The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times

A. Anthony Bloom\textsuperscript{a,b,c,1}, Jean-François Exbrayat\textsuperscript{b,c}, Ivar R. van der Velde\textsuperscript{d}, Liang Feng\textsuperscript{b,c}, and Mathew Williams\textsuperscript{b,c}

\textsuperscript{a}Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109; \textsuperscript{b}School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FF, United Kingdom; \textsuperscript{c}National Centre for Earth Observation, Edinburgh EH9 3FF, United Kingdom; and \textsuperscript{d}Meteorology and Air Quality, Wageningen University, Wageningen 6700 AA, The Netherlands

Edited by William H. Schlesinger, Cary Institute of Ecosystem Studies, Millbrook, NY, and approved December 9, 2015 (received for review July 31, 2015)

The terrestrial carbon cycle is currently the least constrained component of the global carbon budget. Large uncertainties stem from a poor understanding of plant carbon allocation, stocks, residence times, and carbon use efficiency. Imposing observational constraints on the terrestrial carbon cycle and its processes is, therefore, necessary to better understand its current state and predict its future state. We combine a diagnostic ecosystem carbon model with satellite observations of leaf area and biomass (where and when available) and soil carbon data to retrieve the first global estimates, to our knowledge, of carbon cycle state and process variables at a 1° × 1° resolution; retrieved variables are independent from the plant functional type and steady-state paradigms. Our results reveal global emergent relationships in the spatial distribution of key carbon cycle states and processes. Live biomass and dead organic carbon residence times exhibit contrasting spatial features (\textit{r} = 0.3). Allocation to structural carbon is highest in the wet tropics (85–88%) in contrast to higher latitudes (73–82%), where allocation shifts toward photosynthetic carbon. Carbon use efficiency is lowest (0.42–0.44) in the wet tropics. We find an emergent global correlation between retrievals of leaf mass per leaf area and leaf lifespan (\textit{r} = 0.64–0.80) that matches independent trait studies. We show that conventional land cover types cannot adequately describe the spatial variability of key carbon states and processes (multiple correlation median = 0.41). This mismatch has strong implications for the prediction of terrestrial carbon dynamics, which are currently based on globally applied parameters linked to land cover or plant functional types.

\begin{itemize}
  \item Carbon cycle
  \item Biomass
  \item Soil carbon
  \item Allocation
  \item Residence time
\end{itemize}

The terrestrial carbon (C) cycle remains the least constrained component of the global C budget (1). In contrast to a relatively stable increase of the ocean CO\textsubscript{2} sink from 0.9 to 2.7 Pg C yr\textsuperscript{-1} over the past 40 y, terrestrial CO\textsubscript{2} uptake has been found to vary between a net 4.1-Pg C yr\textsuperscript{-1} sink to a 0.4-Pg C yr\textsuperscript{-1} source, and accounts for a majority of the interannual variability in atmospheric CO\textsubscript{2} growth. The complex response of terrestrial ecosystem CO\textsubscript{2} exchanges to short- and long-term changes in temperature, water availability, nutrient availability, and rising atmospheric CO\textsubscript{2} (2–6) remains highly uncertain in C cycle model projections (7). As a result, there are large gaps in our understanding of terrestrial C dynamics, including the magnitude and residence times of the major ecosystem C pools (8, 9) and rates of autotrophic respiration (10). Moreover, the impact of climatic extremes on C cycling, such as recent Amazon droughts (11), highlights the importance of understanding the terrestrial C cycle sensitivity to climate variability. To understand terrestrial CO\textsubscript{2} exchanges in the past, present, and future, we need to better constrain current dynamics of ecosystem C cycling from regional to global scales.

