

**Title:** Carbon cycling in mature and regrowth forests globally: a macroecological synthesis based on the global forest carbon database (ForC)

**Authors:** Kristina J. Anderson-Teixeira<sup>1,2\*</sup> Valentine Herrmann<sup>1</sup> Becky Banbury Morgan Ben Bond-Lamberty Susan Cook-Patton Abigail Ferson Norbert Kunert Jennifer McGarvey Helene C. Muller-Landau<sup>1</sup> Maria Wang

**Author Affiliations:**

1. Conservation Ecology Center; Smithsonian Conservation Biology Institute; National Zoological Park, Front Royal, VA 22630, USA
2. Center for Tropical Forest Science-Forest Global Earth Observatory; Smithsonian Tropical Research Institute; Panama, Republic of Panama

\*corresponding author: teixeirak@si.edu; +1 540 635 6546

## Summary

(The Summary should be no longer than 300 words and divided into the following sections: Background, Methods/Design, Review results/ Synthesis and Discussion. Up to 7 Keywords should also be provided. The total length of the article is flexible.)

*Background.* The fate of Earth's climate closely linked to forests, which strongly influence atmospheric carbon dioxide (CO<sub>2</sub>) and climate through their influential role in the global carbon (C) cycle. Synthetic understanding of global forest C cycles is needed to constrain model estimates of forest feedbacks to climate change and to more accurately quantify the influence of land use decisions on climate.

*Methods/Design.* Here, we draw from the Global Forest C Database, ForC, to provide a macroscopic overview of C cycling in the world's forests, giving special attention to stand age-related variation. Specifically, we draw upon ## records from ## geographic locations representing ## C cycle variables to characterize ensemble C budgets for four broad forest types (tropical broadleaf evergreen, temperate broadleaf, temperate conifer, and taiga), including estimates for both mature and regrowth (age <100 years) forests. For regrowth forests, we quantify age trends for all variables.

*Review Results/ Synthesis.* The rate of C cycling generally increased from boreal to tropical regions, whereas C stocks showed less directional variation. The majority of flux variables, together with most live biomass pools, increased significantly with stand age, and the rate of increase again tended to increase from boreal to tropical regions.

*Discussion.* [Discussion section will interpret results, highlighting new and significant findings, and discuss implications. Tentative headings are “Stand level C cycling in forests globally”, “Age trends in C cycling”, and “Implications for climate change mitigation”.]

