycle modelling. This perspective does not downplay the importance of parameters and structures in driving uncertainties in land carbon cycle modelling. Instead, it highlights how these factors interact with biogeography to influence GPP simulations, offering insights into their interconnected impacts.

To properly assess the relative importance of biogeography and biogeochemistry on GPP modelling, we conducted a preliminary analysis where we used machine learning to develop emulators for DGVMs and then forced emulators with a consistent PFT map from MODIS. We found that by applying the MODIS PFT map to all DGVM emulators (i.e., EM(DGVMs, MODIS) in Supplementary Fig. 7), we reduced GPP

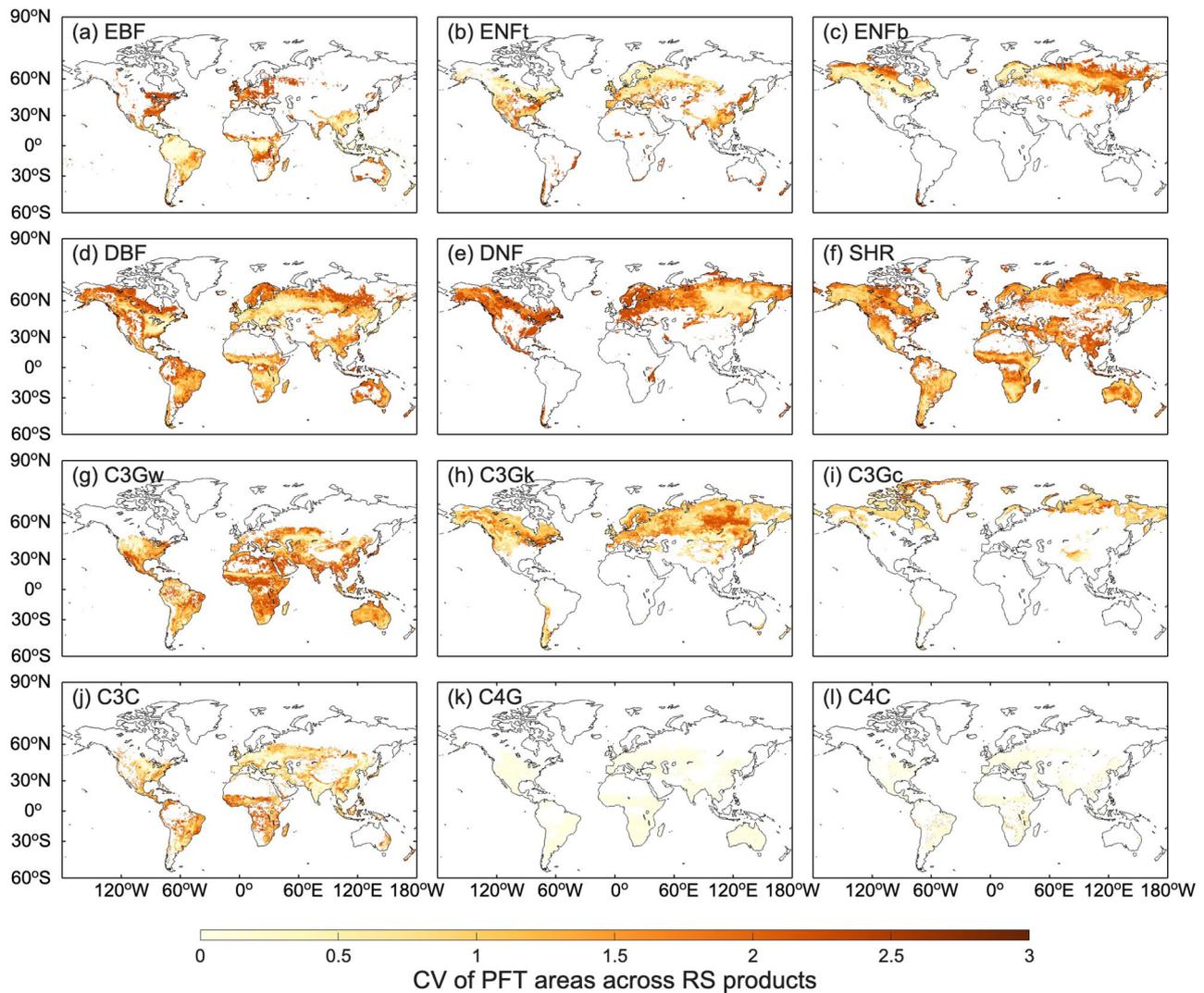


Fig. 3 | The difference in the distribution of PFTs from remote sensing-based (RS) products. The panels represent the coefficient of variation (CV) for **a** evergreen broadleaf forests (EBF), **b** temperate evergreen needleleaf forests (ENFt), **c** boreal evergreen needleleaf forests (ENFb), **d** deciduous broadleaf forests (DBF), **e** deciduous needleleaf forests (DNF), **f** shrubs (SHR), **g** warm C₃ grasses