C uptake, allocation, pool stocks, residence times, respiration, and disturbance together drive net CO\textsubscript{2} exchanges (12) on subdaily to millennial timescales; these C state and process variables also determine the temporal sensitivity of the net C balance to climatic variability. For example, global changes in photosynthetic uptake could lead to a rapid response from short-lived C pools (such as foliage, fine roots, and litter) or a prolonged response from the long-lived C pools (such as woody biomass and soil C), with very different outcomes on ecosystem source–sink behavior. Quantitative knowledge of terrestrial C pathways is, therefore, central to understanding the temporal responses of the major terrestrial C fluxes—including heterotrophic respiration (13), fires (14, 15), and wetland CH\textsubscript{4} emissions (16, 17)—to interannual variations in C uptake.

Although C dynamics have been extensively measured and analyzed at site level (18–21), the respiration and allocation of fixed C and its residence time within the major C pools are difficult and expensive to measure at site level and remain poorly quantified on global scales. As a result, global terrestrial C cycle models rely on land cover type-specific C cycling parameters—based on spatially preassigned plant functional types—to determine C fluxes and C pools (22). Globally spanning C cycle observations can provide a much-needed constraint on the spatial variability and associated dynamics of the terrestrial C cycle. Over the past decade, a growing number of datasets has enhanced understanding of the terrestrial C cycle, including global-scale canopy dynamics [National Aeronautics and Space Administration Moderate Resolution Imaging Spectroradiometer (MODIS) leaf area index (LAI)], empirically derived

\textbf{Significance}

Quantitative knowledge of terrestrial carbon pathways and processes is fundamental for understanding the biosphere’s response to a changing climate. Carbon allocation, stocks, and residence times together define the dynamic state of the terrestrial carbon cycle. These quantities are difficult to measure and remain poorly quantified on a global scale. Here, we retrieve global 1° × 1° carbon state and process variables by combining a carbon balance model with satellite observations of biomass and leaf area (where and when available) and global soil carbon data. Our results reveal emergent continental-scale patterns and relationships between carbon states and processes. We find that conventional land cover types cannot capture continental-scale variations of retrieved carbon variables; this mismatch has strong implications for terrestrial carbon cycle predictions.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Frequently available online through the PNAS open access option.

Data deposition: The outputs reported in this paper have been deposited in an online digital repository of multidisciplinary research datasets produced at the University of Edinburgh, datashare.is.ed.ac.uk/handle/10283/875.

1To whom correspondence should be addressed. Email: abloom@jpl.nasa.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1515160113/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1515160113

PNAS | February 2, 2016 | vol. 113 | no. 5 | 1285–1290
Fig. 1. Diagnostic ecosystem C balance model DALEC2 (19, 35) and datasets used to retrieve 1° × 1° C state and process variables. GPP, a function of climate and foliar C, is partitioned into autotrophic respiration (Ra) and NPP. NPP is partitioned into the live biomass pools. Plant mortality provides input to the DOM pools. Heterotrophic respiration (Rh) is derived from decomposing DOM pools. Fire fluxes are derived from burned area data (35) and all C pools (SI Text, section S2). Within each 1° × 1° grid cell, we use a Bayesian MDF algorithm to retrieve C state/process variables and uncertainties; variables are retrieved without prior land cover type or steady-state assumptions. Data constraints consist of MODIS leaf area, total biomass (24 crops), and soil C (23). Details on the Bayesian fusion approach are provided in Materials and Methods.

global soil C data [Harmonized World Soil Database (HWSD)] (23), satellite-based above- and belowground biomass (ABGB) maps for the tropics (24, 25), and Greenhouse Gases Observing Satellite CO₂ and plant fluorescence (26, 27). These spatially and temporally explicit datasets provide an enhanced view of the terrestrial C cycle and can be used together to retrieve consistent global C state and process variables. Significant efforts in data-driven estimates of the global C fluxes have been made over the past decade. These efforts include estimates based on atmospheric CO₂ concentrations (1, 28, 29), high-resolution global primary production maps (30) based on eddy covariance tower datasets (FLUXNET) (18), mean residence time of terrestrial C (31), ecosystem respiration dependence on temperature based on FLUXNET data (32), and global C cycle data assimilation systems (33).

