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

Authors:

Kristina J. Anderson-Teixeira[1,2]*

Valentine Herrmann[1]

Becky Banbury Morgan

Ben Bond-Lamberty

Susan Cook-Patton

Abigail Ferson

Jennifer McGarvey

Helene C. Muller-Landau[1]

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

Background. The fate of Earth's climate is closely linked to forests, which strongly influence atmospheric carbon dioxide (CO_2) 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. ForC v.XX yielded a fairly comprehensive picture of C cycling in the world's major forest biomes, with broad closure of C budgets. 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. This analysis yields a comprehensive, largely consistent picture of C cycling across the world's forests. *[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

Carbon (C) cycling in forests is central to sustaining life on Earth. Forest ecosystems globally influence climate through their vital role in regulating atmospheric carbon dioxide (CO_2). Their annual gross CO_2 sequestration (gross primary productivity, GPP) is estimated at >69 Gt C yr^{-1} (Bagdley et al. 2019), or >7 times average annual fossil fuel emissions from 2007-2016 (9.4 ± 0.5 Gt C yr-1; Le Quéré et al 2017) (**update**). While most of this enormous C flux is couterbalanced by CO_2 releases to the atmosphere through ecosystem respiration (R_{eco}) or fire, a small portion was retained in ecosystems over recent decades. The resulting in C sink averaged 3.0 ± 0.8 Gt C yr-1 from 2007-2016, offsetting 32% of anthropogenic fossil fuel emissions (Le Quéré et al 2017) (**update, give range**). Moreover, forests contain substantial C stocks: an estimated 92% of terrestrial biomass (Pan et al 2013) and 45% of terrestrial C (biomass and soils; Bonan 2008). Globally, net deforestation (*i.e.*, gross deforestation - regrowth) has been a source of CO_2 emissions, estimated at ~1.1 Gt C yr-1 from YEAR-YEAR (Pan et al 2011), resulting in net C sink of ~1.2-1.7 Gt C yr^{-1} across Earth's forests (Le Quéré et al 2017, Schimel *et al* 2015) (**update, give range**). The future of this sink is dependent both upon forest responses to a broad suite of global change drivers and to future land use decisions, and will strongly influence the course of climate change. Understanding, modeling, and managing forest-atmosphere CO_2 exchange is thereby central to efforts to mitigate climate change (Grassi *et al* 2017) (Friedlingstein et al 2006, Cavaleri et al 2015, Griscom et al 2017; IPCC1.5).

Despite the centrality of forest C cycling in regulating both atmospheric CO_2 and ecosystem function and diversity, we lack a comprehensive understanding of how C cycling varies with across biomes and with stand age. While remote sensing measurements are increasingly useful for global- or regional-scale estimates of forest GPP (Bagdley et al. 2019, (Li and Xiao 2019)), aboveground biomass (B_{ag}) (REFS), woody mortality (*i.e.*, B_{ag} losses to mortality M_{woody}) (Clark *et al* 2004, Leitold *et al* 2018), and to some extent net ecosystem exchange (NEP) (REFS), the remainder of C fluxes and stocks can only be quantified by ground-based measurements. Moreover, data on stand characteristics–including species composition, age, and disturbance history–are critical to interpreting and projecting into the future, but [usually rely on ground-based measurements]. Thus, ground-based measurements are, and will continue to be, central to characterizing forest C cycling.

Tens of thousands of ground-based forest C measurements have been published, but their distribution across literally thousands of scientific articles– along with variation in data formats, units, measurement methods, *etc.*–have made them effectively inaccessible for many global-scale analyses, including those attempting to quantify the role of forests in the global C cycle (*e.g.*, Pan et al 2011), and model evaluation (Clark et al 2017, Luo et al 2012). Important progress has been made in synthesizing data to address how C cycling varies across forest biomes (Luyssaert et al. 2007; REFS) and stand ages (REFS), yet generally consider only a limited set of C variables (but see Anderson-Teixeira *et al* 2016 for the tropics) and do not consider age-by-biome interactions. To address the need for global-scale analyses of forest C cycling, we have developed an open-access Global Forest Carbon database, ForC (Anderson-Teixeira *et al* (2016), Anderson-Teixeira *et al* (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 # (~49,000!) records from # plots and # distinct geographic areas representing all forested biogeographic and climate zones.

Here, we synthesize ForC data (Fig. 1) to provide a macroscopic overview of stand-level carbon cycling of the world's major forest biomes and how it varies with stand age. We address three broad questions:

1. To what extent can we fully represent, and "close", C budgets for each of the world's major forest biomes (*i.e.*, tropical, temperate broadleaf and deciduous, boreal) based on the current ForC data?
2. How do C cycling vary across the worlds major forest biomes?
3. How does C cycling vary with stand age?

While components of these questions have been previously addressed (Luyssaert *et al* 2007, Anderson-Teixeira *et al* 2016, pp @anderson–teixeira_forc_2018, @banbury_morgan_global_nodate), 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.

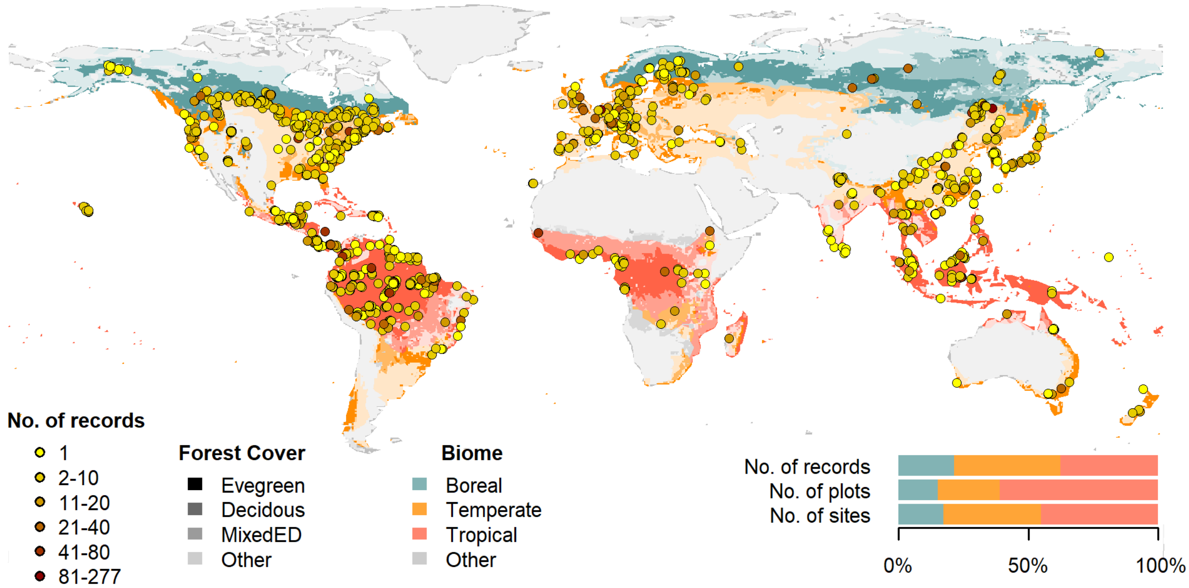


Figure 1 | Map of sites included in this analysis. Symbols are colored according to the number of records at each site. Underlying map shows coverage of evergreen, deciduous, and mixed forests (from SYNMAP; Jung et al. 2006) and biomes. Distribution of sites, plots, and records among biomes is shown in the inset.

Methods/ Design

This review synthesizes data from the ForC database (Fig. 1; https://github.com/forc-db/ForC; Anderson-Teixeira *et al* 2016, pp @anderson–teixeira_forc_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 (SRDB v.##; Bond-Lamberty and Thomson 2010), the Global Reforestation Opportunity Assessment database (GROA v1.0; Cook-Patton et al. in press; ZenodoDOI), and data from several publications (Taylor et al. 2017, Lutz et al. 2018, Johnson_2018_csss; see GitHub repo for complete list) (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: TBD).

