Title

Harmonizing direct and indirect anthropogenic land carbon fluxes indicates a substantial missing sink in the global carbon budget since the early 20th century

Authors

Walker, Anthony P

Obermeier, Wolfgang A

Pongratz, Julia

et al.

Publication Date

2024

DOI

10.1002/ppp3.10619

Received: 15 May 2024 | Revised: 12 November 2024 | Accepted: 12 November 2024
DOI: 10.1002/ppp3.10619

Plants People Planet PPP Open Access

RESEARCH ARTICLE

Harmonizing direct and indirect anthropogenic land carbon fluxes indicates a substantial missing sink in the global carbon budget since the early 20th century

Anthony P. Walker[1] | Wolfgang A. Obermeier[2] | Julia Pongratz[2,3] | Pierre Friedlingstein[4,5] | Charles D. Koven[6] | Clemens Schwingshackl[2] | Stephen Sitch[4] | Michael O'Sullivan[4]

[1]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

[2]Department of Geography, Ludwig-Maximilians-Universität München, Munich, Germany

[3]Climate Dynamics Department, Max Planck Institute for Meteorology, Hamburg, Germany

[4]Faculty of Environment, Science and Economy, University of Exeter, Exeter, UK

[5]Laboratoire de Météorologie Dynamique, Institut Pierre-Simon Laplace, CNRS, École Normale Supérieure, Université PSL, Sorbonne Université, École Polytechnique, Paris, France

[6]Climate Sciences Department, Lawrence Berkeley National Lab, Berkeley, California, USA

Correspondence
Anthony P. Walker, Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA.
Email: walkerap@ornl.gov

Funding information
U.S. Department of Energy; Office of Science; Office of Biological and Environmental Research; European Space Agency, Grant/Award Number: 4000144908/24/I-LR

Societal Impact Statement

The global carbon budget provides annual updates to society on the main cause of climate change—CO_2 emissions—and quantifies carbon-uptake ecosystem services provisioned by the biosphere. We show that more consistent assumptions in the estimates of land-atmosphere carbon exchange results in a global carbon budget that is imbalanced (gains do not equal losses). This imbalance implies that key processes causing land carbon fluxes, especially processes associated with human land management and recovery following abandonment in anthropogenic biomes (anthromes), have been misquantified. This impacts policy for land carbon management across scales and calls for better understanding of carbon cycling in anthromes.

Summary

- Inconsistencies in the calculation of the two anthropogenic land flux terms of the global carbon cycle are investigated. The two terms—the direct anthropogenic flux (caused by direct human disturbance in anthromes, currently a carbon source to the atmosphere) and the indirect anthropogenic flux (caused indirectly by human activities that lead to global change and affecting all biomes, currently an atmospheric carbon sink)—are typically calculated independently, resulting in inconsistent underlying assumptions.
- We harmonize the estimation of the two anthropogenic land flux terms by incorporating previous estimates of these inconsistencies. We recalculate the global carbon budget (GCB) and apply change-point analysis to the cumulative budget imbalance.
- Cumulative over 1850–2018 (1959–2018), harmonization results in a 13% lesser (4% greater) land use source from anthromes and a 20% (23%) lesser land sink. This recalculation yields a greater non-closure of the GCB, indicating a missing carbon sink averaging 0.65 Pg C year^{-1} since the early 20th century. The imbalance likely results from a combination of method discontinuity and structural errors in

the assessment of the direct anthropogenic land use flux, greater ocean carbon uptake, structural errors in land models, and in how these land terms are quantified for the budget.

- We caution against overconfidence in considering the GCB a solved problem and recommend further study of methodological discontinuities in budget terms. We strongly recommend studies that quantify the direct and indirect anthropogenic land fluxes simultaneously to ensure consistency, with a deeper understanding of human disturbance and legacy effects in anthromes.

KEYWORDS

anthromes, bookkeeping models, carbon cycle, dynamic global vegetation models (DGVMs), global carbon budget, land cover and land use change emissions, natural carbon sink

1 | INTRODUCTION

Cumulative anthropogenic CO_2 emissions from fossil fuel combustion and land use change since the industrial revolution are approximately twice the amount of atmospheric CO_2 increase (Friedlingstein et al., 2023). Ocean (25%) and land (31%) CO_2 uptake account for the fate of those emissions not remaining in the atmosphere. Despite substantial uncertainty in the estimation of many of these terms, the global carbon budget (GCB) is mostly balanced (Friedlingstein et al., 2023), which may imbue confidence in all budget terms. The two land flux terms are of opposite sign (Canadell et al., 2007; Friedlingstein et al., 2023; Le Quéré et al., 2013)—net carbon flux from land use and land cover change (currently a source to the atmosphere, caused by direct human activities in anthropogenic biomes—anthromes) and the 'natural' land sink (caused indirectly by human activities via their effects on global change, with the potential to occur in all biomes). Since the 2017 GCB, these two land flux terms have been calculated independently (Le Quéré et al., 2018). However, these two terms cannot be independently verified with observations (Pongratz et al., 2021; Walker et al., 2021) and, while desirable, the independent calculation methods of these two terms use inconsistent assumptions (Dorgeist et al., 2024; Obermeier et al., 2021). Making the calculation of the two land flux terms more consistent (harmonizing) will (1) make estimates of the GCB more consistent, (2) better align these estimates with the development of anthromes over the industrial period, and (3) re-evaluate our confidence in the land flux terms of the GCB.

In this manuscript, we use terminology that is consistent with the Intergovernmental Panel on Climate Change (IPCC) reports on land fluxes (IPCC, 2010)—direct and indirect anthropogenic land fluxes. The GCB uses the terms land use change emissions and terrestrial or 'natural' land sink. Here we want to emphasize that these two land flux terms are "two sides of the same coin," both are caused by human actions (and are thus anthropogenic), and that their observation and estimation are inextricably linked. The commonly termed "natural" land sink (indirect anthropogenic land flux) is natural in that natural processes are affected indirectly by human actions that cause global change, such as effects from rising atmospheric CO_2 concentration, N deposition, and climate change (IPCC, 2010), albeit in addition to natural climate variability. In addition, there exists a truly natural land sink driven by natural processes, and which is important over longer timescales, for example, peatland development and natural climate change over the early-to-mid Holocene (Stocker et al., 2017). The notion of direct and indirect anthropogenic carbon fluxes may also become more important (potentially for the oceans as well) as we move towards a more intentionally managed carbon cycle and the assessment of country-specific or collective climate efforts.

In the GCB, the direct anthropogenic land flux (also known as E_{LUC}) is calculated using three bookkeeping (BK) models (Gasser et al., 2020; Hansis et al., 2015; Houghton & Castanho, 2023). BK models are simple carbon-cycle models that estimate gross and net E_{LUC} based on a set of land cover classes, and a time series of land use and land cover maps that specifies the extent of natural and anthropogenic land cover, land cover transitions, and some classes of land management (such as wood harvest). The land cover classes in two of the three BK models are associated with invariable equilibrium carbon area densities (hereafter, simply carbon densities). Turnover and recovery times are used to calculate the temporal evolution of land carbon following a land cover or land use transition.

The indirect anthropogenic land flux (also known as S_{LAND}) is calculated using Dynamic Global Vegetation Models (DGVMs) (Sitch et al., 2015, 2024). DGVMs are process-based carbon cycle and land surface models with various levels of mechanistic detail in process representation driven by time-varying climate, atmospheric CO_2, and nitrogen deposition data. Land cover is represented by plant functional types in various proportions on a given grid cell. Anthropogenic land cover in the simulations used to estimate S_{LAND} is typically static and uses the land cover in the year 1700 throughout a simulation.

The assumption of 1700 land cover in the DGVMs is clearly inconsistent with the assumption of time-varying land use and land cover in BK models. Given the declining proportion of forest from 1700 to present (Figure 1a–d) and that DGVMs predict the strongest indirect anthropogenic land flux in forests (Friedlingstein et al., 2023), the assumption of static 1700 land cover over time leads to an