Given an increasing number of C cycle observations, what remains an outstanding challenge is to produce a data-consistent analysis of terrestrial C cycling—including retrievals of C fluxes, C pools, autotrophic respiration, allocation fractions, and residence times—based on multiple global-scale earth observations and datasets. Current global-scale terrestrial biosphere models, because of their complexity and structures, are ill-equipped to ingest an ever-increasing volume of earth observations to estimate (rather than prescribe) model parameters based on the currently available observations. To overcome this challenge, we use a model-data fusion (MDF) approach to retrieve terrestrial C state and process variables during the period 2001–2010 without invoking plant functional type or steady-state assumptions. We bring together global MODIS LAI, a tropical biomass map (24), a soil C dataset (23), MODIS burned area (34), and a diagnostic ecosystem C balance model [Data Assimilation Linked Ecosystem Carbon Model version two (DALEC2)] (19, 35) to retrieve C state and process variables by producing a novel data-consistent and spatially explicit analysis of terrestrial C cycling on a global 1° × 1° grid (Fig. 1) [we henceforth refer to this MDF setup as the C data model framework (CARDAMOM)]. Specifically, we address the following questions: How is C uptake partitioned between the live biomass pools and respiration? What is the residence time of C within the major ecosystem C pools? How do estimates of C cycle states and processes vary spatially, and to what degree do emergent variable patterns match land cover maps? We use a Markov Chain Monte Carlo MDF algorithm to retrieve C state and process variables—and their associated uncertainty—within each 1° × 1° grid cell (Materials and Methods). The MDF approach retrieves the state and process variables that minimize the model mismatch against any available C cycle observations. Therefore, in the absence of extratropical biomass data or wintertime MODIS LAI observations, estimates of 2001–2010 C state and process variables are achievable, albeit more uncertain.

Results

Distinct C allocation patterns emerge from our terrestrial C analysis (Fig. 2). Net primary production (NPP) allocation to structural biomass (wood and fine roots) is largely ≥80% (area-weighted 25th to 75th percentile range = 85–88%) in the wet tropics (<23° N/S; annual precipitation >1,500 mm) in contrast to the dry tropics (77–87%) and extratropical regions (73–82%). The highest NPP allocations to foliage (≥30%) spatially coincide with major grassland areas, including the North America prairies, the central Asia steppes, and the Sahel region in Africa. The dry tropics exhibit relatively high NPP allocation to labile C (7–14%) (Fig. S1), which reflects the increasing impact of seasonality on production as precipitation declines, requiring labile C stores for leaf flush. C use efficiency (CUE; equivalent to 1 – autotrophic respiration fraction) is overall lowest within the wet tropics (0.42–0.44) in contrast to dry tropical (0.45–0.50), temperate (23–55 N/S; 0.47–0.50), and high-latitude (>55° N/S; 0.49–0.50) areas.
Live biomass and dead organic C residence times exhibit contrasting spatial features \((r = 0.3)\) (Fig. 3). Within the majority of wet tropical land area (56\%)—especially across most of the Amazon River (76\%) and Congo River (69\%) basins—the longest C residence time occurs within the woody pool (Fig. S1). In the dry tropics and extratropical latitudes, soil C residence times exceed wood C residence time by a median factor of 2.6 (1.6–4.3). Woody residence time is typically shorter in the dry tropics (8–19 y) compared with other biomes (wet tropics: 12–21 y; temperate latitudes: 21–29 y; and high latitudes: 25–28 y). Litter C residence time is typically longer in extratropical ecosystems (0.8–1.6 y) compared with tropical ecosystems (0.4–0.5 y). The longest foliar residence time (or leaf lifespan) occurs in the wet tropics and semiarid regions (Fig. S1).