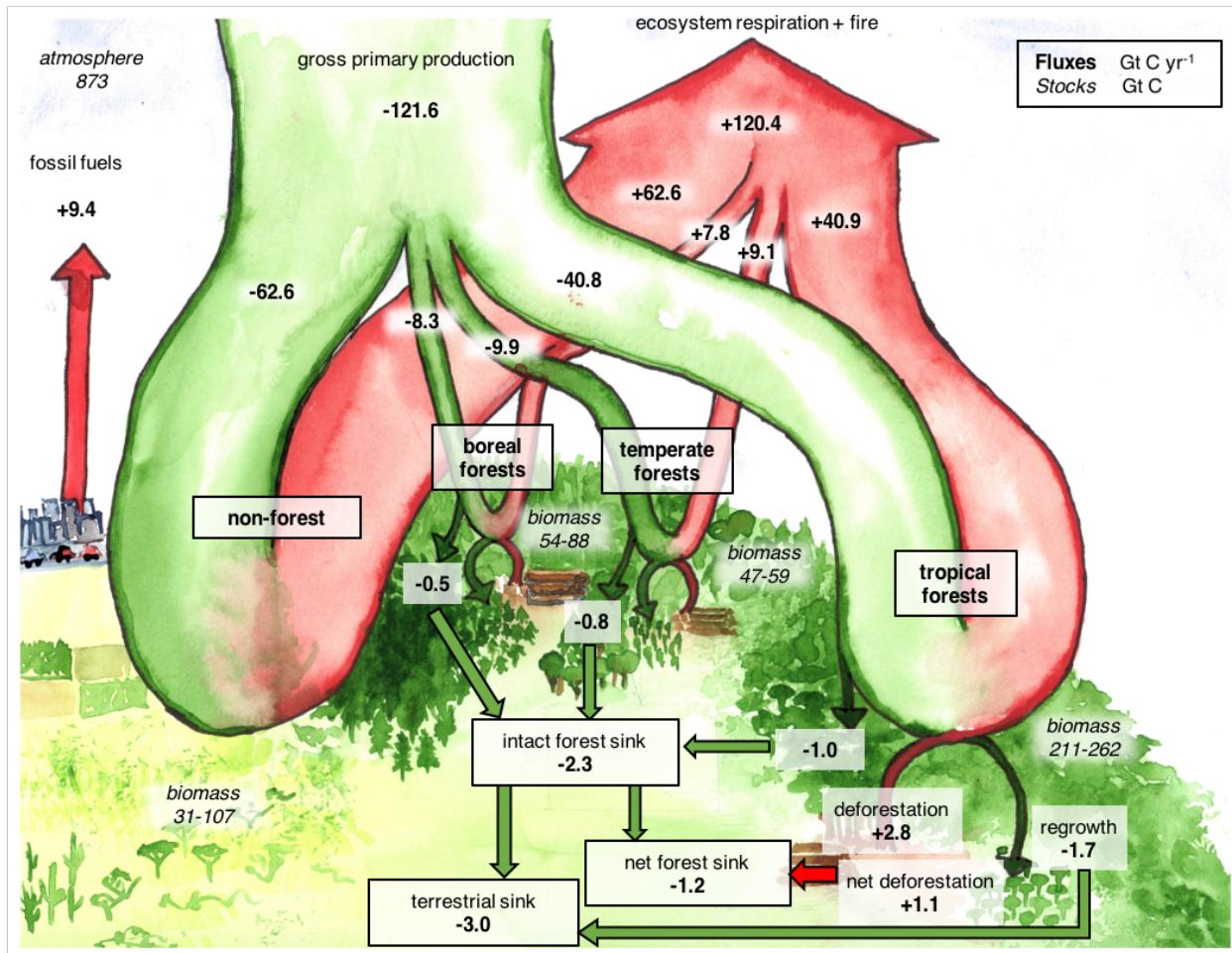
*Key words:* forest ecosystems; carbon cycle; stand age; productivity; respiration; biomass; global

## Background

(The Background section needs to present the rationale for why a systematic review of this topic is needed along with a history of what has been done to date and an expectation of what new will emerge from the review, especially if quantitative meta-analyses of studies are being considered.)

Forest ecosystems globally influence climate through their critical role in the global carbon (C) cycle (Fig. 1). Their annual gross CO<sub>2</sub> sequestration (gross primary productivity, GPP) is estimated at 59 Gt C yr<sup>-1</sup> (Beer et al 2010), or 6.3 times average annual fossil fuel emissions from 2007-2016 ( $9.4 \pm 0.5$  Gt C yr<sup>-1</sup>; Le Quéré et al 2017). A small portion of global terrestrial GPP is retained in ecosystems (mainly forests), resulting in a C sink that averaged  $3.0 \pm 0.8$  Gt C yr<sup>-1</sup> from 2007-2016, offsetting 32% of anthropogenic fossil fuel emissions (Le Quéré et al 2017). The remaining ~98% of global GPP is counterbalanced by ecosystem respiration (Reco) or wildfire. Perturbation to the global GPP- Reco balance can substantially influence the growth rate of atmospheric CO<sub>2</sub>; for example, the 2015-2016 El Niño, which brought historically high temperatures and low precipitation to the tropics, released an extra  $2.5 \pm 0.3$  Gt C to the atmosphere, resulting in the largest recorded atmospheric CO<sub>2</sub> growth rate (Le Quéré et al 2016, Liu et al 2017). In addition, forests contain substantial C stocks that, when disturbed, release significant amounts of CO<sub>2</sub> to the atmosphere. Although they cover only ~30% of the land surface, forests contain an estimated 92% of terrestrial biomass (Pan et al 2013) and 45% of terrestrial C (biomass and soils; Bonan 2008). Globally, gross tropical deforestation averaged 2.8 Gt C yr<sup>-1</sup> from 2000-2007, but ~40% of this was offset by forest regrowth, resulting in a net source of ~1.1 Gt C yr<sup>-1</sup> from tropical land use change (Pan et al 2011). This, coupled with minimal net deforestation in the extratropics and net uptake by intact forests (Pan et al 2011, Schimel et al 2015), resulted in a total gain in forest C of ~1.2-1.7 Gt C yr<sup>-1</sup> (Le Quéré et al 2017, Schimel et al 2015), thereby substantially slowing the rate of increase of atmospheric CO<sub>2</sub>.