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; Rubel and Kottek 2010). 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). Any forests in dry (B) and polar (E) Köppen-Geiger zones were excluded from the analysis. 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" (≥ 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 based on ForC v.2.0 (Anderson-Teixeira *et al* 2018), for which they covered 91% of the primary variable records for forests of known age, or 86% of total records. 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), 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 were reconciled as described in $APPENDIX\ A$. 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_2, 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, for figures 6 and 7, the original values were analyzed via a linear mixed effects model ('lmer' function in 'lme4' R package) with biome as fixed effect and plot nested within geographic.area as random effects on the intercept. When Biome had a significant effect, we looked at a Tukey's pairwise comparison to see which biomes were significantly different from one another. There were enough data to run this analysis for all focal variables but **biomass_ag_understory, BNPP_root.turnover_fine, deadwood_down, total.ecosystem_2.**

For young (<100yrs) forest types, we employed a mixed effects model with biome and log10[stand.age] as fixed effects and plot nested within geographic.area as a random effect on the intercept. 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

Data Coverage

Of the # records in ForC v.#, # met the criteria for inclusion in this study (Fig. 1). *(Give some stats on coverage by age and biome)*

C cycling in mature forests

Average C cycles for tropical broadleaf, temperate broadleaf, temperate conifer, and boreal forests ≥ 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 23 flux and 11 stock variables mapped in these diagrams, ForC contained estimates from ≥ 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 ≥ 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. 5). **update this: https://github.com/forc-db/ERL-review/issues/16**

Table 1. Carbon cycle variables included in this analysis, their sample sizes, and summary of biome differences and age trends.

Variable	Description	N records			biome differences*	age trend†
		records	plots	geographic areas		
Annual fluxes						
NEP	net ecosystem production or net ecosystem exchange (+ indicates C sink)	n	n	n	n.s.	-
GPP	gross primary production ($NPP + R_{auto}$ or $R_{eco} - NEE$)	n	n	n	Tr > TeB = TeN > B	+
NPP	net primary production ($ANPP + BNPP$)	n	n	n	Tr > TeB = TeN > B	+
$ANPP$	aboveground NPP	n	n	n	Tr > TeB ≥ TeN ≥ B	+, xB
$ANPP_{woody}$	woody production ($ANPP_{stem}$ + $ANPP_{branch}$)	n	n	n		
$ANPP_{stem}$	woody stem production	n	n	n		
$ANPP_{branch}$	branch turnover	n	n	n		
$ANPP_{foliage}$	foliage production, typically estimated as annual leaf litterfall	n	n	n		
$ANPP_{litterfall}$	litterfall, including leaves, reproductive structures, twigs, and sometimes branches	n	n	n		
$ANPP_{repro}$	production of reproductive structures (flowers, fruits, seeds)	n	n	n		
$ANPP_{folivory}$	foliar biomass consumed by folivores	n	n	n		
M_{woody}	woody mortality–i.e., B_{ag} of trees that die	n	n	n		
$BNPP$	belowground NPP ($BNPP_{coarse}$+ $BNPP_{fine}$)	n	n	n		
$BNPP_{coarse}$	coarse root production	n	n	n		
$BNPP_{fine}$	fine root production	n	n	n		
R_{eco}	ecosystem respiration (R_{auto}+ R_{het})	n	n	n		
R_{auto}	autotrophic respiration ($R_{auto-ag} + R_{root}$)	n	n	n		
$R_{auto-ag}$	aboveground autotrophic respiration (i.e., leaves and stems)	n	n	n		
R_{root}	root respiration	n	n	n		
R_{soil}	soil respiration ($R_{het-soil} + R_{root}$)	n	n	n		
$R_{het-soil}$	soil heterotrophic respiration	n	n	n		
R_{het-ag}	aboveground heterotrophic respiration	0	0	0		
R_{het}	heterotrophic respiration ($R_{het-ag} + R_{het-soil}$)	0	0	0		
Stocks						
B_{tot}	total live biomass (B_{ag}+B_{root})	n	n	n		
B_{ag}	aboveground live biomass ($B_{ag-wood}$+$B_{foliage}$)	n	n	n		
$B_{ag-wood}$	woody component of aboveground biomass	n	n	n		
$B_{foliage}$	foliage biomass	n	n	n		
B_{root}	total root biomass ($B_{root-coarse}$+$B_{root-fine}$)	n	n	n		
$B_{root-coarse}$	coarse root biomass	n	n	n		
$B_{root-fine}$	fine root biomass	n	n	n		
DW_{tot}	deadwood ($DW_{standing}$+DW_{down})	n	n	n		
$DW_{standing}$	standing dead wood	n	n	n		
DW_{down}	fallen dead wood, including coarse and sometimes fine woody debris	n	n	n		
OL	organic layer / litter/ forest floor	n	n	n		

* Tr: Tropical, TeB: Temperate Broadleaf, TeN: Temperate Needleleaf, B: Boreal, n.s.: no significant differences

† + or -: significant positive or negative trend, xB: significant age x biome interaction, n.s.: no significant age trend

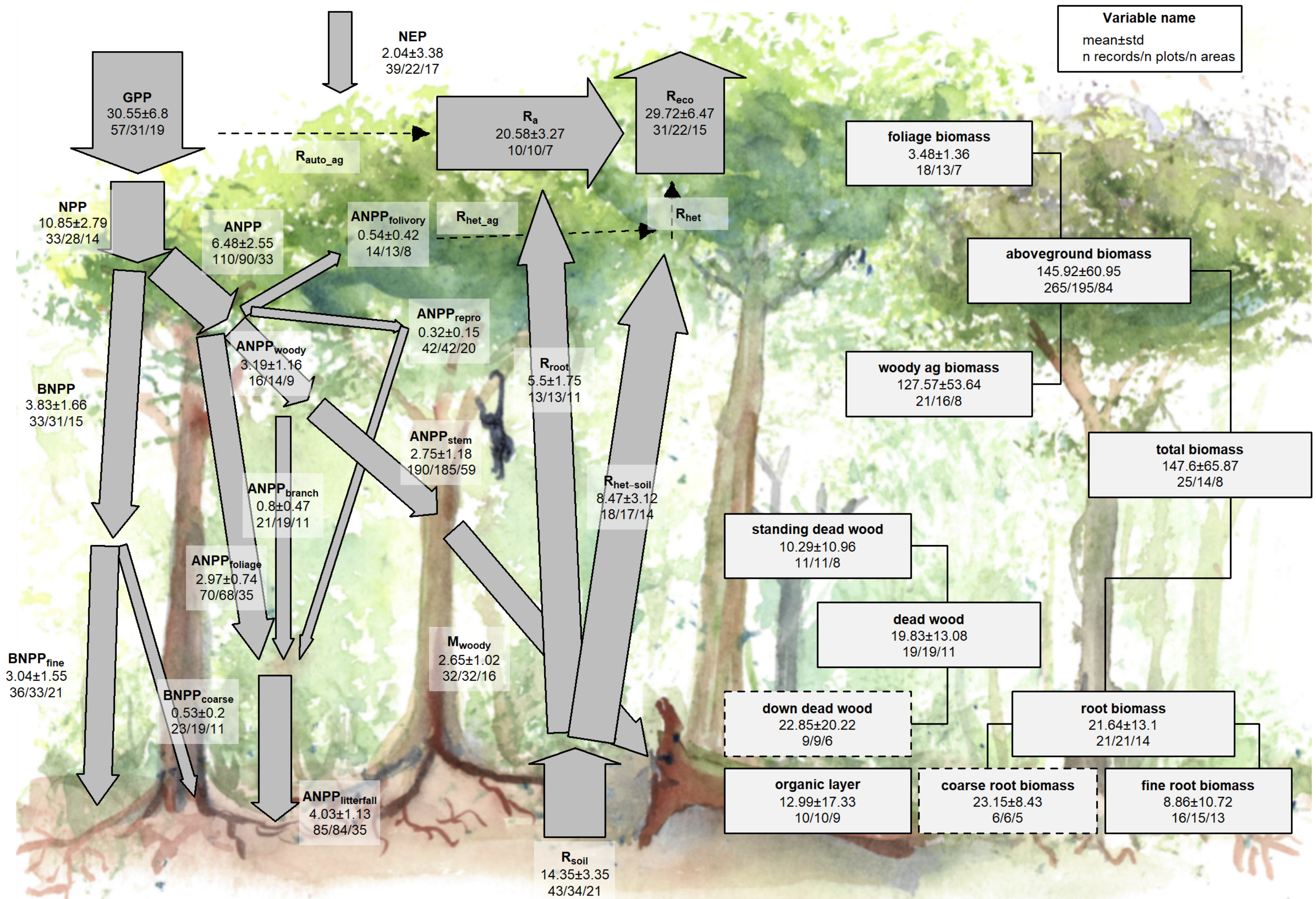


Figure 2 | C cycle diagram for mature tropical broadleaf forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data. Arrows are scaled as the **square root of flux divided by 5**.

