

Package: allodb (via r-universe)

October 17, 2024

Title Tree Biomass Estimation at Extra-Tropical Forest Plots

Version 0.0.1.9000

Description Standardize and simplify the tree biomass estimation process across globally distributed extratropical forests.

License GPL-3

URL <https://docs.ropensci.org/allodb/>,
<https://github.com/ropensci/allodb>

BugReports <https://github.com/ropensci/allodb/issues>

Depends R (>= 3.4)

Imports data.table, ggplot2, rlang, tibble, withr

Suggests covr, knitr, rmarkdown, spelling, testthat (>= 3.0.0), utils

VignetteBuilder knitr

Config/testthat/edition 3

Encoding UTF-8

Language en-US

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Repository <https://ropensci.r-universe.dev>

RemoteUrl <https://github.com/ropensci/allodb>

RemoteRef master

RemoteSha 4207f8616411233f954779789c688433250e651b

Contents

equations	2
est_params	3
genus_family	4

get_biomass	5
gymno_genus	7
illustrate_allodb	7
koppenMatrix	9
missing_values	10
new_equations	10
references	12
resample_agb	13
sabi_stem1	14
shrub_species	15
sitespecies	15
sites_info	16
weight_allom	17

Index 19

equations	<i>Tables of allometric equations and associated metadata</i>
-----------	---

Description

- [equations](#): Table of allometric equations.
- [equations_metadata](#): Explanation of columns for [equations](#) table.

Usage

`equations`

`equations_metadata`

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 570 rows and 47 columns.

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 47 rows and 7 columns.

Details

A compilation of best available allometry equations to calculate tree above-ground biomass (AGB) per species based on extratropical ForestGEO sites.

Source

See [references](#) for equations original sources.

See Also

Other database datasets: [missing_values](#), [references](#), [sites_info](#), [sitespecies](#)

Examples

```
# preview the datasets
print(head(equations))
print(head(equations_metadata))
```

 est_params

Calibrate new allometric equations

Description

This function calibrates new allometric equations from sampling previous ones. New allometric equations are calibrated for each species and location by resampling the original compiled equations; equations with a larger sample size, and/or higher taxonomic rank, and climatic similarity with the species and location in question are given a higher weight in this process.

Usage

```
est_params(
  genus,
  coords,
  species = NULL,
  new_eqtable = NULL,
  wna = 0.1,
  w95 = 500,
  nres = 10000
)
```

Arguments

genus	a character vector, containing the genus (e.g. "Quercus") of each tree.
coords	a numeric vector of length 2 with longitude and latitude (if all trees were measured in the same location) or a matrix with 2 numerical columns giving the coordinates of each tree.
species	a character vector (same length as genus), containing the species (e.g. "rubra") of each tree. Default is NULL, when no species identification is available.
new_eqtable	Optional. An equation table created with the new_equations() function. Default is the compiled <i>allobd</i> equation table.
wna	a numeric vector, this parameter is used in the weight_allom() function to determine the dbh-related and sample-size related weights attributed to equations without a specified dbh range or sample size, respectively. Default is 0.1.
w95	a numeric vector, this parameter is used in the weight_allom() function to determine the value at which the sample-size-related weight reaches 95% of its maximum value (max=1). Default is 500.
nres	number of resampled values. Default is "1e4".

Value

An object of class "data.frame" of fitted coefficients (columns) of the non-linear least-square regression:

$$AGB = a * dbh^b + e, \text{ with } e \sim N(0, \sigma^2)$$

See Also

[weight_allom\(\)](#), [new_equations\(\)](#).

Examples

```
# calibrate new allometries for all Lauraceae species
lauraceae <- subset(scbi_stem1, Family == "Lauraceae")
est_params(
  genus = lauraceae$genus,
  species = lauraceae$species,
  coords = c(-78.2, 38.9)
)
```

genus_family

Genus and family table for selected ForestGEO sites

Description

Genus and their associated family identified in the extratropical ForestGEO sites used in allodb.

Usage

```
genus_family
```

Format

An object of class tbl_df (inherits from tbl, data.frame) with 248 rows and 2 columns.

