The PROFOUND database

Ramiro Silveyra Gonzalez, Christopher Reyer, Florian Hartig

For questions, please contact reyer@pik-potsdam.de

This document is part of the following data publication and should be cited like:

Reyer CPO, Silveyra Gonzalez R, Dolos K, Hartig F, Hauf Y, Noack M, Lasch-Born P, Rötzer T, Pretzsch H, Meesenburg H, Fleck S, Wagner M, Bolte A, Sanders T, Kolari P, Mäkelä A, Vesala T, Mammarella I, Pumpanen J, Matteucci G, Collalti A, D'Andrea E, Foltýnová L, Krejza J, Ibrom A, Pilegaard K, Loustau D, Bonnefond J-M, Berbigier P, Picart D, Lafont S, Dietze M, Cameron D, Vieno M, Tian H, Palacios-Orueta A, Cicuendez V, Recuero L, Wiese K, Büchner M, Lange S, Volkholz J, Kim H, Weedon GP, Sheffield J, Babst F, Vega del Valle I, Suckow F, Horemans J, Martel S, Bohn F, Steinkamp J, Chikalanov A, Mahnken M, Gutsch M, Trotta C, Babst F, Frieler K (2020) The PROFOUND database for evaluating vegetation models and simulating climate impacts on European forests. V.0.3. GFZ Data Services. http://doi.org/10.5880/PIK.2020.006

This dataset (version 0.3) replaces an earlier version published (v0.1.12) as Reyer et al. (2019) by replacing tree and stand data for the site Soro and fixing a range of minor issues documented in the file changelog_Profound-DB_v03.pdf.

Reyer CPO, Silveyra Gonzalez R, Dolos K, Hartig F, Hauf Y, Noack M, Lasch-Born P, Rötzer T, Pretzsch H, Meesenburg H, Fleck S, Wagner M, Bolte A, Sanders T, Kolari P, Mäkelä A, Vesala T, Mammarella I, Pumpanen J, Matteucci G, Collalti A, D'Andrea E, Foltýnová L, Krejza J, Ibrom A, Pilegaard K, Loustau D, Bonnefond J-M, Berbigier P, Picart D, Lafont S, Dietze M, Cameron D, Vieno M, Tian H, Palacios-Orueta A, Cicuendez V, Recuero L, Wiese K, Büchner M, Lange S, Volkholz J, Kim H, Weedon GP, Sheffield J, Vega del Valle I, Suckow F, Horemans J, Martel S, Bohn F, Steinkamp J, Chikalanov A, Frieler K (2019) The PROFOUND database for evaluating vegetation models and simulating climate impacts on forests. V.0.1.12. GFZ Data Services. http://doi.org/10.5880/PIK.2019.008

A full description paper is also in preparation:

Reyer CPO, Silveyra Gonzalez R, Dolos K, Hartig F, Hauf Y, Noack M, Lasch-Born P, Rötzer T, Pretzsch H, Meesenburg H, Fleck S, Wagner M, Bolte A, Sanders T, Kolari P, Mäkelä A, Vesala T, Mammarella I, Pumpanen J, Matteucci G, Collalti A, Trotta C, D'Andrea E, Foltýnová L, Krejza J, Ibrom A, Pilegaard K, Loustau D, Bonnefond J-M, Berbigier P, Picart D, Lafont S, Dietze M, Cameron D, Vieno M, Tian H, Palacios A, Cicuendez V, Recuero L, Wiese K, Büchner M, Lange S, Volkholz J, Kim H, Weedon GP, Sheffield J, Babst F, Vega del Valle I, Suckow F, Horemans J, Martel S, Bohn F, Steinkamp J, Chikalanov A, Mahnken M, Gutsch M, Frieler K (2020) The PROFOUND database for evaluating vegetation models and simulating climate impacts on forests. Earth System Science Data Discussions. https://doi.org/10.5194/essd-2019-220, 2020

2020-04-28

Contents

Data Policy

The PROFOUND Database (DB) is available under the Creative Commons Attribution-NonCommercial 4.0 International license (CC BY-NC 4.0). Further data policy statements of the individual data sets contained in the PROFOUND database are listed in the table below.