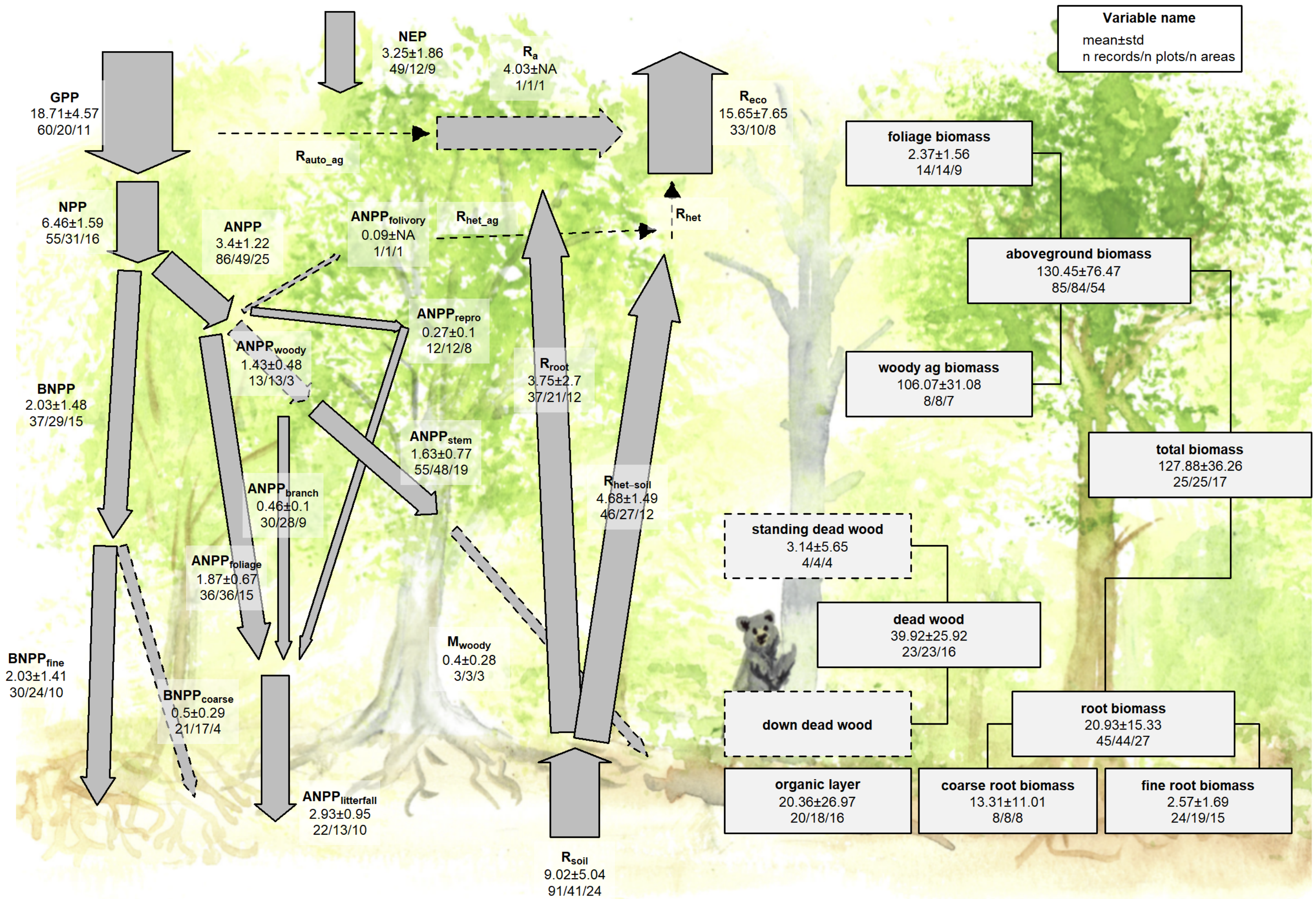


Figure 3 | C cycle diagram for mature temperate broadleaf forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.

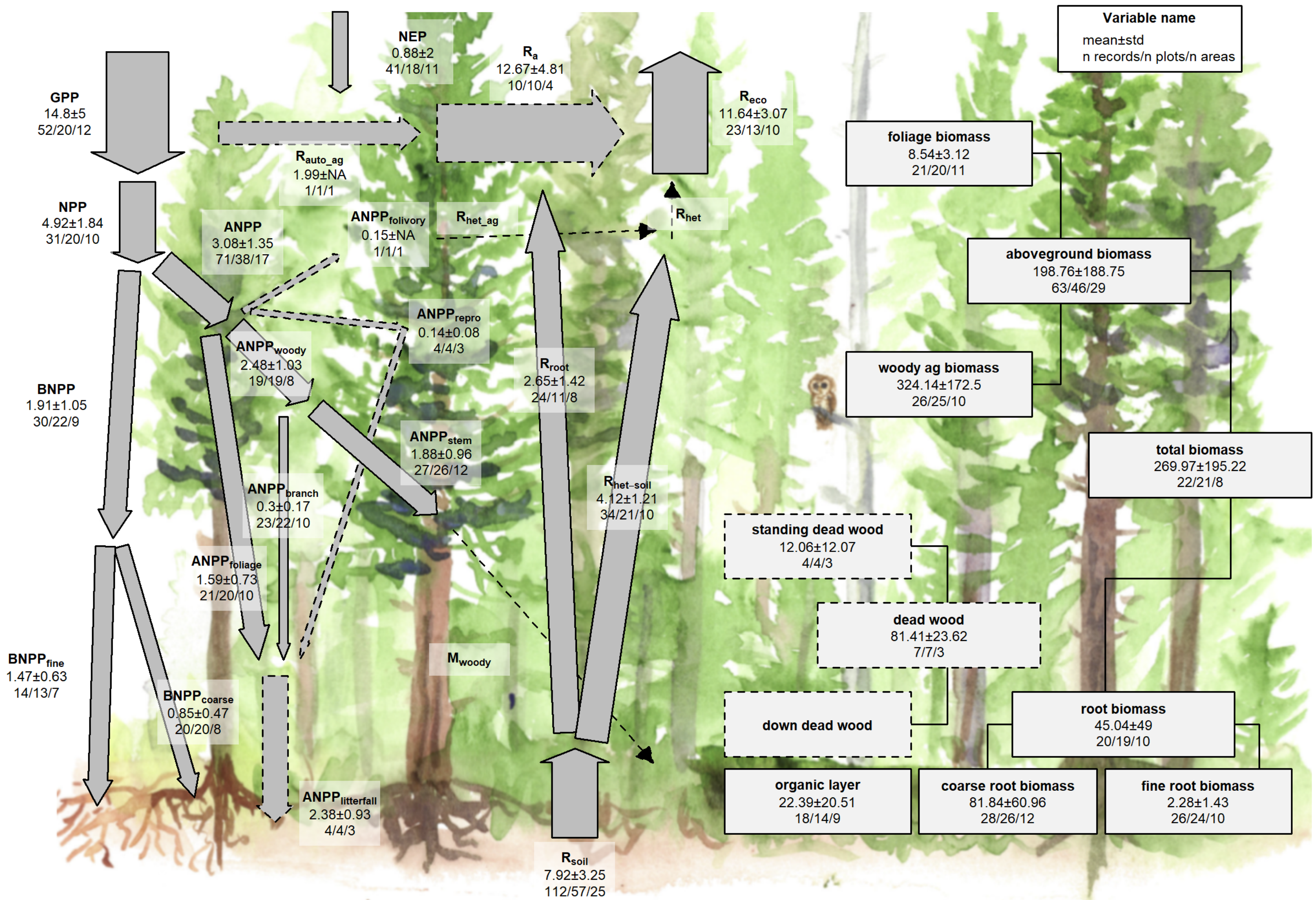


Figure 4 | C cycle diagram for mature temperate conifer forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.