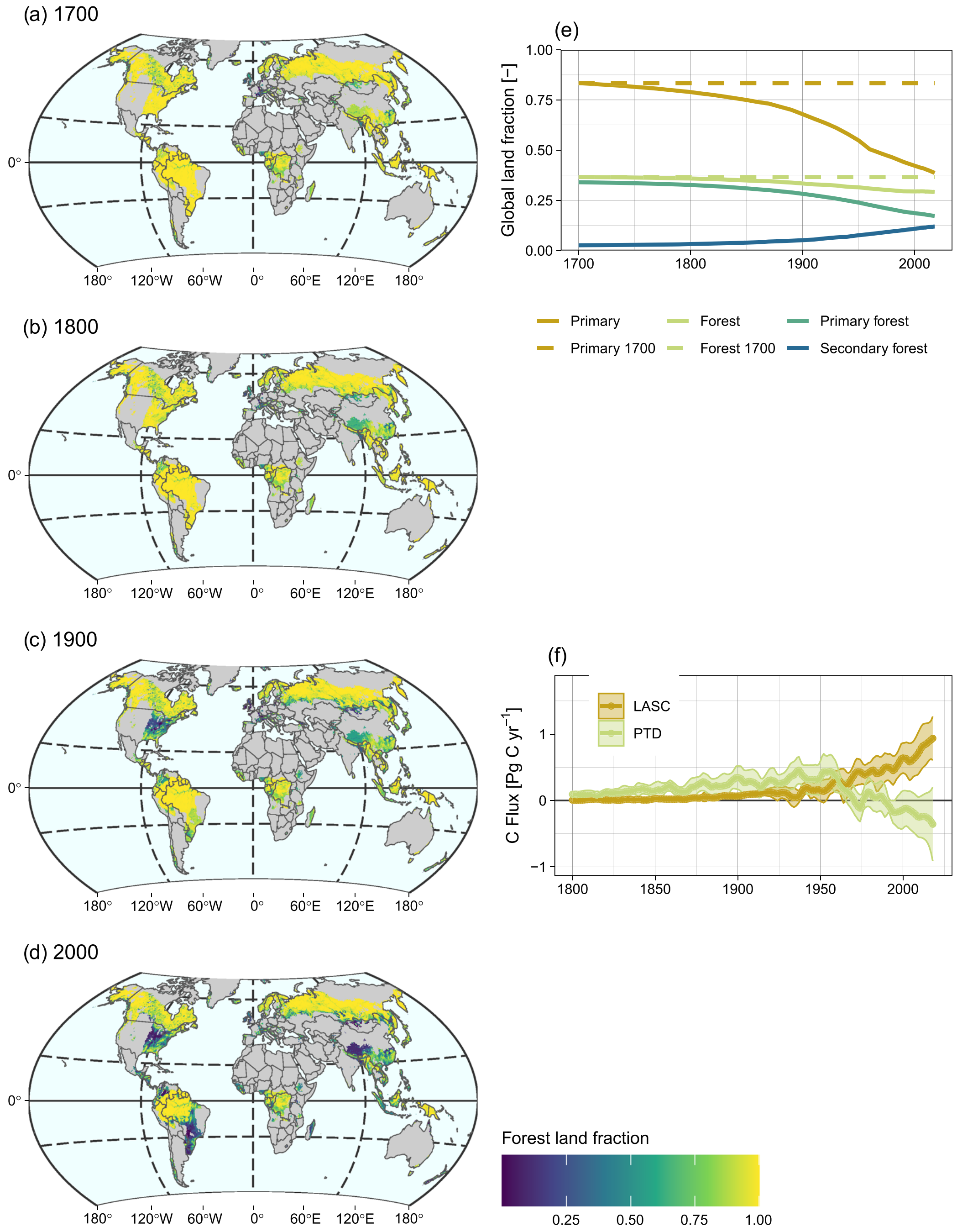

FIGURE 1 Forest cover (primary and secondary) from the LUH2 dataset (Hurtt et al., 2020) used in this study and the 2019 global carbon budget in (a) 1700, the start date of TRENDY simulations, (b) 1800, (c) 1900, (d) 2000, and (e) time series of global proportions of primary land, forest, primary forest, and secondary forest with the proportions of primary land and forest land in 1700 (dashed lines) also shown over the whole time period for comparison. (f) 5-year moving window average LASC and PTD (model mean ± SD) from Obermeier et al. (2021). Non-forest areas in Figure 1a–d shown in gray. LASC, loss of additional sink capacity; LUH2, Land Use Harmonization 2; PTD, present versus transient difference; SD, standard deviation.

overestimate of the indirect flux. DGVMs predict that the indirect land flux has historically been a sink, primarily because of the effects of increasing atmospheric CO_2. Thus, the clearing of forests represents a loss in the potential carbon sink capacity associated with the indirect anthropogenic land fluxes. In the literature, this is typically referred to as the loss of additional sink capacity (LASC), which is not accounted for in current DGVM simulations for the GCB (Dorgeist et al., 2024, Obermeier et al., 2021, Pongratz et al., 2014), albeit this omission is known and acknowledged by the GCB (Friedlingstein et al., 2023). The LASC could also be thought of as a component of the direct human land flux term, as the LASC arises through a combination of both direct and indirect anthropogenic influence on land ecosystems.

Nevertheless, two of the three BK models commonly used to calculate the direct anthropogenic land flux assume that the equilibrium ecosystem carbon density for a given land cover class is constant through time (Hansis et al., 2015; Houghton & Castanho, 2023). However, driven by time-varying climate and environment, the DGVMs predict that equilibrium carbon area densities are not static and are, for the most part and particularly in forests, net increasing with environmental changes (Friedlingstein et al., 2023). The carbon density parameters in BK models are typically based on data from "the present," albeit the most common values used are from Houghton et al. (1983), which were taken from data published in the 1970's (Schlesinger, 1977; Whittaker & Likens, 1973). Static carbon densities through time mean that the BK models are inconsistent with the DGVMs as the DGVMs predict lower equilibrium carbon area densities in previous decades and centuries. This inconsistency is termed the present versus transient difference (PTD) and leads to BK models calculating higher fluxes in the past than DGVMs would if attempting to simulate the direct land flux (Dorgeist et al., 2024; Obermeier et al., 2021).

In this study, we aim to harmonize the anthropogenic land flux terms of the GCB using pre-existing DGVM and BK model data. We adjust the direct and indirect anthropogenic land flux terms to incorporate the LASC and PTD. To evaluate the consequences of this land flux harmonization for our understanding of the GCB, we recalculate the GCB and analyze the budget imbalance (BIM, the budget residual, i.e., when source and sink terms in the budget do not sum to zero).

The BIM became a component of the GCB in 2017 when the indirect anthropogenic land flux was first calculated from DGVMs (Le Quéré et al., 2018). Prior to that, the indirect land flux was calculated simply as the residual of the other terms in the GCB, and so the method determined that the budget was perfectly balanced. However, several of the GCB terms are highly uncertain and thus errors in the terms are expected to lead to non-zero budget residuals (imbalances). If these residuals can be represented by a stationary statistical distribution with mean zero, it could be inferred they are caused by random variability ("aleatory" uncertainty) (Beven, 2016) without further recourse to identifying specific causes and solutions beyond general uncertainty reduction. However, should these residuals exhibit structure, we might infer causes related to the knowledge used in the budget calculations ("epistemic" uncertainty) (Beven, 2016). For example, the existing BIM shows a consistent sign, roughly over the period 1930–1960 (Friedlingstein et al., 2023). This period is associated with a transition in the land use and land cover dataset from estimates derived from population and per capita land use data (pre-1960) to FAO statistics of crop and pasture data (1960 to present) (Klein Goldewijk et al., 2017). In this study, we apply change-point analysis to the cumulative BIM to examine the structure in these residuals and whether this structure might indicate uncertainty in our knowledge related to these anthropogenic land flux terms.

2 | MATERIALS AND METHODS

The GCB is published on an annual basis by the Global Carbon Project (Friedlingstein et al., 2023). Three BK models are used to calculate the direct anthropogenic land flux (i.e., emissions from land use change, E_{LUC}) (Hansis et al., 2015; Gasser et al., 2020; Houghton & Castanho, 2023), and the TRENDY DGVM model ensemble is used to calculate the indirect anthropogenic land flux ('natural' land sink, S_{LAND}). As described earlier, the standard calculations of these land flux terms in the GCB are not consistent. The flux from DGVMs is derived under transient environmental conditions, and thus, ecosystem carbon densities evolve in response to changing environmental conditions, particularly in forested areas. This contrasts with BK models, which use static "present-day" (1970s) carbon densities. In addition, DGVMs use a static 1700 land cover map while BK models use time-evolving land cover.

Here we outline the TRENDY simulation ensemble and the method for calculating the LASC and present versus transient difference (PTD). For further details, see Obermeier et al. (2021), Friedlingstein et al. (2019), and Sitch et al. (2024). The standard TRENDY ensemble consists of four simulations (S0–S3) that sequentially add factors of environmental variation over the industrial period. S0 is a baseline simulation with time-invariant pre-industrial CO_2, climate and land use, S1 adds time-varying CO_2 and nitrogen deposition (hereafter, implicit when CO_2 is mentioned in methods) to S0, S2 adds climate change to S1, and S3 adds land use change to S2 (Table 1).

The indirect land flux (S_{LAND}) is calculated by subtracting the S0 net biome production (NBP_{S0}, which has a mean of zero) time series from NBP in S2 (NBP_{S2}):

TABLE 1 Description of dynamic global vegetation model simulations used in the global carbon budget to estimate the indirect anthropogenic land flux.

Simulation	CO_2 and N deposition	Climate	Land use
S0	Pre-industrial	Pre-industrial	Pre-industrial
S1	Time-varying	Pre-industrial	Pre-industrial
S2	Time-varying	Time-varying	Pre-industrial
S3	Time-varying	Time-varying	Time-varying
S4	Pre-industrial	Pre-industrial	Time-varying
S5	Present	Present	Time-varying
S6	Present	Present	Pre-industrial

$$S_{LAND} = NBP_{S2} - NBP_{S0} \quad \text{(Eq. 1)}$$

To calculate the LASC and PTD used in this study (Obermeier et al., 2024), three additional simulations were run (S4–S6) in TRENDY version 8 used in the 2019 GCB (Friedlingstein et al., 2019; Obermeier et al., 2021). S4 runs with pre-industrial CO_2 and climate and time-varying land use change, S5 runs with present-day CO_2 (2018 value) and climate (repeating 1999–2018 values) and time-varying land use change, and S6 runs with present-day CO_2 and climate and static pre-industrial land cover (Table 1).