Table 1: Additional data policy statements specific to the individual datasets included in the PROFOUND database.

dataset	dataPolicy
CLIMATE_ISIMIP2B	Standard PROFOUND database policy
CLIMATE_ISIMIP2BLBC	Standard PROFOUND database policy
CLIMATE_ISIMIP2A	Standard PROFOUND database policy
CLIMATE_ISIMIPFT	Standard PROFOUND database policy
NDEPOSITION_EMEP	Please consider this data policy statement from EMEP: The Environment Agency - Austria allows the reproduction of its resources, with appropriate citation, for non-commercial purposes provided no other more specific rules apply. The officially reported emission data should be cited as: EMEP/CEIP 2014 Present state of emission data; http://www.ceip.at/webdab_emepdatabase/reported_emissiondata/ or http://www.ceip.at/status_reporting/2014_submissions/.
NDEPOSITION_ISIMIP2B	Standard PROFOUND database policy
CO2_ISIMIP	Standard PROFOUND database policy
ENERGYBALANCE	Please consider this data policy statement from FLUXNET: This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: CarboEuropeIP, CarboItaly and ICOS. The FLUXNET eddy covariance data processing and harmonization was carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices. Tier One data are open and free for scientific and educational purposes and their use will follow the fair use policy, stated here. Data users describe the intended use of the data when they fill out the data-download form; this intended-use statement will be emailed to the data producer(s) and posted on the Fluxdata website (https://fluxnet.fluxdata.org). The fair use policy dictates that (1) data producers are informed of who uses the data and for what purpose (which can be satisfied by the aforementioned mechanism) and (2) that proper acknowledgment and citations are given to all data used in a peer reviewed publication, via the following protocols: The data citation will be either a per-site DOI that is provided with the data download or a citation of a publication for each site. Every publication should use the standard FLUXNET acknowledgment given below. It is requested that every publication specify each site used with the FLUXNET-ID, data-years used, data DOI (in preparation), and brief acknowledgment for funding (if provided by FLUXNET PI) in the text or supplementary material. Finally, all data providers should be informed of forthcoming publications.
FLUX	Please consider this data policy statement from FLUXNET: This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: CarboEuropeIP, CarboItaly and ICOS. The FLUXNET eddy covariance data processing and harmonization was carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices. Tier One data are open and free for scientific and educational purposes and their use will follow the fair use policy, stated here. Data users describe the intended use of the data when they fill out the data-download form; this intended-use statement will be emailed to the data producer(s) and posted on the Fluxdata website (https://fluxnet.fluxdata.org). The fair

	use policy dictates that (1) data producers are informed of who uses the data and for what purpose (which can be satisfied by the aforementioned mechanism) and (2) that proper acknowledgment and citations are given to all data used in a peer reviewed publication, via the following protocols: The data citation will be either a per-site DOI that is provided with the data download or a citation of a publication for each site. Every publication should use the standard FLUXNET acknowledgment given below. It is requested that every publication specify each site used with the FLUXNET-ID, data-years used, data DOI (in preparation), and brief acknowledgment for funding (if provided by FLUXNET PI) in the text or supplementary material. Finally, all data providers should be informed of forthcoming publications.
METEOROLOGICAL	Please consider this data policy statement from FLUXNET: This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: CarboEuropeIP, CarboItaly and ICOS. The FLUXNET eddy covariance data processing and harmonization was carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices. Tier One data are open and free for scientific and educational purposes and their use will follow the fair use policy, stated here. Data users describe the intended use of the data when they fill out the data-download form; this intended-use statement will be emailed to the data producer(s) and posted on the Fluxdata website (https://fluxnet.fluxdata.org). The fair use policy dictates that (1) data producers are informed of who uses the data and for what purpose (which can be satisfied by the aforementioned mechanism) and (2) that proper acknowledgment and citations are given to all data used in a peer reviewed publication, via the following protocols: The data citation will be either a per-site DOI that is provided with the data download or a citation of a publication for each site. Every publication should use the standard FLUXNET acknowledgment given below. It is requested that every publication specify each site used with the FLUXNET-ID, data-years used, data DOI (in preparation), and brief acknowledgment for funding (if provided by FLUXNET PI) in the text or supplementary material. Finally, all data providers should be informed of forthcoming publications.
SOILTS	Please consider this data policy statement from FLUXNET: This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: CarboEuropeIP, CarboItaly and ICOS. The FLUXNET eddy covariance data processing and harmonization was carried out by the ICOS Ecosystem Thematic Center, AmeriFlux Management Project and Fluxdata project of FLUXNET, with the support of CDIAC, and the OzFlux, ChinaFlux and AsiaFlux offices. Tier One data are open and free for scientific and educational purposes and their use will follow the fair use policy, stated here. Data users describe the intended use of the data when they fill out the data-download form; this intended-use statement will be emailed to the data producer(s) and posted on the Fluxdata website (https://fluxnet.fluxdata.org). The fair use policy dictates that (1) data producers are informed of who uses the data and for what purpose (which can be satisfied by the aforementioned mechanism) and (2) that proper acknowledgment and citations are given to all data used in a peer reviewed publication, via the following protocols: The data citation will be either a per-site DOI that is provided with the data download or a citation of a publication for each site. Every publication should use the standard FLUXNET acknowledgment given below. It is requested that every publication specify each site used with the FLUXNET-ID, data-years used, data DOI (in preparation), and brief acknowledgment for funding (if provided by FLUXNET PI) in the text or supplementary material. Finally, all data providers should be informed of forthcoming publications.
MODIS_MOD09A1	When using the data, please cite it as ORNL DAAC 2008 and include the following reference (ORNL DAAC 2008. MODIS Collection 5 Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed June 25, 2016. Subset obtained for MOD09A1 product at various sites in Spatial Range: N=70.00N, S=35.00N, E=50.00E, W=10.00W, time period: 2000-02-18 to 2015-12-27, and subset