See Also

Other datasets: [gymno_genus](#), [koppenMatrix](#), [scbi_stem1](#), [shrub_species](#)

Examples

```
# preview the dataset
print(head(genus_family))
```

get_biomass	<i>Compute tree aboveground biomass (AGB) based on allometric equations</i>
-------------	---

Description

This function calculates the aboveground biomass (or other tree components) of a given tree based on published allometric equations. Users need to provide a table (i.e. dataframe) with DBH (cm), parsed species Latin names, and site(s) coordinates. The biomass of all trees in one (or several) censuses can be estimated using this function.

Usage

```
get_biomass(
  dbh,
  genus,
  coords,
  species = NULL,
  new_eqtable = NULL,
  wna = 0.1,
  w95 = 500,
  nres = 10000
)
```

Arguments

dbh	a numeric vector containing tree diameter at breast height (dbh) measurements, in cm.
genus	a character vector (same length as dbh), containing the genus (e.g. "Quercus") of each tree.
coords	a numeric vector of length 2 with longitude and latitude (if all trees were measured in the same location) or a matrix with 2 numerical columns giving the coordinates of each tree.
species	a character vector (same length as dbh), containing the species (e.g. "rubra") of each tree. Default is NULL, when no species identification is available.
new_eqtable	Optional. An equation table created with the new_equations() function.
wna	a numeric vector, this parameter is used in the weight_allom() function to determine the dbh-related weight attributed to equations without a specified dbh range. Default is 0.1.
w95	a numeric vector, this parameter is used in the weight_allom() function to determine the value at which the sample-size-related weight reaches 95% of its maximum value (max=1). Default is 500.
nres	number of resampled values. Default is "1e4".

Details

allodb estimates AGB by calibrating a new allometric equation for each taxon (arguments `genus` and `species`) and location (argument `coords`) in the user-provided census data. The new allometric equation is based on a set of allometric equations that can be customized using the `new_eqtable` argument. Each equation is then given a weight with the `weight_allom()` function, based on: 1) its original sample size (numbers of trees used to develop a given allometry), 2) its climatic similarity with the target location, and 3) its taxonomic similarity with the target taxon (see documentation of the `weight_allom()` function). The final weight attributed to each equation is the product of those three weights. Equations are then resampled with the `resample_agb()` function: the number of samples per equation is proportional to its weight, and the total number of samples is provided by the argument `nres`. The resampling is done by drawing DBH values from a uniform distribution on the DBH range of the equation, and estimating the AGB with the equation. The couples of values (DBH, AGB) obtained are then used in the function `est_params()` to calibrate a new allometric equation, by applying a linear regression to the log-transformed data. The parameters of the new allometric equations are then used in the `get_biomass()` function by back-transforming the AGB predictions based on the user-provided DBHs.

Value

A "numeric" vector of the same length as `dbh`, containing AGB value (in kg) for every stem.

Warning

The function can run into some memory problems when used on large datasets (usually several hundred thousand observations).

See Also

`weight_allom()`, `new_equations()`

Examples

```
# Estimate biomass of all individuals from the Lauraceae family at the SCBI
# plot
lau <- subset(scbi_stem1, Family == "Lauraceae")
lau$agb <- get_biomass(lau$dbh, lau$genus, lau$species,
  coords = c(-78.2, 38.9)
)
lau

# Estimate biomass from multiple sites (using scbi_stem1 as example with
# multiple coord)
dat <- scbi_stem1[1:100, ]
dat$long <- c(rep(-78, 50), rep(-80, 50))
dat$lat <- c(rep(40, 50), rep(41, 50))
dat$biomass <- get_biomass(
  dbh = dat$dbh,
  genus = dat$genus,
  species = dat$species,
  coords = dat[, c("long", "lat")]
```

```
)
dat
```

gymno_genus

Gymnosperms identified in selected ForestGEO sites

Description

Table with genus and their associated family for Gymnosperms identified in the ForestGEO sites used in allodb.