LASC was calculated by subtracting the land use change flux estimated with pre-industrial environmental conditions ($NBP_{S0} - NBP_{S4}$) from the land use change flux estimated under time-varying environmental conditions ($NBP_{S2} - NBP_{S3}$) (see Equation 4 in Obermeier et al., 2021):

$$LASC = (NBP_{S2} - NBP_{S3}) - (NBP_{S0} - NBP_{S4}) \quad \text{(Eq. 2)}$$

PTD was calculated by subtracting the land use change flux estimated under time-varying environmental conditions ($NBP_{S2} - NBP_{S3}$) from the land use change flux estimated with present-day environmental conditions ($NBP_{S6} - NBP_{S5}$) (see Equation 5 in Obermeier et al., 2021):

$$PTD = (NBP_{S6} - NBP_{S5}) - (NBP_{S2} - NBP_{S3}) \quad \text{(Eq. 3)}$$

The GCB data from Friedlingstein et al. (2023) are used to calculate the BIM:

$$BIM = E_{FF} + E_{LUC} - (S_{LAND} + S_{OCEAN} + G_{ATM}) \quad \text{(Eq. 4)}$$

where, E_{FF} are fossil fuel emissions (including those from cement production and the cement carbonate sink), S_{OCEAN} is the ocean sink, and G_{ATM} is the atmospheric growth rate. All units are in Pg carbon per year. We use the budget estimates of the GCB 2019 despite later updates, because only for that budget were the additional TRENDY S4–S6 simulations run. The harmonized budget imbalance (BIM_H) accounts for the LASC from S_{LAND} by defining the indirect anthropogenic land flux as (S_{LAND} − LASC) and accounts for the present versus transient difference in land use change emissions because of altered ecosystem carbon densities by defining the direct land flux as ($E_{LUC} - PTD$):

$$BIM_H = E_{FF} + (E_{LUC} - PTD) - (S_{LAND} - LASC + S_{OCEAN} + G_{ATM}) \quad \text{(Eq. 5)}$$

All the GCB calculations in this study are from the 2019 GCB (Friedlingstein et al., 2023) and the associated TRENDY version 8. Uncertainty in the BIM was calculated by standard error propagation for arithmetic sums, that is, the sum of errors in quadrature. This results in large uncertainty in the BIM given the large component fluxes and their associated errors, which may overestimate uncertainty in the BIM given a perfect BIM would be equal to zero.

Land cover data in Figure 1 are the Land Use Harmonization 2 (LUH2) dataset (v2h) (Hurtt et al., 2019, 2020) (https://luh.umd.edu/) produced with data up to and including 2018. To calculate forest cover in this dataset, we summed the primary forest ("primf") and secondary forest ("secdf") variables in each grid cell.

Change-point analysis was applied to the cumulative time series of the standard budget imbalance (as derived in the GCBs, Equation 1) and the harmonized budget imbalance (Equation 2). In order to not presuppose a particular number of change points on the time series, a set of candidate models was proposed based on visual inspection of the curves (different sets of candidate models were used for the two cumulative time series). Models were fitted to the data using the default parameters of the "mcp" function in the "mcp" package (Lindeløv, 2020). Model validation and estimate of fit used post hoc, approximate leave-one-out cross-validation from the "loo" package, which uses Pareto smoothed importance sampling to approximate true leave-one-out cross-validation (Vehtari et al., 2016, 2023). Model selection was based on expected log pointwise predictive density (EPLD) using the "loo_compare" function of the "loo" package, with the highest EPLD indicating better predictive accuracy compared to the others. The pareto K estimate was used as a diagnostic for the validity of the post hoc leave-one-out cross-validation. Final model selection was based on the EPLD, pareto K, and visual inspection.

All data analysis and visualization was conducted in R v4.3.2 (R Core Team, 2023) and commonly using "tidyverse" packages (Wickham et al., 2019). Figures were produced using color scales from the "viridis" (Garnier et al., 2023) and "MoMAColors" (Mills 2023) packages. Change point analysis was conducted using Bayesian Markov Chain Monte Carlo (MCMC) with the package "mcp" (Lindeløv, 2020) which uses JAGS (Just Another Gibbs Sampler, https://sourceforge.net/projects/mcmc-jags/) tools via the "rjags" package (Plummer 2023). Where necessary, error propagation was based on standard arithmetic and product error propagation with 95% credible intervals assumed as 95% confidence intervals from normal distribution.

3 | RESULTS

Harmonizing the direct and indirect anthropogenic land flux terms by subtracting PTD from the direct flux and subtracting the LASC from the indirect flux substantially modifies those fluxes. The direct anthropogenic land flux over the most recent decade of the time series, 2009–2018, increases by 18% from 1.51 ± 0.10 (mean ± SD) to 1.78 ± 0.11 Pg C year^{-1} (Table 2), while the indirect anthropogenic land flux decreases by 27% from 3.15 ± 0.57 to 2.31 ± 0.57 Pg C year^{-1}.

Over the full time series, 1850–2018, harmonization of the land fluxes reduces both the cumulative direct estimate by 13% from 203 ± 60 to 177 ± 65 Pg C (mean ± cross-model SD; Table 2, uncertainties from Friedlingstein et al., 2019) and the cumulative indirect estimate by 20% from 197 ± 40 to 157 ± 43 Pg C. Over the modern instrument record, 1959–2018, harmonization of the land fluxes slightly increases the cumulative direct flux estimate by 4% from 80 ± 40 to

TABLE 2 Anthropogenic land fluxes and budget imbalance (BIM) terms for the standard and harmonized budgets.

	Cumulative fluxes				Annual fluxes			
	Pg C				Pg C $year^{-1}$			
	1850–2018		1959–2018		2009–2018		1999–2018	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
E_{LUC}	203	60	80	40	1.51	0.11	1.43	0.14
S_{LAND}	197	40	130	25	3.15	0.60	2.95	0.75
LASC	40	15	31	11	0.84	0.07	0.73	0.12
PTD	27	25	-3	12	−0.27	0.04	−0.21	0.07
S_{LAND} − LASC	157	43	100	27	2.31	0.60	2.22	0.72
E_{LUC} − PTD	177	65	83	42	1.78	0.12	1.65	0.18
BIM	31	78	10	55	0.43	0.55	0.28	0.48
BIM_H (LASC and PTD harmonized)	44	83	44	58	1.53	0.52	1.22	0.54
BIM LASC harmonized	71	79	41	56	1.26	0.52	1.01	0.51
BIM PTD harmonized	5	82	13	56	0.70	0.54	0.49	0.50

Note: Direct and indirect anthropogenic land fluxes calculated in the global carbon budget (E_{LUC} and S_{LAND}, respectively) and the loss of additional sink capacity (LASC) and present transient difference (PTD).

83 ± 42 Pg C and reduces the cumulative indirect estimate by 23% from 130 ± 25 to 100 ± 27 Pg C (Table 2).

Accounting for the LASC and PTD increases, the annual mean budget imbalance, 2009–2018, from 0.43 ± 0.52 to 1.53 ± 0.50 Pg C $year^{-1}$ (over 1999–2018 from 0.28 ± 0.46 to 1.22 ± 0.53 Pg C $year^{-1}$). High interannual variability in the budget imbalance (0.7 Pg C $year^{-1}$ standard deviation in the standard budget over the full timeseries) makes it difficult to discern non-stationarity in these budget residuals and obscures differences among the standard and harmonized budget imbalances (Figure 2a). Analysis of the cumulative imbalances provides a clearer picture of structure in these residuals (Figure 2b).

Over the full time series, harmonization of land fluxes increases the cumulative budget imbalance by 42% from 31.2 to 44.2 Pg C. The harmonized cumulative budget imbalance starts to deviate strongly from the standard budget imbalance in 1875. Similar to the standard budget imbalance, the harmonized budget imbalance starts to increase in the late 1920s, albeit from pronounced lower values. From about 1975, once again the harmonized budget imbalance breaks from the trend of the standard imbalance with a continued increase while the standard imbalance levels off.

By harmonizing the budget with just the LASC or the PTD separately, it can be seen that the initial decline of the harmonized imbalance relative to the standard imbalance is mainly caused by the PTD. In contrast, the second main deviation of the harmonized imbalance relative to the standard imbalance, the sustained post-1975 increase, is predominantly caused by the LASC.