	size: 0.5 x 0.5 km. http://dx.doi.org/10.3334/ORNLDAAC/1241) in the reference list of your publication. Please also acknowledge the data in the following way: The MOD09A1 was (were) retrieved from MODISTools (Tuck et al., 2014), courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/. For MODISTools see Tuck, Sean L. and Phillips, Helen R.P. and Hintzen, Rogier E. and Scharlemann, Jörn P.W. and Purvis, Andy and Hudson, Lawrence N. (2014) MODISTools – downloading and processing MODIS remotely sensed data in R. Ecology and Evolution, (4) 24, 4658–4668. http://dx.doi.org/10.1002/ece3.1273
MODIS_MOD11A2	When using the data, please cite it as ORNL DAAC 2008 and include the following reference (ORNL DAAC 2008. MODIS Collection 5 Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed June 25, 2016. Subset obtained for MOD11A2 product at various sites in Spatial Range: N=70.00N, S=35.00N, E=50.00E, W=10.00W, time period: 2000-03-05 to 2015-12-27, and subset size: 1 x 1 km. http://dx.doi.org/10.3334/ORNLDAAC/1241) in the reference list of your publication. Please also acknowledge the data in the following way: The MOD11A2 was (were) retrieved from MODISTools (Tuck et al., 2014), courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/. For MODISTools see Tuck, Sean L. and Phillips, Helen R.P. and Hintzen, Rogier E. and Scharlemann, Jörn P.W. and Purvis, Andy and Hudson, Lawrence N. (2014) MODISTools – downloading and processing MODIS remotely sensed data in R. Ecology and Evolution, (4) 24, 4658–4668. http://dx.doi.org/10.1002/ece3.1273
MODIS_MOD13Q1	When using the data, please cite it as ORNL DAAC 2008 and include the following reference (ORNL DAAC 2008. MODIS Collection 5 Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed June 25, 2016. Subset obtained for MOD13Q1 product at various sites in Spatial Range: N=70.00N, S=35.00N, E=50.00E, W=10.00W, time period: 2000-02-18 to 2015-12-19, and subset size: 0.25 x 0.25 km. http://dx.doi.org/10.3334/ORNLDAAC/1241) in the reference list of your publication. Please also acknowledge the data in the following way: The MOD13Q1 was retrieved from MODISTools (Tuck et al., 2014), courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/. For MODISTools see Tuck, Sean L. and Phillips, Helen R.P. and Hintzen, Rogier E. and Scharlemann, Jörn P.W. and Purvis, Andy and Hudson, Lawrence N. (2014) MODISTools – downloading and processing MODIS remotely sensed data in R. Ecology and Evolution, (4) 24, 4658–4668. http://dx.doi.org/10.1002/ece3.1273
MODIS_MOD15A2	When using the data, please cite it as ORNL DAAC 2008 and include the following reference (ORNL DAAC 2008. MODIS Collection 5 Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed June 25, 2016. Subset obtained for MOD15A2 product at various sites in Spatial Range: N=70.00N, S=35.00N, E=50.00E, W=10.00W, time period: 2000-02-18 to 2015-12-19, and subset size: 0.25 x 0.25 km. http://dx.doi.org/10.3334/ORNLDAAC/1241) in the reference list of your publication. Please also acknowledge the data in the following way: The MOD15A2 was retrieved from MODISTools (Tuck et al., 2014), courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/. For MODISTools see Tuck, Sean L. and Phillips, Helen R.P. and Hintzen, Rogier E. and Scharlemann, Jörn P.W. and Purvis, Andy and Hudson, Lawrence N. (2014) MODISTools – downloading and processing MODIS remotely sensed data in R. Ecology and Evolution, (4) 24, 4658–4668.