Usage

```
gymno_genus
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 95 rows and 3 columns.

See Also

Other datasets: [genus_family](#), [koppenMatrix](#), [sobi_stem1](#), [shrub_species](#)

Examples

```
# preview the dataset
print(head(gymno_genus))
```

illustrate_allodb

Illustrate the resampling of AGB values used in allodb

Description

This function illustrates the resampling of AGB values used in *allodb*. It creates objects of class "ggplot".

Usage

```
illustrate_allodb(
  genus,
  coords,
  species = NULL,
  new_eqtable = NULL,
  logxy = FALSE,
  neq = 10,
  eqinfo = "equation_taxa",
```

```

    wna = 0.1,
    w95 = 500,
    nres = 10000
  )

```

Arguments

genus	A character value, containing the genus (e.g. "Quercus") of the tree.
coords	A numeric vector of length 2 with longitude and latitude.
species	A character value, containing the species (e.g. "rubra") of the tree. Default is NULL, when no species identification is available.
new_eqtable	Optional. An equation table created with the <code>new_equations()</code> function. Default is the base <i>allodb</i> equation table.
logxy	Logical: should values be plotted on a log scale? Default is FALSE.
neq	Number of top equations in the legend. Default is 10, meaning that the 10 equations with the highest weights are shown in the legend.
eqinfo	Which column(s) of the equation table should be used in the legend? Default is "equation_taxa".
wna	a numeric vector, this parameter is used in the <code>weight_allom()</code> function to determine the dbh-related and sample-size related weights attributed to equations without a specified dbh range or sample size, respectively. Default is 0.1.
w95	a numeric vector, this parameter is used in the <code>weight_allom()</code> function to determine the value at which the sample-size-related weight reaches 95% of its maximum value (max=1). Default is 500.
nres	number of resampled values. Default is "1e4".

Value

An object of class "ggplot" showing all resampled dbh-agb values. The top equations used are shown in the legend. The red curve on the graph represents the final fitted equation.

See Also

`weight_allom()`, `new_equations()`.

Examples

```

illustrate_allodb(
  genus = "Quercus",
  species = "rubra",
  coords = c(-78.2, 38.9)
)

```

koppenMatrix	<i>Koppen climate classification matrix</i>
--------------	---

Description

A table built to facilitate the comparison between the Koppen climate of a site and the allometric equation in question.

Usage

```
koppenMatrix
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 900 rows and 3 columns.

Details

The value of column `we` is the weight given to the combination of Koppen climates in columns `zone1` and `zone2`; the table is symmetric: `zone1` and `zone2` can be interchanged. This weight is calculated in 3 steps: (1) if the main climate group (first letter) is the same, the climate weight starts at 0.4; if one of the groups is "C" (temperate climate) and the other is "D" (continental climate), the climate weight starts at 0.2 because the 2 groups are considered similar enough; otherwise, the weight is 0; (2) if the equation and site belong to the same group, the weight is incremented by an additional value between 0 and 0.3 based on precipitation pattern similarity (second letter of the Koppen zone), and (3) by an additional value between 0 and 0.3 based on temperature pattern similarity (third letter of the Koppen zone). The resulting weight has a value between 0 (different climate groups) and 1 (exactly the same climate classification).

See Also

Other datasets: [genus_family](#), [gymno_genus](#), [scbi_stem1](#), [shrub_species](#)

Examples

```
# preview the dataset
print(head(koppenMatrix))
```

missing_values	<i>Explanations of missing values codes</i>
----------------	---

Description

Explanation of the codes used to indicate missing information in equation table.

Usage

```
missing_values
```

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 4 rows and 3 columns.

See Also

Other database datasets: [equations](#), [references](#), [sites_info](#), [sitespecies](#)

Examples

```
# preview the dataset
print(head(missing_values))
```

new_equations	<i>Modify the original equation table</i>
---------------	---

Description

This function modifies the original equation table to be used in other functions of the package including: subset the original equation table, add new equations, and choose whether to include equations with a height allometry.

