AGU ADVANCING EARTH AND SPACE SCIENCES

# JGR Biogeosciences

DATA ARTICLE
10.1029/2024JG008381

**Key Points:**

- 360,570 biomass estimates between the years 2014 and 2023 for 35 NEON sites in the United States and Puerto Rico
- Inclusion of sapling and young tree biomass estimates using adjustment factors
- Presented as a standalone data set and as an R data package (NEONForestAGB)

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

J. W. Atkins,
jwatkins6@vcu.edu

**Citation:**

Atkins, J. W., Alveshere, B., Breland, S., Meier, C., Langley, M., & Zhai, L. (2025). Individual tree and sapling aboveground biomass (AGB) estimates for 35 NEON terrestrial observation sites for 2014–2023. *Journal of Geophysical Research: Biogeosciences*, *130*, e2024JG008381. https://doi.org/10.1029/2024JG008381

Received 23 JUL 2024
Accepted 11 MAY 2025

**Author Contributions:**

**Conceptualization:** Jeff W. Atkins, Courtney Meier, Lu Zhai
**Data curation:** Jeff W. Atkins, Sabrie Breland, Courtney Meier, Michael Langley
**Formal analysis:** Jeff W. Atkins, Brandon Alveshere
**Methodology:** Jeff W. Atkins, Brandon Alveshere, Courtney Meier, Lu Zhai
**Project administration:** Jeff W. Atkins
**Software:** Jeff W. Atkins
**Supervision:** Jeff W. Atkins
**Validation:** Jeff W. Atkins, Brandon Alveshere, Sabrie Breland, Courtney Meier, Michael Langley, Lu Zhai
**Visualization:** Jeff W. Atkins
**Writing – original draft:** Jeff W. Atkins
**Writing – review & editing:** Jeff W. Atkins, Brandon Alveshere, Sabrie Breland, Courtney Meier, Michael Langley, Lu Zhai



# Individual Tree and Sapling Aboveground Biomass (AGB) Estimates for 35 NEON Terrestrial Observation Sites for 2014–2023

**Jeff W. Atkins[1,2], Brandon Alveshere[2], Sabrie Breland[3], Courtney Meier[4], Michael Langley[1], and Lu Zhai[5]**

[1]USDA Forest Service, Southern Research Station, New Ellenton, SC, USA, [2]Department of Biology, Virginia Commonwealth University, Richmond, VA, USA, [3]USDA Forest Service, Savannah River Site, New Ellenton, SC, USA, [4]National Ecological Observatory Network, Battelle, Boulder, CO, USA, [5]Department of Natural Resource Ecology & Management, Oklahoma State University, Stillwater, OK, USA

**Abstract** Here we present aboveground biomass (AGB) estimates from individual tree diameters scaled to whole-tree biomass estimates using generalized allometric equations for 35 National Ecological Observatory Network (NEON) sites within the United States and Puerto Rico. These data are in both a standalone data file made publicly available via Figshare and as an R data package (NEONForestAGB) that allows for direct import of data into the R statistical computing environment. AGB is an Essential Climate Variable (ECV), yet biomass estimation from large forest inventory data can be cumbersome. Here we seek to provide a useful data set for community use from NEON data. The data set includes 92,281 unique individuals of 478 different species from 1,216 terrestrial observation plots for 360,570 biomass estimates between the years 2014 and 2023.

**Plain Language Summary** We created a data set of individual tree-level estimates of aboveground biomass for 35 NEON sites located in the United States and Puerto Rico. Biomass estimates are based on individual stem diameter measurements taken from plots at those sites between the years 2014 and 2023, which were then scaled mathematically to whole tree biomass using equations from two common data sources that allow for estimation of whole-tree volume based on stem diameter measurements. We provide aboveground biomass estimates for every measured tree that we could find a suitable equation for, noting that some tree species lack biomass scaling equations due to the scarcity of that species or the lack of previous scientific or market interest. These data have been made available via direct download from a data repository (Figshare) and as a data package (NEONForestAGB) that can be directly imported into R to streamline use.

## 1. Introduction

Accurate, constrained, and consistent biomass estimates are important for understanding the global carbon system as well as informing carbon offset and biomass markets. To monitor aboveground carbon storage and inform forest management, it is often necessary to have individual tree-level biomass estimates. Biomass is a recognized Global Climate Observing System (GCOS) Essential Climate Variable (ECV) and an important input to Earth system models (Herold et al., 2019). At the individual level, tree biomass is commonly determined by allometric equations which relate measures of tree diameter and/or height to overall tree volume and biomass (Hulshof et al., 2015). Allometric equations are based on dimensional analyses and built from scaling relationships of tree stem diameter measurements to overall tree volume and mass determined through standardized destructive sampling methods (Chojnacky et al., 2014; Jenkins et al., 2003; Ter-Mikaelian & Korzukhin, 1997). Stem diameter measurements for allometric scaling are most often taken at "breast-height" (diameter-at-breast-height or DBH) which is at approximately 130–140 cm stem height, although stem diameter measurements taken below breast-height may be adjusted using stem taper functions. Allometric equations for biomass estimation can be species- and region-specific (Duncanson et al., 2015) or can be generalized—that is, based on groupings of species that share morphological, phenological, and taxonomic characteristics (Jenkins et al., 2003).