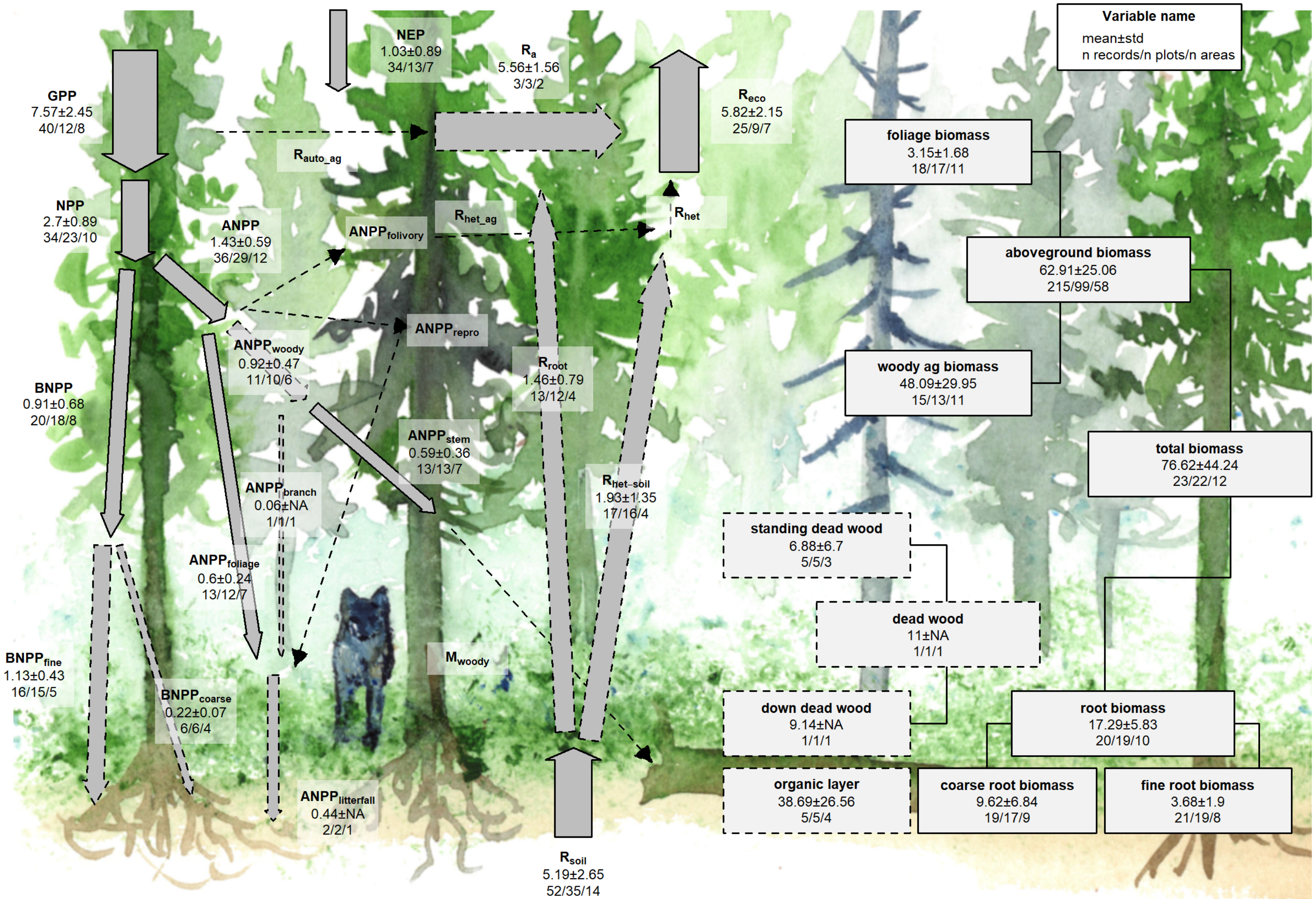


Figure 5 | C cycle diagram for mature boreal conifer forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.

(**check paragraph with latest data**) The largest C fluxes—including GPP, NPP, $ANPP$, $BNPP$, R_{soil_het}, R_{soil}, R_{eco}, — were highest in tropical forests, intermediate in temperate (broadleaf or conifer) forests, and lowest in boreal forests (ForC_variable_averages_per_Biome) (Fig. 6). The same held true for some of the subsidiary fluxes: ANPP_foliage, ANPP_woody, [**OTHERS?**]. Other subsidiary fluxes—including NPP_woody, ANPP_repro, ANPP_stem, ANPP_branch, woody.mortality, BNPP_coarse, BNPP_fine, [**OTHERS?**]—deviated from this pattern and/or lacked data for some biomes. Net ecosystem production (NEP) did not follow this pattern, with no significant differences across biomes but the largest average in temperate broadleaf forests, followed by temperate conifer, boreal, and tropical forests. Thus, C cycling rates generally decreased from tropical to temperate to boreal forests, but with less apparent trends for some of the subsidiary fluxes and an important exception in the overall C balance (NEP).

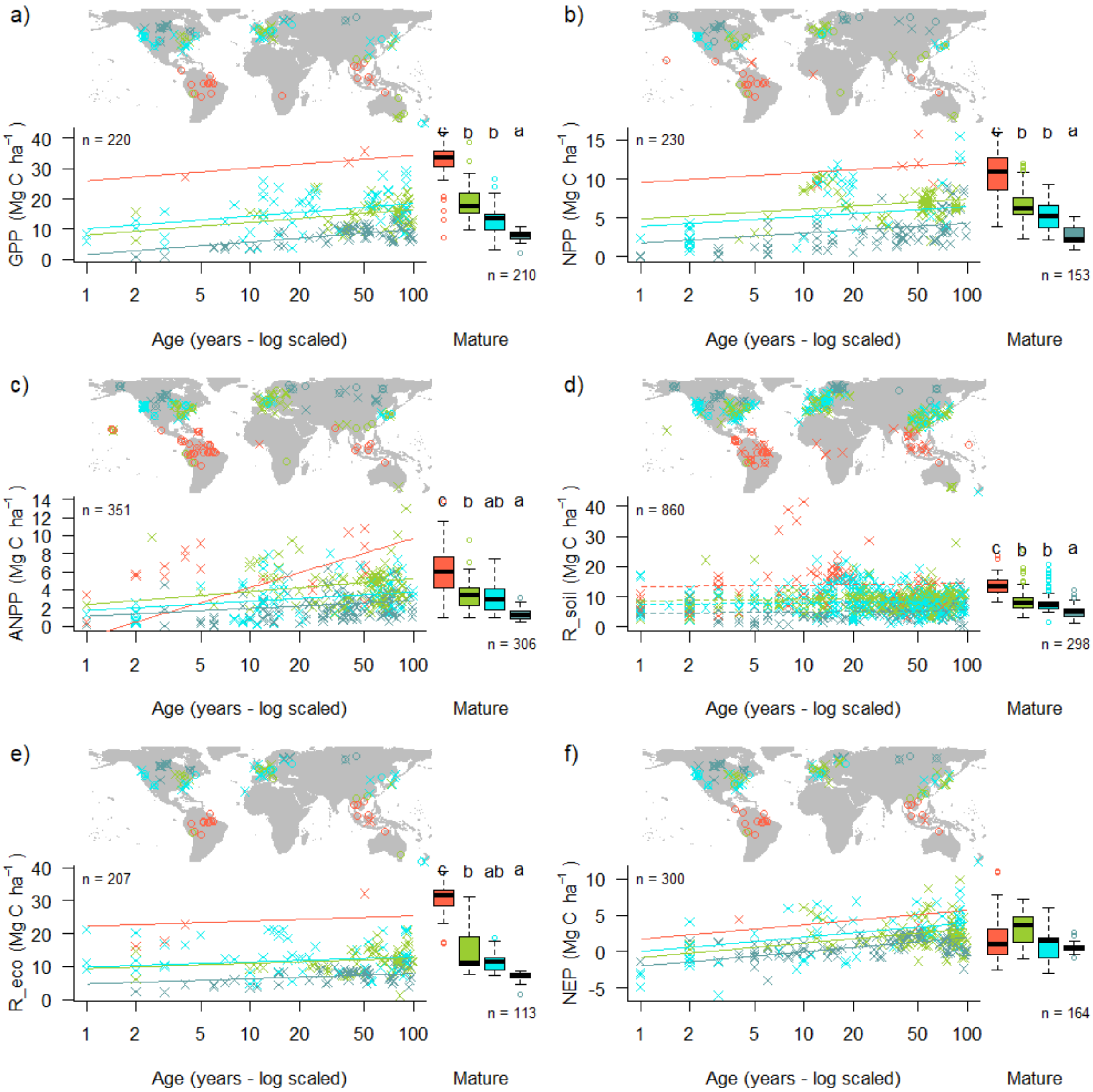


Figure 6 | Age trends and biome differences in some of the major C fluxes: (a) GPP, (b) NPP, (c) $ANPP$, (d) R_{soil}, (e) R_{eco}, and (f) NEP. Map shows data sources (x and o indicate young and mature stands, respectively). Lines... .