	http://dx.doi.org/10.1002/ece3.1273
MODIS_MOD17A2	When using the data, please cite it as ORNL DAAC 2008 and include the following reference (ORNL DAAC 2008. MODIS Collection 5 Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed June 25, 2016. Subset obtained for MOD17A2 product at various sites in Spatial Range: N=70.00N, S=35.00N, E=50.00E, W=10.00W, time period: 2000-02-18 to 2015-12-19, and subset size: 0.25 x 0.25 km. http://dx.doi.org/10.3334/ORNLDAAC/1241) in the reference list of your publication. Please also acknowledge the data in the following way: The MOD17A2 was retrieved from MODISTools (Tuck et al., 2014), courtesy of the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/. For MODISTools see Tuck, Sean L. and Phillips, Helen R.P. and Hintzen, Rogier E. and Scharlemann, Jörn P.W. and Purvis, Andy and Hudson, Lawrence N. (2014) MODISTools – downloading and processing MODIS remotely sensed data in R. Ecology and Evolution, (4) 24, 4658–4668. http://dx.doi.org/10.1002/ece3.1273
TREE	See site specific policy
STAND	See site specific policy
SOIL	See site specific policy
CLIMATE_LOCAL	See site specific policy

Database structure

The PROFOUND database is a relational SQLite database and it is made of several independent tables (Fig. 1). From these tables views are created that can be accessed and downloaded by users with the ProfoundData R-package.

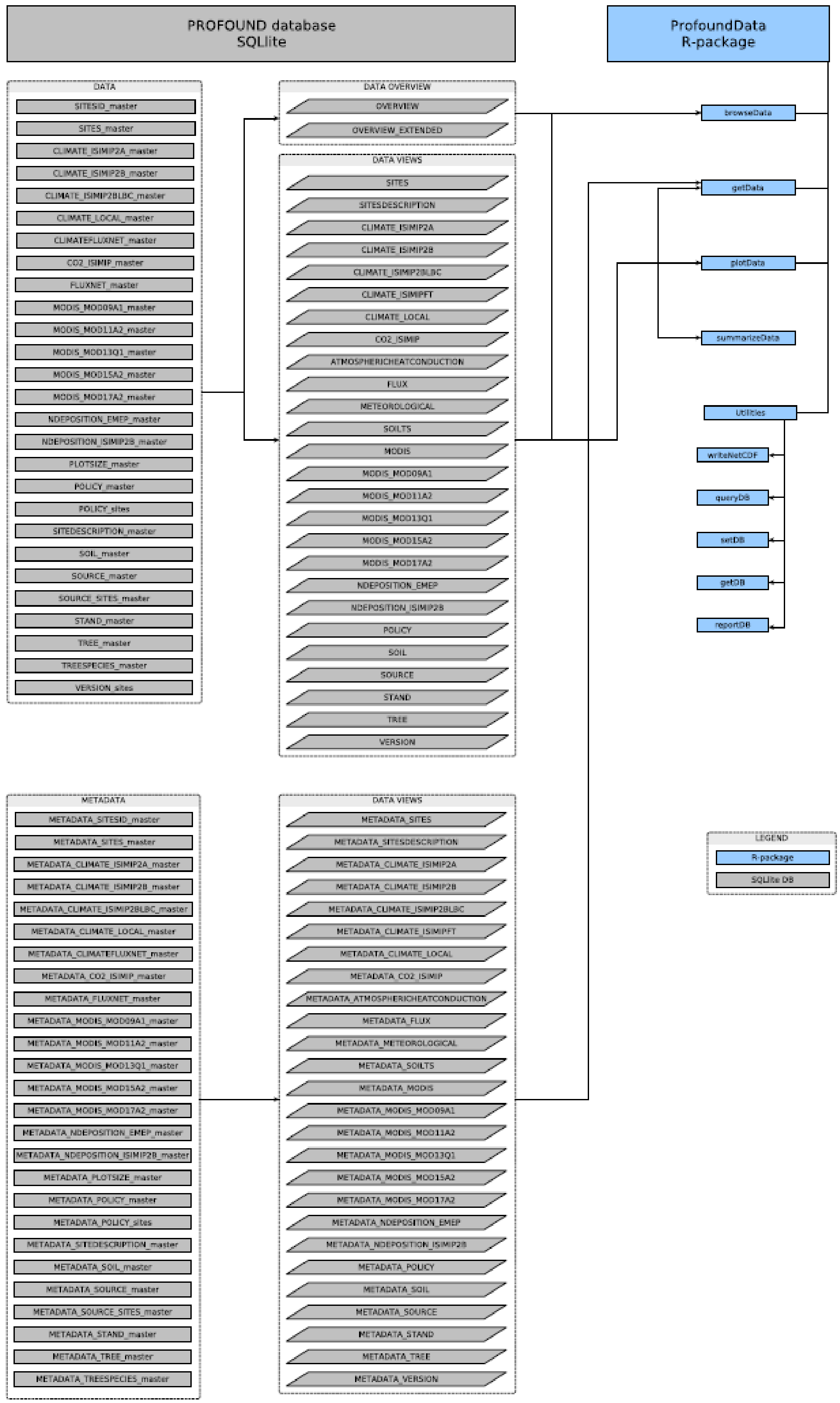

Figure 1: Overview on the PROFOUND database and the R package.