Here we provide estimates of total aboveground biomass (AGB) for sapling, small, single-bole, and multi-bole trees from 35 National Ecological Observatory Network (NEON) terrestrial observation sites within the United States and Puerto Rico using two generalized allometries (Chojnacky et al., 2014; Jenkins et al., 2003) and one sapling-based allometry (Annighöfer et al., 2016). Estimates reflect total biomass per individual, inclusive of

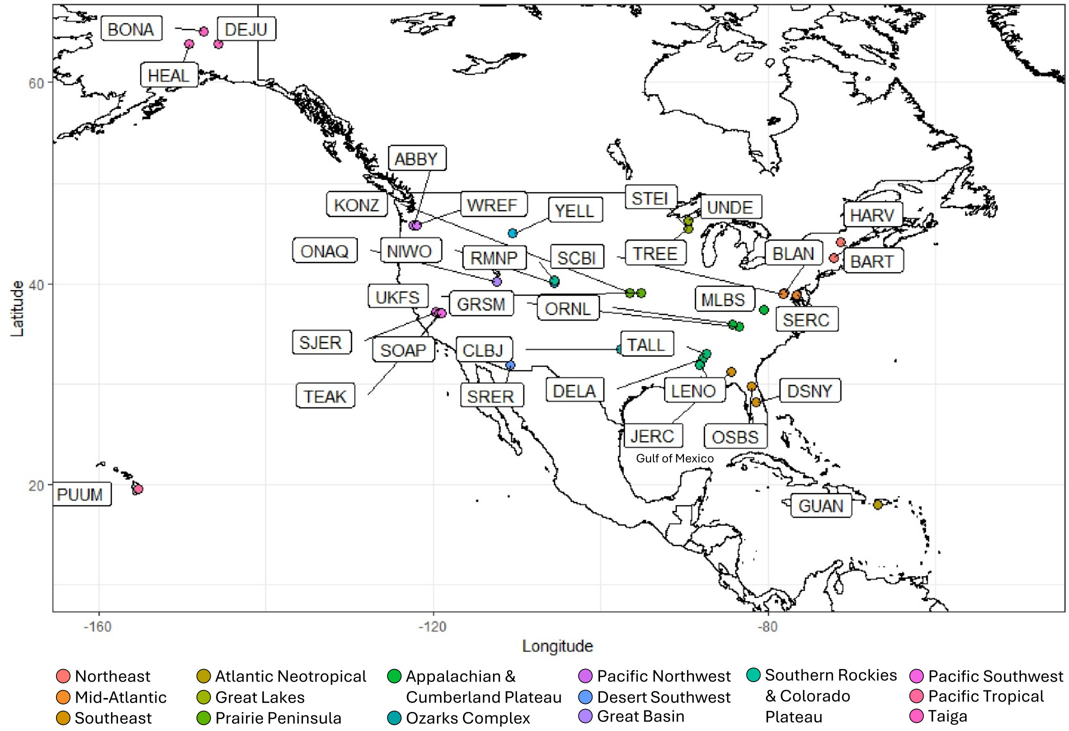


**Figure 1.** Mapped locations of National Ecological Observatory Network (NEON) sites with each site labeled using NEON four-letter "siteID" codes. Colors represent NEON ecoclimatic domains, which were determined based on clustering analyses of shared vegetation, landforms, and ecosystem dynamics (Hargrove & Hoffman, 1999, 2004). Sites chosen for this analysis span 15 of the 20 total NEON ecoclimatic domains, thus capturing a wide range of ecological and climatic conditions. For more detailed information see https://www.neonscience.org/field-sites/about-field-sites.

all aboveground components (e.g., boles, branches, leaves, etc.). AGB estimates are made at the individual level and do not reflect either plot- or site-level AGB estimates. These data are made available for direct download via Figshare (https://doi.org/10.6084/m9.figshare.25625352.v2) as well as in the form of a data package (NEON-ForestAGB; https://zenodo.org/doi/10.5281/zenodo.11193519) for the R statistical and computing environment (R Core Team, 2023). All data are machine-readable and formatted for additional analyses and data product generation. A full description of the variables in the data set is included in Table S1 of Supporting Information S1. This analysis is also informed by additional data and information from the USDA PLANTS Database ("USDA Plants Database", 2024) and from the USDA Forest Service Forest Inventory and Analysis program (FIA) documentation (Woudenberg et al., 2010) to provide greater interoperability with other data.

## 2. Materials and Methods

### 2.1. NEON Site Selection

There are 81 NEON sites in total, including 47 terrestrial observation sites—the remaining 34 are aquatic observation sites. Collectively, these sites represent 20 distinct ecoclimatic domains of the US and Puerto Rico (Hargrove & Hoffman, 1999, 2004; Figure 1) and represent 7 of the 9 major biomes found in the world—with the exception of tundra and tropical rain forest biomes (Ricklefs, 2008; Figure 2). The 35 NEON terrestrial sites selected for these analyses were chosen based solely on the presence or absence of trees and represent 15 of the ecoclimatic domains.

### 2.2. NEON Vegetation Structure Data Overview

The NEON vegetation structure (VST) data are the source data for these estimates (National Ecological Observatory Network, 2021; NEON Product ID: DP1.10098.001). When accessed, VST data include several files (described more fully at https://data.neonscience.org/data-products/DP1.10098.001). To create biomass

