TropForC Database
Anderson-Teixeira et al., 201X

Global forest carbon database (ForC-db)

Source: Anderson-Teixeira, K.J. et al. 201X. Carbon dynamics of mature and regrowth tropical forests derived from a pantropical database (TropForC). *Global Change Biology.* XXX.

Database description:

The Tropical Forest C database (TropForC-db) is the tropical component of ForC-db, which is a global C database that is currently under development (Anderson-Teixeira et al., unpublished). Its structure is derived from that of BETY-db (www.betydb.org; LeBauer et al. 2010). In brief, the database consists of a series of cross-referenced data tables describing (1) sites, (2) plots and their history, (3) measurements of C cycle variables, (4) variables, (5) disturbance/history event type, and (6) species/ plant functional types (PFTs). We obtained data from multiple previous compilations and individual published studies, as detailed in Anderson-Teixeira et al. (201X). Original citations associated with all data are given in the measurements table (3).

Throughout the database, missing data were coded to indicate the reason for the missing values (Table 1).

Overview of database structure and data files:

Data file	Description	Links to other data files
(1) '**sites**' TropForC_sites.csv	Geographic, climatic, and edaphic site data	(2) and (3)
(2) '**plothistory**' TropForC_plothistory.csv	Known history of each plot or set of replicate plots	(1), (3), and (5)
(3) '**measurements**' TropForC_measurements.csv	Records of ecosystem-level measurements relevant to C cycling	(1), (2), (4), and (6)
(4) '**variables**' TropForC_variables.csv	Definitions of C cycle variables and covariates.	Defines codes for columns 'variables.name', 'covariate_1', 'covariate_2', and 'covariate_3' in (3).
(5) '**disttype**' TropForC_disttype.csv	Definition of disturbance, management or regeneration history event types.	Defines codes for column 'disttype' in (2)
(6) '**pft'** TropForC_pft.csv	Definitions of plant functional codes	Defines codes for columns 'acceptedsymbol' and 'dominantveg' in (3)

TropForC Database
Anderson-Teixeira et al., 201X

Data File Contents:

(1) TropForC_sites.csv - '**sites**'

Column	Description	Units
siteID	Unique numerical identifier for each record in the sites table.	-
sites.sitename	Site identifier, sufficient to uniquely identify the site within the paper	-
city	Nearest city	-
state	State	-
country	Country	-
lat	Latitude in decimal form	-
lon	Longitude in decimal form	-
masl	Elevation (meters above sea level) (reported)	m
mat	Mean annual temperature (reported)	°C
min.temp	Mean temperature of the coldest month (reported)	°C
max.temp	Mean temperature of the warmest month (reported)	°C
map	Mean annual precipitation (reported)	mm
soil	USDA soil classification (based on % sand-silt-clay)	-
som	Soil classification (USDA soil taxonomy or FAO soil classification)	-
soilnotes	Soil details not included above	-
sitenotes	Site details not included above.	-
siteref	Additional references used to obtain site data.	-
area	Groups sites that are geographically proximate and edaphically similar. Defined as a group of sites where no site is removed from the rest in its group by >0.25 degrees latitude or longitude.	-
biogeog	Biogeographical region, extracted from map of Olson *et al.* (2001)	-
Koeppen	Köppen-Geiger classification, extracted from the ESRI Köppen-Geiger map (http://www.arcgis.com/home/item.html?id=7a53584fa55643df969f93cec83788e1).	-
FAOecozone	FAO global ecological zones classification, extracted from FAO's GeoNetwork (http://www.fao.org:80/geonetwork).	-

(2) TropForC_plothistory.csv - '**plothistory**'

Column	Description	Units
plothistoryID	Unique numerical identifier for each record in the history table.	-
sites.sitename	Unique site identifier, sufficient to identify the site within the original publication.	-
plot.name	Unique plot name, sufficient to identify the plot within the original publication.	-
plotarea	Area of plot or combined area of replicate plots.	ha
date	Date of plot history event.	-
dateloc	Level of confidence in date (codes given in Table 2).	-
distcat	Plot history event categories, namely 'Disturbance' (includes natural and anthropogenic disturbances, management), 'Establishment', 'No.info' (data not available), and 'Regrowth'.	-
disttype	Plot history event types. See **disttype** table for code definitions.	-
level	Level of history event type.	-
units	Level units.	-
distnotes	History event details not included above.	-

(3) TropForC_measurements.csv - ‘**measurements**’