Given their vital role in regulating atmospheric CO<sub>2</sub>, the future of Earth's forests will strongly influence the course of climate change. If forests globally respond to the suite of contemporary and future global change drivers—including elevated CO<sub>2</sub>, climate change, and atmospheric deposition—with increased productivity and net biomass increases, they could continue to act as a significant buffer against anthropogenic emissions.



**Figure 1 | The role of forests in the global carbon cycle.** Values apply to the period 2000–2018. Sources are as follows: biomass- (Pan et al 2013, Baccini et al 2012) IPCC 2000; intact forest sink, net forest sink, tropical deforestation and regrowth-(Pan et al 2011); terrestrial sink- (Le Quéré et al 2017); GPP- (Beer et al 2010); respiration+fire: calculated here; fossil fuel emissions- (Le Quéré et al 2017); atmospheric CO<sub>2</sub> (2018 value from <https://scripps.ucsd.edu/programs/keelingcurve/>). \*\*see issue 19: <https://github.com/forc-db/ERL-review/issues/19>\*\*

In contrast, if factors such as higher temperatures and drought stress lead to net C losses, the C sink of forests could be lost and even reversed. Global coupled climate models vary substantially in predictions regarding the future of the global forest C sink, with uncertainty in end-of-century terrestrial NEP exceeding current annual anthropogenic CO<sub>2</sub> emissions (Friedlingstein et al 2006, Cavalieri et al 2015). Anthropogenic land use decisions will also strongly influence the future course of atmospheric CO<sub>2</sub>. In the absence of measures to reduce deforestation, this will continue to be a significant component of anthropogenic CO<sub>2</sub> emissions. On the other hand, forest conservation, reforestation, and forest management hold strong potential to help stabilize and eventually reduce total anthropogenic CO<sub>2</sub> emissions (estimated total of 4.4 Gt C yr<sup>-1</sup>; Griscom et al 2017; *Cook-Patton et al.; other refs?*) and thereby contribute towards the goals of the Paris Climate Agreement (UNFCCC 2015, Houghton et al 2015; **REFS**). Indeed, forest-based climate mitigation was a key component of Paris Climate Agreement commitments, totaling approximately one-fourth of nations' planned emissions reductions (Grassi et al 2017; IPCC1.5). Thus, forest will strongly influence the future course of climate change both through their responses to global change and through forest-related anthropogenic land use decisions.

While forests will play a critical role in the future course of climate change, there remain significant uncertainties as to current and future C cycling the world's forests. Even as the climate changes and forest age distributions shift, there are important gaps in our fundamental understanding of how C cycling varies with climate and stand age. A relative dearth of data on C cycling in secondary forests, particularly in the tropics (Anderson-Teixeira et al 2016), is problematic in that almost 2/3 of the world's forests were secondary as of 2010 (FAO 2010), implying an under-filled need to characterize age-related trends in forest C cycling. Recent studies attempting to quantify the role of forests in the global C cycle (e.g., Pan et al 2011) have been hampered by insufficient data on regrowth rates of tropical forests worldwide, C in non-living pools (i.e., dead wood, litter, soil), and source/sink status of forests in some regions. Moreover, as discussed above, model representation of forest responses to global change remain highly uncertain (Friedlingstein et al 2006, Cavalieri et al 2015), with efforts to reduce model uncertainty in part limited by ready availability of appropriate benchmark data (Clark et al 2017, Luo et al 2012). To address the need for global-scale analyses of forest C cycling, we have developed an open-access Global Forest Carbon database, ForC (<https://github.com/forc-db/ForC>; Anderson-Teixeira et al 2016, 2018). ForC contains data on forest ecosystem C stocks and annual fluxes (>50 variables) and associated data required for interpretation (e.g., stand history, measurement methods) amalgamated from numerous previous data compilations and directly from original publications. ForC currently contains # records from # plots and # distinct geographic areas representing all forested biogeographic and climate zones.

