**Title:** Informing forest carbon inventories under the Paris Agreement using the Global Forest Carbon Database (ForC v4.0)

**Authors:**

Kristina J. Anderson-Teixeira1,2\*

Valentine Herrmann1

Madison Williams1

Teagan Rogers1

Rebecca Banbury Morgan1,3

Ben Bond-Lamberty4

Susan Cook-Patton5

**Affiliations:**

1. Center for Conservation Ecology, Smithsonian’s National Zoo & Conservation Biology Institute, Front Royal, VA, United States
2. Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Panama, Republic of Panama
3. School of Geography, University of Leeds, Leeds, UK
4. Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, United States
5. The Nature Conservancy; Arlington VA 22203, USA

\*corresponding author: [teixeirak@si.edu](mailto:teixeirak@si.edu)

# Summary

*(NEEDS REVISION)*

Forests are critical for climate change mitigation and constitute a substantial portion of planned net emissions reductions under the 2015 Paris Agreement. However, the efficacy of greenhouse gas mitigation planning and reporting is dependent upon the quality of available emission factors data, including forest carbon (C) stocks and changes therein. Tens of thousands of relevant forest C estimates have been published, yet are not readily accessible to the practitioners compiling national greenhouse gas inventories. Many of these data have, however, been compiled in the Global Forest C database (ForC; <https://forc-db.github.io/>) and stand to be of value to greenhouse gas inventories if made available through the Emission Factor Database (EFDB) of the International Panel on Climate Change (IPCC). Here, we develop and document a process for semi-automated submission of data from ForC into the EFDB, assess the data available and submitted to date, and provide recommendations for improving forest data collection, analysis, and reporting to improve inventories of forest-sector greenhouse gas emissions and removals. We begin by reconciling terminology and mapping ForC fields into EFDB. This process required some updates to the ForC database structure, leading to the release of a new version of ForC (v4.0; described here). As of October 26, 2023, ForC contained ~19316 independent records relevant to EFDB, 1438 of which have undergone necessary review and been submitted to date. Among the data in ForC, there is disproportionate representation of biomass (particularly aboveground) stocks, with far fewer records for dead organic matter and soil C, and relatively few or no records for net annual increments or C fluxes into or out of the IPCC-defined C pools. Geographic representation is also quite uneven, with the highest densities of relevant records in temperate forests, and with relatively scant representation of tropical forests in Africa and Asia. In the future, forest C estimates in EFDB can be improved through targeted research to fill critical gaps, reporting of information required by IPCC, and continued submission of data from scientific publications to the EFDB. Given that climate change is rapidly impacting the world’s forests, timely reporting of recent estimates will be especially critical to accurate forest C inventories.

**Keywords:**

# Societal Impact Statement

# 1 Introduction

Forests are critical to management of the atmospheric concentration of the greenhouse gas carbon dioxide (CO2), and thereby climate change. In recent decades, CO2 uptake by forests, woodlands, and savannas has exceeded releases from deforestation and other severe disturbances, resulting in a net carbon CO2 sink of ~0.88 Gt C yr-1 (all biomes with trees, Xu *et al.*, 2021) to ~1.6 Gt C yr-1 (forests only, Harris *et al.*, 2021). This has offset an estimated 10% to 18% of anthropogenic CO2 emissions from fossil fuels and cement (Xu *et al.*, 2021; Harris *et al.*, 2021), dramatically slowing the pace of atmospheric CO2 accumulation and associated climate change. The future of this important CO2 sink is highly uncertain, and depends upon both forest responses to climate change, which are likely to reduce the sink strength (McDowell *et al.*, 2020; Hammond *et al.*, 2022), and human conservation, restoration, and management of forests (IPCC, 2019a, 2022a).

Forests play a substantial role in international plans for climate change mitigation under the Paris Agreement (UNFCCC, 2015). Forest conservation, reforestation, and improved sustainable management all have significant – and relatively cost-effective – potential as climate change mitigation options (Roe *et al.*, 2021), with conservation and reforestation having the fourth and fifth largest net emission reduction potentials or all mitigation options (IPCC, 2022b). As of 2016, forest-based mitigation accounted for 26% of total planned greenhouse gas mitigation within Nationally Determined Contributions under the Paris Agreement (Grassi *et al.*, 2017). Yet, envisioned forest-based climate change mitigation initiatives do not always correspond to actual emission reductions through on-the-ground implementation (e.g., Badgley *et al.*, 2022). One critical need for ensuring that forest-based climate change mitigation initiatives are effective is realistic planning and reporting, underlain by solid scientific data (Deng *et al.*, 2021; Anderson-Teixeira & Belair, 2022).

The International Panel on Climate Change (IPCC) provides guidance for national greenhouse gas inventories for reporting to the United Nations Framework Convention on Climate Change (UNFCCC, IPCC, 2006a; IPCC, 2019b). Under this guidance, greenhouse gas inventories include all managed land, including most of the world’s forest land (Ogle *et al.*, 2018). The IPCC inventory guidelines include specific instructions for inventories for greenhouse gas (mainly CO2) exchanges between forest land and the atmosphere (IPCC, 2006b, 2019b). A tiered approach is employed, where the lowest tier (Tier 1) represents the simplest approach and relies on default parameter values – for example, forest carbon (C) stocks values by ecozone (FAO, 2012) and forest age class derived as the average of published estimates (IPCC, 2019b; Rozendaal *et al.*, 2022). Tier 1 values have improved over the years as more of the relevant underlying data has become available (Requena Suarez *et al.*, 2019; Rozendaal *et al.*, 2022), but there remains room for improvement as datasets grow and become more widely accessible. For example, the year following the release of the latest IPCC guidelines, a more thorough analysis of C accumulation in regrowth forests found that IPCC’s Tier 1 default failed to capture eight-fold variation within ecozones (Cook-Patton *et al.*, 2020). In addition, it was revealed that C stocks in mature African tropical montane forests were two-thirds higher than the IPCC Tier 1 values for these forests (Cuni-Sanchez *et al.*, 2021). This rapid evolution of scientific information on C cycling in forests is valuable for informing climate change mitigation efforts but requires improved mechanisms for communicating the latest information from scientific researchers to the practitioners who need reliable estimates for greenhouse gas mitigation planning. Moreover, high variability of forest C cycling within ecozones (e.g., Cook-Patton *et al.*, 2020; Cuni-Sanchez *et al.*, 2021) implies that it is useful for practitioners to have access to locally-specific information, when available.