Overall, the wet tropics are characterized by relatively high structural C \((>100 \text{ tC ha}^{-1})\) and photosynthetic C \((>2.5 \text{ tC ha}^{-1})\) (Fig. 4); in contrast, the dry tropics and extratropical regions exhibit less structural and/or photosynthetic C. Foliar C stocks are typically larger in the wet tropics (2.8–4.7 tC ha\(^{-1}\)) relative to other biomes (0.2–0.6 tC ha\(^{-1}\)); similarly, fine root stocks are also greater in the wet tropics (4.0–5.3 tC ha\(^{-1}\)) compared with other biomes (0.8–2.7 tC ha\(^{-1}\)). Root:shoot (fine root C:leaf C) is lowest in the wet tropics (1.1–1.5) followed by the dry tropics (1.6–1.9) and extratropics (1.8–2.1). We find larger woody C uncertainties \((1^\circ \times 1^\circ \text{90\% confidence range/median})\) in the extratropics (1.8–4.6) in contrast to tropical woody C (1.4–1.6) because of the latitudinal limits of the total ABGB map (24). Litter C is greater in high latitudes (2.4–3.4 tC ha\(^{-1}\)) relative to temperate (0.6–2.4 tC ha\(^{-1}\)) and tropical (0.2–2.6 tC ha\(^{-1}\)) regions. High-latitude ecosystems have higher labile C stocks linked to seasonal leaf expansion (0.2–0.5 tC ha\(^{-1}\)) relative to temperate (0.1–0.3 tC ha\(^{-1}\)) and tropical (0.1–0.3 tC ha\(^{-1}\)) ecosystems.

We find high leaf C mass per leaf area (LCMA) values in the wet tropics (85–97 gC m\(^{-2}\)) and semiarid regions, such as the Sahel, southwestern United States, and the Australian continent (typically >100 gC m\(^{-2}\)) (Fig. 5); LCMA estimates are lower (typically <80 gC m\(^{-2}\)) in high latitudes and the dry tropics. We find a positive correlation between leaf lifespan and LCMA in LCMA in high-altitude \((r = 0.79)\), temperate \((r = 0.80)\), dry tropical \((r = 0.78)\), and wet tropical \((r = 0.64)\) areas.

Global gross primary production (GPP; global 25th to 75th percentile = 91–134 Pg C y\(^{-1}\)), ecosystem respiration (91–137 Pg C y\(^{-1}\)), and fires (1.3–2.0 Pg C y\(^{-1}\)) are broadly consistent with a terrestrial C model ensemble (22), data-driven estimates (36), and bottom-up inventories (37) (Fig. S2). The net C exchange uncertainty (–8 to +13 Pg C y\(^{-1}\)) is an order of magnitude greater than mode net C exchange (NCE; –2 Pg C y\(^{-1}\)); NCE latitudinal uncertainty is larger but comparable with the terrestrial C model ensemble range. Global atmospheric model CO\(_2\) concentrations based on CARDAMOM mode NCE fluxes are seasonally consistent \([r^2 = 0.93; \text{root-mean-square error (RMSE)} = 0.53 \text{ ppm CO}_2]\) with mean total column CO\(_2\) measurements (38) (Fig. S3). The mean integrated C residence time in ref. 31 is within the range of individual pool residence times at locations B, T, D, and W (Fig. 3). The 2001–2010 CARDAMOM analysis spatial and temporal LAI variability is consistent with the MODIS LAI constraints \([r^2 = 0.8; \text{RMSE} = 0.6 \text{ m}^2/\text{m}^2]\). When alternative GPP (36), alternative model structure, or biased data constraints (±20%) are imposed at locations B, T, D, and W, 88% of median sensitivity analysis estimates are within ±50% of median C state and process variable retrievals (Fig. S4).
Retrieved C cycle variables are broadly consistent with a range of in situ measurements (Table S1). Estimates of CUE within the Amazon River basin are comparable with the upper bound of recent measurements (0.32–0.47) (39). Recent estimates of extratropical forest C density (40) are, on average, 38% lower than CARDAMOM total biomass estimates within forested areas (although the forest biomass estimates are typically within the CARDAMOM 1°×1° uncertainty). Estimates of mean Amazon woody C residence times (15–21 y) are lower but comparable with aboveground woody C residence times derived from site-level measurements (~20–70 y) (20).