Column	Description	Units
measurementID	Unique identifier for each record in the measurements table.	-
sites.sitename	Unique site identifier name for the site at which the measurement was made. Links to **sites** and **history** tables.	-
plot.name	Unique plot identifier name for the plot in which the measurement was made. Links to **plothistory** table.	-
stand.age	Age of stand in years as reported in the original publication or calculated based on the date of initiation of forest regrowth. When the publication reports a range of ages, the mean is recorded.	-
dominantveg	Code used to identify plant functional trait of dominant vegetation. Identical to **acceptedsymbol** for multi-species forests. For plantations, this is the the appropriate **pftcode** based on the planted species.	-
acceptedsymbol	Acronym for plant functional type (pft; multi-species stands) or species listed in scientificname (monospecific stands—i.e., plantations). Plant functional type code, **pftcode**, is as defined in the the **pft** table. For single species, species name acronyms are as listed in the USDA Plants Checklist (http://plants.usda.gov/dl_all.html), or listed as 'nocode' if not available for the species.	-
scientificname	Genus and species for single dominant species, if applicable (plantations only).	-
variables.name	Code name for variable sampled. Codes defined in **variables** table.	-
date	Date on which measurement was made. Usually found as the year of measurement in the methods section, or, when unavailable, approximated as the year before publication.	-
dateloc	Level of confidence in date. See Table 4 in the BETY-DB Data Entry Workflow Documentation for codes behind level of confidence.	-
start_date	Date on which measurements were started.	-
start_dateloc	Level of confidence in date. See dateloc above.	-
end_date	Date on which measurements ended.	-
end_dateloc	Level of confidence in date. See dateloc above.	-
mean	Mean value of measurement recorded. Units dependent on **variables.name**. See **variables** table for the associated units.	as defined in variables table
n	Number of experimental replicates used to estimate mean and statistical summary.	-
statname	Name of reported statistic (SE- standard error; SD- standard deviation; 95%CI- 95% confidence interval).	-

Column	Description	Units
stat	Value of reported statistic.	-
notes	Details of study.	-
covariate_#	Code name for covariate associated with the variable being measured. Codes defined in **variables** table.	-
coV#_value	Value of reported covariate.	-
coV#_units	Units of reported covariate.	-
dupcode	Marks entries that meet one of the following criteria: - Duplicate estimates of the same variable but with different values (**code D** for duplicate) - Replicate sampling of the same variables within the same study (**code R** for replicate) - Multiple estimates of the same variables within the same study but using different methods (**code M** for different measurement methods)	-
dupnum	Assigns a number to every entry in each set of duplicates/replicates, in chronological order according to the citation.year. If two sources were published in the same year, the primary source (as opposed to a synthesis) comes first.	-
citations.doi	Digital object identifier unique to publication, if available.	-
citations.author	First author's last name of publication.	-
citations.year	Year of publication.	-
citations.title	Complete title of publication.	-
loaded_from	Reference to data compilations from which data and reference to primary source were obtained, if applicable.	-

(4) TropForC_variables.csv - '**variables**'

Column	Description	Units
variables.name	Code name for variable sampled. Used in **measurements** table.	-
units	Units of the variable.	-
description	Definition and notes about the variable.	-
associated_covariate	Covariates associated with the variable listed. The codes for these covariates are also defined in this table.	-

(5) TropForC_disttype.csv - '**disttype**'

Column	Description	Units
distcat	Plot history event categories, i.e. Disturbance (includes natural and anthropogenic disturbances, management), Establishment, No.info (data not available), Regrowth. Used in **plothistory** table.	
disttype	Code name for plot history event types. Used in **plothistory** table.	-
description	Definition and notes about the disturbance, management, or regeneration type.	-
units	Units of the variable.	-

(6) TropForC_pft.csv - '**pft**'

Column	Description	Units
pftcode	Code name for plant functional traits. Used in **measurements** table. Codes match those in BETY-db (LeBauer *et al.*, 2010).	-
description	Definition of the **pftcode**.	-
notes	Details.	-

Additional Tables:

Table 1. Missing data codes

Code	Definition	Description
NA	Not Applicable	-
NAC	Not Acquired	Information may be available but has not been acquired.
NRA	Not Readily Available	Information was not readily available to the authors (e.g., publication not readily available).
NI	No Information	Publication does not contain the required information.

Table 2. Date level of confidence (**dateloc**) numbering convention, used in **plothistory** and **measurements** tables. Adapted from Table 4 in the BETY-db Data Entry Workflow (available at URL: http://pecan.ncsa.illinois.edu/bety/node21.html#252; LeBauer *et al.*, 2010).

Dateloc*	Definition
9	nodata
8	year
7	season
6	month
5	day
4	time of day

* When the dateloc is followed by '.5', it indicates that a range of dates (year/month/day) were given; e.g. "late 1990s" would be given a dateloc of 8.5 for the year "1999".

References:

LeBauer D, Wang D, Dietze M (2010) Biofuel Ecophysiological Traits and Yields Database Version 1.0. Energy Biosciences Institute, Urbana, IL http://ebi-forecast.igb.uiuc.edu/bety/.

Olson DM, Dinerstein E, Wikramanayake ED et al. (2001) Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, **51**, 933.

How to cite:

Publications using these data should cite both the original publication and the Dryad data package:

Anderson-Teixeira, K.J. et al. **201**X. Carbon dynamics of mature and regrowth tropical forests derived from a pantropical database (TropForC). *Global Change Biology.* XXX.

Anderson-Teixeira, K.J. et al. **201**X. Data from: Carbon dynamics of mature and regrowth tropical forests derived from a pantropical database (TropForC). Dryad Digital Repository. http://dx.doi.org/XXXXX