To improve data accessibility for preparing greenhouse gas estimates, the IPCC created the Emission Factor Database (EFDB; <https://www.ipcc-nggip.iges.or.jp/EFDB/main.php>), which is intended as a recognized library of emission factors and other parameters that can be used for estimating greenhouse gas emissions and removals. The EFDB can be used both for efforts to tally a nation’s intended or accomplished greenhouse gas reductions, or as a basis of comparison for external parties to evaluate these inventories. The EFDB encourages researchers to submit estimates of emission factors or other related parameters from peer-reviewed journal papers or other accepted sources for inclusion in the database. In the case of forests, emission factors include C stocks, net annual increments, and annual fluxes for various pools (IPCC, 2006a; IPCC, 2019b).

The Global Forest Carbon Database, ForC (<https://forc-db.github.io/>), is the largest collection of published estimates of forest C stocks, increments, and annual fluxes (Anderson-Teixeira *et al.*, 2018, 2021; Anderson-Teixeira *et al.*, 2023). ForC includes data ingested from individual publications and relevant databases, including the Global Reforestation Opportunity Assessment (GROA) database (Cook-Patton *et al.*, 2020, database doi: 10.5281/zenodo.3983644), and the global soil respiration database (SRDB-V5, Bond-Lamberty & Thomson, 2010; Jian *et al.*, 2021). As of October 26, 2023, ForC contained 39848 records from 10589 plots in 1535 distinct geographical areas, along with records of stand age and disturbance history. As such, ForC is positioned to improve forest-related estimates of CO2 emissions and removals through the submission of data to the EFDB. The purpose of this publication is to document that process and provide recommendations for future improvements.

Here, we (1) review IPCC methods and definitions applied to estimate CO2 emissions and removals from forest in the context of typical forest C estimation methodologies; (2) describe mapping of ForC to IPCC’s EFDB; (3) describe updates to ForC (ForC v4.0), most of which were implemented to facilitate data submission to EFDB; (4) summarize the data in ForC relevant to EFDB and records that have been submitted to date; and (5) provide recommendations as to how the scientific community can better provide useful data for forest C inventories under the Paris Agreement.

# 2 Materials and Methods

## 2.1 Mapping ForC to EFDB

The mapping of ForC fields into EFDB fields is summarized in Table S2. For the majority of fields, contents of the field in ForC was copied directly into an EFDB field, either as the only contents of that field or as part of a composite record. For example, ten ForC fields describing site location, climate, and edaphic properties all mapped into the EFDB field *Region/Regional conditions* (Table S2). In cases where original studies did not present 95% confidence intervals (required by IPCC when available) but did present information required to calculate these (standard error or n and standard deviation), we calculated the 95% confidence intervals and populated the EFDB field with this information (noting the calculation in the EFDB field *Comments from Data Provider*). For some fields, simple conditional logic was used to populate EFDB fields based on ForC records. For example, for stock variables presented in the original publication in units of dry organic matter mass (as opposed to C), several greenhouse gasses (CO2, CO, CH4, NO, NO2, N2O) were entered in the EFDB field indicating the greenhouse gases to which the record could be pertinent (*Gases* field) because these values could be used in calculations of greenhouse gas emissions from biomass burning (IPCC, 2006a); otherwise, the only pertinent greenhouse gas would be CO2. There were two cases in which more complex mapping was required: (1) mapping of C cycle variables (section 4.1.1) and (2) land classification (section 4.1.2).

### 2.1.1 Carbon cycle variables

With input from the IPCC’s Technical Support Unit, we reviewed the list of ForC variables to identify those that were relevant to EFDB and to appropriately map them into EFDB (Fig. 1). For each C pool (Table 1), we identified variables representing organic matter or C stocks, net annual increments, influxes (a.k.a. “gross annual increments” by IPCC), and outfluxes. As described in section 3.2, we also defined 15 new EFDB-relevant variables that were not previously represented in ForC. It is important to note that the correspondence of ForC variables to IPCC criteria often depends upon measurement protocols (“important sources of estimate variation” in Table 1). For example, ForC records of biomass and dead wood vary in the minimum stem diameter censused, such that some records would match the IPCC criteria whereas others would not. Information on minimum diameters censused and other important sources of methodological variation are recorded as covariates in ForC and mapped into the EFDB field *Other Properties* (Table S2). Details on the mapping of ForC variables to EFDB – including associated covariates, IPCC pools (Table 1) and relevant equations (IPCC, 2006a) – are documented in the file ForC\_variables\_mapping.csv in the GitHub repository associated with this publication IPCC-EFDB-integration repository in ForC-db organization (<https://github.com/forc-db/IPCC-EFDB-integration>).



**Figure 1. Schematic illustrating the carbon pools defined under IPCC Guidelines for national greenhouse gas inventories; corresponding ForC variables, and relationships among them.** For each C pool, we show ForC variables corresponding to the stock, net annual increment, influx, and outflux. Most, but not all, EFDB-relevant ForC variables are shown here. Correspondence of ForC variables to IPCC criteria often depends upon measurement protocols (e.g., minimum stem diameter censused). Additional caveats are as follows: (a,b) branch fall and mortality of stems below the minimum stem diameter censused, which are necessary for a full accounting of dead organic matter production but typically assumed negligible for calculations of biomass change, are excluded by common measurement practice (a) or ForC variable definition (b); (c) assumes that leaf production equals leaf fall, or that changes in foliage biomass are negligble; (d,e) belowground components excluded by common measurement practice (d) or ForC variable definition (e); (f) excludes movement of dead wood into litter through breakage or size reduction; (g) measurements often limited to litter horizon (OL) and may exclude larger branches and stems classified as litter and/or the more decomposed layers of the O horizon. \*\*This variable is techically EFDB-relevant but not selected for submission because their is no corresponding influx variable.