Usage

```
new_equations(
  subset_taxa = "all",
  subset_climate = "all",
  subset_region = "all",
  subset_ids = "all",
  subset_output = c("Total aboveground biomass", "Whole tree (above stump)"),
  new_taxa = NULL,
  new_allometry = NULL,
  new_coords = NULL,
  new_min_dbh = NULL,
```

```

new_max_dbh = NULL,
new_sample_size = NULL,
new_unit_dbh = "cm",
new_unit_output = "kg",
new_input_var = "DBH",
new_output_var = "Total aboveground biomass",
use_height_allom = TRUE
)

```

Arguments

subset_taxa	character vector with taxa to be kept. Default is "all", in which case all taxa are kept.
subset_climate	character vector with Koppen climate classification to be kept. Default is "all", in which case all climates are kept.
subset_region	character vector with name of location(s) or country(ies) or broader region(s) (eg. "Europe", "North America") to be kept. Default is "all", in which case all regions/countries are kept.
subset_ids	character vector with equation IDs to be kept. Default is "all", in which case all equations are kept.
subset_output	What dependent variable(s) should be provided in the output? Default is "Total aboveground biomass" and "Whole tree (above stump)", other possible values are: "Bark biomass", "Branches (dead)", "Branches (live)", "Branches total (live, dead)", "Foliage total", "Height", "Leaves", "Stem (wood only)", "Stem biomass", "Stem biomass (with bark)", "Stem biomass (without bark)", "Whole tree (above and belowground)". Be aware that currently only a few equations represent those other variables, so estimated values might not be very accurate.
new_taxa	character string or vector specifying the taxon (or taxa) for which the allometry has been calibrated.
new_allometry	a character string with the allometric equation.
new_coords	a vector or matrix of coordinates (longitude, latitude) of the calibration data.
new_min_dbh	numerical value, minimum DBH for which the equation is valid (in cm). Default is NULL (nothing is added).
new_max_dbh	numerical value, maximum DBH for which the equation is valid (in cm). Default is NULL (nothing is added).
new_sample_size	number of measurements with which the allometry was calibrated. Default is NULL (nothing is added).
new_unit_dbh	character string with unit of DBH in the equation (either cm, mm or inch). Default is "cm".
new_unit_output	character string with unit of equation output (either "g", "kg", "Mg" or "lbs" if the output is a mass, or "m" if the output is a height).
new_input_var	independent variable(s) needed in the allometry. Default is "DBH", other option is "DBH, H".

`new_output_var` dependent variable estimated by the allometry. Default is "Total aboveground biomass".

`use_height_allom` a logical value. In *allodb* we use Bohn et al. (2014) for European sites. User need to provide height allometry when needed to calculate AGB. Default is TRUE.

Value

An object of class "data.frame" of new equations.

Examples

```
new_equations(
  new_taxa = "Faga",
  new_allometry = "exp(-2+log(dbh)*2.5)",
  new_coords = c(-0.07, 46.11),
  new_min_dbh = 5,
  new_max_dbh = 50,
  new_sample_size = 50
)
```

references

Equation references and associated metadata

Description

- [references](#): A data frame listing all references used in equation table.
- [references_metadata](#): Metadata for reference table.

Usage

references

references_metadata

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 57 rows and 6 columns.

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 7 rows and 4 columns.

Details

Bibliographical information for sourced equations. Links to the [equations](#) table by `ref_id`.

See Also

Other database datasets: [equations](#), [missing_values](#), [sites_info](#), [sitespecies](#)

Examples

```
# preview the datasets
print(head(references))
print(head(references_metadata))
```

resample_agb

Resample allodb equations to calibrate new allometries

Description

After attributing a weight to each equation in *allodb* using the [weight_allom\(\)](#) function, equations are then resampled within their original DBH range using [resample_agb\(\)](#): the number of resampled values for each equation is proportional to its weight. It creates S3 objects of class "numeric".