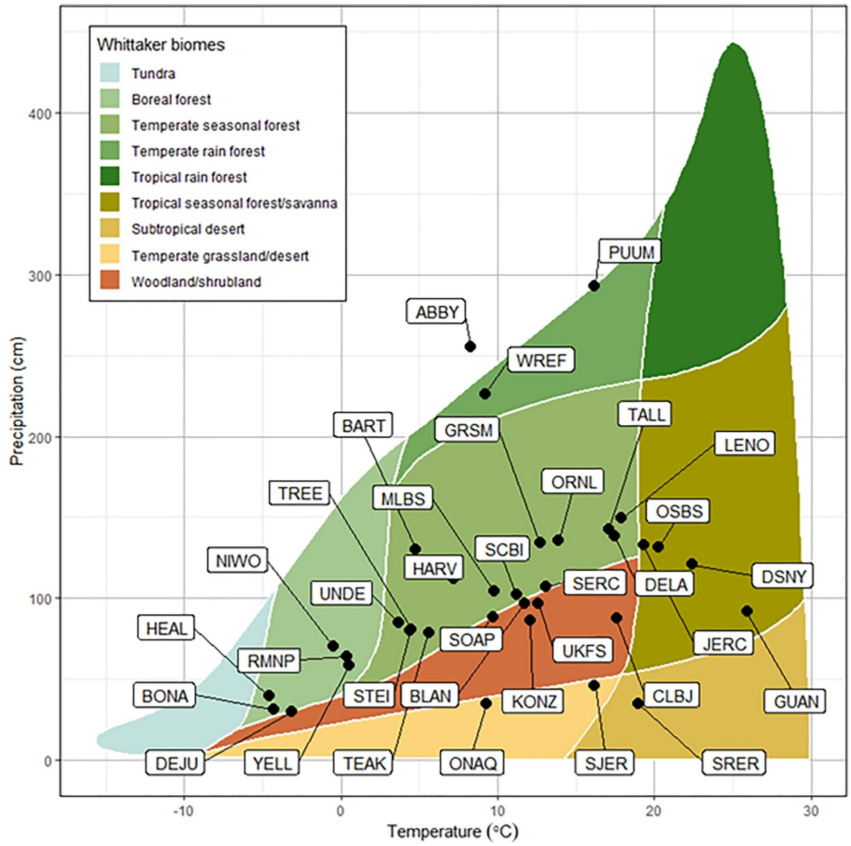


**Figure 2.** Distribution of the 35 National Ecological Observatory Network sites in this analysis in climate space based on mean annual temperature (MAT; *x*-axis) and mean annual precipitation (MAP; *y*-axis) from WorldClim2 data (Fick & Hijmans, 2017). Colors indicate the Whittaker biomes that typically occur for a given MAP or MAT region (Ricklefs, 2008; Ştefan & Levin, 2018).

estimates, we used the "vst_mappingandtagging" and "vst_apparentindividual" files. Each file contains the variable "individualID," a universal identifier for each tagged and measured plant within the NEON VST data product. The "individualID" (e.g., NEON.PLA.D16.ABBY.02745) is a concatenation of specific identifying variables including the data set of origin (e.g., PLA for plants), domain (e.g., D16 for Domain 16—Pacific Northwest), site (e.g., ABBY, NEON terrestrial site Abby Road) and a unique, sequential number (e.g., 02745), which allows different data files within the data product (i.e., tables) to be merged. The "vst_mappingandtagging" table contains invariant location and taxonomic information for each "individualID," including the four-letter USDA PLANTS code ("USDA Plants Database", 2024) corresponding to the species/subspecies of the tree ("taxonID"), the identified species ("scientificName"), and plot location information such as distance ("stemDistance") and azimuth ("stemAzimuth") relative to a plot point ("pointID") as determined from high-resolution GPS data. The "vst_apparentindividual" table contains information for each individual per measurement event, including three broad groups of information important for this analysis: (a) identification data—this includes the "individualID" as well as the measurement date ("date") and corresponding plot information ("plotID," "subplotID"); (b) plant condition metadata—the plant growth form or habit ("growthForm"), position within the canopy ("canopyPosition"), and health status ("plantStatus"); and (c) structural measurements—measured diameter ("stemDiameter" or "basalStemDiameter") the height at which the diameter was taken ("measurementHeight" or "basalStemDiameterMsrmntHeight") and plant height ("height"). The NEON standard operating procedure (SOP) for VST sampling defines 7 distinct growth forms ("growthForm") (Meier & Jones, 2023):

1. Lianas—non-self-supporting woody stems with diameter-at-breast height (DBH) ≥ 1 cm.
2. Single-bole tree (SBT)—Self-supporting individual with a single bole ≥10 cm DBH, typically taller than 5 m.

3. Multi-bole tree (MBT)—Similar to SBT, but with at least one bole ≥10 cm DBH and additional boles that qualify for measurement based on standardized criteria (see NEON.DOC.000987, “Measurement of Vegetation Structure” protocol).
4. Small tree (SMT)—Woody individual with potential to grow into either single- or multi-bole tree; at least one stem 1 cm ≤ DBH <10 cm, typically shorter than 5 m.
5. Sapling (SAP)—Self-supporting individual, potential to grow into SBT or MBT, basal diameter (at 10 cm height) >1 cm, DBH <1 cm or height <130 cm.
6. Single shrub (SIS)—Self-supporting individual, single or multiple primary stems, DBH of at least one stem 1 cm ≤ DBH <10 cm, typically shorter than 5 m.
7. Small shrub (SMS)—Similar to SIS, but basal diameter (at 10 cm) ≥1 cm for at least one stem and DBH <1 cm or height <130 cm. Individuals that are SMS may grow into a SIS, although this category also includes smaller-stature species that rarely or never exceed 130 cm in height.

Given the limited availability and applicability of biomass allometries for lianas and shrubs, we chose to focus only on sapling and tree growth forms for this work—that is, SBT, MBT, SMT, and SAP. There is a difference in sampling, classification, and nomenclature between NEON and FIA data and workflows. First, in FIA protocols, stem diameters are measured at a standard DBH of 137 cm height. Only trees with DBH values above a 12.7 cm threshold were included as trees and considered in biomass estimates for FIA. Second, FIA protocols classify trees with DBH between 2.54 and 12.7 cm as “saplings” and trees with DBH below 2.54 cm were not surveyed or sampled. Typically, seedlings and saplings (trees with DBH <12.7 cm) are omitted from AGB estimations, although sometimes measured individuals that do not reach the 2.54 cm DBH threshold may also be classified as “seedlings” (Annighöfer et al., 2016) and included in some AGB estimates. In NEON protocols, measurement height may vary, but for trees (SBT, MBT, SMT) is 130 cm—an approximation of the 137 cm FIA height. NEON's stem diameter range for SMT is 1–10 cm, and 10 cm is the stem diameter threshold for trees (SBT and MBT). NEON also supplies stem diameter measurements at the root-collar or basal height (typically 10 cm above the ground), as well as height measurements for many NEON-defined saplings (SAP; stem diameter <1 cm). We have included estimates for both the small tree and sapling AGB pools for a selection of taxa for which suitable allometries could be found.