We find that 88–99% of C state and process variability is accounted for by eight empirical orthogonal basis functions (EOFs) (Fig. 6); in other words, retrieved C state and process variables are largely explained by eight modes of spatial variability (Fig. S5). On average, the Global Land Cover Map (GLOBCOVER) land cover classifications (41) (e.g., deciduous forests, evergreen forests, and grasslands) account for ~50% of C state and process variability (median multiple correlation coefficient R = 0.41); GLOBCOVER land cover types best describe spatial variations in C stocks (0.5 ≤ R ≤ 0.8) followed by LCMA (R = 0.4), residence times (0.3 ≤ R ≤ 0.5), and allocation fractions (0.1 ≤ R ≤ 0.4).

Discussion

Typically, C allocation and residence time parameters are based on land cover types in global-scale terrestrial C cycle studies (refs. 9 and 22 among others); here, spatially broad allocation and residence patterns emerge instead as a result of the MDF approach. For example, high-biomass ecosystems throughout the wet tropics display similar C allocation, residence time, and LCMA configurations (Figs. 2–5). Similarly, we find that dead organic matter (DOM) C residence is generally longer in high latitudes (Fig. 3). Compared with conventional land cover types, EOFs 1–4 account for a larger degree of the spatial structures in retrieved C variables (Fig. 6); for most variables, the two dominant EOF modes—which together reflect first-order variations in latitude and global precipitation patterns (Fig. S5)—explain more spatial variability than GLOBCOVER land cover types. The mismatch between land cover types and retrieved variables has major implications for the estimation and prediction of terrestrial C cycling, which is currently based on small sets of globally applied parameters linked to land cover types. The importance of climate, biodiversity, fire, and anthropogenic disturbance in generating these mismatches needs to be explored in additional research (42).

It also is clear that plant traits vary across biomes (Figs. 2–4 and Fig. S1), not just at biome boundaries (43), and that there are continental-scale tradeoffs and correlations among traits (44). Our analysis is consistent with these viewpoints: for example, the emergent relationship between LCMA (proportional to leaf mass per area) and leaf lifespan (Fig. 5) matches the positive correlation found in global plant trait datasets (45). Evaluating global plant trait patterns emerging from CARDAMOM provides a novel opportunity for connections to theoretical and functional biodiversity research and a route to integrating this knowledge into predictive terrestrial C cycle modeling.

The residence times of major C stocks provide substantial insights into the sensitivity and potential future trajectories of the terrestrial C cycle. For example, land cover changes in the wet tropics may result in rapid DOM C losses given the relatively short DOM residence times (<30 y) (Fig. 3). In contrast, high-latitude C residence times are an order of magnitude higher (30–300 y), and therefore, shifts in C allocation or turnover rates are likely to result in long-lived C flux responses. Overall, given the predominant role of C residence times in future terrestrial up-take responses (9), the derived residence times provide a first-order estimate of ecosystem response times as a result of changes in C cycling regimes. However, we note that model structure is likely to be a major source of uncertainty in long-lived (>10 y) C flux predictions. For example, although reduced complexity models can capture some of the principal long-term (>10 y) DOM dynamics represented in earth system models (8), systematic errors in DOM dynamics can arise because of the underrepresentation of processes controlling DOM residence times (46, 47). We also note that our decadal analysis is unlikely to be able to capture slow feedback processes acting on longer
timescales, such as permafrost remobilization and priming (48). The large allocation and stocks and short residence time of wood in the wet tropics indicate the potentially rapid postdisturbance regrowth and C accumulation (49). We note that fires are less frequent but major events within boreal ecosystems (50), and therefore, longer time periods are required for retrievals to fully account for the effect of fires on high-latitude C residence times. C state and process variable retrievals are sensitive to the uncertainty characteristics of C cycle observations (35) and the prior parameter ranges (Table S2). We highlight that the current coverage and accuracy of C cycle observations (24, 51) remain major limiting factors in our approach. For example, extratropical C stock and residence time uncertainties are higher because of the absence of biomass observations. Undoubtedly, future estimates of globally spanning biomass density will provide a major constraint on CARDAMOM estimates of extratropical C state and process variables (52).