Here, we synthesize ForC data to provide a big-picture overview of stand-level carbon cycling of the world's major forest biomes and how it varies with stand age. We address two broad questions:

1. How do C cycling vary across the worlds major forest biomes (i.e., tropical, temperate, boreal)?
2. How does C cycling vary with stand age?

While components of these questions have been previously addressed (e.g., REFS(Luyssaert et al 2007, Anderson et al 2006, Anderson-Teixeira et al 2013), our analysis represents by far the most comprehensive analysis of C cycling in global forests, and thereby stands to serve as a foundation for improved understanding of global forest C cycling.

## Methods/ Design

(The Methods/Design section needs to describe how articles in the reviews were identified and what criteria were used for justifying inclusion in the review. While traditional reviews relying on past experience and expert knowledge are acceptable, editors should encourage reviews that are set up as so-called 'systematic reviews' (see the Cochrane Review procedure which initiated the process in the medical and health science: <http://community.cochrane.org/about-us/evidence-based-health-care>.)

This review synthesizes data from the ForC database (<https://github.com/forc-db/ForC>; Anderson-Teixeira et al 2016, 2018). ForC amalgamates numerous intermediary data sets (e.g., REFS) and original studies. Original publications were referenced to check values and obtain information not contained in intermediary

data sets, although this process has not been completed for all records. The database was developed with goals of understanding how C cycling in forests varies across broad geographic scales and as a function of stand age. As such, there has been a focus on incorporating data from regrowth forests (e.g., Anderson et al 2006, Martin et al 2013, Bonner et al 2013) and obtaining stand age data when possible (83% of records in v.2.0; Anderson-Teixeira et al 2018). Particular attention was given to developing the database for tropical forests (Anderson-Teixeira et al 2016), yet these represented only approximately one-third of records in ForC v.2.0 (Anderson-Teixeira et al 2018). Since publication of ForC v.2.0, we added the following data to ForC: the Global Database of Soil Respiration Database (Bond-Lamberty and Thomson 2010; *hopefully Cook-Patton*; see GitHub list) .... We note that there remains a significant amount of relevant data that is not yet included in ForC, particularly biomass data from national forest inventories (e.g.: REFS). The database version used for this analysis has been tagged as a new release on Github (XX) and assigned a DOI through Zenodo (DOI: XX).

For this analysis, we grouped forests into four broad biome types (tropical broadleaf, temperate broadleaf, temperate needleleaf, and boreal needleleaf) and two age classifications (young and mature). Climate zones were defined according to Köppen-Geiger zones, which were extracted based on site geographic coordinates from the ESRI Köppen-Geiger map (downloaded June 2014 from

[http://maps3.arcgisonline.com/ArcGIS/rest/services/A-16/Köppen-Geiger\\_Observed\\_and\\_Predicted\\_Climate\\_Shifts/MapServer](http://maps3.arcgisonline.com/ArcGIS/rest/services/A-16/Köppen-Geiger_Observed_and_Predicted_Climate_Shifts/MapServer); Rubel and Kottek 2010) and are recorded in the Koeppen field of the ForC SITES table. Tropical climates were defined to include all equatorial (A) zones, temperate climates were defined to include all warm temperate (C) zones and warmer snow climates (Dsa, Dsb, Dwa, Dwb, Dfa, and Dfb), and boreal climates were defined to include the colder snow climates (Dsc, Dsd, Dwc, Dwd, Dfc, and Dfd). Leaf type (broadleaf / needleleaf) was defined based on descriptions in original publications (prioritized) or values extracted from a global map based on satellite observations (SYNMAP; Jung et al 2006) and recorded in the dominant.veg field of the ForC MEAUREMENTS table. Stand age was obtained from the stand.age field of MEASUREMENTS table, and forests classified as “young” (< 100 years) or “mature” ( $\geq$  100 years or classified as “mature”, “old growth”, “intact”, or “undisturbed” in original publication). Records for which stand age was unknown were excluded from the analysis. These groupings were defined in March 2018, at which point they covered 91% of the primary variable records for forests of known age, or 86% of total records, in ForC (v.2.0; Anderson-Teixeira et al 2018). The most well-represented forest types excluded were boreal broadleaf and boreal and temperate mixed broadleaf-needleleaf, each with <400 records total for stands of any age.