(C3Gw), **h** cool C₃ grasses (C3Gk), **i** cold C₃ grasses (C3Gc), **j** C₃ crops (C3C), **k** C₄ grasses (C4G), and **l** C₄ crops (C4C) across RS products. The CV is calculated as the ratio of the standard deviation to the mean PFT area, with higher CV values indicating greater uncertainties in PFT simulations.

uncertainty by 46% from the original DGVMs (Supplementary Fig. 7), compared to a 75% reduction when using the relationship between PFT-specific total GPP and total area (i.e., DGVM_{RS}). This analysis sheds light on how DGVMs simulate PFT distribution – an accurate distribution of PFTs involves correctly modelling both the location and area of PFTs. In the DGVM_{RS}, we adjusted only the PFT area, achieving a 75% reduction in GPP uncertainty (Fig. 1; Fig. 2). In contrast, the emulator approach (EM(DGVMs, MODIS)) used both consistent PFT area and location, yet the uncertainty was reduced only by 46%, with the remaining uncertainty attributed to biogeochemistry processes. When PFT locations were consistent, uncertainty in GPP (i.e., EM(DGVMs, MODIS)) increased compared to DGVM_{RS}, implying model parameters and structures on biogeochemistry are inducing substantial changes in GPP estimates at the PFT locations. Overall, the EM(DGVMs, MODIS) approach perhaps provides a more realistic and conservative estimate of biogeography's role in GPP modelling, and indicates that improvements in biogeography in GPP modelling should be accompanied by improvements in biogeochemistry.

We acknowledge that variables like leaf area index (LAI) could also serve as observables in reducing the uncertainty of GPP estimates,

given GPP's strong dependence on LAI, and global LAI can be obtained through remote sensing. However, we used the relationship between GPP and area, not the relationship between GPP and LAI, for several reasons: biogeography and LAI are closely connected, as LAI simulations and associated parameters are likely PFT-specific (e.g., EBF typically has higher LAI than GRA); biogeography is more broadly linked to major carbon fluxes and stocks in the land carbon cycle (Supplementary Fig. 8b), while LAI primarily relates only to GPP. Lastly, we found the variation in PFT GPP is more closely linked to variation in PFT area than that in LAI or leaf photosynthetic rates (Supplementary Fig. 8a). Using other approaches (i.e., multivariate analysis and variance decomposition analysis), we reaffirmed that PFT area is more important than LAI and leaf photosynthetic rates in controlling the inter-model GPP variation for each PFT (Supplementary Table 5). This suggests that biogeography is likely a more important factor influencing the global GPP than LAI and leaf biogeochemistry.

This framework also helped explain the variability in the strength of GPP-area relationships among PFTs (i.e., R² and slope values in Fig. 1) – the relatively lower R² values, such as those for C₃ grasses and crops, indicate potentially greater inter-model differences in LAI or leaf

photosynthetic rates for C₃ grasses and crops compared to other PFTs. Previous studies indeed reported a slightly larger inter-model spread in the maximum carboxylation rate of photosynthesis for C₃ grasses than those for forest types⁵⁰ and a likely greater inter-model LAI discrepancy for grasses^{51,52}. Consistently, our statistical analysis shows that LAI differences in C₃ grasses contribute strongly to inter-model GPP variation (Supplementary Table 5), implying a considerable difference in grass LAI among models. However, without a comprehensive comparison of model protocols and structures, it remains difficult to pinpoint the specific biogeochemical processes or parameterizations responsible for the varying strength of the GPP–area relationship across PFTs. The greater slope in the GPP–area relationship for a given PFT is likely related to the greater leaf photosynthetic capacity or leaf area abundance prescribed for it in the models^{53,54}. Under our statistical framework, the future changes in carbon fluxes are driven by biogeographic changes (i.e., PFT area) and the changes in biogeography–carbon cycle relationships (e.g., PFT area–GPP relationships). While our results emphasize the role of PFT area in modulating future GPP, we acknowledge that the changes in the slopes of PFT area–GPP relationships might also cause future GPP changes. This is likely because the slopes essentially reflect impacts of PFT LAI, leaf photosynthetic rate, and other biogeochemical processes, which are known to change with climate and elevated CO₂^{55,56}.

Our study reported that climate change and elevated CO₂ caused changes in biogeography, which in turn affected the land carbon cycle. The climate change-driven PFT changes have been reported by some studies, mostly characterised by an increase in deciduous forests replacing evergreen forests in boreal zones⁵⁷. As for CO₂-driven PFT changes, the most significant changes include the decrease in C₄ vegetation distribution^{29,58} and the increase in C₃ grasses and woody plants⁵⁹. Elevated CO₂ can also reduce transpiration and alleviate the water demand for plants, indirectly influencing their distribution in arid environments in the future⁶⁰. Compared to the widely reported land cover changes driven by human activities, such as deforestation for crop expansion^{61,62}, the impacts of elevated CO₂ and climate change on PFT changes have been less examined, though our estimates on several DGVMs suggest they have substantial impacts on the land carbon cycle (Fig. 2d).