### 2.1.2 Land classification

Determination of the IPCC land-use category (i.e., Forest Land, Grassland, Wetlands, Cropland, Settlements, or Other Land; section 2.2) was made based on the categorical ForC field *dominant.life.form*, sometimes drawing upon stand age. Records with “woody” *dominant.life.form* were classified as Forest Land. Those with *dominant.life.form* of “woody+grass”, which in ForC is indicative of anything from a shrub-encroached grassland to a tree-dominated savanna, were given dual classification of Forest Land and Grassland. This dual classification indicates that records may be relevant to either category depending on the definition of forest applied (varies by country). For (rare) cases where *dominant.life.form* was grass and stand age was greater than zero, indicative of early successional vegetation, we assigned a classification of Forest Land, consistent with the IPCC definition that Forest Land includes land expected to succeed to forest. Cases where *dominant.life.form* was grass or crop and stand age was zero were indicative of a control for studies of forest regrowth following agricultural abandonment, and were classified as Grassland and Cropland, respectively.

Classification into sub-categories was dependent upon stand age and site history (section 2.2). For Forest Land ≥ 20 years old or of unknown (relatively mature) age, or Forest Land < 20 years old that was forest prior to a stand-clearing disturbance, the past land-use category was Forest Land, making the sub-category “Forest Land Remaining Forest land”. For forests <20 years old with history including cultivation/ tillage or grazing, past land-use categories were Cropland and Grassland, respectively, making land-use subcategories were “Cropland converted to Forest Land” and “Grassland converted to Forest Land”, respectively. For forests <20 years old with unspecified previous agricultural use, we assigned the sub-category “Land Converted to Forest land”. Forests <20 years old with unknown land use prior to the study date were simply classified as “Forest Land”. The same logic was applied for savannas, but including both Forest Land and Grassland as potentially relevant categories.

Given the lack of public information needed to determine whether lands are classified as managed (Ogle *et al.*, 2018; Deng *et al.*, 2021), and because the IPCC’s definition of managed land is more expansive than is commonly applied in the scientific literature and hence in ForC, we did not include any classification of land management status from ForC in the records submitted to EFDB. However, we did provide auxiliary information that should be useful in making this determination, including geographical location and notable disturbance events.

## 2.2 Updating ForC

Previous versions of ForC (Anderson-Teixeira *et al.*, 2016, 2018, 2021) contained most of the information required by EFDB, and, more broadly, to inform C stock change calculations under the Paris Agreement. However, modest changes to the structure and contents of ForC were needed in order to provide all information required by EFDB and to improve ForC’s capacity to serve as a repository of valuable information for forest C inventories under IPCC guidelines. To support export of data to EFDB, and to improve the overall quality of the ForC database, we added or modified 18 fields (Table S1), defined 15 new variables, implemented enhanced quality control, manually reviewed >1963 records to obtain additional required information, and added 329 new records. Having implemented these changes to ForC v3.0 (Anderson-Teixeira *et al.*, 2021), which are described in detail in Methods S1, we released a new major version: ForC v4.0.

## 2.3 Submission of ForC data to EFDB

To submit complete, reviewed ForC records into EFDB, we created R scripts to restructure ForC records and populate EFDB’s bulk import form (“EFDB bulk import.xlsx”). Criteria for data submission were that (1) records had been checked against the original study and determined to be complete and correct, and as originally presented, (2) the original study presented values in tables or text, as opposed to the values having been digitized from graphs or calculated based on related variables, and (3) the records had not previously been submitted to EFDB. Once converted into EFDB format, the records were reviewed and then sent to the IPCC’s Technical Support Unit for submission to EFDB. Complete records needed to be reviewed by the EFDB editorial board and then posted in the database – a process that lags behind submission of records and had not yet been completed for all records sent as of October 26, 2023.

# 3 Results

## 3.1 ForC v4.0 contents

As of October 26, 2023, ForC (v4.0) contained 32686 independent records (39848 total), 19316 of which were for the 42 variables relevant to EFDB (Fig. 1). These records were distributed across all forested continents and ecozones, albeit unevenly (Fig. 2). The largest number of records came from Asia, followed by North America, South America, and Europe, with relatively few records from Africa, Australia, and Oceania (Fig. 3c). Categorized by FAO ecozone, the greatest numbers of records came from subtropical humid forests, temperate mountain systems, and tropical rain forests, each with >2,000 independent records (Fig. 3b). Boreal coniferous forests, temperate continental forests, subtropical mountain systems, and tropical moist deciduous forests had >1,000 independent records each, while other ecozones all had <1,000 records. The most widely represented forest type was needleleaf evergreen, followed by broadleaf deciduous and broadleaf evergreen (Fig. 3a). In terms of stand age, the most represented age class was 20-100 years, followed by <20 years and then >100 years (Fig. 3d).



**Figure 2. Map of sites in ForC shaded by number of independent records relevant to (circles) and submitted to (triangles) EFDB.** Symbols are colored according to the number of records at each site. Underlying map shows FAO ecozones, which are coded as follows: Ba-Boreal coniferous forest, Bb-Boreal tundra woodland, BM-Boreal mountain systems, P-Polar, SBSh-Subtropical steppe, SBWh-Subtropical desert, SCf-Subtropical humid forest, SCs-Subtropical dry forest, SM-Subtropical mountain systems, TAr-Tropical rain forest, TAwa-Tropical moist deciduous forest, TAwb-Tropical dry forest, TBSh-Tropical shrubland, TBWh-Tropical desert, TeBSk-Temperate steppe, TeBWk-Temperate desert, TeDc-Temperate continental forest, TeDo-Temperate oceanic forest, TeM-Temperate mountain systems, TM-Tropical mountain systems.