### 2.3. AGB Data Processing

To calculate AGB for each individual plant, NEON VST data (National Ecological Observatory Network (NEON) 2023) were downloaded and imported directly into R 4.3.2 (R Core Team, 2023) using the *neonstore* package (Boettiger et al., 2021). Data were then filtered first by site based on presence or absence of trees, retaining all data if any value in the “growthForm” column included at least one tree or sapling (i.e., SBT, MBT, SAP, SMT). Then two key data files (“vst_mappingandtagging” and “vst_apparentindividual”) from the NEON VST data product were merged by the shared “individualID” column, thus creating a singular data frame containing the variables: “individualID,” “plotID,” “siteID,” “domainID,” “date,” “plantStatus,” “growthForm,” “canopyPosition,” “measurementHeight,” “height,” “stemDiameter,” “basalStemDiameter,” “basalStemDiameterMsrmntHeight,” “remarks,” “taxonID,” and “scientificName.”

### 2.4. Allometric Equations

To begin individual biomass estimation, data were merged by the variable “taxonID” with allometric equation information from both Jenkins et al. (2003) and Chojnacky et al. (2014) based on a constructed look up table (see file “MasterTaxonList.csv”). Each of these sets of allometric equations are generalized allometries, meaning they are suitable for application for one or more species, genera, or family, depending on the equation set or specific allometry. Jenkins groups ∼100 species or genera into nine equations, whereas Chojnacky groups ∼130 species or genera into 35 equations—building on the Jenkins allometries by additionally considering interspecific differences in specific wood gravity. The NEON VST data set includes ∼420 different tree species—with the caveat that these are field identified trees, and in some instances, workers were unable to determine the specific species and listed the tree at the genus level (e.g., reporting *Abies* spp.), or if that was not possible, the tree was identified with a “taxonID” of “2PLANT,” “2PLANT-H,” or “2PLANT-S” which stands for unidentified tree, unidentified hardwood, and unidentified softwood, respectively.