To more rigorously and quantitatively compare the standard and the harmonized budgets, we apply change-point analysis to the trends in the cumulative budget imbalances. For the standard budget (Figure 2c), change-point model selection indicates that the budget was balanced for three decades from 1850 to 1880s (1875–1883, 95% CI), followed by two decades of imbalance until the end of the 19th century (1892–1902) at a mean rate of −0.46 (−0.28 to −0.65) Pg C $year^{-1}$, then was balanced through 1930s (1933–1935), then did not balance for three decades to the mid-1960s (1962–1967) at a mean rate of 1.07 (1.00–1.14) Pg C $year^{-1}$, and finally through to the present day has been in balance.

For the harmonized budget (Figure 2d), change-point model selection indicates that the budget was balanced (no trend in cumulative imbalance) for about two decades from 1850 to 1870 or so (1864–1874, 95% CI) (Table 3), then was not balanced through to the 1910s (1908–1927) at a mean rate of −0.44 (−0.37 to −0.51) Pg C $year^{-1}$, then was balanced through to the early 1930s (1929–1935), and finally has not been balanced through to the present day (2018) at a mean rate of 0.65 (0.63–0.67) Pg C $year^{-1}$.

The standard and harmonized budgets share similar features: (1) an initial period of balance; (2) a period of negative imbalance at similar mean rates, representing either an overestimated global carbon sink or an underestimated carbon source; (3) another period of balance; and (4) followed by a period of positive imbalance representing an underestimated global carbon sink or an overestimated source.

The cumulative harmonized budget imbalance differs from the standard as follows (with the numbered list corresponding to those in the preceding paragraph): (1) duration of initial balanced period, about two decades in the harmonized budget and three decades in the standard budget. (2) Duration of period of negative imbalance, about five decades harmonized versus about two decades standard, which despite similar rates (−0.44 versus −0.46 Pg C $year^{-1}$) leads to a difference in the cumulative imbalance of the period, −21.6 ± 2.7 Pg C versus −8.3 ± 4.2 Pg C, respectively. (3) Duration of central period of balance, about two decades versus about four decades. (4) Duration of period of positive imbalance, about nine decades

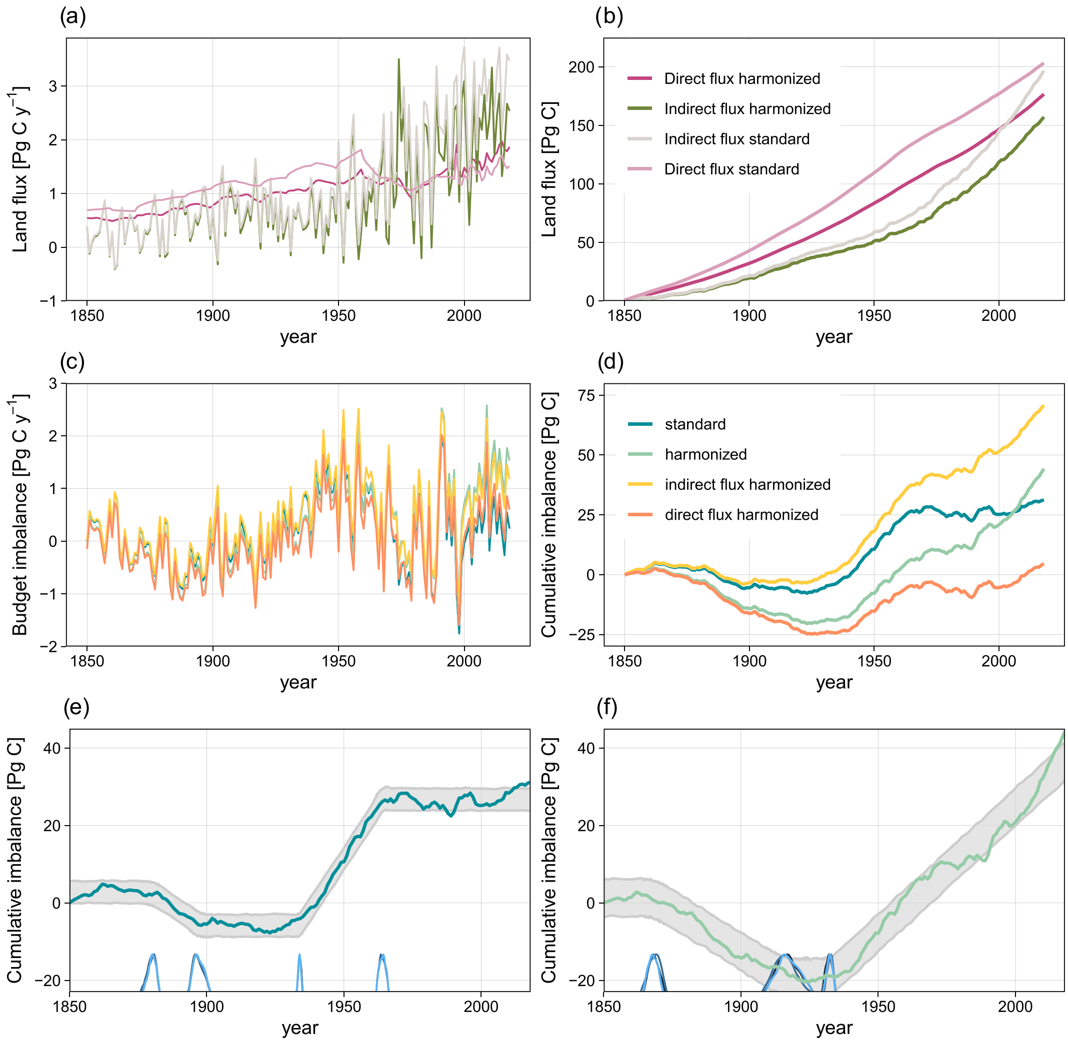

FIGURE 2 GCB standard and harmonized direct and indirect land carbon fluxes (top row: [a] annual; [b] cumulative). Standard and harmonized budget imbalance (middle row: [c] annual; [d] cumulative), including the BIM calculated while using only the harmonized direct flux (orange) and only the harmonized indirect flux (yellow). Change point models (bottom row) of (e) standard and (f) harmonized cumulative budget imbalance (colored lines), the 95% predictive interval of the change point models are shown in gray polygon, and the posterior distributions of the change points on each MCMC chain are shown as blue lines of different shades (these distributions do not correspond to the *y*-axis). BIM, budget imbalance; GCB, global carbon budget; MCMC, Markov Chain Monte Carlo.

years harmonized versus about three decades standard (with a further five decades of balance in the standard budget), and rate, 0.65 Pg C year^{-1} versus 1.07 Pg C year^{-1}. The differences in duration and rate in the final period of each budget lead to a cumulative imbalance difference during those periods of 55.9 ± 2.2 Pg C from the early 20th century to present in the harmonized budget versus 32.1 ± 3.9 Pg C in the middle third of the 20th century in the standard budget.

4 | DISCUSSION

In this analysis, we have shown that harmonizing assumptions in estimates of the direct and indirect anthropogenic land fluxes (E_{LUC} and S_{LAND}) of the GCB results in a longer period of budget imbalance during the 20thcentury (close to a century versus three decades) and a greater cumulative imbalance since the early 20th century (56 Pg C versus 32 Pg C) (Figure 2 and Table 2). Harmonization also indicates an imbalance of a longer duration in the later third of the 19th and early 20th centuries. The additional imbalance of the harmonized budget during the later two-thirds of the twentieth century is mainly driven by the LASC (Figure 2b), that is because of increasing expansion of anthromes and anthropogenic forest loss that reduced the proportion of ecosystems with the capacity to increase their carbon densities in response to accelerating global change. In other words, DGVMs predict the greatest CO_2-driven increases in carbon stocks in forests, and thus anthropogenic forest loss reduces the global capacity of the land sink to respond to increasing CO_2. When correcting for this loss of carbon uptake capacity, we find that instead of reducing uncertainty or closing the imbalance of the GCB, the correction results in a greater budget imbalance.

TABLE 3 Comparison of final change-point models for the standard and harmonized cumulative global carbon budget (GCB) imbalance.

GCB	Parameter	Mean	95% CI lower	95% CI upper	R hat
Standard	Change point 1	1880	1875	1883	1.0055
	Change point 2	1898	1893	1903	1.0020
	Change point 3	1934	1933	1935	1.0030
	Change point 4	1964	1961	1967	1.0001
	Intercept 1	2.81	2.30	3.36	1.0014
	Intercept 2	−5.82	−6.30	−5.30	1.0016
	Intercept 3	26.74	26.35	27.12	1.0000
	Sigma	1.43	1.27	1.58	1.0000
	Slope 1	−0.46	−0.65	−0.29	1.0062
	Slope 2	1.07	1.001	1.14	1.0027
Harmonized	Change point 1	1868	1864	1873	1.0108
	Change point 2	1917	1908	1927	1.0057
	Change point 3	1932	1929	1935	1.0325
	Intercept 1	1.33	0.14	2.43	1.0056
	Intercept 2	−19.45	−21.29	−17.80	1.0276
	Sigma	2.42	2.16	2.70	1.0003
	Slope 1	−0.44	−0.51	−0.37	1.0123
	Slope 2	0.65	0.63	0.67	1.0036

Note: Mean parameter values are shown with 95% credible interval (CI lower and upper) and assessment of convergence (R hat, potential scale reduction factor, values <1.05 indicate convergence).