**Figure 3. Histograms of number of independent records in ForC relevant to (grey) and submitted to (black) EFDB, organized by (a) dominant vegetation type, (b) FAO ecozone, (c) continent, and (d) stand age.** For dominant vegetation (a), ‘Other’ includes deciduous needleleaf, mixed broadleaf- needleleaf, non-woody vegetation (e.g., early successional), and incompletely classified or mixed forest types. For FAO ecozones (b), codes are as listed in the caption of Figure 2.

ForC contained records for 29 of the 42 variables (or closely-related variable groups) relevant to EFDB (Table 2, Fig. 1). The records were very unevenly distributed across variables. The variable with most records was aboveground biomass, representing 42% of all independent records relevant to EFDB, and aboveground biomass components (woody biomass or foliage) representing an additional 5%. A total of 27% of relevant records were for root biomass (including fine and coarse root components), while 4% described total biomass. The non-living pools were less represented, with 4% of relevant were for dead wood (including standing and fallen components), 0.4% for litter, 0.3% for total ecosystem C excluding soils, and 2.1% for soil carbon.

Increment and flux variables were poorly represented (Table 2). The increment variable with most records was the aboveground biomass increment, representing 0.8% of all independent records relevant to EFDB. The only other relevant increment variable with any records was the O horizon (litter) increment, with just 4 records. Relevant flux variable records (n=2751) together constituted 14% of ForC’s independent records relevant to EFDB.

## 3.2 Data submissions to EFDB

As of October 26, 2023, we had reviewed or added 2292 EFDB-relevant records, 1438 records of which were submitted to EFDB, and 376 of which have been reviewed, accepted, and posted (Figs. 2-3, Table 2). The 37% attenuation between records reviewed and those sent to EFDB was attributable to the presence of digitized records and records where a variable’s value had been calculated as the sum or difference of related variables rather than presented directly in the text. The discrepancy between the number of records sent and that posted to EFDB is primarily attributable to the time required for the IPCC to review and post the records, and also because a minority of records were deemed not applicable to EFDB by the review panel.

The ForC records submitted to EFDB were broadly distributed across Earth’s forests (Fig. 2). However, the density of these records was very unevenly distributed across continents, biomes, and forest types and was not proportional to the numbers of relevant records in ForC (Fig. 3). Rather, the largest number of records came from North America, followed by Asia, South America, and Africa (Fig. 3c), with the most represented FAO ecozones being boreal coniferous forest, temperate continental forest, and temperate mountain systems, followed by tropical rain forests and moist deciduous forests (Fig. 3b). In terms of dominant vegetation, by far the most records came from needleleaf evergreen forests, followed by broadleaf deciduous and broadleaf evergreen (Fig. 3b). The largest records came from mature forests (>100 years), followed by young and intermediate-aged stands (Fig. 3d).

In terms of variables, records were submitted for 19 variables (or closely-related variable groups), including variables from each C pool (Table 2). The majority (82%) of records sent were for C stocks, including 3% for total biomass, 53% for aboveground biomass, 4% for components of aboveground biomass (wood or foliage), 4% for root biomass, 2% for components of root biomass (coarse or fine roots), 5% for dead wood, 3% for components of dead wood (standing or fallen), 4% for litter (entire O horizon or OL layer component), and 4% for SOM/ SOC. Increment records totaled 9% of records sent, virtually all for aboveground biomass (excepting 4 records for delta.O.horizon). The remaining 9% of records sent described fluxes, all of which were either inputs or outputs to the aboveground biomass pool, a subset of which also described inputs to the dead wood or litter pool (Table 2, Fig. 1).

# 4 Discussion

\*\* RECOMENDATIONS \*\*

Based on our experience contributing forest C data to EFDB via ForC, we make several recommendations as to how scientists can improve forest C records in EFDB through database work (section 6.1), new data collection and analysis (section 6.2), and reporting (section 6.3).

## 4.1 Database needs

There is vast potential to expand forest C data in EFDB by completing the process of reviewing and submitting data that are already in ForC (Figs. 2-3). So far, only ~7% of the EFDB-relevant data in ForC have been submitted to EFDB. Although this process requires manual review of records, the submission of new records to EFDB is hugely facilitated by the fact that most pertinent information for each record is already entered in ForC and can be easily prepared for submission to EFDB using the system developed here. Future efforts to review studies for submission should optimize for representation across geographic regions, forest types, and variables, giving priority to those from currently under-represented regions and forest types (Figs. 2-3, Table 2), to records from countries relying on existing data for their greenhouse gas inventories (Tier 1 or 2 methodology), to the variables most needed by EFDB users, and to the more contemporary records.

In addition to the large potential to expand EFDB using records already in ForC, there are innumerous published EFDB-relevant forest C data that are not currently included in ForC, with more being published on a nearly daily basis. Coverage of particular variables or regions could be vastly improved through systematic review of the literature. Indeed, recent efforts have compiled large databases of relevant data from monoculture plantation forests (Bukoski *et al.*, 2022) and mixed species plantation forests (Warner *et al.*, 2022; Feng *et al.*, 2022), and such a compilation is in works for agroforestry (Susan Cook-Patton, unpublished data). Beyond expanding collections of relevant forest C records, such reviews are valuable for assessing the availability of published records and identifying variables and regions that require additional data collection and analysis.

## 4.2 Data collection and analysis needs

New data collection and analysis is needed to fill notable knowledge gaps. While aboveground biomass stocks in particular have received – and continue to receive – by far the most research attention (NISAR, 2018; Quegan *et al.*, 2019; Dubayah *et al.*, 2020; Table 2, Anderson-Teixeira *et al.*, 2021), production of an accurate global map of forest C stocks remains an ongoing challenge (Araza *et al.*, 2023). Other pools and variables remain very poorly quantified (Table 2, Anderson-Teixeira *et al.*, 2021), introducing substantive uncertainties into global forest C budgets (Pan *et al.*, 2011; Harris *et al.*, 2021). Furthermore, data distribution is uneven across forest types and geographical regions (Figs. 2-3). For instance, data on C cycling of tropical forests – particularly in Africa – remains relatively sparse, in large part due to substantial barriers to data collection and distribution (de Lima *et al.*, 2022). Significant investment in research and researchers focused on ground-based measurement of forest C in such regions will be important to filling knowledge gaps in forest C cycling (de Lima *et al.*, 2022; Labrière *et al.*, 2023; Araza *et al.*, 2023).