Our study provided a framework to quantify the magnitude of land carbon cycle using RS PFT maps, however, we must emphasize the considerable difference among RS PFT maps (Fig. 1; Fig. 3). A long line of research^{63,64} has indicated large disagreement between RS PFT maps, and those differences are still challenging to reconcile. The key difference could be relevant to the fraction of specific PFT types in the world (e.g., especially the EBF and C₃ crops). For example, MODIS suggests the EBF area is $21.56 \pm 1.23 \times 10^6$ km², but ESA CCI suggests it is only $14.23 \pm 2.85 \times 10^6$ km². The discrepancy in RS PFTs can propagate into the global GPP estimates, leading to an approximately 20 Pg C difference (Supplementary Fig. 2). Nevertheless, the uncertainty in RS PFTs (Fig. 3) is still smaller than models PFTs (Supplementary Fig. 5), making it feasible to use RS PFTs to constrain the land carbon cycle in DGVMs.

In addition to the inter-product difference in RS PFT mapping, we also note that for individual RS PFT maps, there was also considerable uncertainty incurred in its production process^{28,65} (e.g., uncertainty in the metrics used for classification, cross-walking from land cover to PFT). However, we were unable to consider the uncertainty associated with the production of each RS PFT map, as there was often no provision of such assessment or not all sources of uncertainty in the production process were reported^{66–68} (see Methods). Therefore, the uncertainty in RS PFT area we quantified, which refers to the inter-product difference in PFT area, was not comprehensive (Fig. 1). However, by examining the magnitude of production uncertainty for MODIS and ESA CCI PFT maps (Supplementary Table 3; Supplementary Fig. 2), we note that the inter-product difference in PFT area is

often greater than the production uncertainty of each RS PFT map. We also note that while recent advances in using traits to replace PFT maps for carbon cycle modelling^{56,69} are critical steps toward improving the modelling, they are not necessarily free from biogeography due to the need for PFT maps to produce accurate trait maps^{70,71}. To improve biogeography process and land carbon cycle simulations, we suggest prioritizing efforts to clear hotspots of biogeography uncertainty (e.g., those identified in Fig. 3) and to develop a more reliable global PFT map.

In conclusion, our study found that the uncertainty in modelling the land carbon cycle is largely associated with the uncertainty in global biogeography, and integrating remote sensing-based PFT maps can effectively reduce the uncertainty. Our findings highlight the crucial role of biogeography in contributing to the uncertainty in land carbon cycle simulation and potentially climate prediction, and indicate the urgency to advance biogeography studies to support carbon and climate sciences.

Methods

Dynamic global vegetation models (DGVMs)

To examine the relationship between the PFT total GPP and the corresponding PFT area, we used simulations from an ensemble of DGVMs participating in the TRENDY v9 project⁷². The DGVMs were coordinated to perform simulations under four different scenarios (S0, S1, S2, and S3), following the TRENDY protocol (<https://blogs.exeter.ac.uk/trendy/protocol/>). S0 is the control simulation, with no changes in the external forcings for DGVMs (i.e., time-invariant pre-industrial CO₂, climate and land use mask). In S1, only atmospheric CO₂ concentration varied when forcing DGVMs. In S2, atmospheric CO₂ and climate varied. In S3, atmospheric CO₂, climate and land use varied when forcing DGVMs.

We mainly used the simulations from the S2 scenario to examine the relationship between PFT-specific total GPP and area. Changing the simulations to other scenarios, such as S3, would not affect the relationship between PFT total GPP and PFT area across models (Supplementary Fig. 9), however, we would like to exclude the consideration of land use and land cover change in this analysis. A total of 11 models with varying spatial resolutions provide the required PFT-specific outputs (Supplementary Table 1). We obtained the annual total GPP and total area for each PFT for the period 2001 to 2019, and examined the relationships between them using the multi-year averages. Since not all models adopted the same PFT classification system, we harmonized the PFT classifications in each DGVM into twelve common PFT types, which cover the major functional differences in plant photosynthetic pathways, climate adaptation, and phenological cycles. These PFTs include evergreen broadleaf forests (EBF), temperate evergreen needleleaf forests (ENFt) and boreal evergreen needleleaf forests (ENFb), deciduous broadleaf forests (DBF), deciduous needleleaf forests (DNF), shrubs (SHR), warm C₃ grasses (C3Gw), cool C₃ grasses (C3Gk), and cold C₃ grasses (C3Gc), C₄ grasses (C4G), C₃ crops (C3C), and C₄ crops (C4C). The availability of PFT types in each DGVM is presented in Supplementary Table 1. For those DGVMs that only provide evergreen needleleaf forests (ENF) and C₃ grasses (C3G) without subtypes, we used a set of bioclimatic boundaries⁷³ to divide ENF into ENFt and ENFb, and to divide C3G into C3Gw, C3Gk, and C3Gc (Supplementary Table 2), as distinct climates otherwise might mask the relationship between PFT area and GPP (Supplementary Fig. 1). The climate data is obtained from CRU JRA v2.1⁷⁴, which is also the climate forcing data used for DGVMs in the TRENDY v9 project.