The change-point analysis on cumulative data was used in this study to investigate broad trends and minimize the influence of inter-annual variability. We avoid over-interpretation of the estimated start and end dates of trends for methodological reasons described below. We primarily focus the remaining discussion around the later period of budget imbalance from the early 20th century with an average budget imbalance of 0.65 Pg C year^{-1}. The positive imbalance of the harmonized budget since the early 20th century indicates an underestimated sink term (atmosphere, ocean and land), an overestimated source term (fossil fuel and net land use change), or further missing components and inconsistencies in the budget calculation (analogous to those investigated in this study). In the following sections, we discuss various alternative explanations, noting that any explanation needs to account for why the imbalance started in the early 20th century and persists through to the present at a relatively constant rate.

4.1 | Potentially underestimated sink terms since the early 20th century

Of the sink terms in the GCB (atmosphere, ocean and land), the recent growth in the atmosphere's CO_2 pool is tightly constrained, with an annual growth rate of 5.2 ± 0.02 Pg C year^{-1} over the past decade (Friedlingstein et al., 2023). Atmospheric CO_2 has been measured by a network of flasks since 1959. However, for the years prior to 1959, atmospheric CO_2 is estimated from merged ice-core data (Joos & Spahni, 2008). Because of this methodological discontinuity, the dating accuracy and precision of the concentrations before and after 1959 are different. The air trapped in ice is a little older than the ice itself and represents an average of atmospheric composition over several years, depending on the varying rate of ice accumulation across sites (King et al., 2024; MacFarling Meure et al., 2006). For some of the cores used in Joos and Spahni (2008), the air age estimate is thought to be sufficiently precise to resolve sub-decadal (4–5 year) variability but not inter-annual variability (Etheridge et al., 1996; MacFarling Meure et al., 2004; Trudinger et al., 2002). Where both ice-core and flask records exist, these different methods generally show good agreement over multi-year timescales but less good for individual years (Etheridge et al., 1996; MacFarling Meure et al., 2006). Thus, overall, we suggest that the atmosphere is an unlikely candidate for an additional sink.

For the ocean sink, recent data-constrained estimates have demonstrated a roughly 10% greater sink than unconstrained ocean-model estimates, equivalent to about 0.25 Pg C year^{-1} over 1990–2020 (Terhaar et al., 2022). The addition of these revised data-driven estimates to the latest GCB increases the cumulative ocean sink (1959–2018) by 9.3 Pg C (Friedlingstein et al., 2023) compared with the 2019 budget. So the ocean sink is a likely candidate for some of the missing sink since the early 20th century identified within. However, it is unlikely to be solely responsible given the ocean sink is primarily driven by atmospheric CO_2 concentrations, which have increased significantly since the early 20th century (MacFarling Meure et al., 2006).

The indirect anthropogenic land flux (land sink, S_{LAND}), driven by CO_2 and climate change, is highly uncertain (0.8 Pg C year^{-1}) (Arora et al., 2020; Friedlingstein et al., 2023) and is a potential candidate for the greater sink. This sink is primarily driven by the effect of increasing CO_2 on plant photosynthesis and subsequent downstream effects on biomass production, standing biomass, and soil organic matter formation. While the CO_2 response of terrestrial ecosystems is represented in the TRENDY model ensemble, it has been suggested that the CO_2 response could be significantly greater given a particular set of ecological assumptions (Haverd et al., 2020) which align historical gross primary production (GPP) predictions with global scale inference from ice-core measurements of atmospheric carbonyl sulfide (OCS) (Campbell et al., 2017). However, the empirical GPP inference from OCS cannot be directly linked to increasing CO_2, and if the GPP trend were solely a CO_2 response, it would mean the relative global GPP increase is directly proportional to the relative atmospheric CO_2 increase (i.e., 1:1), which is not what is expected from leaf and canopy-scale understanding (Walker et al., 2021). The effects of greater carbon availability on tree growth (from CO_2-stimulated GPP) are further limited by various factors which are less well represented in DGVMs (Walker et al., 2021), which makes the indirect flux an unlikely candidate to explain the large BIM. At the landscape scale, models assume an equally strong impact of the CO_2 effect in managed and unmanaged lands, which is unlikely (see Grassi et al., 2023 for further discussion). Multi-scale (primarily satellite) estimates of recent live biomass trends (2000–2019) indicate a net land vegetation sink at about 50% of the standard GCB net land sink (Xu et al., 2021), albeit these trends are highly uncertain with potential biases (Araza et al., 2023) and do not include changes in soil carbon.

The CO_2 response of plants is further an unlikely cause of the imbalance since the early 20th century as the imbalance rate is constant over a long period while the CO_2 response (or the indirect land flux more broadly) has increased significantly over the same period.

4.2 | Potentially overestimated source terms since the early 20th century

Of the sources, fossil fuel emissions estimates are well constrained from national energy-use statistics (Briggs et al., 2023; Gilfillan & Marland, 2021), albeit estimates become increasingly less well constrained the further back in time the data go. Also, because of the large flux, the absolute uncertainty in fossil fuel emissions is substantial (20 Pg C cumulative 1960–2022), equating to about half the uncertainty in the direct and indirect land flux terms over the same period (45 and 35 Pg C, respectively) (Friedlingstein et al., 2023). Occasionally, revisions in the energy mix of a nation and thus conversion factors used to calculate the carbon intensity of energy production impact estimates of global fossil fuel emissions (e.g., Liu et al., 2015). Emissions reduction goal-setting could incentivize countries to report inflated figures but the evidence suggests “there are no indications of systematic manipulation or bias” (Briggs et al., 2023). Overall, overestimated fossil fuel emissions are an unlikely candidate to explain the large BIM, albeit they are such a large flux that small relative changes can be substantial. Further, as with the ocean flux and indirect anthropogenic land flux, there is no why fossil fuels would account for a constant imbalance of 0.65 Pg C year^{-1} since the early 20th century when fossil fuel emissions have grown 10-fold over the same time period.

The direct anthropogenic flux (land use change, E_{LUC}) is highly uncertain (0.7 Pg C year^{-1}) (Friedlingstein et al., 2023). This flux has been much more consistent since the early 20th century, 1.4 Pg C year^{-1} in the 1930s to 1.3 Pg C year^{-1} in the most recent decade (Friedlingstein et al., 2023). The direct flux is the net sum of two terms: carbon losses (e.g., from deforestation, peat drainage) and carbon gains (e.g., from land abandonment and subsequent forest regrowth). These gross terms are both larger than the net direct anthropogenic flux; thus, the budget imbalance is relatively smaller compared to the gross terms and could arise from either a lower loss of carbon from ecosystems than previously estimated or a greater gain as ecosystems recover from past land use.

The sources of land cover information used to generate the direct land flux estimates have an important change in their methods in 1960. Post 1960, FAO land cover statistics were used while pre-1960 land cover information was reconstructed from population estimates and historical reviews (Houghton, 2007; Houghton et al., 1983; Klein Goldewijk et al., 2017). The 1930 to 1960 period was of huge upheaval across the world, propagated in the Global North, with the great depression, the second world war, and the post-WWII great acceleration as new economic institutions and the cold war drove industrialization. It was a period of widespread land abandonment and reorganization of land management practices (McNeill & Engelke, 2016), including the transition from draught-animal power to the internal combustion engine in many places and the beginnings of industrial fertilizer and pesticide use.

The transition in land cover methods from population-based estimates to FAO data is a likely candidate for the budget imbalance in both the standard and the harmonized GCB. The standard budget only shows an imbalance from 1934 to 1964, which spans the method discontinuity and the period of social and economic upheaval. However, the continued imbalance to the present in the harmonized budget may be picking up legacy effects of this land cover and land use upheaval. One key legacy effect of land use change is carbon uptake because of forest regrowth and recovery following land abandonment and protection. In the bookkeeping models used to calculate the direct anthropogenic land flux, recovery is usually governed by a single rate parameter for a given land cover class. A recent study using DGVMs showed that across the northern hemisphere, models underestimate NBP by around 50% in forests under 80 years old, which contributes to a 1.1 ± 0.8 Pg C year^{-1} mismatch in model and atmospheric inversion estimates of the northern carbon sink (O'Sullivan et al., 2024). Simple representation of secondary forest regrowth in BK models (and DGVMs) is a potentially strong candidate for the missing sink and requires further attention.

4.3 | Remaining conceptual and epistemic inconsistencies

There are remaining inconsistencies in the budget calculation, even after LASC and PTD harmonization, resulting in part from using DGVMs to make these calculations and their inability to cleanly separate all the key terms of the terrestrial carbon budget (Dorgeist et al., 2024). While our harmonization resolves an important part of the inconsistencies in the current calculation of terrestrial fluxes, the obtained direct and indirect anthropogenic fluxes are not yet fully consistent: The LASC term also includes the impact of changing environmental conditions on land use fluxes (Dorgeist et al., 2024), which should rather be attributed to the direct anthropogenic flux—whereas in our harmonization, it is attributed to the indirect anthropogenic flux. The resulting effect of these inconsistencies is likely a small overestimation of both the indirect anthropogenic land sink and the direct anthropogenic carbon emissions, still remaining after the harmonization. However, these inconsistencies cannot be quantified with the currently available DGVM simulations, and their magnitude thus remains unspecified. Additionally, all budget terms can be subject to revisions to the GCB estimation methods. For example, as well as the upward-revised ocean sink discussed earlier, other budget terms have changed (1959–2018) in the 2023 GCB and the cement carbonation sink has been added (Friedlingstein et al., 2023).