Several EFDB-relevant variables have not been calculated and presented as frequently as would be possible given existing forest census data and minimal extra research effort. For example, aboveground woody mortality (*woody.mortality\_ag*) and aboveground biomass increment (*delta.agb*) can be calculated from the same census data as aboveground woody productivity (*ANPP\_woody*), yet the latter has received far more research attention, and correspondingly has far more records in ForC (Table 2, Anderson-Teixeira *et al.*, 2021; but see Piponiot *et al.*, 2022). Similarly, live coarse root biomass, total biomass, and changes in both of these pools could in theory be estimated in parallel with aboveground biomass, with the greatest barrier being that allometric models for estimating root biomass are not as reliable or easily available as are those for aboveground biomass (Chave *et al.*, 2014; Réjou-Méchain *et al.*, 2017; Gonzalez-Akre *et al.*, 2022). However, while equations for estimating root (and thereby total) biomass require improvement, they do exist for many forest types (Mokany *et al.*, 2006; e.g., Brassard *et al.*, 2011; Chojnacky *et al.*, 2014; Waring & Powers, 2017). In addition, standing dead trees are captured in most forest censuses and could be used to estimate standing dead wood, although additional data on breakage would be needed for accurate estimates. We recommend that, when possible, researchers calculate and report these variables, following the reporting guidelines specified in section 6.3.

Filling knowledge gaps in other EFDB-relevant variables will require more effort, but this effort is warranted given their importance for estimating forest C stock chnages. Although aboveground biomass is the most studied variable considered here (Table 2) and is the target of satellite missions (NISAR, 2018; Quegan *et al.*, 2019; Dubayah *et al.*, 2020), significant ground-based research effort is required to create accurate global maps of forest biomass and changes therein (Duncanson *et al.*, 2019; Calders *et al.*, 2022; Labrière *et al.*, 2023). Given observations of increasing tree mortality in some forested regions (McDowell *et al.*, 2020), better characterization of forest dead wood will be critical. Additionally, C stocks in forest organic horizons and soils can be quite substantial and highly uncertain in many parts of the world (Tifafi *et al.*, 2018). Significant investment in ground-based forest research will be critical to filling these gaps.

## 4.3 Data reporting needs

We recommend that, in order to make research valuable to estimate C stock changes according to methods provided in the IPCC guidelines, researchers calculate and report results according to IPCC good practice (Table 3). It is particularly noteworthy that simple decisions on the presentation of results will determine whether the records meet the criteria for inclusion in EFDB. Some examples are as follows: (1) presenting data only in a figure makes them ineligible for inclusion in EFDB, whereas presentation in a table or supplementary data file allows inclusion while supporting FAIR goals (**stall\_make\_2019?**); (2) direct presentation of all relevant variables allows inclusion, whereas presenting only components of variables of interest (e.g., parsing litter into fine woody debris, OL, OF, and OH layers) or requiring simple mathematical operations to obtain a variable of interest (e.g., *delta.agb* = *ANPP\_woody* - *woody.mortality.agb*) disqualifies records from inclusion; (3) matching IPCC-defined thresholds for defining C pools (Table 1), which may vary by country, can make the data far more relevant estimating forest C stock changes according to IPCC guidelines (e.g., using a 10 cm cutoff between dead wood and litter, presenting soil C to a depth of 30 cm). It should also be emphasized that reporting of 95% confidence intervals (or other metrics of error), when applicable, is highly desirable and makes the data more relevant to IPCC.

For those compiling published records (e.g., for meta-analyses), the data set can have added value if all information required by EFDB is extracted from original publications. This includes – but is not limited to – retaining original values as presented without modification or rounding, noting whether data were digitized, recording confidence intervals, and recording all required fields (as indicated in the EFDB’s bulk import template). The significant effort required to map a database into EFDB has been accomplished here (Table S2), and we welcome other researchers to use the ForC template.

Once EFDB-relevant data are available in peer-reviewed publications, they may be submitted directly to EFDB or may use the ForC - EFDB data pipeline developed here. For individual publications, the former option will generally be more efficient. However, data incorporated into ForC as well as EFDB will be more broadly useful; for example, these data may be used for basic science (e.g., Banbury Morgan *et al.*, 2021; Anderson-Teixeira *et al.*, 2021), analyses of forest-based climate change mitigation potential (Goldstein *et al.*, 2020; e.g., Cook-Patton *et al.*, 2020), and model benchmarking (Fer *et al.*, 2021).

# 5 Conclusions

The ForC database contains large numbers of records that could potentially be useful for estimating C stock changes applying methodological guidance provided by the IPCC. Here we have developed a framework for submitting these records to the EFDB, thus making those data more accessible for reporting CO2 emissions and removals from forest land consistent with good practice in the IPCC guidelines (IPCC, 2006a; IPCC, 2019b). As of October 26, 2023, we have submitted 1438 records to EFDB. Although this represents just 7% of relevant records in ForC, it substantially increases the number of forest land records in EFDB. The records submitted to EFDB and present in ForC are very unevenly distributed across variables, regions, and forest types (Figs. 2-3, Table 2), reflecting broader patterns in allocation of research effort.

Going forward, forest researchers can make their research more useful for forest C inventories under IPCC guidelines by calculating and reporting results in ways that are consistent with methodologies provided in the IPCC guidelines (Tables 1, 3). In addition, substantial investments in research and researchers focused on ground-based measurement of forest C will be required to fill knowledge gaps and thereby increase the accuracy of forest CO2 inventories for forest lands under teh Paris Agreement. This challenge is heightened by the fact that forests are changing rapidly (e.g., McDowell *et al.*, 2020), and data collected a decade or more in the past may no longer be relevant. This heightens the need for an efficient system of making forest C data accessible for national greenhouse gas inventories. We view the system developed here for submitting ForC data to the IPCC EFDB as one important step towards that goal.