There were less distinct trends in C stocks across biomes (Fig. 7). Biome means for live aboveground and total biomass followed the same general trend as the major fluxes, with biomass of tropical= temperate broadleaf $\geq$ temperate conifer> boreal forests (Fig. 7a). However, the relative differences in these means were much smaller than for the major C fluxes, and there was significant variation within biomes. Maximum aboveground biomass values followed a very different trend than the means: temperate broadleaf> temperate conifer>boreal>tropical. There were some statistically significant biome differences in less frequently sampled C stocks (*e.g.*, woody biomass, foliage biomass, deadwood), but given high within-biome variability in C stocks and relatively low sample sizes, these were likely attributable to sampling biases and methodological differences than to true differences across biomes.

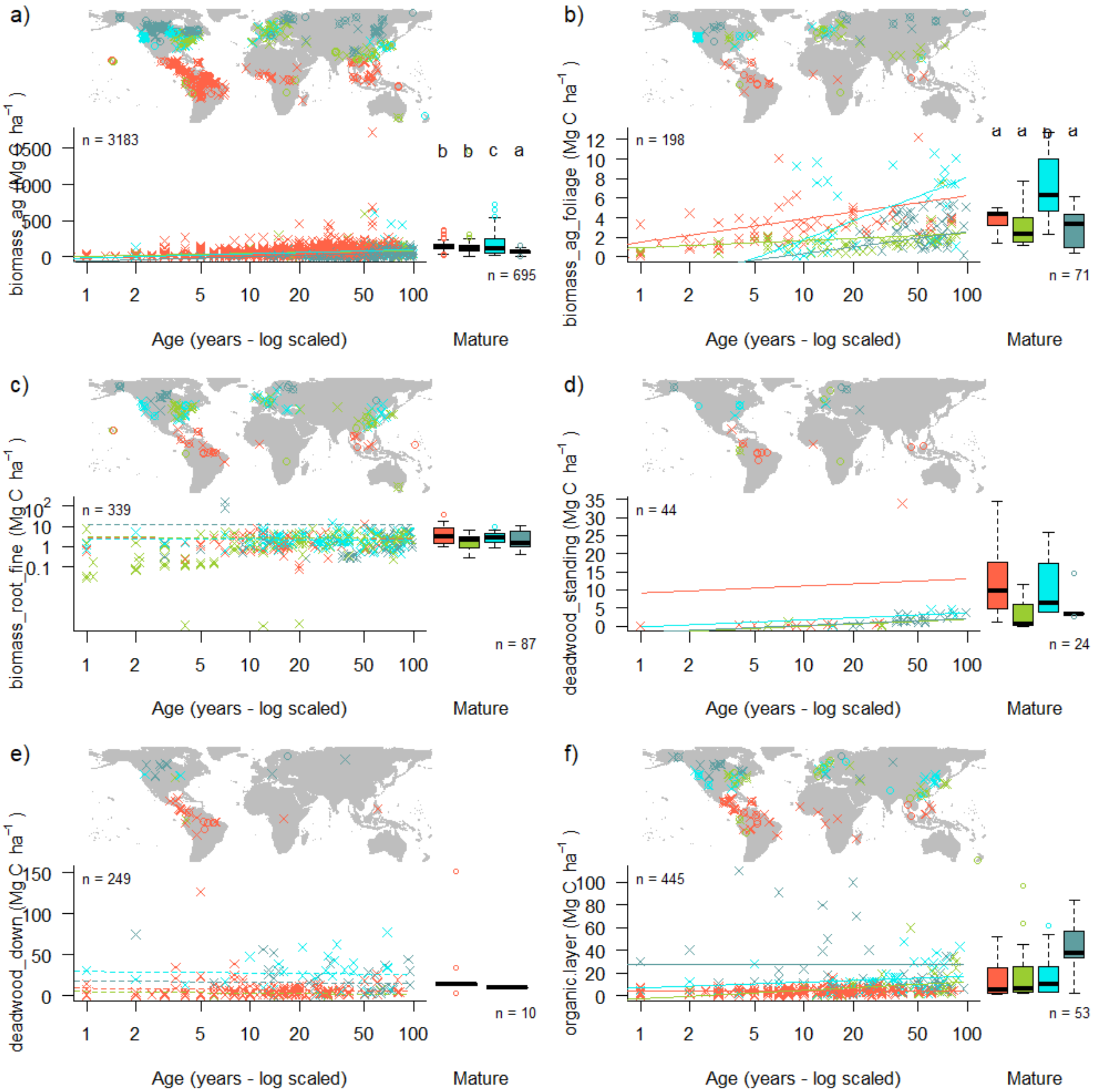


Figure 7 | Age trends and biome differences in some of the major forest C stocks: (a) aboveground biomass, (b) foliage, (c) fine roots, (d) dead wood.

C cycling in young forests

(**check paragraph with latest data**) Average C cycles for forests <100 years old are presented in Figures 8-11. In general, ForC contained roughly comparable information on C fluxes in regrowth vs mature forests, with the notable exception of tropical forests, for which there were no fluxes that had been measured in ≥ 7 different geographic areas. C stocks were better documented for regrowth stands, with a total of # records from # distinct geographic areas. Both C stocks and fluxes commonly displayed significant trends with stand age for within-biome analyses (Fig. 5-10; detailed below). Differences across biomes typically paralleled those observed for mature forests, with C cycling generally most rapid in the tropics and slowest in boreal forests

(Figs. 6-7).

(**check paragraph with latest data**) ForC contained 14 flux variables with sufficient data for cross-biome analyses of age trends in regrowth forests (see Methods) (Fig. 6-7 and **S#- SI figures including plots for all variables**). Of these, 9 increased significantly with log10[stand.age]: GPP, NPP, $ANPP$, $ANPP_{foliage}$, $ANPP_{woody}$, $ANPP_{woody-stem}$, $BNPP$, $BNPP_{root-fine}$, R_{eco}, and net C sequestration (NEP). The remaining five–$ANPP_{woody-branch}$, $BNPP_{root-coarse}$, $R_{soil-het}$, and $R_{soil-het}$-–displayed no significant relationship to stand age, although all displayed a positive trend. In terms of C stocks, 10 variables had sufficient data to test for age trends. Six of these—total biomass, aboveground biomass, aboveground woody biomass, foliage biomass, root biomass, and coarse root biomass—increased significantly with $log10[stand.age]$. The remaining four displayed non-significant positive trends: fine root biomass, total dead wood, standing dead wood, and organic layer.