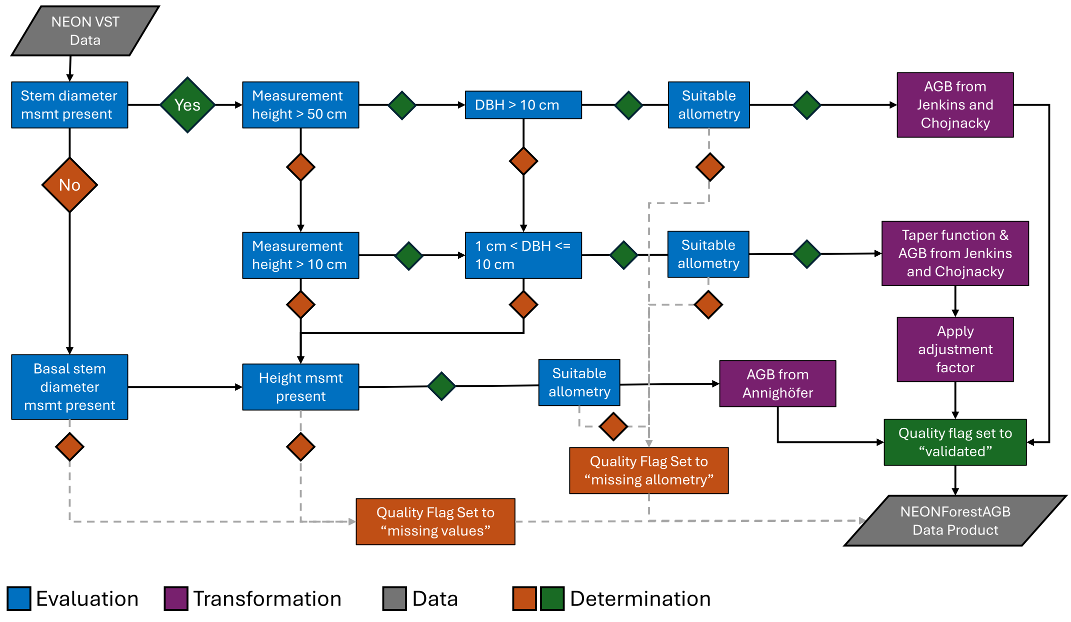


**Figure 3.** AGB workflow showing evaluation and determination points through the entire process.

To calculate AGB as comprehensively as possible, we associated each "taxonID" without a corresponding Jenkins or Chojnacky model classification with the most appropriate match by carefully considering the taxonomy and morphology for each unmatched "taxonID." First we used the USDA Plants ("USDA Plants Database," 2024) database and the Flora of North America (Flora of North America Editorial Committee, 2023) to classify growth habit at the species, genus, or family level if possible. Growth habit is the form that individuals from that taxon have the potential to adopt as opposed to the growth form observed in the field—measured in the NEON data as "growthForm." Taxa were then classified by "growthForm" as herb/other, liana, shrub, tree, or tree/shrub. This allowed us to identify whether an unmatched "taxonID" was associated with a tree or not. Then, with additional information from the FIA Database User's Manual (Woudenberg et al., 2010) which provides detailed information on species as classified by their "FIA.code" (an equivalent identifier to the USDA PLANTS/NEON "taxonID"), we matched some "taxonID" values to corresponding Jenkins allometries based on growth habit, genus, and family taxonomic information. Then, for some "taxonID" values that could not be matched to Jenkins allometries via the previous step, specific wood gravity values were taken from published data tables and resources (Alden, 1995; Miles, 2009) that allowed for matching with Chojnacky allometries based on the subcategorization of certain taxa using differences in specific wood gravity. The data included 737 taxa in total ("taxonID") for the 35 NEON sites used in this analysis. The total number of taxa included all growth forms measured by NEON and of those 737 taxa, 422 had resolvable matching allometries from Jenkins et al. (2003) and 193 had resolvable allometries from Chojnacky et al. (2014), leaving 309 taxa with no matching allometry from which to estimate biomass. These assignments are all contained in the "MasterTaxonList.csv" for evaluation and future revisions. When no suitable allometry could be determined, those data remain as full records with no biomass determined—this issue is primarily due to limitations in taxonomic classification or insufficient data from which to generate suitable allometries and could be resolved with additional effort from domain experts.

## 2.5. AGB Workflow

To determine AGB, data were first sorted by measurement height and then by stem (SD) or basal stem diameter (BSD) as demonstrated in the workflow diagram (Figure 3). Table 1 shows details of the data by diameter and measurement height classes. Although this data product is meant to streamline the use of these data so that users do not need to work directly with NEON forest inventory data, it is important to understand that "stemDiameter" as a variable is not synonymous with the common variable of diameter-at-breast-height (DBH), but rather is the measured diameter on the plant at the "measurementHeight." For example, a measurement height of 130 cm would be a common DBH measurement analogous to FIA standard height of 137 cm, whereas a measurement at 10 cm height would correspond to a BD measurement. From 2018 and onward, most BD data are reported in the

**Table 1**
*Description of National Ecological Observatory Network Vegetation Structure Data by Measurement Heights of Stem and Basal Stem Diameters*

| | | Measurement height class | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Missing Msmt. Ht. | Basal height | Stem diameter 0–10 cm | 10 cm | 10–50 cm | 50–130 | 130 cm | Greater than 130 cm |
| Diameter class | Missing diameter | 4,302 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Sapling (w/basal stem Msmt) | 0 | 60,138 | 2 | 2 | 0 | 0 | 9 | 0 |
| | Sapling (SD < 1 cm) | 0 | 15 | 0 | 21 | 0 | 0 | 12 | 1 |
| | Small tree (1 < SD < 10 cm) | 0 | 0 | 37 | 9,245 | 63 | 1,109 | 63,980 | 2,595 |
| | Tree (SD > 10 cm) | 0 | 0 | 31 | 24 | 117 | 2,706 | 98,969 | 6,169 |

*Note.* SD—stem diameter and BSD—basal stem diameter.

separate "basalStemDiameter" field, which was used here to calculate sapling biomass, but older records may have basal diameters reported via the "stemDiameter" and "measurementHeight" fields. Users of the data are encouraged to take note. Within the data, some individuals have both BSD and SD measurements—different data fields as described above. There were 394 instances for 263 unique individuals (UI) where this was the case. AGB estimation for those data were determined based on the workflow (Figure 3) as outlined.

#### 2.5.1. Tree Biomass (SD > 10 cm)

To estimate biomass for trees with DBH >10 cm and measurement height >50 cm, SD measurements were passed to the primary allometric functions. Both Jenkins et al. (2003) and Chojnacky et al. (2014) include species- or taxon-specific coefficients (i.e., β0 and β1) that convert SD (also known as diameter at breast height or DBH) using the same equation:

$$\text{AGB} = \exp(\beta_0 + \ln \beta_1 \text{DBH}) \tag{1}$$

Within our workflow, AGB was calculated using this approach with both Jenkins and Chojnacky for all individuals that met criteria: (a) DBH >10 cm; (b) measurement height >50 cm; and (c) presence of suitable allometry. This was done regardless of the recorded plant status ("plantStatus"; e.g., live, dead, or broken).

#### 2.5.2. Small Trees (1 < SD < 10 cm)

For smaller trees with SD between 1 and 10 cm and measurement height above 50 cm, biomass estimation followed the same steps as outlined above with the added step that after biomass was estimated using Equation 1 with an additional adjustment factor (Woudenberg et al., 2010) used by FIA. These adjustment factors, which are genera- or species-specific, are published in Woudenberg et al. (2010) and included in both the final data file ("NEONForestAGBv2.csv) and the source taxon data ("MasterTaxonList.csv").

For smaller trees with SD between 1 and 10 cm and measurement height between 10 and 50 cm, we applied a taper function (Chojnacky et al., 2014) that estimates diameter-at-breast height (DBH) measurements from the diameter-at-root-collar (DRC) measurement using a log-log function:

$$\log(\text{DBH}) = -0.35031 + 1.03991 \log(\text{DRC}) \tag{2}$$

In this equation, we used the measured diameter of the tree taken within the approximate 10–50 cm range as DRC —with the assumption based on field protocols that these are DRC measures. This approach was informed based on regression of 205 trees including multiple conifer and deciduous species with an $R^2$ = 0.987 (Chojnacky et al., 2014) between DRC and actual DBH measurements. The estimated DBH (~130 cm) value (named "est. dbh" within the data frame) is then passed to Equation 1 above, with AGB determined in the same manner. Then, the small tree adjustment described above was used to calibrate those biomass estimates to account for species/genus level differences in biomass allocation (Woudenberg et al., 2010).