# 6 Acknowledgements

We gratefully acknowledge the substantial contributions of Valentyna Slivinska and Sandro Federici for collaboration on the conception, design, and technical review of this project. Thank you to all researchers who collected and published the data contained in ForC, and to all research assistants and collaborators who have contributed to building the database. Thank you to Avni Malhotra for helpful comments on an earlier draft of this manuscript. Funding for this study was provided by the Smithsonian (Forest Global Earth Observatory, Smithsonian Working Land and Seascapes); a Bezos Earth Fund grant to the Nature Conservancy, with a sub-grant to NZCBI; and the Institute for Global Environmental Strategies.

# 7 Author Contribution

KAT and VH conceived and designed the project; VH wrote the scripts for database management, data submission to EFDB, and the analyses presented here; MW, TR, and RBM added and reviewed ForC data, BBL and SCP contributed large databases to ForC (EFDB and GROA, respectively); CP provided methodological expertise; KAT, VH, and MW prepared the first draft of the manuscript; all authors reviewed the results and approved the final version of the manuscript.

# 8 Data Availability Statement

All code and data are openly available. The ForC database and associated code are available via the ForC repository within the ForC-db organization on GitHub (<https://github.com/forc-db/ForC>), and the version used here (ForC v4.0) is archived in Zenodo (Anderson-Teixeira *et al.*, 2023, DOI: 10.5281/zenodo.8020861). The data and code associated with data submission to EFDB and preparation of this manuscript are available via the the IPCC-EFDB-integration repository within the ForC-db organization on GitHub (<https://github.com/forc-db/IPCC-EFDB-integration>) and archived in Zenodo (DOI: 10.5281/zenodo.8021474).

# References

**Anderson-Teixeira KJ, Belair EP**. **2022**. [Effective forest-based climate change mitigation requires our best science](https://doi.org/10.1111/gcb.16008). *Global Change Biology* **28**: 1200–1203.

**Anderson-Teixeira K, Herrmann V, Morgan BB, Actions-User, Williams M, Rogers T, McGregor I, Hua MWM, Ferson A, Bond-Lamberty B, *et al.*** **2023**. [Forc-db/ForC: First version with EFDB integration](https://doi.org/10.5281/ZENODO.8020861).

**Anderson-Teixeira KJ, Herrmann V, Morgan RB, Bond-Lamberty B, Cook-Patton SC, Ferson AE, Muller-Landau HC, Wang MMH**. **2021**. [Carbon cycling in mature and regrowth forests globally](https://doi.org/10.1088/1748-9326/abed01). *Environmental Research Letters* **16**: 053009.

**Anderson-Teixeira KJ, Wang MMH, McGarvey JC, Herrmann V, Tepley AJ, Bond-Lamberty B, LeBauer DS**. **2018**. [ForC : A global database of forest carbon stocks and fluxes](https://doi.org/10.1002/ecy.2229). *Ecology* **99**: 1507–1507.

**Anderson-Teixeira KJ, Wang MMH, McGarvey JC, LeBauer DS**. **2016**. [Carbon dynamics of mature and regrowth tropical forests derived from a pantropical database (TropForC-db)](https://doi.org/10.1111/gcb.13226). *Global Change Biology* **22**: 1690–1709.

**Araza A, Herold M, de Bruin S, Ciais P, Gibbs DA, Harris N, Santoro M, Wigneron J-P, Yang H, Málaga N, *et al.*** **2023**. [Past decade above-ground biomass change comparisons from four multi-temporal global maps](https://doi.org/10.1016/j.jag.2023.103274). *International Journal of Applied Earth Observation and Geoinformation* **118**: 103274.

**Badgley G, Freeman J, Hamman JJ, Haya B, Trugman AT, Anderegg WRL, Cullenward D**. **2022**. [Systematic over-crediting in California’s forest carbon offsets program](https://doi.org/10.1111/gcb.15943). *Global Change Biology* **28**: 1433–1445.

**Banbury Morgan R, Herrmann V, Kunert N, Bond-Lamberty B, Muller-Landau HC, Anderson-Teixeira KJ**. **2021**. [Global patterns of forest autotrophic carbon fluxes](https://doi.org/10.1111/gcb.15574). *Global Change Biology* **27**: 2840–2855.

**Bond-Lamberty B, Thomson A**. **2010**. [A global database of soil respiration data](https://doi.org/10.5194/bg-7-1915-2010). *Biogeosciences* **7**: 1915–1926.

**Brassard BW, Chen HYH, Bergeron Y, Paré D**. **2011**. [Coarse root biomass allometric equations for Abies balsamea, Picea mariana, Pinus banksiana, and Populus tremuloides in the boreal forest of Ontario, Canada](https://doi.org/10.1016/j.biombioe.2011.06.045). *Biomass and Bioenergy* **35**: 4189–4196.

**Bukoski JJ, Cook-Patton SC, Melikov C, Ban H, Chen JL, Goldman ED, Harris NL, Potts MD**. **2022**. [Rates and drivers of aboveground carbon accumulation in global monoculture plantation forests](https://doi.org/10.1038/s41467-022-31380-7). *Nature Communications* **13**: 1–13.

**Calders K, Verbeeck H, Burt A, Origo N, Nightingale J, Malhi Y, Wilkes P, Raumonen P, Bunce RGH, Disney M**. **2022**. [Laser scanning reveals potential underestimation of biomass carbon in temperate forest](https://doi.org/10.1002/2688-8319.12197). *Ecological Solutions and Evidence* **3**: e12197.

**Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, *et al.*** **2014**. [Improved allometric models to estimate the aboveground biomass of tropical trees](https://doi.org/10.1111/gcb.12629). *Global Change Biology* **20**: 3177–3190.