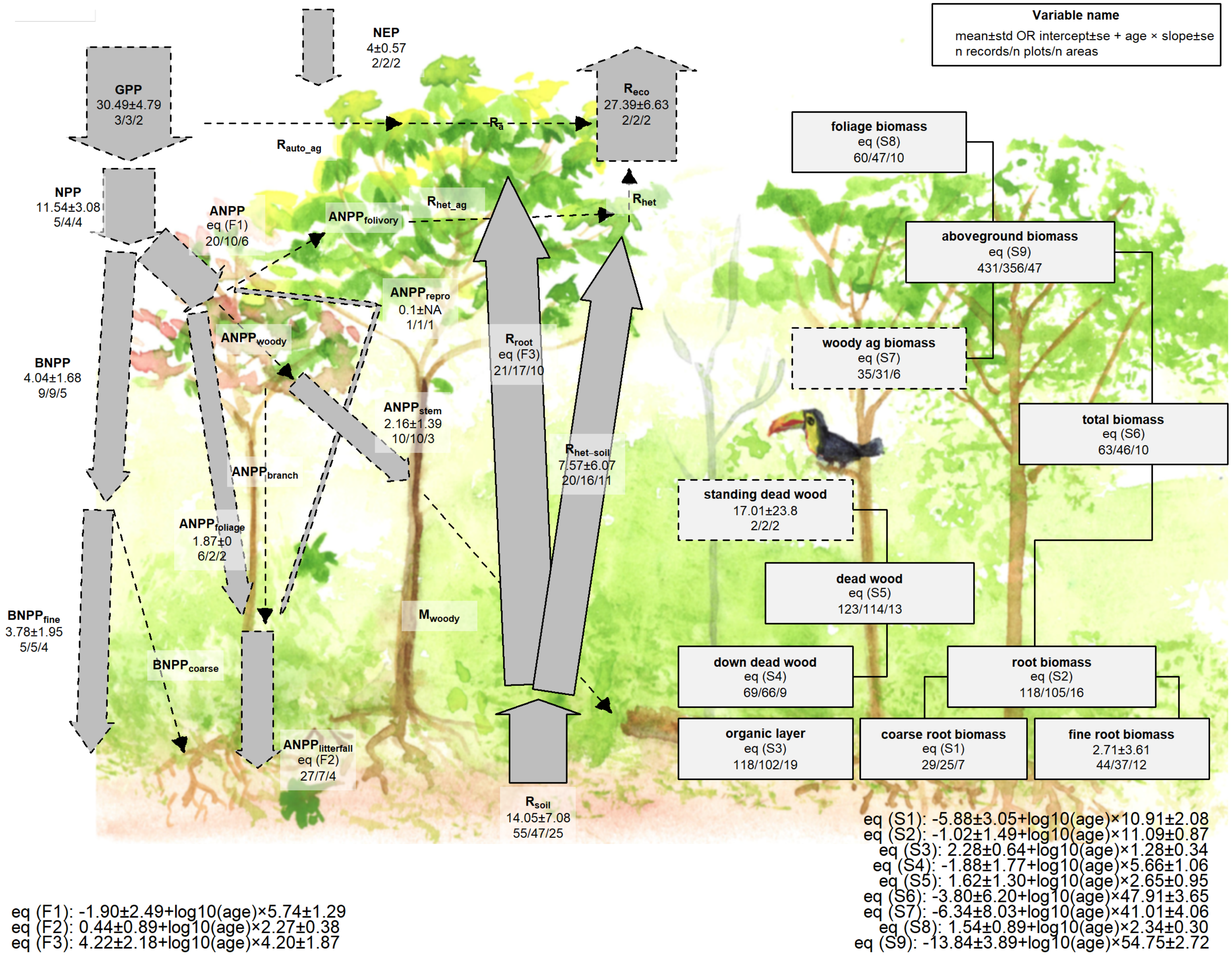


Figure 8 | C cycle diagram for young tropical broadleaf forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. When age trends are signficant, they are presented with numbered equations; otherwise, means are presented. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.

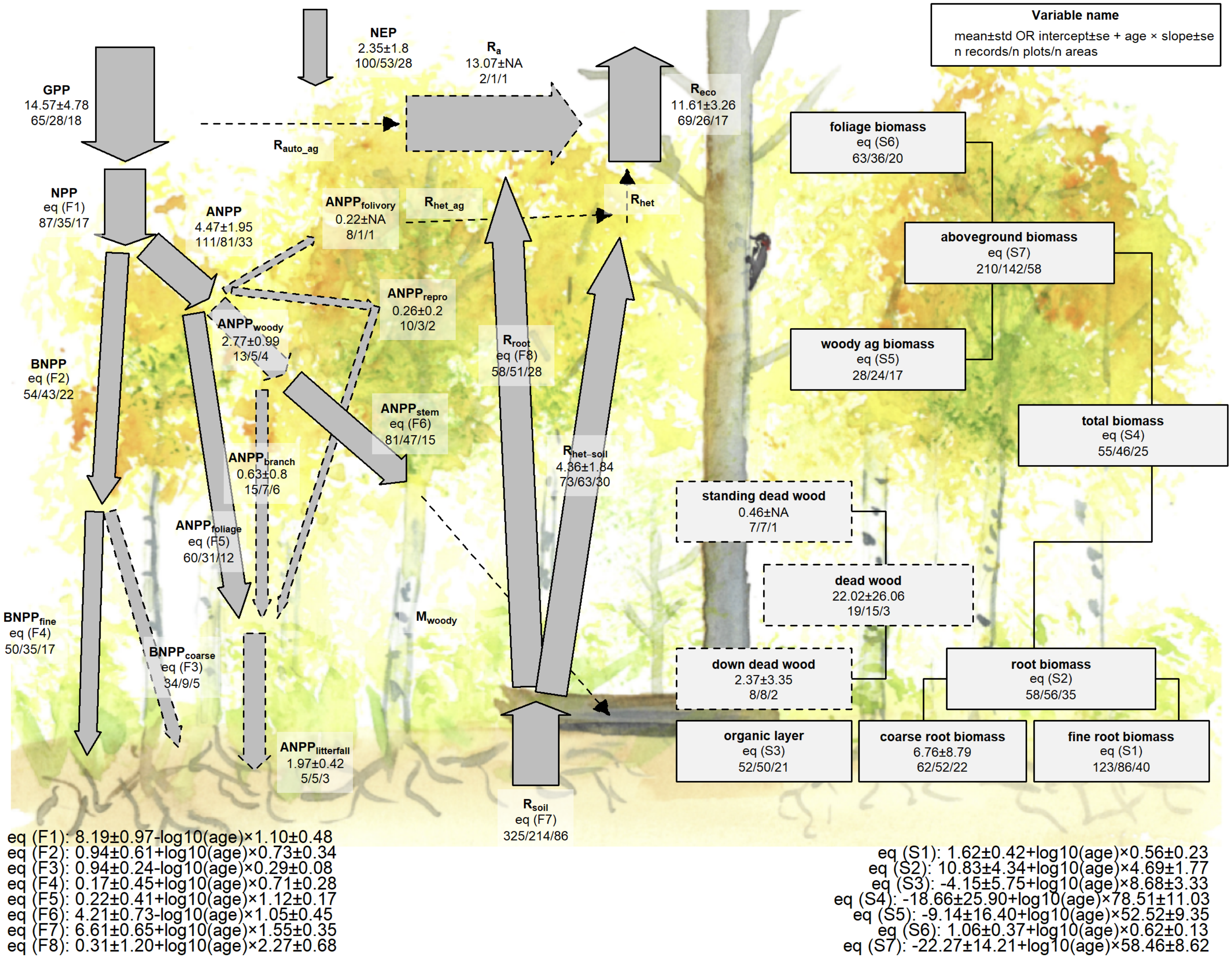


Figure 9 | C cycle diagram for young temperate broadleaf forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. When age trends are signficant, they are presented with numbered equations; otherwise, means are presented. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.