We drew upon records for # annual flux and # C stock variables (Table 1). For this analysis, we combined some of ForC’s specific variables (e.g., multiple variables for net primary productivity including various components) into more broadly defined variables (Table 1, this table)). Although ForC contains information that may be used to standardize or control for methodological differences (e.g., area sampled, min stem diameter sampled, allometric equations applied; Anderson-Teixeira et al 2018), for this analysis we included all relevant data in ForC. Throughout ForC, for all measurements drawing from tree census data (e.g., biomass, productivity), the minimum stem diameter sampled was  $\leq$  10cm. All records were measured directly or derived from field measurements (as opposed to modeled).

Analyses drew from ForC-simplified ([https://github.com/forc-db/ForC/blob/master/ForC\\_simplified](https://github.com/forc-db/ForC/blob/master/ForC_simplified)), which is a rearrangement of ForC intended to facilitate analyses. In generating ForC-simplified, all measurements originally expressed in units of dry organic matter (*OM*) were converted to units of C using the IPCC default of  $C = 0.47 * OM$  (IPCC 2006). Duplicate or otherwise conflicting records (indicated in conflicts field of MEASUREMENTS table) were reconciled as follows. Replicate measurements (i.e., replicates from within a single study) were averaged. Records that subsumed others—*i.e.*, the time period included that of  $\geq$  2 other records or dates were unknown and therefore conflicted with  $\geq$  2 other records—were removed. For each group of duplicate records—*i.e.*, measurements of the same variable in the same plot at the same time—one record was assigned precedence (recorded in D.precedence field). When measurement periods overlapped or were not specified, precedence was given first to records representing longer measurement periods (*i.e.*, end.date - start.date) and then to more recently published values. We manually reviewed duplicates that differed only in methodology, assigning precedence to the record employing a more comprehensive approach (e.g., inclusion of understory, lianas, or bamboo as opposed to just trees) or using a favored methodology.

Table 1. ForC variables included in this analysis. Shown are variable names and descriptions, the associated variable name(s) in the database, and number of plots with records. Complete list of variables with full definitions is available at [GitHub url].

Variable	Description	records	plots	geographic areas
<b>Annual fluxes</b>				
<i>NEE</i>	net ecosystem exchange or net ecosystem production (- indicates C sink)	n	n	n
<i>GPP</i>	gross primary production ( <i>NPP</i> + $R_{auto}$ or $R_{eco}$ - <i>NEE</i> )	n	n	n
<i>NPP</i>	net primary production ( <i>ANPP</i> + <i>BNPP</i> )	n	n	n
<i>ANPP</i>	aboveground <i>NPP</i>	n	n	n
<i>ANPP<sub>woody</sub></i>	woody production (stem growth + branch turnover)	n	n	n
<i>ANPP<sub>stem</sub></i>	stem production	n	n	n
<i>ANPP<sub>branch</sub></i>	branch turnover	n	n	n
<i>ANPP<sub>foliage</sub></i>	foliage production, typically estimated as annual leaf litterfall	n	n	n
<i>ANPP<sub>litterfall</sub></i>	litterfall, including leaves, reproductive structures, twigs, and sometimes branches	n	n	n
<i>ANPP<sub>repo</sub></i>	production of reproductive structures (flowers, fruits, seeds)	n	n	n
<i>ANPP<sub>folivory</sub></i>	foliar biomass consumed by folivores	n	n	n
<i>M<sub>woody</sub></i>	aboveground biomass of trees that died	n	n	n
<i>BNPP</i>	belowground NPP ( <i>BNPP<sub>coarse</sub></i> + <i>BNPP<sub>fine</sub></i> )	n	n	n
<i>BNPP<sub>coarse</sub></i>	coarse root production	n	n	n
<i>BNPP<sub>fine</sub></i>	fine root production	n	n	n
<i>R<sub>eco</sub></i>	ecosystem respiration	n	n	n
<i>R<sub>auto</sub></i>	autotrophic respiration ( $R_{auto-ag}$ + $R_{root}$ )	n	n	n
<i>R<sub>auto-ag</sub></i>	aboveground autotrophic respiration (i.e., leaves and stems)	n	n	n
<i>R<sub>root</sub></i>	root respiration	n	n	n
<i>R<sub>soil</sub></i>	soil respiration ( $R_{het-soil}$ + $R_{root}$ )	n	n	n
<i>R<sub>het-soil</sub></i>	soil heterotrophic respiration	n	n	n
<i>R<sub>het-ag</sub></i>	aboveground heterotrophic respiration	0	0	0
<i>R<sub>het</sub></i>	heterotrophic respiration ( $R_{het-ag}$ + $R_{het-soil}$ )	0	0	0
<b>Stocks</b>				
biomass	total live biomass	n	n	n
aboveground biomass	aboveground live biomass	n	n	n
woody aboveground biomass	woody component of aboveground biomass	n	n	n
foliage biomass	foliage biomass	n	n	n
root biomass	total root biomass	n	n	n
coarse root biomass	coarse root biomass	n	n	n
fine root biomass	fine root biomass C	n	n	n
deadwood	deadwood (standing+down)	n	n	n
standing deadwood	standing dead wood	n	n	n
down deadwood	fallen dead wood, including coarse and sometimes fine woody debris	n	n	n
organic layer	organic layer / litter/ forest floor	n	n	n