Including the LASC, as implied by the name, leads to a reduction in the land sink predicted by DGVMs as anthropogenic land use change reduces forest cover. Lower forest cover reduces the land sink as models predict the greatest sink in forests (Friedlingstein et al., 2023), many of which are still considered primary forests in the model simulations. This model prediction implies that in equilibrium forests, biomass production can still be increased by higher available photosynthate carbon and soil carbon can be increased by greater litter inputs. However, there is little observational evidence to support this (Jiang et al., 2020; Lajtha et al., 2018; Walker et al., 2021), which implies the modeled land sink in primary forests may be overestimated.

Adjusting the BK estimates for the emissions from land use change by the PTD (as derived from DGVMs) reduces the cumulative estimates for the period 1850 until 2018 but increases the more recent estimates for the period 2009–2018 (Table 2). This is because transient simulations with DGVMs predict increasing ecosystem carbon density with historical environmental change. BK estimates based on transient carbon densities also find larger emissions in recent periods (Dorgeist et al., 2024). The standard BK model setup used by GCB instead assumes time-independent static carbon densities for a given land cover class. The PTD harmonization implies that primary forests had substantially lower carbon densities during pre-industrial times, yet again there is little direct evidence. A number of studies suggest that tropical primary forests have been accumulating carbon over the past four decades (Hubau et al., 2020), but it is likely that many of these forests were subject to multiple, cryptic anthropogenic disturbances in their past (McMichael et al., 2017), with impacts on the carbon cycle that are unclear.

5 | CONCLUSIONS

While independent estimates of budget terms are desirable and can provide confidence in those estimates when they all sum to balance the budget, independent estimates can also allow inconsistent assumptions to creep into the method. Current estimates of the GCB are based on terms that are, to varying degrees, incompatible with each other, as has been seen in similar studies that reconcile model estimates of the land terms and national greenhouse gas inventory reporting (Grassi et al., 2018, 2023). We want to be clear that this conclusion is not a criticism of the GCB efforts. The budget is an exemplar of community science (coordinating >100 scientists) that rigorously integrates and updates many data streams on an annual basis with little central support. Much of the work performed on the project is on a voluntary basis. Our goal here is to highlight that global carbon budgeting is not as solved as might be implied by a relatively closed budget, as currently is the case for the standard budget (Friedlingstein et al., 2023). Harmonizing the budget, as we have carried out here, requires adding simulations to the annual GCB project cycle (Table 1, S4–S6). To avoid additional simulations acting as a barrier to participation, these simulations could be added to the protocol on an optional basis to harmonize the budget with a subset of models. Understanding where on the land surface and for which land cover types harmonization makes the most difference is also essential for reconciling national greenhouse gas inventories with the GCB, facilitating the tracking of climate mitigation by country and sector (Grassi et al., 2018, 2023). Disaggregation by land cover requires tracking net biome production, and thus heterotrophic respiration, separately for each land cover type within a grid cell. We recommend modeling teams develop the capability to disaggregate heterotrophic respiration by land cover class.

Harmonization of anthropogenic land flux terms indicates a greater budget imbalance since the early 20th century. A budget imbalance of similar period and magnitude has been indicated in a multiscale observation and constrained modeling study (Lienert & Joos, 2018). While they did not specifically diagnose the causes of this imbalance, they found that adding additional processes related to the direct anthropogenic flux was as responsible for as much variation in cumulative NBP as model parameter uncertainty. Dorgeist et al. (2024) also find a similar increase in the BIM (1960–2021) with similar improved accounting for the direct and indirect fluxes and suggest underestimates of the ocean sink as a good candidate for re-missing sink. Uncertainties exist in all budget terms including method discontinuities in the atmospheric CO_2 and direct land flux terms, model structural errors in land and ocean terms. Because of the timing of the budget imbalance and its consistency since the early 20th century, the most plausible hypotheses for causes of the imbalance include as follows: the discontinuity of methods in the direct anthropogenic land flux and structural errors associated with the calculation of this flux, remaining inconsistencies in the harmonization method, and with a likely contribution from a larger ocean sink in more recent years.

Given the large uncertainties in budget terms and remaining inconsistencies in the method, we are not overly confident in the exact quantification of the missing sink (or smaller source) of 56 Pg C since the early 20th century. Rather than motivating a hunt for these exact numbers, we hope to motivate the community to take a deeper dive into understanding the historical terrestrial carbon cycle. In particular, increased efforts to reconcile 1900 to present day data for atmospheric CO_2 concentrations and human land cover and land use will be beneficial. We also highly recommend studies that simultaneously assess uncertainties and structural errors in both the direct and indirect anthropogenic land fluxes. Too often studies assess only one or the other, especially those assessing the indirect flux.

Current observational constraints cannot separate the direct and indirect terms, and the harmonization method used in this study shows that these two terms are intimately related. This is a critical point, as currently no observation can constrain the direct and indirect fluxes individually, only their sum can be constrained. Thus any study evaluating one or the other flux against an observational dataset is inherently accepting the errors embedded within the flux that is not of primary interest to that study. Further, if one flux is revised following the application of a new method, the other flux would require revision in the opposite direction (assuming their original sum matched the observational constraint).

Harmonizing two assumptions that are made when calculating the land carbon terms of the GCB leads to a greater budget imbalance. As with any budget, from a household budget to that of a nation or large multinational corporation, if the calculations do not balance, we know that we need to revisit our calculations and try to find the error. What these results indicate is that we need to continue to deepen our understanding and more finely resolve our estimates and associated uncertainties of the components of the GCB, in particular the highly uncertain land terms (direct and indirect anthropogenic fluxes). A more refined understanding of these two fluxes is essential for improving model predictions of the future land sink, which is in turn critical to estimate the remaining carbon budget as well as Nationally Determined Contribution to meet internationally agreed upon global temperature targets (Gidden et al., 2023). A more refined understanding and ability to model anthropogenic disturbance regimes in anthromes will lead to a greater ability to diagnose sources and sinks in the GCB and predict the carbon cycle under future scenarios.

AUTHOR CONTRIBUTIONS

Anthony P. Walker, Wolfgang A. Obermeier, and Julia Pongratz planned and designed the research. Anthony P. Walker, Wolfgang A. Obermeier, and Michael O'Sullivan synthesized and analyzed the data. Anthony P. Walker wrote the manuscript with contributions and feedback from Wolfgang A. Obermeier, Julia Pongratz, Pierre Friedlingstein, Charles D. Koven, Stephen Sitch, Michael O'Sullivan, and Clemens Schwingshackl.

ACKNOWLEDGMENTS

We would like to thank Fortunat Joos for discussions around the ice-core atmospheric CO_2 data and budget imbalance. We also thank two anonymous reviewers for their time and thoughtful feedback. We acknowledge the Global Carbon Project, which is responsible for the Global Carbon Budget (GCB). We thank the GCB TRENDY land modeling, ocean modeling, and fCO2-mapping groups for producing and making available their model outputs and fCO2-product output. This research was supported as part of the FACE Model Data Synthesis project, funded by the U.S. Department of Energy, Office of Science, and Office of Biological and Environmental Research. ORNL is managed by UT-Battelle, LLC, for the DOE under contract DE-AC05-1008 00OR22725. SS, MS, PF acknowledge support of the European Space Agency, CLIMATE-SPACE RECCAP2 project (Contract 4000144908/24/I-LR).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The global carbon budget data are made available on the Global Carbon Project website (https://www.globalcarbonproject.org/) for the 2019 and the 2023 budget here (https://www.globalcarbonproject.org/carbonbudget/archive.htm). The Land Use Harmonization data are made available here (https://luh.umd.edu/data.shtml). Data on the LASC and PTD used in this study are published online on the ESS-DIVE (https://data.ess-dive.lbl.gov/view/doi:10.15485/2477409) (Obermeier et al., 2024).