**Table 2**

*Tallies of Number of Unique Individuals, Measurement Moments, and Aboveground Biomass Estimations From the National Ecological Observatory Network Vegetation Structure Data From 2014 to 2023: Total Sample (n)*

| Growth form | Total Msmt moments[a] | Total SD and BSD Msmts[a] | No. of unique individuals (UI) | Total with AGB (Jenkins) | Total with AGB (Chojnacky) | Total with AGB (Annighöfer) |
|---|---|---|---|---|---|---|
| MBT | 17,131 | 16,941 | 6,631 | 16,472 | 13,108 | - |
| SBT | 93,213 | 92,668 | 30,401 | 91,995 | 81,765 | 2 |
| SMT | 65,095 | 64,677 | 25,209 | 63,338 | 52,118 | 23 |
| SAP | 71,123 | 67,498 | 37,040 | 8,138 | 5,897 | 33,804 |
| Total trees | 246,562 | 241,784 | 92,281[b] | 179,943 | 152,798 | 33,829 |

[a]The "Total Msmt Moments" column exceeds "Total SD and BSD Msmts"—stem diameter (SD) and basal stem diameter (BSD) measurements may not be taken at every measurement interval due to field-related issues with sampling preventing measurements form being taken at that time, though those data have been retained.
[b]Column sum does not match the sum of all rows because some individuals change growth form over time or between sampling events and are therefore counted multiple times if grouped by growth form.

### 2.6. Sapling Biomass

Sapling biomass was determined for 15 genera—those corresponding to allometries determined from matching genera in European forests (Annighöfer et al., 2016)—for which there were both BSD measurements at or below 10 cm height (typically the "basalStemDiameter" values) and measurements of sapling height ("height"). For individuals meeting those criteria and for which there were matching allometries, the BD and height (H) measurements were then passed to the following equation to determine AGB:

$$\mathrm{AGB} = \beta_1 \left(\mathrm{BD}^2\mathrm{H}\right)^{\beta_2} \tag{3}$$

## 3. Data

### 3.1. Accessing and Using the Data

In promotion of FAIR data principles, we have made the data available as both a direct download from a public data repository (Figshare; https://doi.org/10.6084/m9.figshare.25625352.v2) and as a data package for the R Statistical and Computing Environment (NEONForestAGB). Data may be updated annually as NEON updates VST data based on annual data collections. Efforts will be made to keep the "NEONForestAGB" product as current as possible, with package versions reflecting any updates—users should consult associated metadata for version information. When using these data for analysis and publications, please cite this manuscript, the data product used (either NEONForestAGB or the Figshare repository), and the original NEON data source (NEON Product ID: DP1.10098.001).

### 3.2. Data Overview and Summary

For the 35 sites and 4 growth forms (MBT, SBT, SMT, and SAP) for the years 2014–2023 for which we estimated AGB, there were are a total of 92,281 UI based on distinct occurrences in the "individualID" field resulting in a total of 246,562 measurement moments—number of UI times the number of measurement dates—and a total of 241,784 SD and BSD measurements (Table 2). There are fewer diameter measurements than measurement moments due to senescence and mortality or missed measurements during measurement time periods (e.g., workers could not find the individual to take a measurement). Table 2 shows detailed tallies, but in brief, we were able to estimate AGB for ~99% of the available stem diameter data for trees (i.e., SBT, SMT, and MBT) using the Jenkins et al. (2003) allometries and for ~84% of trees using both Jenkins et al. (2003) and Chojnacky et al. (2014; *All taxa with a Chojnacky equation had a corresponding Jenkins equation, but not vice versa, hence the difference). We were able to estimate AGB for ~50% of the saplings from the Annighöfer et al. (2016) allometric set —this proportion increases to ~62% when sapling estimates from Jenkins et al. (2003) are included. Inconsistencies in this table may be attributed to some individuals changing growth forms as they age. Additionally, mortality is high and thus saplings often disappear between sampling events, but for the purpose of productivity estimation, NEON creates and retains records for saplings that are not subsequently found.

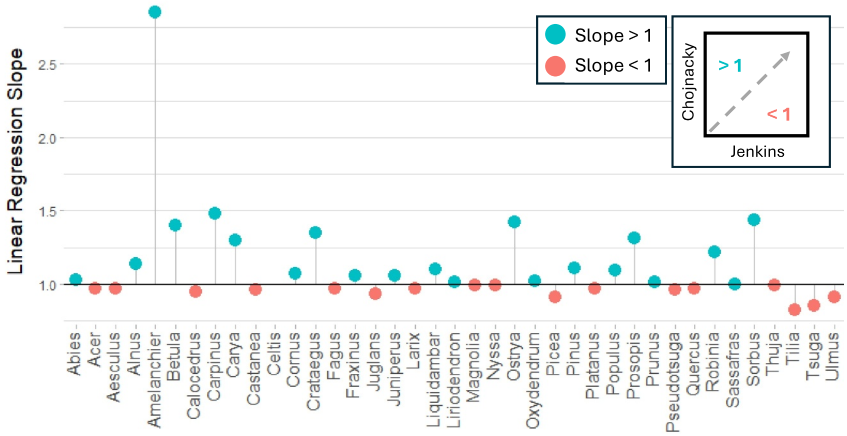


**Figure 4.** Illustration of regression slopes comparing Chojnacky to Jenkins AGB estimates for six taxa. Slopes greater than one (in blue) indicate that AGB estimates from Chojnacky are higher than Jenkins, whereas slopes less than one (in red) indicate that AGB estimates from Jenkins are higher than Chojnacky. A 1:1 line (dark gray, dashed) is included for reference purposes.