Records were filtered to remove plots that had undergone significant anthropogenic management or major disturbance since the most recent stand establishment (i.e., that reflected by stand.age). Specifically, we removed all plots flagged as managed in ForC-simplified (managed field). This included plots with any record of managements manipulating CO<sub>2</sub>, temperature, hydrology, nutrients, or biota, as well as any plots whose site or plot name contained the terms “plantation”, “planted”, “managed”, “irrigated”, or “fertilized”. Plots flagged as disturbed in ForC-simplified included stands that had undergone anthropogenic thinning or partial harvest (“Cut” or “Harvest” codes) unless this was very minor (percent.mortality= “minor”). We retained sites that were grazed or had undergone low severity natural disturbances (<10% mortality) including droughts, major storms, fires, and floods. We also removed all plots for which no stand history information had been retrieved.

Data were analyzed to produce basic summaries of C cycle patterns across biomes and stand ages following an approach similar to that of Anderson-Teixeira et al (2016). For mature forests, to obtain the values reported in the C cycle schematics, we first averaged any repeated measurements within a plot, weighting flux measurements according to the length of measurement periods (i.e., end.date - start.date). Values were then averaged across plots clustered within 25 km of one another (geographic.area field of SITES table, sensu Anderson-Teixeira et al 2018), weighting by area.sampled (MEASUREMENTS table) or plot.area (PLOTS table) if available for all records. This step was taken to avoid pseudo-replication and to combine any records from sites with more than one name in ForC. Finally, we computed statistics with geographic.area as the unit of replication. To compare across biomes, [Valentine, please describe]. There were enough data to run this analysis for all focal variables but XX.

For young (<100yrs) forest types, we employed a mixed effects model (XX in R) with biome and log<sub>10</sub>[stand.age] as fixed effects and plot nested within geographic.area as a random effect. When the effect

of stand.age was significant at  $p \leq 0.05$  and when each biome had records for stands of at least 10 different ages, a biome - stand.age interaction was included in the model. In the C cycle schematics for young forests, we report equations based on these models. In cases where there was no significant effect of stand.age, records were averaged as for mature stands.