Site overview

The PROFOUND database currently includes 9 forest sites. They are listed in the table below.

Table 2: Forest sites included in the database.

site_id	site	lat	lon	epsg	country	aspect_deg	elevation_masl	slope_percent
3	bily_kriz	49.3	18.32	4326	Czech Republic	180	875	12.5
5	collelongo	41.85	13.59	4326	Italy	252	1560	10
12	hyytiala	61.85	24.3	4326	Finland	180	185	2
13	kroof	48.25	11.4	4326	Germany	1.8	502	2.1
14	le_bray	44.72	-0.769	4326	France	–	61	0
16	peitz	51.92	14.35	4326	Germany	–	50	0
20	solling_beech	51.77	9.57	4326	Germany	225	504	1
21	soro	55.49	11.64	4326	Denmark	–	40	0
25	solling_spruce	51.77	9.58	4326	Germany	90	508	1

There is an overview table to provide the information on which data is available for each site. The table is created by combining all existing tables in the database.

Table 3: Overview of sites and datasets (1= data available, 0= data unavailable)

site_id	site	SITES	TREE	STAND	SOIL	CLIMATE_LOCAL	CLIMATE_ISIMIP2B	CLIMATE_ISIMIP2BLBC	CLIMATE_ISIMIP2A	CLIMATE_ISIMIPFT	METEOROLOGICAL	FLUX	ATMOSPHERICHEATCONDUCTION	SOILTS	NDEPOSITION_EMEP	NDEPOSITION_ISIMIP2B	CO2_ISIMIP	MODIS_MOD09A1	MODIS_MOD15A2	MODIS_MOD11A2	MODIS_MOD13Q1	MODIS_MOD17A2	MODIS
3	bily_kriz	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	collelongo	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	hyytiala	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	kroof	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1
14	le_bray	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	peitz	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1
20	solling_beech	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1
21	soro	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	solling_spruce	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1

Datasets

Dataset SITES

The sites parameters were provided by the local site responsibles. The sites variables that we included in the database are listed in the table below.

Table 4: Description of SITES variables included in the database.

variable	type	units	description
site	TEXT	adimensional	Site name
site2	TEXT	adimensional	Additional site name
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
aspect_deg	REAL	degree	Direction of slope inclination. Degrees against North. No Value indicates no exposition.
country	TEXT	adimensional	Country
elevation_masl	REAL	m	Elevation above sea level as recorded by PI
epsg	INTEGER	adimensional	EPSG Coordinate System
lat	REAL	degree decimal	Latitude
lon	REAL	degree decimal	Longitude
natVegetation_code1	TEXT	adimensional	Code of the vegetation mapping unit group in the “Map of the Natural Vegetation of Europe”. BOHN, U.; GOLLUB, G. & HETTWER, C. (2000) Karte der natuerlichen Vegetation Europas. Massstab 1:2.500.000 Karten und Legende. Teil 1-3.. Bundesamt fuer Naturschutz, Bonn, Germany.
natVegetation_code2	TEXT	adimensional	Code of the vegetation mapping unit in the “Map of the Natural Vegetation of Europe”. BOHN, U.; GOLLUB, G. & HETTWER, C. (2000) Karte der natuerlichen Vegetation Europas. Massstab 1:2.500.000 Karten und Legende. Teil 1-3.. Bundesamt fuer Naturschutz, Bonn, Germany.
natVegetation_description	TEXT	adimensional	Description of natVegetation_code2. BOHN, U.; GOLLUB, G. & HETTWER, C. (2000) Karte der natuerlichen Vegetation Europas. Massstab 1:2.500.000 Karten und Legende. Teil 1-3.. Bundesamt fuer Naturschutz, Bonn, Germany.
slope_percent	REAL	percent	Mean slope within the plot

Dataset SITEDESCRIPTION

The sites description were provided by the local site responsibles. The site description variables that we included in the database are listed in the table below.

Table 5: Description of SITEDESCRIPTION variables included in the database.

variable	type	units	description
site	TEXT	adimensional	Site name
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
description	TEXT	adimensional	Ecological description of the site
reference	TEXT	adimensional	Publications referring to the site description and the site datasets

Dataset TREE

The individual tree data were provided by the local site responsibles. The tree variables that we included in the database are listed in the table below.