We also used PFT-specific GPP outputs from DGVM, which was used to establish PFT-specific relationship between total GPP and total area (Fig. 1). However, since DGVMs under TRENDY v9 did not supply PFT-specific ER, we were unable to establish PFT-specific relationship between total ER and area. We were only able to establish a relationship between global GPP and global ER in our analysis (Fig. 2c).

Remote sensing estimates and uncertainties of global PFT distribution

We adopted five widely-used remote sensing-based (RS) PFT products, including the MODIS³³, ESA CCI³⁴, CGLS_LC100³⁵, GLC_FCS30D³⁶, and GLASS-GLC³⁷, to constrain the GPP estimates based on the relationship between PFT-specific total GPP and total area.

The MODIS and ESA CCI datasets directly provide PFT maps (MCD12Q1 v6.1 and ESA CCI v2.0.8), with some assessments of their production uncertainties^{28,65}. Specifically, the MODIS PFT map was translated from land cover classification within the same product at a spatial resolution of 500 metres using a series of climate-based rules^{33,75}. ENFt and ENFb were classified from the ENF land cover type, with ENFt requiring a minimum temperature of the coldest month (T_c) above -19°C and an accumulated growing degree-days above 5°C (GDD_5) exceeding 1200°C , while ENFb had colder conditions⁷⁵. C_3 and C_4 grass were distinguished within grassland cover with $GDD_5 > 1000^\circ\text{C}$, $T_c > 22^\circ\text{C}$ and monthly precipitation $> 25\text{ mm}$ as C_4 grass, others as C_3 grass⁷⁵. C_3 grass was further classified to C3Gw, C3Gk, and C3Gc subtypes using DGVM rules (Supplementary Table 2). EBF, DBF, DNF, and SHR were mapped from their corresponding land cover types from MODIS³³.

The ESA CCI PFT map was translated from the ESA CCI land cover product at a 300-metre resolution from 1992 to 2020, with high-resolution auxiliary datasets such as tree cover and tree canopy height data^{34,76,77}. Pixels classified as tree types (e.g., EBF, DBF, ENF, and DNF) in the land cover maps were assigned to corresponding PFTs, with tree cover percentage determined using a 30-metre resolution tree cover dataset. For shrubland classes (a mix of tree and shrub woody vegetation types), a tree canopy height dataset was used to separate shrubs from trees^{34,75}.

The CGLS_LC100 provided the annual fraction of land cover classes, including EBF, ENF, DNF, DBF, SHR, mixed forest (MF), unknown forest (UNF), cropland and grassland, at a 100-m resolution from 2015 to 2019³⁵. The GLC_FCS30D provided the categorical information for the same classes at 30-m grid from 1985 to 2022, updated every five years before 2000 and annually thereafter³⁶. The GLASS-GLC provided annual distribution of forest (F), SHR, grassland and cropland at 5-km resolution from 1982 to 2015³⁷. After aggregating to 0.5-degree resolution for each land cover map, the MF fractions were translated to forest PFTs and SHR according to the translation rules (Supplementary Table 2). The UNF and F were translated to EBF, ENF, DBF, and DNF PFTs according to the same criteria. To make the PFTs map compatible with those from DGVMs, we further divided the RS PFTs using environment constraints⁷⁸ when needed (i.e., such as from ENF to ENFt and ENFb).

Since all the RS PFT maps lack explicit information on C_4/C_3 grass and C_4/C_3 crop classification, we used a newly developed C_4 vegetation map³⁰ to divide croplands and grasslands into their C_3 and C_4 counterparts. The C_4 vegetation map was generated using a global database of photosynthetic pathways, satellite observations, and a photosynthetic optimality theory. It contains C_4 grass and C_4 crop distribution over the globe at a 0.5-degree resolution from 2001 to 2019²⁹.

We quantified the uncertainty in RS PFT maps as the one standard deviation of PFT area among MODIS, ESA CCI, CGLS_LC100, GLC_FCS30D and GLASS-GLC maps. We did not consider the production uncertainty of individual RS PFT maps as there was a lack of assessments for most products (Supplementary Table 3)^{28,65}.