Usage

```
resample_agb(
  genus,
  coords,
  species = NULL,
  new_eqtable = NULL,
  wna = 0.1,
  w95 = 500,
  nres = 10000
)
```

Arguments

genus	a character value, containing the genus (e.g. "Quercus") of the tree.
coords	a numeric vector of length 2 with longitude and latitude.
species	a character value, containing the species (e.g. "rubra") of the tree. Default is "NULL", when no species identification is available.
new_eqtable	Optional. An equation table created with the new_equations() function. Default is the original <i>allodb</i> equation table.
wna	a numeric vector, this parameter is used in the weight_allom() function to determine the dbh-related and sample-size related weights attributed to equations without a specified dbh range or sample size, respectively. Default is 0.1.
w95	a numeric vector, this parameter is used in the weight_allom() function to determine the value at which the sample-size-related weight reaches 95% of its maximum value (max=1). Default is 500.
nres	number of resampled values. Default is "1e4".

Value

An object of class "data.frame" of resampled DBHs and associated AGB from the equation table; the number of resampled DBHs is proportional to the weight provided by the `weight_allom()` function.

See Also

`weight_allom()`, `new_equations()`.

Examples

```
resample_agb(  
  genus = "Quercus",  
  species = "rubra",  
  coords = c(-78.2, 38.9)  
)
```

scbi_stem1

Tree census data from SCBI ForestGEO plot

Description

A table with tree data from the Smithsonian Conservation Biology Institute, USA (SCBI) ForestGEO dynamics plot. This dataset is an extract from the first tree census in 2008, only covering 1 hectare (SCBI plot is 25.6 ha). DBH in cm.

Usage

```
scbi_stem1
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 2287 rows and 6 columns.

Source

Full datasets for tree census data at SCBI can be requested through the ForestGEO portal (<https://forestgeo.si.edu/>). Census 1, 2, and 3 can also be accessed at the public GitHub repository for SCBI-ForestGEO Data (<https://github.com/SCBI-ForestGEO>).

See Also

Other datasets: `genus_family`, `gymno_genus`, `koppenMatrix`, `shrub_species`

Examples

```
# preview the datasets  
print(head(scbi_stem1))
```

shrub_species	<i>Shrub species identified in selected ForestGEO sites</i>
---------------	---

Description

Genus and species of shrubby plants identified in the 24 extratropical ForestGEO sites used in allodb.

Usage

```
shrub_species
```

Format

An object of class character of length 1.

See Also

Other datasets: [genus_family](#), [gymno_genus](#), [koppenMatrix](#), [scci_stem1](#)

Examples

```
# preview the dataset
print(head(shrub_species))
```

sitespecies	<i>Sites and tree species used in allodb and associated metadata</i>
-------------	--

Description

- [sitespecies](#): Table of extratropical ForestGEO sites in allodb (n=24) and their tree species.
- [sitespecies_metadata](#): Metadata for [sitespecies](#) table.

Usage

```
sitespecies
```

```
sitespecies_metadata
```

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 1113 rows and 10 columns.

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 10 rows and 4 columns.

See Also

Other database datasets: [equations](#), [missing_values](#), [references](#), [sites_info](#)

Examples

```
# preview the datasets
print(head(sitespecies))
print(head(sitespecies_metadata))
```

sites_info	<i>ForestGEO sites used in allodb</i>
------------	---------------------------------------

Description

Table with geographical information for extratropical ForestGEO sites used in allodb (n=24).

Usage

```
sites_info
```

Format

An object of class `spec_tbl_df` (inherits from `tbl_df`, `tbl`, `data.frame`) with 24 rows and 6 columns.

Details

More details on geographical aspects of these ForestGEO sites can be found in the accompanying manuscript.

See Also

Other database datasets: [equations](#), [missing_values](#), [references](#), [sitespecies](#)

Examples

```
# preview the datasets
print(head(sites_info))
```

weight_allom	<i>Attribute weights to equations</i>
--------------	---------------------------------------

Description

This function attributes a weight to each equation based on its sampling size, and taxonomic and climatic similarity with the species/site combination considered.

Usage