Land to atmospheric C flux estimates could be used to further constrain CARDAMOM C fluxes (Fig. S2) and C cycle variables associated to nonsteady C states. For example, soil C residence time samples are negatively correlated with corresponding mean 2001–2010 NCE samples at locations B (r = −0.3), T (r = −0.4), D (r = −0.5), and W (r = −0.3); therefore, regional- or grid-scale estimates of NCE could provide a much-needed additional constraint on soil C residence time. CARDAMOM flux magnitude and uncertainty can be used as prior information in a global atmospheric CO2 inversion; in turn, the assimilation of Greenhouse Gases Observing Satellite (26) and Orbiting Carbon Observatory 2 atmospheric CO2 observations (53) should further constrain CARDAMOM NCE estimates and their associated uncertainties. In this manner, nonsteady-state C fluxes can ultimately be reconciled with ecosystem state and process variables, such as C stocks and residence times.

The CARDAMOM approach provides a framework to test alternative model structures (54); in this manner, combined C cycle model parametric and structural uncertainties can be characterized, while ensuring consistency between models and global-scale datasets. This assessment would amount to a major step forward from conventional C cycle model intercomparison studies. Ultimately, an ensemble of models can be used to determine the degree to which retrievals of key C state and process variables are model-dependent. Moreover, alternative model structures could be used in CARDAMOM to assimilate globally spanning plant traits related to C cycling (55) and satellite observations, such as solar-induced fluorescence (27), vegetation optical depth (56), soil moisture (57, 58), and changes in aboveground biomass (25, 59, 60). We anticipate that the incorporation of additional datasets and alternative model structures into CARDAMOM will generate quantifiable reductions in retrieved C variable uncertainties and new ecological insights on the state of the terrestrial C cycle.

Materials and Methods

We grid MODIS LAI, ABGB (24), and HWSO topsoil and subsoil (0–100 cm) C density (23) at a 1° × 1° resolution (SI Text, section S1). DALEC2 is analytically described in ref. 35; an overview of DALEC2 C fluxes and pools is shown in Fig. 6. All CARDAMOM datasets presented in this study can be downloaded under The Netherlands Organization for Scientific Research Project VIDI: 803. We also used a principal component analysis on C state and process variables to retrieve the primary 1° × 1° EOFs. The details of the CARDAMOM results evaluation and analyses are fully described in SI Text, sections S5, S6, and S8. The Pearson’s correlation coefficient is abbreviated as r throughout the text. All spatially derived r and RMSE values reported in the text are area-weighted. Retrieved C variable ranges—reported as area-weighted 25th to 75th percentile range—are derived from 1° × 1° mean allocation and C stocks, log-based mean C residence times (Fig. S1), and median LEMA C values (Fig. 6). All CARDAMOM datasets presented in this study can be downloaded from dataset.is.ed.ac.uk/handle/10283/875.

ACKNOWLEDGMENTS. A.A.B., J.-F.E., L.F., and M.W. were funded by the NERC National Centre for Earth Observation. I.R.D.V. was financially supported under The Netherlands Organization for Scientific Research Project VIDI: 803. We used data made available by the European Space Agency (ESA) under The Netherlands Organization for Scientific Research Project VIDI: 803. The work was carried out at the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA.