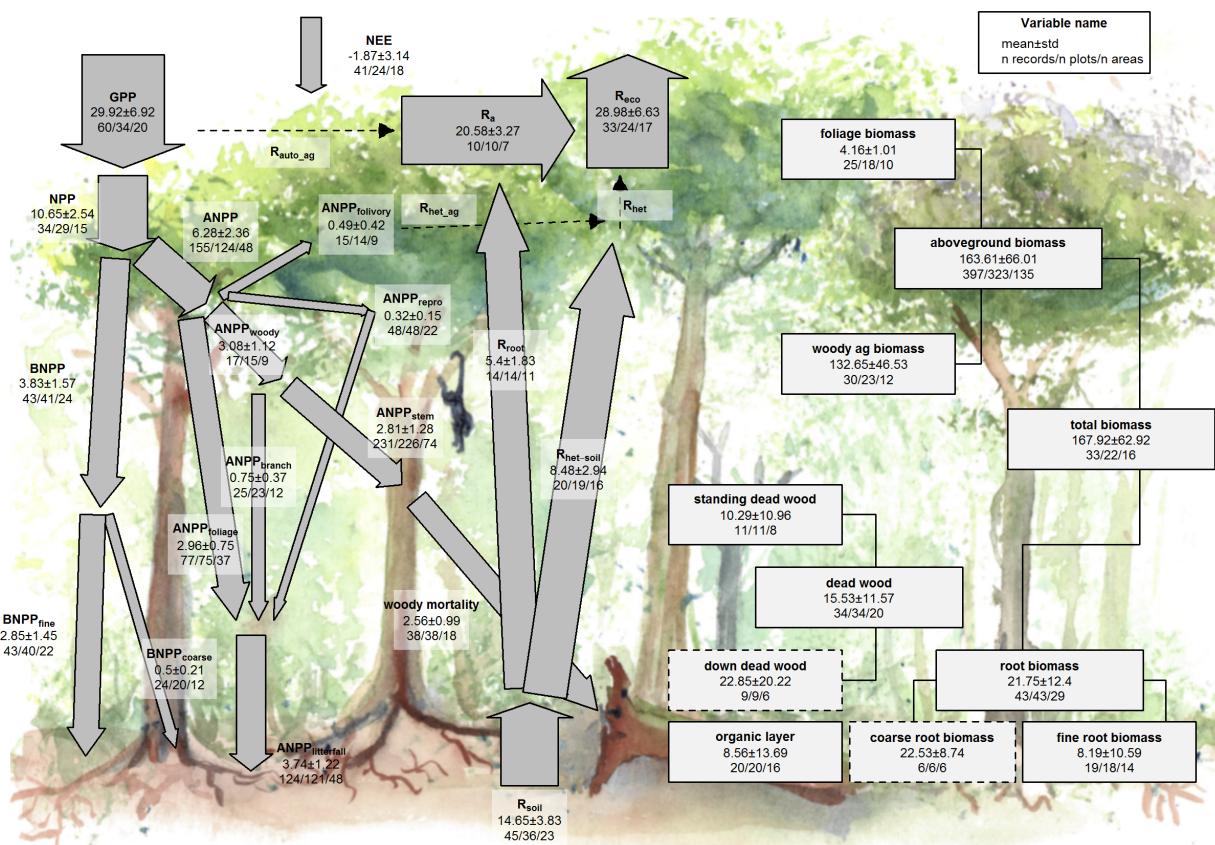
All database manipulation, analyses, and figure production were fully automated in R (version, citation). Materials required to fully reproduce these analyses, including data, R scripts, and image files, are archived in Zenodo (DOI: TBD]. Data, scripts, and results presented here are also available through the open-access ForC GitHub repository (<https://github.com/forc-db/ForC>), where many will be updated as the database develops.

## Review Results/ Synthesis

(*The Review Results/Synthesis section presents the findings of the analyses.*)

### C stocks and fluxes of mature forests

Average C cycles for tropical broadleaf, temperate broadleaf, temperate conifer, and boreal forests  $\geq 100$  years old and with no known major disturbance or significant anthropogenic management are presented in Figures 2-5, and statistics for each biome type are also summarized at: [GitHub URL]. Of the # flux and # stock variables mapped in these diagrams, ForC contained estimates from  $\geq 7$  distinct geographic areas for # fluxes and # stocks in tropical broadleaf forests, # fluxes and # stocks in temperate broadleaf forests, #fluxes and # stocks in temperate conifer forests, and fluxes and # stocks in boreal forests. For variables with records from  $\geq 7$  distinct geographic areas, these ensemble C budgets were generally consistent. That is, component fluxes and stocks summed to within 1 std of more inclusive fluxes in all but one instance (in temperate conifer forests, *abovegroundwoodybiomass + foliagebiomass > abovegroundbiomass + 1std*; Fig. 4). **This is based on my quick scan of the figures. Need to add a formal automated test:**  
<https://github.com/forc-db/ERL-review/issues/16>



**Fig. 2 | C cycle diagram for mature tropical broadleaf forests.** Presented are mean $\pm$ std, where geographically distinct areas are treated as the unit of replication. All units are Mg C ha $^{-1}$  yr $^{-1}$  (fluxes) or Mg C ha $^{-1}$  (stocks). Arrows indicate fluxes, boxes indicate stocks. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.