```
weight_allom(
  genus,
  coords,
  species = NULL,
  new_eqtable = NULL,
  wna = 0.1,
  w95 = 500
)
```

Arguments

genus	a character value, containing the genus (e.g. "Quercus") of the tree.
coords	a numeric vector of length 2 with longitude and latitude.
species	a character vector (same length as genus), containing the species (e.g. "rubra") of the tree. Default is NULL, when no species identification is available.
new_eqtable	Optional. An equation table created with the <code>new_equations()</code> function.
wna	a numeric vector, this parameter is used in the <code>weight_allom()</code> function to determine the sample-size related weights attributed to equations without a specified sample size. Default is 0.1.
w95	a numeric vector, this parameter is used to determine the value at which the sample-size-related weight reaches 95% of its maximum value (max=1). Default is 500.

Details

Each equation is given a weight by the function `weight_allom()`, calculated as the product of the following components:

(1) sample-size weight, calculated as:

$$1 - \exp(-n * (\log(20)/w95))$$

where n is the sample size of the equation; the weight given to equations with no sample size information is determined by argument wna (0.1 by default).

(2) climate weight, based on the similarity between the climatic conditions of the equation site and the target location, using the three-letter system of Koppen's climate scheme. Climate weights

associated with each combination of two Koppen climates are provided in `data("koppenMatrix")`. The resulting weight has a value between $1e-6$ (different climate groups) and 1 (exactly the same climate classification). When an equation was calibrated with trees from several locations with different Koppen climates, the maximum value out of all pairwise equation-site climate weights is used.

(3) taxonomic weight: equal to 1 for same species equations, 0.8 for same genus equations, 0.5 for same family equations and for equations calibrated for the same broad functional or taxonomic group (e.g. shrubs, conifers, angiosperms). All other equations are given a low taxonomic weight of $1e-6$: these equations will have a significant relative weight in the final prediction only when no other more specific equation is available.

Value

A named "numeric" vector with one weight for each equation.

See Also

`get_biomass()`, `new_equations()`.

Examples

```
x <- weight_allom(
  genus = "Acer",
  species = "negundo",
  coords = c(-78.2, 38.9)
)
str(x)
head(x)
```

Index

- * **database datasets**
 - equations, [2](#)
 - missing_values, [10](#)
 - references, [12](#)
 - sites_info, [16](#)
 - sitespecies, [15](#)
- * **datasets**
 - equations, [2](#)
 - genus_family, [4](#)
 - gymno_genus, [7](#)
 - koppenMatrix, [9](#)
 - missing_values, [10](#)
 - references, [12](#)
 - scbi_stem1, [14](#)
 - shrub_species, [15](#)
 - sites_info, [16](#)
 - sitespecies, [15](#)
- equations, [2](#), [2](#), [10](#), [12](#), [13](#), [16](#)
- equations_metadata, [2](#)
- equations_metadata (equations), [2](#)
- est_params, [3](#)
- est_params(), [6](#)
- genus_family, [4](#), [7](#), [9](#), [14](#), [15](#)
- get_biomass, [5](#)
- get_biomass(), [6](#), [18](#)
- gymno_genus, [4](#), [7](#), [9](#), [14](#), [15](#)
- illustrate_allodb, [7](#)
- koppenMatrix, [4](#), [7](#), [9](#), [14](#), [15](#)
- missing_values, [2](#), [10](#), [13](#), [16](#)
- new_equations, [10](#)
- new_equations(), [3–6](#), [8](#), [13](#), [14](#), [17](#), [18](#)
- references, [2](#), [10](#), [12](#), [12](#), [16](#)
- references_metadata, [12](#)
- references_metadata (references), [12](#)
- resample_agb, [13](#)
- resample_agb(), [6](#), [13](#)
- scbi_stem1, [4](#), [7](#), [9](#), [14](#), [15](#)
- shrub_species, [4](#), [7](#), [9](#), [14](#), [15](#)
- sites_info, [2](#), [10](#), [13](#), [16](#), [16](#)
- sitespecies, [2](#), [10](#), [13](#), [15](#), [15](#), [16](#)
- sitespecies_metadata, [15](#)
- sitespecies_metadata (sitespecies), [15](#)
- weight_allow, [17](#)
- weight_allow(), [3–6](#), [8](#), [13](#), [14](#), [17](#)