### 3.3. Differences in AGB Estimates by Allometries

Individual AGB estimates from Jenkins and Chojnacky generally agree, yet noticeable differences in the magnitude of AGB estimates between the two equation sets do appear for certain tree species or genera (Figure 4; Table S2 in Supporting Information S1). To illustrate bias between each equation set we used the *lm_table* function from the *forestmangr* package (Braga et al., 2023) in *R* to run linear regressions by taxon and then generated regression statistics (e.g., $R^2$, slope, intercept, SE; Figure 4; Table S2 in Supporting Information S1). *Tilia* and *Tsuga* each had slopes below 0.90, a strong negative bias indicating that Chojnacky et al. (2014) produce lower AGB estimates of these taxa relative to Jenkins et al. (2003) (Figure 4). Many taxa had slope values above 1.1, with *Amelanchier*, the most notable at 2.85—a strong positive bias, whereas *Alnus, Betula, Carpinus, Carya, Crataegus, Ostrya, Pinus, Prosopis, Robinia, and Sorbus* were all also above 1.1 (Figures 4 and 5). All the others fell between 0.9 and 1.1. This analysis only demonstrates the presence or absence of bias and in which direction that bias is, if present. A full examination of these differences was outside of the purview of this work but is presented as a description of the data and notice to potential users.

## 4. Results and Discussion

The provided NEONForestAGB data set and data package allow users to address science and policy questions related to aboveground forest biomass with NEON data more easily and quickly, without the need to work with NEON stem diameter and inventory data directly. These data enable the creation of additional data products (e.g., maps of plot-level biomass, carbon sequestration rates) of utility to the research community. The generalizable nature of both the Jenkins et al. (2003) and Chojnacky et al. (2014) equation sets makes them ideal for biomass estimation over large areas and for continuity among and between studies, inventories, and assessments. The underlying principle for each is the scaling relationship between stem diameter and overall tree biomass from a mathematical equation specific to a grouping of species/genera based on taxonomy and shared traits (e.g., wood specific gravity, growth form). We have sought to match as many species as possible to a corresponding allometry—resulting in a matching of ~99% of all stem diameter measurements in the NEON data set for these 35 sites. However, there remain many tree species for which no corresponding allometry could be conclusively determined based on our survey of the existing literature and data. We found a total of 309 species missing Jenkins et al. (2003) and 538 species missing Chojnacky et al. (2014) allometries (Table S2 in Supporting Information S1). To resolve this issue would require more information about these species, specifically measurements of wood specific gravity and destructive sampling to establish diameter-biomass scaling relationships, enabling the derivation of allometric equations. Another possible means to address this is through non-destructive approaches to establish allometric relationships, such as terrestrial laser scanning (Stovall et al., 2018). It is likely that the scarcity of species without existing allometries and their non-commercial viability are the major contributors to the lack of information.

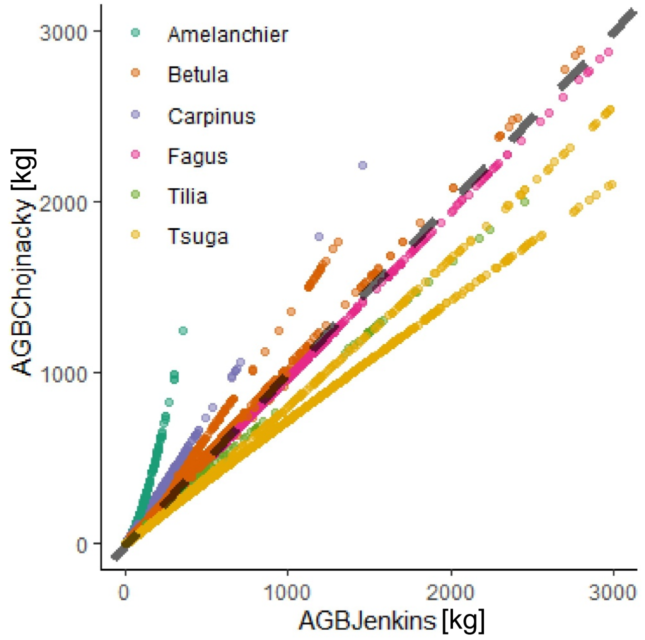


**Figure 5.** AGB (in kg) as estimated by both Jenkins ($x$-axis) and Chojnacky ($y$-axis) for six taxa to illustrate the range of bias present between the two allometric equation sets at the taxon-level. Amelanchier (in green) has the strongest bias, with AGB estimates from Chojnacky much higher than those from the corresponding Jenkins equation for the taxon. Conversely, Tsuga (in yellow) has the strongest bias toward Jenkins. Other genera presented to illustrate the range of agreement.

One novel addition this data set provides is the inclusion of AGB from saplings—an AGB pool that often goes unconsidered or unmeasured. Capturing the entire magnitude and flux of carbon via AGB estimates, however, necessitates consideration of this biomass pool. However, there are major considerations when using the sapling AGB estimates. Many saplings may appear only once in the data and as they were not remeasured. There are valid reasons for this, that is, saplings that have died between sampling events but have their records retained in "vst_apparentindvidiual" in the next sampling event to enable productivity estimates. Generally, sapling biomass estimates remain elusive simply because there is often no need for such estimates. Shrub biomass is currently not included in these data but could be added in the future to facilitate more comprehensive estimates of total AGB. Additionally, this would allow for the addition of many NEON terrestrial sites where shrubs are the primary or only AGB pool. However, generalized shrub biomass allometries at the continental scale currently do not exist. We have included a look-up table and processing scripts that would allow for AGB estimates from species missing those estimates currently if suitable allometries are found. As all methods and scripts are open-source, this allows for community interaction and participation—which we welcome.

AGB estimates in this data product are at the individual tree level. For purposes of making estimates of net primary productivity, areal biomass estimates, or other derived data product from these data, it is imperative that users consult the appropriate data fields for required calculations for their purpose. Tree-level estimates do not represent overall biomass for a site, nor is every plot at every NEON site sampled annually. Users should consult NEON sampling design documents and associated peer-reviewed literature included in the references and cited within this manuscript during any subsequent use of this data product.

## Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

## Data Availability Statement

Data supporting this manuscript include the NEONForestAGB data set (Atkins et al., 2024) and supporting R data package (Atkins, 2025)—both of which are derived from source field inventory data from NEON (NEON, 2023).

**Acknowledgments**
The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle. This material is based in part upon work supported by the National Science Foundation through the NEON Program (Award 1724433). The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

## References

Alden, H. A. (1995). *Hardwoods of North America*. U.S. Department of Agriculture, Forest Service.

Annighöfer, P., Ameztegui, A., Ammer, C., Balandier, P., Bartsch, N., Bolte, A., et al. (2016). Species-specific and generic biomass equations for seedlings and saplings of European tree species. *European Journal of Forest Research*, *135*(2), 313–329. https://doi.org/10.1007/s10342-016-0937-z

Atkins, J. W. (2025). NEONForestAGB, R package version 1.0.2. Retrieved from https://zenodo.org/records/12795497

Atkins, J. W., Meier, C., Alveshere, B., Breland, S., Langley, M., & Zhai, L. (2024). NEONForestAGBv2 [Dataset]. *Figshare*. https://doi.org/10.6084/m9.figshare.25625352.v2

Boettiger, C., Thomas, Q., Laney, C., Lunch, C., & Ross, N. (2021). Neonstore: NEON data store (version 0.4.4). Retrieved from https://cran.r-project.org/web/packages/neonstore/index.html

Braga, S. R., de Oliveira, M. L. R., & Gorgens, E. B. (2023). forestmangr: Forest mensuration and management (version 0.9.6). Retrieved from https://cran.r-project.org/web/packages/forestmangr/index.html

Chojnacky, D. C., Heath, L. S., & Jenkins, J. C. (2014). Updated generalized biomass equations for North American tree species. *Forestry: An International Journal of Forest Research*, *87*(1), 129–151. https://doi.org/10.1093/forestry/cpt053

Duncanson, L. I., Dubayah, R. O., & Enquist, B. J. (2015). Assessing the general patterns of forest structure: Quantifying tree and forest allometric scaling relationships in the United States. *Global Ecology and Biogeography*, *24*(12), 1465–1475. https://doi.org/10.1111/geb.12371

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Flora of North America Editorial Committee. (2023). *Flora of North America North of Mexico [ONLINE]* (Vol. 1–25). Oxford University Press. Retrieved from http://beta.floranorthamerica.org

Hargrove, W. W., & Hoffman, F. M. (1999). Using multivariate clustering to characterize ecoregion borders. *Computers in Science and Engineering*, *1*(4), 18–25. https://doi.org/10.1109/5992.774837

Hargrove, W. W., & Hoffman, F. M. (2004). The potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management*, *34*(5), S39–S60. https://doi.org/10.1007/s00267-003-1084-0

Herold, M., Carter, S., Avitabile, V., Espejo, A. B., Jonckheere, I., Lucas, R., et al. (2019). The role and need for space-based forest biomass-related measurements in environmental management and policy. *Surveys in Geophysics*, *40*(4), 757–778. https://doi.org/10.1007/s10712-019-09510-6

Hulshof, C. M., Swenson, N. G., & Weiser, M. D. (2015). Tree height–diameter allometry across the United States. *Ecology and Evolution*, *5*(6), 1193–1204. https://doi.org/10.1002/ece3.1328

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, *49*(1), 12–35. https://doi.org/10.1093/forestscience/49.1.12

Meier, C., & Jones, K. (2023). *TOS protocol and procedure: Measurement of vegetation structure* (NEON.DOC.000987vM). NEON (National Ecological Observatory Network).

Miles, P. D. (2009). *Specific gravity and other properties of wood and bark for 156 tree species found in North America*. U.S. Department of Agriculture, Forest Service.

National Ecological Observatory Network (NEON). (2021). Woody plant vegetation structure (DP1.10098.001): RELEASE-2021 (version RELEASE-2021) [Csv]. *National Ecological Observatory Network (NEON)*. https://doi.org/10.48443/E3QN-XW47

NEON. (2023). NEON.DOC.000987-TOS protocol and procedure: VST measurement of vegetation structure. Retrieved from https://data.neonscience.org/documents/10179/1883155/NEON.DOC.000987vM/4ff8197e-3858-6cd8-e3f7-6babf96b2482

R Core Team. (2023). R: A language and environment for statistical computing. (Version 4.3.2). Retrieved from http://www.R-project.org/

Ricklefs, R. E. (2008). *The economy of nature*. W. H. Freeman and Company. (Chapter 5, Biological Communities, The biome concept).

Ştefan, V., & Levin, S. (2018). *Plotbiomes: R package for plotting whittaker biomes with ggplot2 (v1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.7145245

Stovall, A. E., Anderson-Teixeira, K. J., & Shugart, H. H. (2018). Assessing terrestrial laser scanning for developing non-destructive biomass allometry. *Forest Ecology and Management*, *427*, 217–229. https://doi.org/10.1016/j.foreco.2018.06.004

Ter-Mikaelian, M. T., & Korzukhin, M. D. (1997). Biomass equations for sixty-five North American tree species. *Forest Ecology and Management*, *97*(1), 1–24. https://doi.org/10.1016/S0378-1127(97)00019-4

USDA Plants Database. (2024). USDA plants database. Retrieved from https://plants.usda.gov/home

Woudenberg, S. W., Conkling, B. L., O'Connell, B. M., LaPoint, E. B., Turner, J. A., & Waddell, K. L. (2010). The forest inventory and analysis database: Database description and users manual version 4.0 for phase 2. Gen. Tech. Rep. RMRS-GTR-245. *U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station*, *336*, 245. https://doi.org/10.2737/RMRS-GTR-245