Table 6: Description of TREE variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site	TEXT	adimensional	Site name
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
species	TEXT	adimensional	Species name
species_id	TEXT	adimensional	Species text code
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)
dbh1_cm	REAL	cm	Diameter at breast height
height1_m	REAL	m	Tree height
size_m2	REAL	m2	Plot size

Dataset STAND

The stand data were provided by the local site responsibles. In some cases, the data were derived using the function *summarizeData* included in the ProfoundData R-package. The stand variables that we included in the database are listed in the table below.

Table 7: Description of STAND variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site	TEXT	adimensional	Site name
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
species	TEXT	adimensional	Species name
species_id	TEXT	adimensional	Species text code
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)
aboveGroundBiomass_kgha	REAL	kg ha-1	Above ground biomass
age	INTEGER	years	Mean stand age
ba_m2ha	REAL	m2 ha-1	Basal area per hectare
branchesBiomass_kgha	REAL	kg ha-1	Branches biomass
dbhArith_cm	REAL	cm	Arithmetic mean diameter
dbhBA_cm	REAL	cm	Average diameter weighted by basal area calculated as dbhBA = (ba1*dbh1 + ba2*dbh2 + … + bak*dbhk) / (ba1 + ba2+ … + bak), where bai and dbhi are the basal area and dbh, respectively, of the tree i, and i = 1, 2, . . , k
dbhDQ_cm	REAL	cm	Mean squared diameter or quadratic mean diameter calculated as dbhDQ = sqrt((dbh1$^{2 + dbh2}$2+ … + dbhk^2) / N), where dbhi is the diameter at breat height of tree i, i = 1, 2, . . , k, N is the total number of trees, and sqrt is the square root
density_treeha	REAL	tree ha-1	Number of tree per ha
foliageBiomass_kgha	REAL	kg ha-1	Foliage biomass
heightArith_m	REAL	m	Arithmetic mean height
heightBA_m	REAL	m	Average height weighted by basal area or Loreys height calculated as heightBA = (ba1*h1 + ba2*h2 + … + bak*hk) / (ba1 + ba2+ … + bak), where bai and hi are the basal area and height, respectively, of the tree i, and i = 1, 2, . . , k
lai	REAL	adimensional	Leaf Area Index
rootBiomass_kgha	REAL	kg ha-1	Root biomass
stemBiomass_kgha	REAL	kg ha-1	Stem biomass
stumpCoarseRootBiomass_kgha	REAL	kg ha-1	Stump and coarse roots biomass

Dataset SOIL

The soil data were provided by the local site responsibles. The variables that we included in the database are listed in the table below. The data is very heteregenous, therefore not all variables are available for each site.

Table 8: Description of SOIL variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site	TEXT	adimensional	Site name
site_id	INTEGER	adimensional	Site code as decimal number (1-99)
date	TEXT	adimensional	Unformatted date of inventory as provided for the inventory. See site specific metadata for further information on date.
bs_percent	REAL	percent	Percentage of alkaline and earth alkaline metals at CEC
cMax_percent	REAL	percent	Maximum soil carbon content
cMin_percent	REAL	percent	Minimum soil carbon content
cOrgSigma_percent	REAL	percent	Soil organic carbon content error estimate as standard deviation
cOrg_gcm3	REAL	g cm-3	Soil organic carbon content
cOrg_percent	REAL	percent	Soil organic carbon content
cSigma_kgm2	REAL	kg m-2	Soil carbon content error estimate as standard deviation
c_kgm2	REAL	kg m-2	Soil carbon content
c_percent	REAL	percent	Soil carbon content
cec_µeqg	REAL	µeq g-1	Soil cation exchange capacity
claySigma_percent	REAL	percent	Soil clay particle content error estimate as standard deviation
clay_percent	REAL	percent	Soil clay particle content
cn	REAL	adimensional	Soil C:N ratio
densitySigma_gcm3	REAL	g cm-3	Soil bulk density content error estimate as standard deviation
density_gcm3	REAL	g cm-3	Soil bulk density
fcapv_percent	REAL	percent	Soil field capacity
fineRoot_percent	REAL	percent	Distribution of fine roots accross soil horizons
gravel_percent	REAL	percent	Soil gravel particle content