The relationship between PFT total GPP and PFT area

In this study, we identified the relationships between the simulated total GPP and total area for each PFT across DGVMs (i.e., GPP-area relationships). We obtained the multi-year average total GPP and total area for each PFT over the period 2001-2019, and examined the

relationships between them.

$$GPP_{PFT,i} = \beta_i \times area_{PFT,i} \quad (1)$$

where, β_i is the relationship between total GPP ($GPP_{PFT,i}$) and area ($area_{PFT,i}$) across DGVMs for i^{th} PFT. The i ranged between 1 to 12, representing EBF, ENFt, ENFb, DNF, DBF, SHR, C3Gw, C3Gk, C3Gc, C4G, C3C, and C4C, respectively. We forced the intercepts of the linear models to zero, based on the assumption that PFT-specific GPP should be zero if the PFT area is zero. Meanwhile, we also examined the relationships without this assumption. We estimated the uncertainties in β_i through a bootstrapping method^{73,79}, where one of the DGVMs was removed each time for regression to get the upper and lower boundary of the relationships.

Based on the GPP-area relationships and area of each PFT from RS PFT maps ($RSarea_{PFT,i}$), we estimate the constrained global GPP ($GPP_{global,C}$) over the globe.

$$GPP_{global,C} = \int_{i=1}^{12} \beta_i \times RSarea_{PFT,i} \quad (2)$$

We also identified a relationship between global ecosystem respiration (ER) and global GPP across DGVMs,

$$ER_{global} = \gamma \times GPP_{global} \quad (3)$$

where, γ is the slope, ER_{global} and GPP_{global} stand for the global ER and global GPP in each DGVM. The global ER in each model is calculated by the sum of the available autotrophic respiration and heterotrophic respiration. Combining γ and the constrained global GPP ($GPP_{global,C}$), we estimated the constrained global ER ($ER_{global,C}$).

$$ER_{global,C} = \gamma \times GPP_{global,C} \quad (4)$$

Machine learning emulators

To assess the relative importance of biogeography and biogeochemistry in GPP modelling, we developed a series of emulators for DGVMs by training random forest models globally for each PFT of each DGVM. Specifically, each emulator was trained to predict PFT GPP rate on grid cell level using annual climate forcings (temperature, precipitation, radiation, pressure, and wind speed) from the CRU JRA v2.1 datasets⁷⁴ from 2001 to 2019, which was consistent with the climate forcing data for DGVMs in the TRENDY v9 project. We excluded the LPJ-GUESS model from our emulator analysis, as in this model, the pixel PFT GPP rate is not available to support the training. We conducted the training process in MATLAB R2024b, with automatic tuning of key parameters, including the number of trees and the minimum leaf size. These emulators were independently validated via cross-validation, demonstrating consistently high predictive accuracy (mean R^2 values > 0.8 for all PFTs, Supplementary Table 4). After developing these emulators, we used them to estimate PFT-specific GPP rate within each grid cell and obtain PFT-specific GPP by using a consistent PFT fraction map from MODIS³³. We then aggregated the PFT-specific GPP of all grid cells and calculated global GPP.

The contribution of PFT changes to GPP changes

Using the relationships between PFT-specific GPP and area, we also estimated the contribution of elevated CO_2 - and climate change-driven PFT changes to the GPP changes. Specifically, we first obtained the simulated PFT area from DGVMs under S0 (time-invariant CO_2 and climate), S1 (varying CO_2 and time-invariant climate), and S2 (varying CO_2 and climate) scenarios. By subtracting the PFT areas of S0 from S1, we calculated the PFT area changes driven by elevated CO_2 . Similarly, by subtracting the PFT areas of S1 from S2,

we achieved the changes in PFT driven by climate change. Then we used the PFT area changes in the GPP-area relationship (Fig. 1) to infer the associated change in GPP. These PFT-driven changes in GPP were further compared against the total GPP changes driven by elevated CO₂ and climate change. Please note that in this analysis, we only used the four DGVMs (i.e., JULES, LPJ-GUESS, LPX-Bern, and SDGVM) that dynamically adjust PFT distribution over time (Supplementary Table 1).

Data availability

DGVM simulations from TRENDY v9 were obtained from Dr. S. Sitch (s.a.sitch@exeter.ac.uk) and are now available at <https://mdosullivan.github.io/GCB/>. The remote sensing-based PFT maps (MODIS, ESA CCI, GLASS-GLC, GLC_FCS30D and CGLS_LC100) are accessible at <https://e4ftl01.cr.usgs.gov/MOTA/MCD12Q1.061/> for MODIS; <https://esa-landcover-cci.org/> for ESA CCI; <https://doi.org/10.1594/PANGAEA.913496> for GLASS-GLC; <https://doi.org/10.5281/zenodo.8239305> for GLC_FCS30D and <https://doi.org/10.5281/zenodo.3939050> for CGLS_LC100. The GPP dataset upscaled from eddy covariance fluxes using machine learning is available at <https://doi.org/10.18160/SNZG-JMJE>. The GPP dataset estimated using SIF datasets are available at http://data.globalecology.unh.edu/data/GOSIF-GPP_v2/ and in the literature^{80,81}. The CRU JRA v2.1 climate data was downloaded from <https://catalogue.ceda.ac.uk/uuid/863a47a6d8414b6982e1396c69a9efe8>. The processed data are available at <https://zenodo.org/records/17718653>.