ORCID

Anthony P. Walker https://orcid.org/0000-0003-0557-5594
Wolfgang A. Obermeier https://orcid.org/0000-0002-7094-8011
Julia Pongratz https://orcid.org/0000-0003-0372-3960
Pierre Friedlingstein https://orcid.org/0000-0003-3309-4739
Charles D. Koven https://orcid.org/0000-0002-3367-0065
Clemens Schwingshackl https://orcid.org/0000-0003-4048-3011
Stephen Sitch https://orcid.org/0000-0003-1821-8561
Michael O'Sullivan https://orcid.org/0000-0002-6278-3392

REFERENCES

Araza, A., Herold, M., de Bruin, S., Ciais, P., Gibbs, D. A., Harris, N., Santoro, M., Wigneron, J.-P., Yang, H., Málaga, N., Nesha, K., Rodriguez-Veiga, P., Brovkina, O., Brown, H. C. A., Chanev, M., Dimitrov, Z., Filchev, L., Fridman, J., García, M., … Hein, L. (2023). Past decade above-ground biomass change comparisons from four multi-temporal global maps. *International Journal of Applied Earth Observation and Geoinformation*, *118*, 103274. https://doi.org/10.1016/j.jag.2023.103274

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., Schwinger, J., Bopp, L., Boucher, O., Cadule, P., Chamberlain, M. A., Christian, J. R., Delire, C., Fisher, R. A., Hajima, T., Ilyina, T., Joetzjer, E., Kawamiya, M., Koven, C. D., … Ziehn, T. (2020). Carbon-concentration and carbon-climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences*, *17*, 4173–4222. https://doi.org/10.5194/bg-17-4173-2020

Beven, K. (2016). Facets of uncertainty: Epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, *61*, 1652–1665. https://doi.org/10.1080/02626667.2015.1031761

Briggs, C., Gilfillan, D., Hefner, M., Marland, E., & Marland, G. (2023). Annual estimates of global and national CO_2 emissions from fossil fuels: Tracking revisions to the United Nations energy statistics database input energy data. *Environmental Data Science*, *2*, e40. https://doi.org/10.1017/eds.2023.38

Campbell, J. E., Berry, J. A., Seibt, U., Smith, S. J., Montzka, S. A., Launois, T., Belviso, S., Bopp, L., & Laine, M. (2017). Large historical growth in global terrestrial gross primary production. *Nature*, *544*, 84–87. https://doi.org/10.1038/nature22030

Canadell, J. G., Le Quéré, C., Raupach, M. R., Field, C. B., Buitenhuis, E. T., Ciais, P., Conway, T. J., Gillett, N. P., Houghton, R. A., & Marland, G. (2007). Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of the National Academy of Sciences*, *104*, 18866–18870. https://doi.org/10.1073/pnas.0702737104

Dorgeist, L., Schwingshackl, C., Bultan, S., & Pongratz, J. (2024). A consistent budgeting of terrestrial carbon fluxes. *Nature Communications*, *15*, 7426. https://doi.org/10.1038/s41467-024-51126-x

Etheridge, D. M., Steele, L. P., Langenfelds, R. L., Francey, R. J., Barnola, J.-M., & Morgan, V. I. (1996). Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn. *Journal of Geophysical Research: Atmospheres*, *101*, 4115–4128. https://doi.org/10.1029/95JD03410

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., … Zaehle, S. (2019). Global Carbon Budget 2019. *Earth System Science Data*, *11*(4), 1783–1838. https://doi.org/10.5194/essd-11-1783-2019

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., … Zheng, B. (2023). Global carbon budget 2023. *Earth System Science Data*, *15*, 5301–5369. https://doi.org/10.5194/essd-15-5301-2023

Garnier, S., Ross, N., Rudis, R., Camargo, P. A., Sciaini, M., et al. (2023). viridis (Lite) - Colorblind-Friendly Color Maps for R.

Gasser, T., Crepin, L., Quilcaille, Y., Houghton, R. A., Ciais, P., & Obersteiner, M. (2020). Historical CO_2 emissions from land use and land cover change and their uncertainty. *Biogeosciences*, *17*, 4075–4101. https://doi.org/10.5194/bg-17-4075-2020

Gidden, M. J., Gasser, T., Grassi, G., Forsell, N., Janssens, I., Lamb, W. F., Minx, J., Nicholls, Z., Steinhauser, J., & Riahi, K. (2023). Aligning climate scenarios to emissions inventories shifts global benchmarks. *Nature*, *624*, 102–108. https://doi.org/10.1038/s41586-023-06724-y

Gilfillan, D., & Marland, G. (2021). CDIAC-FF: Global and national CO_2 emissions from fossil fuel combustion and cement manufacture: 1751–2017. *Earth System Science Data*, *13*, 1667–1680. https://doi.org/10.5194/essd-13-1667-2021

Grassi, G., House, J., Kurz, W. A., Cescatti, A., Houghton, R. A., Peters, G. P., Sanz, M. J., Viñas, R. A., Alkama, R., Arneth, A., Bondeau, A., Dentener, F., Fader, M., Federici, S., Friedlingstein, P., Jain, A. K., Kato, E., Koven, C. D., Lee, D., … Zaehle, S. (2018). Reconciling global-model estimates and country reporting of anthropogenic forest CO_2 sinks. *Nature Climate Change*, *1*, 914–920. https://doi.org/10.1038/s41558-018-0283-x

Grassi, G., Schwingshackl, C., Gasser, T., Houghton, R. A., Sitch, S., Canadell, J. G., Cescatti, A., Ciais, P., Federici, S., Friedlingstein, P., Kurz, W. A., Sanz Sanchez, M. J., Abad Viñas, R., Alkama, R., Bultan, S., Ceccherini, G., Falk, S., Kato, E., Kennedy, D., … Pongratz, J. (2023). Harmonising the land-use flux estimates of global models and national inventories for 2000–2020. *Earth System Science Data*, *15*, 1093–1114. https://doi.org/10.5194/essd-15-1093-2023

Hansis, E., Davis, S. J., & Pongratz, J. (2015). Relevance of methodological choices for accounting of land use change carbon fluxes. *Global Biogeochemical Cycles*, *29*, 1230–1246. https://doi.org/10.1002/2014GB004997

Haverd, V., Smith, B., Canadell, J. G., Cuntz, M., Mikaloff-Fletcher, S., Farquhar, G., Woodgate, W., Briggs, P. R., & Trudinger, C. M. (2020). Higher than expected CO_2 fertilization inferred from leaf to global observations. *Global Change Biology*, *26*, 2390–2402. https://doi.org/10.1111/gcb.14950

Houghton, R. A. (2007). Balancing the global carbon budget. *Annual Review of Earth and Planetary Sciences*, *35*, 313–347. https://doi.org/10.1146/annurev.earth.35.031306.140057

Houghton, R. A., & Castanho, A. (2023). Annual emissions of carbon from land use, land-use change, and forestry from 1850 to 2020. *Earth System Science Data*, *15*, 2025–2054. https://doi.org/10.5194/essd-15-2025-2023

Houghton, R. A., Hobbie, J. E., Melillo, J. M., Moore, B., Peterson, B. J., Shaver, G. R., & Woodwell, G. M. (1983). Changes in the carbon content of terrestrial biota and soils between 1860 and 1980: A net release of CO_2 to the atmosphere. *Ecological Monographs*, *53*, 236–262. https://doi.org/10.2307/1942531

Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., Daniels, A. K., Ewango, C. E. N., Fauset, S., Mukinzi, J. M., Sheil, D., Sonké, B., Sullivan, M. J. P., Sunderland, T. C. H., Taedoumg, H., Thomas, S. C., White, L. J. T., Abernethy, K. A., Adu-Bredu, S., … Zemagho, L. (2020). Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*, *579*(7797), 80–87. https://doi.org/10.1038/s41586-020-2035-0

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., … Zhang, X. (2020). Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geoscientific Model Development*, *13*(11), 5425–5464. https://doi.org/10.5194/gmd-13-5425-2020

Hurtt, G., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J., Krisztin, T., Lawrence, D., Lawrence, P., … Zhang, X. (2019). Harmonization of Global Land Use Change and Management for the Period 850-2015 (Version 20230213) [Data set]. *Earth System Grid Federation*. https://doi.org/10.22033/ESGF/INPUT4MIPS.10454

IPCC. (2010). Meeting Report, 5-7 May, 2009, INPE, São José dos Campos, Brazil: Revisiting the use of managed land as a proxy for estimating national anthropogenic emissions and removals. In H. S. Eggleston, N. Srivastava, K. Tanabe, & J. Baasansuren (Eds.). Institute for Global Environmental Strategies (IGES).