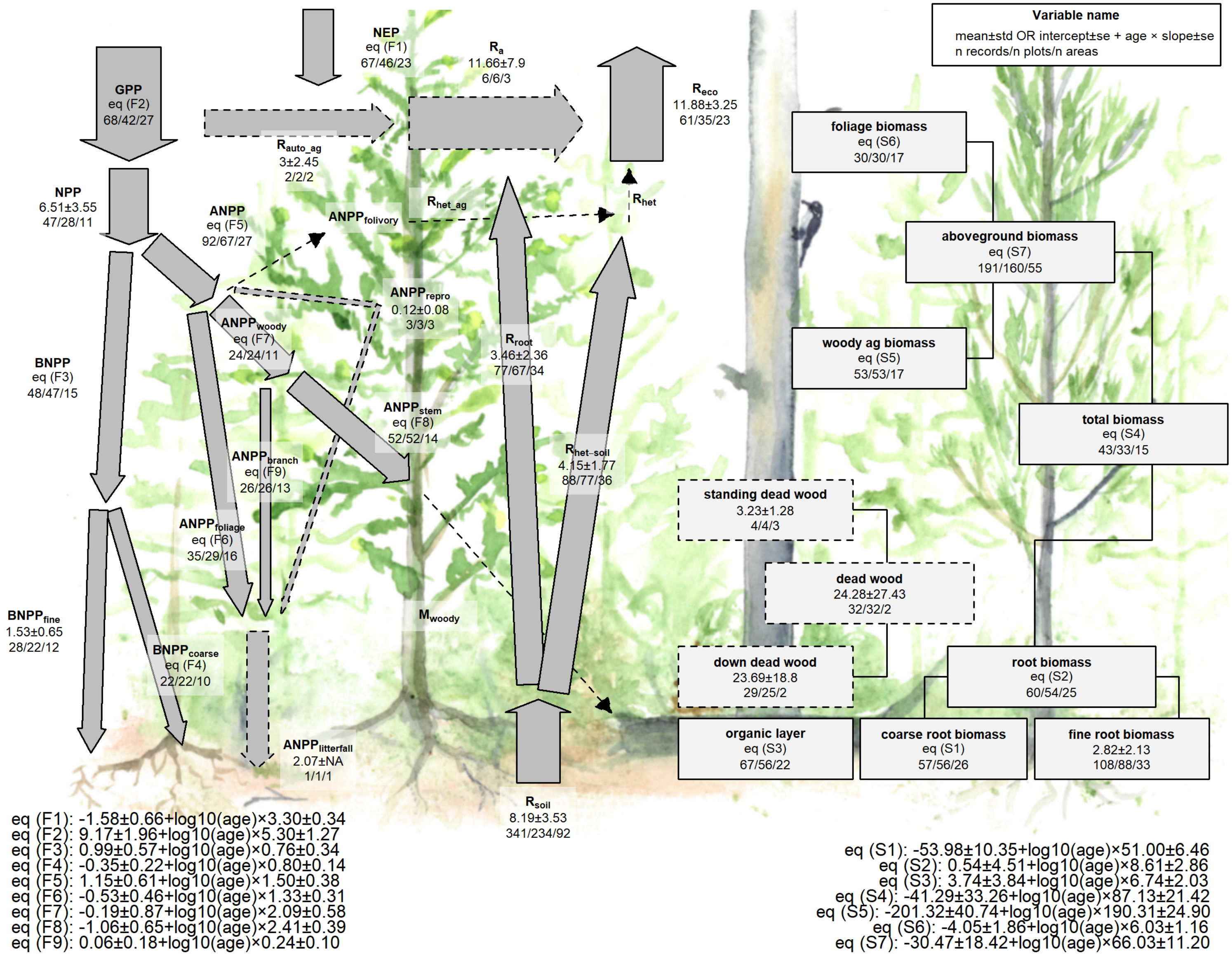
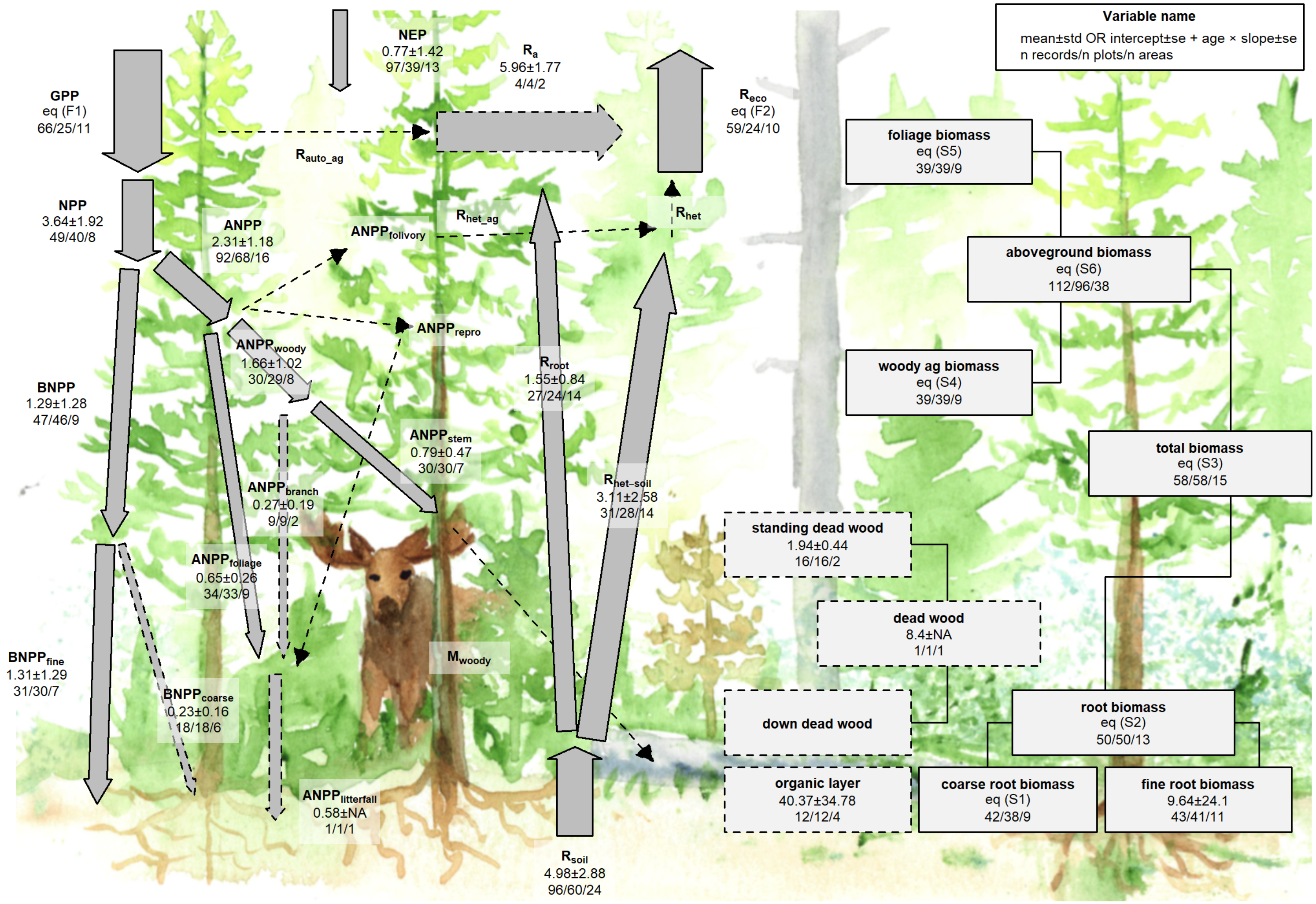


Figure 10 | C cycle diagram for young temperate conifer forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. When age trends are signficant, they are presented with numbered equations; otherwise, means are presented. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.



eq (S1): -9.63±5.51+log10(age)×9.70±2.82
eq (S2): -28.97±7.57+log10(age)×23.86±4.16
eq (S3): -72.38±33.63+log10(age)×74.16±18.41
eq (S4): -86.21±59.47+log10(age)×73.61±33.56
eq (S5): -1.36±1.79+log10(age)×2.19±1.00
eq (S6): -41.46±16.53+log10(age)×50.97±9.11

eq (F1): 3.54±1.22+log10(age)×2.96±0.63
eq (F2): 4.47±1.12+log10(age)×1.76±0.54

Figure 11 | C cycle diagram for young boreal conifer forests. All units are Mg C ha^{-1} yr^{-1} (fluxes) or Mg C ha^{-1}. Presented are mean ± std, where geographically distinct areas are treated as the unit of replication. Arrows indicate fluxes, boxes indicate stocks. When age trends are signficant, they are presented with numbered equations; otherwise, means are presented. Dashed shape outlines indicate variables with records from <7 distinct geographic areas, and dashed arrows indicate fluxes with no data.

Discussion

ForC v.XX yielded a fairly comprehensive and internally consistent picture of C cycling in the world's major forest biomes. Carbon cycling rates generally increased from boreal to tropical regions and with stand age. Specifically, the major C fluxes were highest in tropical forests, intermediate in temperate (broadleaf or conifer) forests, and lowest in boreal forests – a pattern that generally held for regrowth as well as mature forests (Figs. 6-7). In contrast to C fluxes, there was little directional variation in mature forest C stocks across biomes (Figs. 2-5, 7). The majority of flux variables, together with most live biomass pools, increased significantly with stand age (Figs. 6-11). Together, these results indicate that, moving from cold to tropical climates and from young to old stands, there is a general acceleration of C cycling, whereas C stocks of mature forests are influenced by a different set of drivers.

C variable coverage and budget closure

ForC provides fairly good coverage of most major variables. *(discuss how this improves upon previous data compilations/ for which variables does ForC make the greatest difference (e.g., not AGB or NEP/GPP/Reco, but by far the latest data compilation for dead wood, [variables]) (Noteable holes include: fluxes: R_auto_ag, woody mortality, folivary/ herbivory and respiration of herbivores (and therefore total R_het), ANPP_repro; also fluxes in tropical regrowth forests) For the C stocks considered here, the most poorly covered is dead wood (none in E hemisphere!), despite a focused effort on this variable that has resulted in ForC being by far the largest collection of these data.* Thus, overall, we're lacking coverage of fluxes to herbivores and higher consumers, along with the woody mortality and dead wood. Geographically, all variables poorly covered in Africa and Siberia.