**Chojnacky DC, Heath LS, Jenkins JC**. **2014**. [Updated generalized biomass equations for North American tree species](https://doi.org/10.1093/forestry/cpt053). *Forestry* **87**: 129–151.

**Cook-Patton SC, Leavitt SM, Gibbs D, Harris NL, Lister K, Anderson-Teixeira KJ, Briggs RD, Chazdon RL, Crowther TW, Ellis PW, *et al.*** **2020**. [Mapping carbon accumulation potential from global natural forest regrowth](https://doi.org/10.1038/s41586-020-2686-x). *Nature* **585**: 545–550.

**Cuni-Sanchez A, Sullivan MJP, Platts PJ, Lewis SL, Marchant R, Imani G, Hubau W, Abiem I, Adhikari H, Albrecht T, *et al.*** **2021**. [High aboveground carbon stock of African tropical montane forests](https://doi.org/10.1038/s41586-021-03728-4). *Nature* **596**: 536–542.

**de Lima RAF, Phillips OL, Duque A, Tello JS, Davies SJ, de Oliveira AA, Muller S, Honorio Coronado EN, Vilanova E, Cuni-Sanchez A, *et al.*** **2022**. [Making forest data fair and open](https://doi.org/10.1038/s41559-022-01738-7). *Nature Ecology & Evolution*.

**Deng Z, Ciais P, Tzompa-Sosa ZA, Saunois M, Qiu C, Tan C, Sun T, Ke P, Cui Y, Tanaka K, *et al.*** **2021**. [Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions](https://doi.org/10.5194/essd-2021-235). *Earth System Science Data Discussions*: 1–59.

**Dubayah R, Blair JB, Goetz S, Fatoyinbo L, Hansen M, Healey S, Hofton M, Hurtt G, Kellner J, Luthcke S, *et al.*** **2020**. [The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth’s forests and topography](https://doi.org/10.1016/j.srs.2020.100002). *Science of Remote Sensing* **1**: 100002.

**Duncanson L, Armston J, Disney M, Avitabile V, Barbier N, Calders K, Carter S, Chave J, Herold M, Crowther TW, *et al.*** **2019**. [The Importance of Consistent Global Forest Aboveground Biomass Product Validation](https://doi.org/10.1007/s10712-019-09538-8). *Surveys in Geophysics* **40**: 979–999.

**FAO**. **2012**. *Global Ecological Zones for FAO Forest Reporting: 2010 update*. Rome: FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS.

**Feng Y, Schmid B, Loreau M, Forrester DI, Fei S, Zhu J, Tang Z, Zhu J, Hong P, Ji C, *et al.*** **2022**. [Multispecies forest plantations outyield monocultures across a broad range of conditions](https://doi.org/10.1126/science.abm6363). *Science* **376**: 865–868.

**Fer I, Gardella AK, Shiklomanov AN, Campbell EE, Cowdery EM, Kauwe MGD, Desai A, Duveneck MJ, Fisher JB, Haynes KD, *et al.*** **2021**. [Beyond ecosystem modeling: A roadmap to community cyberinfrastructure for ecological data-model integration](https://doi.org/10.1111/gcb.15409). *Global Change Biology* **27**: 13–26.

**Goldstein A, Turner WR, Spawn SA, Anderson-Teixeira KJ, Cook-Patton S, Fargione J, Gibbs HK, Griscom B, Hewson JH, Howard JF, *et al.*** **2020**. [Protecting irrecoverable carbon in Earth’s ecosystems](https://doi.org/10.1038/s41558-020-0738-8). *Nature Climate Change* **10**: 287–295.

**Gonzalez-Akre E, Piponiot C, Lepore M, Herrmann V, Lutz JA, Baltzer JL, Dick CW, Gilbert GS, He F, Heym M, *et al.*** **2022**. [Allodb: An R package for biomass estimation at globally distributed extratropical forest plots](https://doi.org/10.1111/2041-210X.13756). *Methods in Ecology and Evolution* **13**: 330–338.

**Grassi G, House J, Dentener F, Federici S, den Elzen M, Penman J**. **2017**. [The key role of forests in meeting climate targets requires science for credible mitigation](https://doi.org/10.1038/nclimate3227). *Nature Climate Change* **7**: 220–226.

**Hammond WM, Williams AP, Abatzoglou JT, Adams HD, Klein T, López R, Sáenz-Romero C, Hartmann H, Breshears DD, Allen CD**. **2022**. [Global field observations of tree die-off reveal hotter-drought fingerprint for Earth’s forests](https://doi.org/10.1038/s41467-022-29289-2). *Nature Communications* **13**: 1761.

**Harris NL, Gibbs DA, Baccini A, Birdsey RA, Bruin S de, Farina M, Fatoyinbo L, Hansen MC, Herold M, Houghton RA, *et al.*** **2021**. [Global maps of twenty-first century forest carbon fluxes](https://doi.org/10.1038/s41558-020-00976-6). *Nature Climate Change*: 1–7.

**IPCC**. **2006a**. *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. And Tanabe K. (eds).* Japan: IGES.

**IPCC**. **2006b**. Agriculture, Forestry, and Other Land Use. In: Eggleston S, Buendia L, Miwa K, Ngara T, Tanabe K, eds. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Hayama, Japan: Institute for Global Environmental Strategies.

**IPCC**. **2019b**. 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. In: Calvo Buendia E, Tanabe K, Baasansuren J, Fukuda M, Ngarize S, Osako A, Pyrozhenko Y, Shermanau P, Federici S, eds. Switzerland: IPCC.

**IPCC**. **2019a**. *Climate Change and Land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.- O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. Van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]*.

**IPCC**. **2022a**. [Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. Van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA.](https://doi.org/10.1017/9781009157926.001)

**IPCC**. **2022b**. Summary for Policymakers. Policymakers [P.R. Shukla, J. Skea, A. Reisinger, R. Slade, R. Fradera, M. Pathak, A. Al Khourdajie, M. Belkacemi, R. Van Diemen, A. Hasija, G. Lisboa, S. Luz, J. Malley, D. McCollum, S. Some, P. Vyas, (eds.)]. In: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. Van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA. Doi: 10.1017/9781009157926.001.