Jiang, M., Medlyn, B. E., Drake, J. E., Duursma, R. A., Anderson, I. C., Barton, C. V. M., Boer, M. M., Carrillo, Y., Castañeda-Gómez, L., Collins, L., Crous, K. Y., de Kauwe, M. G., dos Santos, B. M., Emmerson, K. M., Facey, S. L., Gherlenda, A. N., Gimeno, T. E., Hasegawa, S., Johnson, S. N., … Ellsworth, D. S. (2020). The fate of carbon in a mature forest under carbon dioxide enrichment. *Nature*, *580*, 227–231. https://doi.org/10.1038/s41586-020-2128-9

Joos, F., & Spahni, R. (2008). Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years. *Proceedings of the National Academy of Sciences*, *105*, 1425–1430. https://doi.org/10.1073/pnas.0707386105

King, A. C. F., Bauska, T. K., Brook, E. J., Kalk, M., Nehrbass-Ahles, C., Wolff, E. W., Strawson, I., Rhodes, R. H., & Osman, M. B. (2024). Reconciling ice core CO_2 and land-use change following New World-Old World contact. *Nature Communications*, *15*, 1735. https://doi.org/10.1038/s41467-024-45894-9

Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene – HYDE 3.2. *Earth System Science Data*, *9*, 927–953. https://doi.org/10.5194/essd-9-927-2017

Lajtha, K., Bowden, R. D., Crow, S., Fekete, I., Kotroczó, Z., Plante, A., Simpson, M. J., & Nadelhoffer, K. J. (2018). The detrital input and removal treatment (DIRT) network: Insights into soil carbon stabilization. *Science of the Total Environment*, *640–641*, 1112–1120. https://doi.org/10.1016/j.scitotenv.2018.05.388

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., … Zeng, N. (2013). The global carbon budget 1959–2011. *Earth System Science Data*, *5*, 165–185. https://doi.org/10.5194/essd-5-165-2013

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., DCE, B., Barbero, L., Becker, M., Betts, R. A., Bopp, L., … Zhu, D. (2018). Global carbon budget 2017. *Earth System Science Data*, *10*, 405–448. https://doi.org/10.5194/essd-10-405-2018

Lienert, S., & Joos, F. (2018). A Bayesian ensemble data assimilation to constrain model parameters and land-use carbon emissions. *Biogeosciences*, *15*, 2909–2930. https://doi.org/10.5194/bg-15-2909-2018

Lindeløv, J. K. (2020). Mcp: An R package for regression with multiple change points. *OSF Preprints*. 1–23. https://doi.org/10.31219/osf.io/fzqxv

Liu, Z., Guan, D., Wei, W., Davis, S. J., Ciais, P., Bai, J., Peng, S., Zhang, Q., Hubacek, K., Marland, G., Andres, R. J., Crawford-Brown, D., Lin, J., Zhao, H., Hong, C., Boden, T. A., Feng, K., Peters, G. P., Xi, F., … He, K. (2015). Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature*, *524*, 335–338. https://doi.org/10.1038/nature14677

MacFarling Meure, C. (2004). The Variation of atmospheric carbon dioxide, methane and nitrous oxide during the Holocene from ice core analysis. PhD Thesis, University of Melbourne.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A., & Elkins, J. (2006). Law dome CO_2, CH_4 and N_2O ice core records extended to 2000 years BP. *Geophysical Research Letters*, *33*, L14810. https://doi.org/10.1029/2006GL026152

McNeill, J. R., & Engelke, P. (2016). *The great acceleration*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9wcc

McMichael, C. N. H., Matthews-Bird, F., Farfan-Rios, W., & Feeley, K. J. (2017). Ancient human disturbances may be skewing our understanding of Amazonian forests. *Proceedings of the National Academy of Sciences*, *114*(3), 522–527. https://doi.org/10.1073/pnas.1614577114

Mills, B. R. (2023). MoMAColors: Color Palettes Inspired by Artwork at the Museum of Modern Art in New York City.

Obermeier, W. A., Nabel, J. E. M. S., Loughran, T., Hartung, K., Bastos, A., Havermann, F., Anthoni, P., Arneth, A., Goll, D. S., Lienert, S., Lombardozzi, D., Luyssaert, S., McGuire, P. C., Melton, J. R., Poulter, B., Sitch, S., Sullivan, M. O., Tian, H., Walker, A. P., … Pongratz, J. (2021). Modelled land use and land cover change emissions—A spatio-temporal comparison of different approaches. *Earth System Dynamics*, *12*, 635–670. https://doi.org/10.5194/esd-12-635-2021

Obermeier, W. A., Walker, A. P., Nabel, J., Loughran, T., Hartung, K., Bastos, A., Havermann, F., Anthoni, P., Arneth, A., Goll, D., et al. (2024). Global carbon budget 1800–2018 loss of additional sink capacity and present transient difference. In *Free air CO_2 enrichment model data synthesis (FACE-MDS), ESS-DIVE repository. Dataset*. https://doi.org/10.15485/2477409

O'Sullivan, M., Sitch, S., Friedlingstein, P., Luijkx, I. T., Peters, W., Rosan, T. M., Arneth, A., Arora, V. K., Chandra, N., Chevallier, F., Ciais, P., Falk, S., Feng, L., Gasser, T., Houghton, R. A., Jain, A. K., Kato, E., Kennedy, D., Knauer, J., … Zaehle, S. (2024). The key role of forest disturbance in reconciling estimates of the northern carbon sink. *Communications Earth & Environment*, *5*, 705. https://doi.org/10.1038/s43247-024-01827-4

Plummer, M. (2023). rjags: Bayesian Graphical Models using MCMC.

Pongratz, J., Reick, C. H., Houghton, R., & House, J. (2014). Terminology as a key uncertainty in net land use and land cover change carbon flux estimates. *Earth System Dynamics*, *5*, 177–195. https://doi.org/10.5194/esd-5-177-2014

Pongratz, J., Schwingshackl, C., Bultan, S., Obermeier, W., Havermann, F., & Guo, S. (2021). Land use effects on climate: Current state, recent Progress, and emerging topics. *Current Climate Change Reports*, *7*, 99–120. https://doi.org/10.1007/s40641-021-00178-y

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Schlesinger, W. H. (1977). Carbon balance in terrestrial detritus. *Annual Review of Ecology and Systematics*, *8*, 51–81. https://doi.org/10.1146/annurev.es.08.110177.000411

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., … Myneni, R. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*, 653–679. https://doi.org/10.5194/bg-12-653-2015

Sitch, S., O'Sullivan, M., Robertson, E., Friedlingstein, P., Albergel, C., Anthoni, P., Arneth, A., Arora, V. K., Bastos, A., Bastrikov, V., Bellouin, N., Canadell, J. G., Chini, L., Ciais, P., Falk, S., Harris, I., Hurtt, G., Ito, A., Jain, A. K., … Zaehle, S. (2024). Trends and drivers of terrestrial sources and sinks of carbon dioxide: An overview of the TRENDY project. *Global Biogeochemical Cycles*, *38*, e2024GB008102. https://doi.org/10.1029/2024GB008102

Stocker, B. D., Yu, Z., Massa, C., & Joos, F. (2017). Holocene peatland and ice-core data constraints on the timing and magnitude of CO2 emissions from past land use. *Proceedings of the National Academy of Sciences*, *114*, 1492–1497. https://doi.org/10.1073/pnas.1613889114

Terhaar, J., Frölicher, T. L., & Joos, F. (2022). Observation-constrained estimates of the global ocean carbon sink from earth system models. *Biogeosciences*, *19*, 4431–4457. https://doi.org/10.5194/bg-19-4431-2022

Trudinger, C. M., Enting, I. G., Rayner, P. J., & Francey, R. J. (2002). Kalman filter analysis of ice core data 2. Double deconvolution of CO_2 and δ13C measurements. *Journal of Geophysical Research: Atmospheres*, *107*, ACH 5-1-ACH 5-24. https://doi.org/10.1029/2001JD001112

Trudinger, C. M., Etheridge, D. M., Rayner, P. J., Enting, I. G., Sturrock, G. A., & Langenfelds, R. L. (2002). Reconstructing atmospheric histories from measurements of air composition in firn. *Journal of Geophysical Research: Atmospheres*, *107*, ACH 15-1-ACH 15-13. https://doi.org/10.1029/2002JD002545

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., & Gelman, A. (2023). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models.

Walker, A. P., Kauwe, M. G. D., Bastos, A., Belmecheri, S., Georgiou, K., Keeling, R. F., McMahon, S. M., Medlyn, B. E., Moore, D. J. P., Norby, R. J., Zaehle, S., Anderson-Teixeira, K. J., Battipaglia, G., Brienen, R. J. W., Cabugao, K. G., Cailleret, M., Campbell, E., Canadell, J. G., Ciais, P., … Zuidema, P. A. (2021). Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric CO_2. *New Phytologist*, *229*, 2413–2445. https://doi.org/10.1111/nph.16866

Whittaker, R. H., & Likens, G. E. (1973). Carbon in the biota. Brookhaven symposia in biology. In G. M. Woodwell & E. V. Pecan (Eds.), *Carbon and the biosphere: Proceedings of the 24th Brookhaven symposium in biology, may 16–18, 1972, Upton, New York* (pp. 281–302). Washington DC, USA.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., Pongratz, J., Bloom, A. A., Bowman, K., Worden, J., Liu, J., Yin, Y., Domke, G., McRoberts, R. E., Woodall, C., Nabuurs, G. J., de-Miguel, S., Keller, M., Harris, N., Maxwell, S., & Schimel, D. (2021). Changes in global terrestrial live biomass over the 21st century. *Science Advances*, *7*, eabe9829. https://doi.org/10.1126/sciadv.abe9829

How to cite this article: Walker, A. P., Obermeier, W. A., Pongratz, J., Friedlingstein, P., Koven, C. D., Schwingshackl, C., Sitch, S., & O'Sullivan, M. (2024). Harmonizing direct and indirect anthropogenic land carbon fluxes indicates a substantial missing sink in the global carbon budget since the early 20th century. *Plants, People, Planet*, 1–14. https://doi.org/10.1002/ppp3.10619