Closure of the C cycle budgets for mature forests (Figs. 2-5) is fairly good. However, SD's are often large, reflective of significant within-biome variation. This makes the standard for closure relatively loose. Lack of closure, in the few instances where it occurs, is probably more reflective of differences in the representation of forest types (e.g., disproportionate representation of US Pacific NW for aboveground woody biomass relative to AGB; Fig. 4) than of methodological accuracy. Thus, overall, a high degree of closure implies that ForC gives a consistent picture of C cycling within biomes. While these means are unlikely to be accurate representations of C cycling within any particular forest, they provide a useful baseline for comparison.

C cycling across biomes

THere is a general acceleration of carbon cycling from the tropics to the high latitudes. For mature forests, this is consistent with lots of previous work, most recently Banbury Morgan *et al* (n.d.). For regrowth forests, we see faster C accumulation in the tropics (*I assume!- analysis not yet complete*), consistent with the observation that biomass accumulation decreases with latitude (Anderson 2006, Cook-Patton et al. 2020). For most C flux variables, this analysis is the first to examine flux trends in regrowth forests across biomes (i.e., age x biome interaction). Data remain sparse, but for better-represented variables, we often see faster acceleration of C cycling in the warmer climates. Further work will be required to explore age x climate interactions, but our broad-brush overview indicates that C cycling of regrowth forests is not only higher in the tropics, parallel to fluxes in mature forests (Banbury Morgan *et al* n.d.), but also that it accelerates more rapidly with stand age in the tropics, consistent with more rapid biomass accumulation.

Age trends in C cycling (Just some rough notes at this point)

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.

Moreover, as disturbances increase, understanding the carbon dynamics of regrowth forests will be increasingly important.

NEP increases with log(age) to 100 –> strongest C sinks are established secondary forests.

Relevance for climate change prediction and mitigation

The future of forest C cycling will shape trends in atmospheric CO_2 and the course of climate change. For a human society seeking to understand and mitigate climate change, the data contained in ForC and summarized here can help to meet two major challenges.

First, improved representation of forest C cycling in models is essential to improving predictions of the future course of climate change. To ensure that models are giving the right answers for the right reasons, it is important benchmark against multiple components of the C cycle. By making tens of thousands of records readily available in standardized format, ForC makes it feasible for the modeling community to draw upon these data to benchmark models. Integration of ForC with models is a goal (Fer et al., in revision). On a more cursory level, the values summarized here can serve as a sanity check for modelers to determine whether model predictions for multiple C cycle variables are reasonable–i.e., within the range of previous observations.

Second, ForC can serve as a pipeline through which forest science can inform forest-based climate change mitigation efforts. Such efforts will be most effective when informed by the best available data, yet it is not feasible for the individuals and organizations designing such efforts to sort through literature, often behind paywalls, with data reported in varying units, terminology, etc. One goal for ForC is to serve as a pipeline through which information can flow efficiently from forest researchers to decision-makers working to implement forest conservation strategies at global, national, or landscape scales. This is already happening! ForC has already contributed to updating the IPCC guidelines for carbon accounting in forests (IPCC 2019, Requena-Suarez et al 2019, Rozendaal et al in prep), mapping C accumulation potential from natural forest regrowth globally (Cook-Patton et al. 2020), and informing ecosystem conservation priorities (Goldstein et al. 2020).

There remain numerous data needs for improved accounting of forest carbon stocks and fluxes in GHG accounting. AGB is the largest stock, and most of the emphasis is on this variable. Remote sensing, with calibration based on high-quality ground-based data (Schepaschenko et al. 2019, Chave et al. 2019), is the best approach for mapping forest carbon (REFS). However, it is limited in that it is not associated with stand age and disturbance history, except in recent decades (e.g., Hansen et al.; REFS). ForC is therefore valuable in defining age-based trajectories in biomass, as in Cook-Patton et al. 2020.

Other variables cannot be remotely sensed. In terms of C stocks, there is a paucity of data on dead wood and organic layer (Pan et al. ?). These can be significant. *(give some stats/ cite figures)*. ForC does not include soil carbon, which is covered by other efforts (REFS). For fluxes, Fluxnet is the keeper of the best data on NEE, GPP, Reco (REFS), and SRDB remains the authority on soil respiration (REFS). ForC includes recent data from both, but is not continuously integrated. For C is the best source for most of the subsidiary fluxes: NPP, woody mortality…

We recommend that use of ForC data go to the original database, as opposed to useing "off-the-shelf" values from this publication. This is because (1) ForC is constantly being updated, (2) analyses should be designed to match the application, (3) age equations presented here all fit a single functional form that is not necessarily the best possible for all the variables.

As climate change accelerates, understanding and managing the carbon dynamics of forests will be critical to forecasting, mitigation, and adaptation. The C data in ForC, as summarized here, will be valuable to these efforts.

Acknowledgements

All researchers whose data is included in ForC and this analysis. Thanks to Norbert Kunert and [Helene's intern] for helpful input at an earlier phase. A Smithsonian Scholarly Studies grant to KAT and HML. WLS grant to KAT.

References

Anderson-Teixeira K J, Wang M M H, McGarvey J C, Herrmann V, Tepley A J, Bond-Lamberty B and LeBauer D S 2018 ForC: A global database of forest carbon stocks and fluxes *Ecology* **99** 1507–7 Online: http://doi.wiley.com/10.1002/ecy.2229

Anderson-Teixeira K J, Wang M M H, McGarvey J C and LeBauer D S 2016 Carbon dynamics of mature and regrowth tropical forests derived from a pantropical database (TropForC-db) *Global Change Biology* **22** 1690–709 Online: http://onlinelibrary.wiley.com/doi/10.1111/gcb.13226/abstract

Banbury Morgan B, Herrmann V, Kunert N, Bond-Lamberty B, Muller-Landau H C and Anderson-Teixeira K J Global patterns of forest autotrophic carbon fluxes *Global Change Biology*

Bond-Lamberty B and Thomson A 2010 A global database of soil respiration data *Biogeosciences* **7** 1915–26 Online: http://www.biogeosciences.net/7/1915/2010/

Clark D B, Castro C S, Alvarado L D A and Read J M 2004 Quantifying mortality of tropical rain forest trees using high-spatial-resolution satellite data *Ecology Letters* **7** 52–9 Online: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1461-0248.2003.00547.x

Grassi G, House J, Dentener F, Federici S, Elzen M den and Penman J 2017 The key role of forests in meeting climate targets requires science for credible mitigation *Nature Climate Change* **7** 220–6 Online: https://www.nature.com/articles/nclimate3227

Leitold V, Morton D C, Longo M, dos-Santos M N, Keller M and Scaranello M 2018 El Niño drought increased canopy turnover in Amazon forests *New Phytologist* **219** 959–71 Online: http://doi.wiley.com/10.1111/nph.15110

Li X and Xiao J 2019 Mapping Photosynthesis Solely from Solar-Induced Chlorophyll Fluorescence: A Global, Fine-Resolution Dataset of Gross Primary Production Derived from OCO-2 *Remote Sensing* **11** 2563 Online: https://www.mdpi.com/2072-4292/11/21/2563

Luyssaert S, Inglima I, Jung M, Richardson A D, Reichstein M, Papale D, Piao S L, Schulze E-D, Wingate L, Matteucci G, Aragao L, Aubinet M, Beer C, Bernhofer C, Black K G, Bonal D, Bonnefond J-M, Chambers J, Ciais P, Cook B, Davis K J, Dolman A J, Gielen B, Goulden M, Grace J, Granier A, Grelle A, Griffis T, Grünwald T, Guidolotti G, Hanson P J, Harding R, Hollinger D Y, Hutyra L R, Kolari P, Kruijt B, Kutsch W, Lagergren F, Laurila T, Law B E, Maire G L, Lindroth A, Loustau D, Malhi Y, Mateus J, Migliavacca M, Misson L, Montagnani L, Moncrieff J, Moors E, Munger J W, Nikinmaa E, Ollinger S V, Pita G, Rebmann C, Roupsard O, Saigusa N, Sanz M J, Seufert G, Sierra C, Smith M-L, Tang J, Valentini R, Vesala T and Janssens I A 2007 CO2 balance of boreal, temperate, and tropical forests derived from a global database *Global Change Biology* **13** 2509–37 Online: http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2486.2007.01439.x

Schimel D, Pavlick R, Fisher J B, Asner G P, Saatchi S, Townsend P, Miller C, Frankenberg C, Hibbard K and Cox P 2015 Observing terrestrial ecosystems and the carbon cycle from space *Global Change Biology* **21** 1762–76 Online: http://doi.wiley.com/10.1111/gcb.12822