**Jian J, Vargas R, Anderson-Teixeira K, Stell E, Herrmann V, Horn M, Kholod N, Manzon J, Marchesi R, Paredes D, *et al.*** **2021**. [A restructured and updated global soil respiration database (SRDB-V5)](https://doi.org/10.5194/essd-13-255-2021). *Earth System Science Data* **13**: 255–267.

**Labrière N, Davies SJ, Disney MI, Duncanson LI, Herold M, Lewis SL, Phillips OL, Quegan S, Saatchi SS, Schepaschenko DG, *et al.*** **2023**. [Toward a forest biomass reference measurement system for remote sensing applications](https://doi.org/10.1111/gcb.16497). *Global Change Biology* **n/a**.

**McDowell NG, Allen CD, Anderson-Teixeira K, Aukema BH, Bond-Lamberty B, Chini L, Clark JS, Dietze M, Grossiord C, Hanbury-Brown A, *et al.*** **2020**. [Pervasive shifts in forest dynamics in a changing world](https://doi.org/10.1126/science.aaz9463). *Science* **368**.

**Mokany K, Raison RJ, Prokushkin AS**. **2006**. [Critical analysis of root : Shoot ratios in terrestrial biomes](https://doi.org/10.1111/j.1365-2486.2005.001043.x). *Global Change Biology* **12**: 84–96.

**NISAR**. **2018**. *NASA-ISRO SAR (NISAR) Mission Science Users’ Handbook*. NASA Jet Propulsion Laboratory.

**Ogle SM, Domke G, Kurz WA, Rocha MT, Huffman T, Swan A, Smith JE, Woodall C, Krug T**. **2018**. [Delineating managed land for reporting national greenhouse gas emissions and removals to the United Nations framework convention on climate change](https://doi.org/10.1186/s13021-018-0095-3). *Carbon Balance and Management* **13**: 9.

**Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG, *et al.*** **2011**. A Large and Persistent Carbon Sink in the World’s Forests. *Science* **333**: 988–993.

**Piponiot C, Anderson-Teixeira KJ, Davies SJ, Allen D, Bourg NA, Burslem DFRP, Cárdenas D, Chang-Yang C-H, Chuyong G, Cordell S, *et al.*** **2022**. [Distribution of biomass dynamics in relation to tree size in forests across the world](https://doi.org/10.1111/nph.17995). *New Phytologist* **n/a**.

**Quegan S, Le Toan T, Chave J, Dall J, Exbrayat J-F, Minh DHT, Lomas M, D’Alessandro MM, Paillou P, Papathanassiou K, *et al.*** **2019**. [The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space](https://doi.org/10.1016/j.rse.2019.03.032). *Remote Sensing of Environment* **227**: 44–60.

**Réjou-Méchain M, Tanguy A, Piponiot C, Chave J, Hérault B**. **2017**. [Biomass: An r package for estimating above-ground biomass and its uncertainty in tropical forests](https://doi.org/10.1111/2041-210X.12753). *Methods in Ecology and Evolution* **8**: 1163–1167.

**Requena Suarez D, Rozendaal DMA, Sy VD, Phillips OL, Alvarez-Dávila E, Anderson-Teixeira K, Araujo-Murakami A, Arroyo L, Baker TR, Bongers F, *et al.*** **2019**. [Estimating aboveground net biomass change for tropical and subtropical forests: Refinement of IPCC default rates using forest plot data](https://doi.org/10.1111/gcb.14767). *Global Change Biology* **25**: 3609–3624.

**Roe S, Streck C, Beach R, Busch J, Chapman M, Daioglou V, Deppermann A, Doelman J, Emmet-Booth J, Engelmann J, *et al.*** **2021**. [Land-based measures to mitigate climate change: Potential and feasibility by country](https://doi.org/10.1111/gcb.15873). *Global Change Biology* **27**: 6025–6058.

**Rozendaal DMA, Suarez DR, Sy VD, Avitabile V, Carter S, Yao CYA, Alvarez-Davila E, Anderson-Teixeira K, Araujo-Murakami A, Arroyo L, *et al.*** **2022**. [Aboveground forest biomass varies across continents, ecological zones and successional stages: Refined IPCC default values for tropical and subtropical forests](https://doi.org/10.1088/1748-9326/ac45b3). *Environmental Research Letters* **17**: 014047.

**Tifafi M, Guenet B, Hatté C**. **2018**. [Large Differences in Global and Regional Total Soil Carbon Stock Estimates Based on SoilGrids, HWSD, and NCSCD: Intercomparison and Evaluation Based on Field Data From USA, England, Wales, and France](https://doi.org/10.1002/2017GB005678). *Global Biogeochemical Cycles* **32**: 42–56.

**UNFCCC**. **2015**. Adoption of the Paris Agreement. : 31.

**Waring BG, Powers JS**. **2017**. [Overlooking what is underground: Root:shoot ratios and coarse root allometric equations for tropical forests](https://doi.org/10.1016/j.foreco.2016.11.007). *Forest Ecology and Management* **385**: 10–15.

**Warner E, Cook-Patton SC, Lewis OT, Brown N, Koricheva J, Eisenhauer N, Ferlian O, Gravel D, Hall JS, Jactel H, *et al.*** **2022**. [Higher aboveground carbon stocks in mixed-species planted forests than monocultures a meta-analysis](https://doi.org/10.1101/2022.01.17.476441). : 2022.01.17.476441.

**Xu L, Saatchi SS, Yang Y, Yu Y, Pongratz J, Bloom AA, Bowman K, Worden J, Liu J, Yin Y, *et al.*** **2021**. [Changes in global terrestrial live biomass over the 21st century](https://doi.org/10.1126/sciadv.abe9829). *Science Advances* **7**: eabe9829.