Code availability

The code used for analysis and figure generation is available at <https://zenodo.org/records/17718653>. Figure 3 in the study was generated with the assistance of the ‘M_Map’ package in MATLAB, whose instruction is available at <https://www.eoas.ubc.ca/~rich/map.html>.

References

- Friedlingstein, P. et al. Global carbon budget 2023. *Earth Syst. Sci. Data* **15**, 5301–5369 (2023).
- Cox, P. M. et al. Emergent constraints on carbon budgets as a function of global warming. *Nat. Commun.* **15**, 1885 (2024).
- Matthews, H. D., Gillett, N. P., Stott, P. A. & Zickfeld, K. The proportionality of global warming to cumulative carbon emissions. *Nature* **459**, 829–832 (2009).
- Anav, A. et al. Spatiotemporal patterns of terrestrial gross primary production: A review. *Rev. Geophys.* **53**, 785–818 (2015).
- Baldocchi, D., Ryu, Y. & Keenan, T. Terrestrial carbon cycle variability. *F1000Research* **5**, 2371 (2016).
- 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).
- Jones, C. D. & Friedlingstein, P. Quantifying process-level uncertainty contributions to TCRE and carbon budgets for meeting Paris Agreement climate targets. *Environ. Res. Lett.* **15**, 74019 (2020).
- Bonan, G. B. et al. Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). *Glob. Biogeochem. Cycle* **33**, 1310–1326 (2019).
- Arora, V. K., Seiler, C., Wang, L. & Kou-Giesbrecht, S. Towards an ensemble-based evaluation of land surface models in light of uncertain forcings and observations. *Biogeosciences* **20**, 1313–1355 (2023).
- Poulter, B., Frank, D. C., Hodson, E. L. & Zimmermann, N. E. Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO₂ airborne fraction. *Biogeosciences* **8**, 2027–2036 (2011).
- Knauer, J. et al. Higher global gross primary productivity under future climate with more advanced representations of photosynthesis. *Sci. Adv.* **9**, eadh9444 (2023).
- Shi, Z., Crowell, S., Luo, Y. & Moore, B. Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nat. Commun.* **9**, 2171 (2018).
- Sabot, M. E. B. et al. One stomatal model to rule them all? Toward improved representation of carbon and water exchange in global models. *J. Adv. Model. Earth Syst.* **14**, e2021MS002761 (2022).
- Dantas, D. E. et al. Including the phosphorus cycle into the LPJ-GUESS dynamic global vegetation model (v4.1, r10994) – global patterns and temporal trends of N and P primary production limitation. *Geosci. Model Dev.* **18**, 2249–2274 (2025).
- Walker, A. P. et al. The impact of alternative trait-scaling hypotheses for the maximum photosynthetic carboxylation rate (V_{cmax}) on global gross primary production. *N. Phytol.* **215**, 1370–1386 (2017).
- Zaehle, S., Sitch, S., Smith, B. & Hatterman, F. Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. *Glob. Biogeochem. Cycle* **19**, GB3020 (2005).
- Tao, F. et al. Convergence in simulating global soil organic carbon by structurally different models after data assimilation. *Glob. Change Biol.* **30**, e17297 (2024).
- Le Quéré, C. et al. The global carbon budget 1959–2011. *Earth Syst. Sci. Data* **5**, 165–185 (2013).
- Li, X. et al. Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: first global analysis based on OCO-2 and flux tower observations. *Glob. Change Biol.* **24**, 3990–4008 (2018).
- 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).
- Nelson, J. A. et al. X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework. *FLUXCOM-X. Biogeosci.* **21**, 5079–5115 (2024).
- Medlyn, B. E. Comment on “Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* **333**, 1093 (2011).
- Pierrat, Z. et al. Diurnal and seasonal dynamics of solar-induced chlorophyll fluorescence, vegetation indices, and gross primary productivity in the boreal forest. *J. Geophys. Res. Biogeosci.* **127**, e2021JG006588 (2022).
- De Kauwe, M. G., Keenan, T. F., Medlyn, B. E., Prentice, I. C. & Terrer, C. Satellite based estimates underestimate the effect of CO₂ fertilization on net primary productivity. *Nat. Clim. Chang.* **6**, 892–893 (2016).
- Keenan, T. F. et al. A constraint on historic growth in global photosynthesis due to rising CO₂. *Nat. Clim. Chang.* **13**, 1376–1381 (2023).
- Sitch, S. et al. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Glob. Change Biol.* **9**, 161–185 (2003).
- Poulter, B. et al. Plant functional type classification for earth system models: results from the European Space Agency’s Land Cover Climate Change Initiative. *Geosci. Model Dev.* **8**, 2315–2328 (2015).
- Hartley, A. J., MacBean, N., Georgievski, G. & Bontemps, S. Uncertainty in plant functional type distributions and its impact on land surface models. *Remote Sens. Environ.* **203**, 71–89 (2017).
- Luo, X. et al. Mapping the global distribution of C4 vegetation using observations and optimality theory. *Nat. Commun.* **15**, 1219 (2024).
- Krinner, G. et al. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Glob. Biogeochem. Cycle* **19**, GB1015 (2005).
- Oberpriller, J. et al. Climate and parameter sensitivity and induced uncertainties in carbon stock projections for European forests (using LPJ-GUESS 4.0). *Geosci. Model Dev.* **15**, 6495–6519 (2022).
- Piao, S. et al. Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO₂ trends. *Glob. Change Biol.* **19**, 2117–2132 (2013).