horizon	TEXT	adimensional	Name of soil horizon
humus_tCha	REAL	tC ha-1	Humus carbon content
hydCondSat_cmd1	REAL	cm d-1	Soil hydraulic conductivity at saturation
layer_id	INTEGER	adimensional	Layer code as decimal number (1-99)
lowerDepth_cm	REAL	cm	Lower soil horizon limit
mbCSigma_mgg	REAL	mg C g-1 dry soil	Soil microbial biomass carbon error estimate as standard deviation
mbC_mgg	REAL	mg C g-1 dry soil	Soil microbial biomass carbon
mbNSigma_mgNg	REAL	mg N g-1 dry soil	Soil microbial biomass nitrogen error estimate as standard deviation
mbN_mgNg	REAL	mg N g-1 dry soil	Soil microbial biomass nitrogen
minRSigma_mgkgh	REAL	mg N kg-1 h-1	Soil mineralisation rate error estimate as standard deviation
minR_mgkgh	REAL	mg N kg-1 h-1	Soil mineralisation rate
nMax_percent	REAL	percent	Maximum soil nitrogen content
nMin_percent	REAL	percent	Minimum soil nitrogen content
nOrgSigma_percent	REAL	percent	Soil organic nitrogen content error estimate as standard deviation
nOrg_percent	REAL	percent	Soil organic nitrogen content
n_kgm2	REAL	kg m-2	Soil nitrogen content
n_percent	REAL	percent	Soil nitrogen content
ofhC_percent	REAL	percent	The organic fermentative-humic (Ofh) subhorizon consists of forest litter (leaves, bark, twigs etc) showing considerable decay.
ofhN_percent	REAL	percent	Carbon content in a gram of OFH sample
ofh_gDWm2	REAL	g DW m-2	Litter layer (leaves not decomposed)
ol_gDWm2	REAL	g DW m-2	Nitrogen content in a gram of OFH sample
phSigma_h2o	REAL	adimensional	Soil pH determined with H2O error estimate as standard deviation
phSigma_kcl	REAL	adimensional	Soil pH determined by KCl error estimate as standard deviation
ph_cacl2	REAL	adimensional	Soil pH determined with CaCl2
ph_h2o	REAL	adimensional	Soil pH determined with H2O
ph_kcl	REAL	adimensional	Soil pH deterimed with KCl
porosity_percent	REAL	percent	Soil water content at saturation in the bulk soil

rainGroundWater	REAL	adimensional	Whether the soil is mostly influenced by rain or ground water
sandSigma_percent	REAL	percent	Soil sand particle content error estimate as standard deviation
sand_percent	REAL	percent	Soil sand particle content
siltSigma_percent	REAL	percent	Soil silt particle content error estimate as standard deviation
silt_percent	REAL	percent	Soil silt particle content
table_id	INTEGER	adimensional	Table code as decimal number (1-99)
texture	TEXT	adimensional	Soil texture
thicknesSigma_cm	REAL	cm	Soil thickness error estimate
thickness_cm	REAL	cm	Soil thickness
type_fao	TEXT	adimensional	Soil type after ISSS-ISRIC-FAO (1998) World reference basis for soil resources. World Soil Resources Reports 84. FAO, Rome. 92 p.
type_ka5	TEXT	adimensional	Soil type after AG Boden (2005) Bodenkundliche Kartieranleitung. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover
upperDepth_cm	REAL	cm	Upper soil horizon limit
whcSigma_mm	REAL	mm	Soil water holding capacity error estimate
whc_mm	REAL	mm	Soil water holding capacity
whcp_percent	REAL	percent	Water holding capacity for plant available water
wiltp_percent	REAL	percent	Soil wilting point

Dataset CLIMATE

The climate data contains daily measurements of the following variables: min, max and mean temperature, precipitation, relative humidity, air pressure, global radiation and wind speed.

Table 9: Description of CLIMATE variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site	TEXT	adimensional	Site name
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
date	TEXT	adimensional	Date in format YYYY-MM-DD
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)
mo	INTEGER	MM	Month as decimal number (01-12)
day	INTEGER	DD	Day of the month as decimal number (01-31)
airpress_hPa	REAL	hPa	Mean daily air pressure
p_mm	REAL	mm	Total daily precipitation
rad_Jcm2	REAL	J cm-2	Total daily global radiation
relhum_percent	REAL	percent	Mean daily relative humidity
tmax_degC	REAL	degree Celsius	Maximum daily temperature
tmean_degC	REAL	degree Celsius	Mean daily temperature
tmin_degC	REAL	degree Celsius	Minimum daily temperature
wind_ms	REAL	m s-1	Mean daily wind speed

Dataset CLIMATE_LOCAL

The CLIMATE LOCAL data refers to climate data measured at each forest site or meteorological stations close to the site of the forest site. For those forest sites for which the data has been derived from half-hourly FLUXNET2015 data, we also provide the original half-hourly data in the table METEOROLOGICAL.