33. Sulla-Menashe, D. & Friedl, M. A. User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product. *USGS: Reston, VA, USA*, 1–18 (2018).
34. Harper, K. L. et al. A 29-year time series of annual 300 m resolution plant-functional-type maps for climate models. *Earth Syst. Sci. Data*. **15**, 1465–1499 (2023).
35. Tsendbazar, N. et al. Towards operational validation of annual global land cover maps. *Remote Sens. Environ.* **266**, 112686 (2020).
36. Zhang, X. et al. GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. *Earth Syst. Sci. Data*. **16**, 1353–1381 (2024).
37. Liu, H. et al. Annual dynamics of global land cover and its long-term changes from 1982 to 2015. *Earth Syst. Sci. Data*. **12**, 1217–1243 (2020).
38. MacBean, N. et al. Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Sci. Rep.* **8**, 1973 (2018).
39. Jian, J. et al. Historically inconsistent productivity and respiration fluxes in the global terrestrial carbon cycle. *Nat. Commun.* **13**, 1733 (2022).
40. Lai, J. et al. Terrestrial photosynthesis inferred from plant carbonyl sulfide uptake. *Nature* **634**, 855–861 (2024).
41. Way, D. A., Katul, G. G., Manzoni, S. & Vico, G. Increasing water use efficiency along the C3 to C4 evolutionary pathway: a stomatal optimization perspective. *J. Exp. Bot.* **65**, 3683–3693 (2014).
42. Eggels, S., Blankenagel, S., Schön, C. & Avramova, V. The carbon isotopic signature of C4 crops and its applicability in breeding for climate resilience. *Theor. Appl. Genet.* **134**, 1663–1675 (2021).
43. Whittaker, R. H. *Communities and ecosystems*. (Macmillan, New York, USA, 1970).
44. Haxeltine, A. & Prentice, I. C. BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. *Glob. Biogeochem. Cycle* **10**, 693–709 (1996).
45. 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).
46. Fisher, R. A. et al. Vegetation demographics in Earth System Models: A review of progress and priorities. *Glob. Change Biol.* **24**, 35–54 (2018).
47. Scheiter, S., Langan, L. & Higgins, S. I. Next-generation dynamic global vegetation models: learning from community ecology. *N. Phytol.* **198**, 957–969 (2013).
48. Hantson, S. et al. The status and challenge of global fire modelling. *Biogeosciences* **13**, 3359–3375 (2016).
49. Koppa, A. et al. Dryland self-expansion enabled by land–atmosphere feedbacks. *Science* **385**, 967–972 (2024).
50. Rogers, A. The use and misuse of Vc, max in Earth System Models. *Photosynth. Res.* **119**, 15–29 (2014).
51. Anav, A. et al. Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth system models. *J. Clim.* **26**, 6801–6843 (2013).
52. Song, X., Wang, D., Li, F. & Zeng, X. Evaluating the performance of CMIP6 Earth system models in simulating global vegetation structure and distribution. *Adv. Clim. Chang. Res.* **12**, 584–595 (2021).
53. Rogers, A. et al. A roadmap for improving the representation of photosynthesis in Earth system models. *N. Phytol.* **213**, 22–42 (2017).
54. Hu, Z. et al. Joint structural and physiological control on the inter-annual variation in productivity in a temperate grassland: A data-model comparison. *Glob. Change Biol.* **24**, 2965–2979 (2018).
55. Piao, S. et al. Characteristics, drivers and feedbacks of global greening. *Nat. Rev. Earth Environ.* **1**, 14–27 (2020).
56. Walker, A. P. et al. Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric CO₂. *N. Phytol.* **229**, 2413–2445 (2021).
57. Mekonnen, Z. A., Riley, W. J., Randerson, J. T., Grant, R. F. & Rogers, B. M. Expansion of high-latitude deciduous forests driven by interactions between climate warming and fire. *Nat. Plants* **5**, 952–958 (2019).
58. Lavergne, A., Harrison, S., Atsawawanunt, K., Dong, N. & Prentice, I. Recent C4 vegetation decline is imprinted in atmospheric carbon isotopes. (2024).
59. Stevens, N., Lehmann, C. E. R., Murphy, B. P. & Durigan, G. Savanna woody encroachment is widespread across three continents. *Glob. Change Biol.* **23**, 235–244 (2017).
60. Zhang, M. et al. Elevated CO₂ moderates the impact of climate change on future bamboo distribution in Madagascar. *Sci. Total Environ.* **810**, 152235 (2022).
61. Winkler, K., Fuchs, R., Rounsevell, M. & Herold, M. Global land use changes are four times greater than previously estimated. *Nat. Commun.* **12**, 2501 (2021).
62. Song, X. et al. Global land change from 1982 to 2016. *Nature* **560**, 639–643 (2018).
63. Fritz, S. et al. Highlighting continued uncertainty in global land cover maps for the user community. *Environ. Res. Lett.* **6**, 44005 (2011).
64. Jung, M., Henkel, K., Herold, M. & Churkina, G. Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sens. Environ.* **101**, 534–553 (2006).
65. Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P. & Friedl, M. A. Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. *Remote Sens. Environ.* **222**, 183–194 (2019).
66. Wang, L., Arora, V. K., Bartlett, P., Chan, E. & Curasi, S. R. Mapping of ESA’s Climate Change Initiative land cover data to plant functional types for use in the CLASSIC land model. *Biogeosciences* **20**, 2265–2282 (2023).
67. Xu, P. et al. Comparative validation of recent 10 m-resolution global land cover maps. *Remote Sens. Environ.* **311**, 114316 (2024).
68. Valle, D., Izbicki, R. & Leite, R. V. Quantifying uncertainty in land-use land-cover classification using conformal statistics. *Remote Sens. Environ.* **295**, 113682 (2023).
69. Anderegg, L. D. L. et al. Representing plant diversity in land models: An evolutionary approach to make “Functional Types” more functional. *Glob. Change Biol.* **28**, 2541–2554 (2022).
70. Dechant, B., Kattge, J. & Pavlick, R. Intercomparison of global foliar trait maps reveals fundamental differences and limitations of upscaling approaches. *Remote Sens. Environ.* **311**, 114276 (2023).
71. Butler, E. E. et al. Mapping local and global variability in plant trait distributions. *Proc. Natl. Acad. Sci.* **114**, (2017).
72. Friedlingstein, P. et al. Global carbon budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019).
73. Yu, K. et al. Field-based tree mortality constraint reduces estimates of model-projected forest carbon sinks. *Nat. Commun.* **13**, (2022).
74. Harris, I. C. CRU JRA v2.1: A forcings dataset of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data. *Centre for Environmental Data Analysis* <https://catalogue.ceda.ac.uk/uuid/10d2c73e5a7d46f4ada08b0a26302ef7> (2020).
75. Bonan, G. B., Levis, S., Kergoat, L. & Oleson, K. W. Landscapes as patches of plant functional types: an integrating concept for climate and ecosystem models. *Glob. Biogeochem. Cycle* **16**, 1021 (2002).
76. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

77. Potapov, P. et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **253**, 112165 (2021).
78. Prentice, I. C. et al. Special paper: a global biome model based on plant physiology and dominance, soil properties and climate. *J. Biogeogr.* **19**, 117–149 (1992).
79. Cox, P. M. Emergent constraints on climate-carbon cycle feedbacks. *Curr. Clim. Change Rep.* **5**, 275–281 (2019).
80. Joiner, J. et al. Estimation of terrestrial global gross primary production (GPP) with satellite data-driven models and eddy covariance flux data. *Remote Sens.* **10**, 1346 (2018).
81. Norton, A. J. et al. Estimating global gross primary productivity using chlorophyll fluorescence and a data assimilation system with the BETHY-SCOPE model. *Biogeosciences* **16**, 3069–3093 (2019).

Acknowledgements

R.Z. and X.L. are supported by a NUS Foresight Grant awarded to X.L. (A-8002864-00-00) and a Tier 2 Academic Research Fund from the Singapore Ministry of Education (MOE-T2EP50222-0006). L.P.K. is supported by the Singapore National Research Foundation and Agency for Science, Technology and Research LCER Phase 2 funding programme (U2208D4101). ORNL is managed by UT-Battelle, LLC, for the DOE under contract DE-AC05-1008 00OR22725. F.H. is supported by the RUBISCO Science Focus Area, which is sponsored by Earth and Environmental Systems Sciences Division (EESD) of the Office of Biological and Environmental Research (BER) in the US Department of Energy Office of Science.

Author contributions

X.L. conceived the idea. X.L. and R.Z. designed the study. R.Z. carried out the analysis and led the writing, with input from X.L. A.P.W., F.H., and L.K. provided comments at various stages of discussions. All authors contributed to the interpretation of the results and writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67636-1>.

Correspondence and requests for materials should be addressed to Ruiying Zhao or Xiangzhong Luo.

Peer review information *Nature Communications* thanks Andrew Hartley, Kailiang Yu and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025