When relative humidity was not part of the original data, we calculated it from the vapour pressure deficit and the daily temperatures as

$$relhum\_percent = (1 - VPD\_F/es) * 100$$

where

$$es = (es(tmax\_degC) - es(tmin\_degC))/2$$

and

$$es(Ta) = 0.6108e^{(17.27*Ta)/(Ta+237.3)}$$

Besides the variables listed in Table 9, the CLIMATE_LOCAL dataset contains additional variables to indicate the data quality.

Table 10: Description of CLIMATE_LOCAL quality variables from FLUXNET included in the database.

variable	type	units	description
site	TEXT	adimensional	Site name
airpress_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
p_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
rad_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
relhum_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
tmax_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
tmean_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
tmin_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value
wind_qc	REAL	adimensional	fraction between 0-1; indicating percentage of measured and good quality gapfill half-hourly data used to create the daily value

Dataset CLIMATE_ISIMIP

There are several climatic datasets based on the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP). For each forest site, we extracted the climate data from the corresponding gridcell in the ISIMIP data.

- **CLIMATE_ISIMIPFT:** Based on climate data described here: Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., Schewe, (2013) The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP), Proceedings of the National Academy of Sciences,111,9, 3228-3232. https://doi.org/10.1073/pnas.1312330110, https://www.isimip.org/gettingstarted/#input-data-bias-correction

- **CLIMATE_ISIMIP2A:** Based on climate data described here: https://www.isimip.org/gettingstarted/#input-data-bias-correction

- **CLIMATE_ISIMIP2B:** Based on climate data described here: Frieler K., R. Betts, E. Burke, P. Ciais, S. Denvil, D. Deryng, K. Ebi, T. Eddy, K. Emanuel, J. Elliott, E. Galbraith, S.N. Gosling, K. Halladay, F. Hattermann, T. Hickler, J. Hinkel, V. Huber, C. Jones, V. Krysanova, S. Lange, H.K. Lotze, H. Lotze-Campen, M. Mengel, I. Mouratiadou, H. Müller Schmied, S. Ostberg, F. Piontek, A. Popp, C.P.O. Reyer, J. Schewe, M. Stevanovic, T. Suzuki, K. Thonicke, H. Tian, D.P. Tittensor, R. Vautard, M. van Vliet, L. Warszawski, F. Zhao (accepted pending revisions) Assessing the impacts of 1.5°C global warming - simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). Geoscientific Model Development (https://www.isimip.org/gettingstarted/#input-data-bias-correction)

- **CLIMATE_ISIMIP2BLBC:** Based on climate data described here: Frieler K, R Betts, E Burke, P Ciais, S Denvil, D Deryng, K Ebi, T Eddy, K Emanuel, J Elliott, E Galbraith, SN Gosling, K Halladay, F Hattermann, T Hickler, J Hinkel, V Huber, C Jones, V Krysanova, S Lange, HK Lotze, H Lotze-Campen, M Mengel, I Mouratiadou, H

Müller Schmied, S Ostberg, F Piontek, A Popp, CPO Reyer, J Schewe, M Stevanovic, T Suzuki, K Thonicke, H Tian, DP Tittensor, R Vautard, M van Vliet, L Warszawski, F Zhao (accepted pending revisions) Assessing the impacts of 1.5°C global warming - simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). Geoscientific Model Development, https://www.isimip.org/gettingstarted/#input-data-bias-correction

ISIMIP climatic datasets contain additionally one or both of the variables *forcingCondition* and *forcingDataset*.

Table 11: Description of CLIMATE_ISIMIP addtional variables included in the database.

variable	type	units	description
site	TEXT	adimensional	Site name
forcingCondition	TEXT	adimensional	This category refers to the conditions underlying the climatic forcing, e.g. following historical CO2 time series, preindustrial picontrol runs or representative concentration pathways (rcp).
forcingDataset	TEXT	adimensional	This category refers to data taken from bias-corrected general circulation models (e.g. hadgem) or historical global meteorological forcing data based on bias-corrected reanalysis data (e.g. watch)

Dataset NDEPOSITION

The nitrogen deposition data contain annual measurements of the variables listed in the table below.

Table 12: Description of NDEPOSITION variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site	TEXT	adimensional	Name of the site
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)
nhx_gm2	REAL	g m-2	Total deposition of reduced nitrogen (Dry+Wet RdN)
noy_gm2	REAL	g m-2	Total deposition of oxidized nitrogen (Dry+Wet oxN)

Dataset NDEPOSITION_EMEP

The NDEPOSITION_EMEP data were obtained from EMEP/CEIP 2014 Present state of emissions as used in EMEP models (http://www.ceip.at/webdab_emepdatabase/emissions_emepmodels/).

Dataset NDEPOSITION_ISIMIP2B

The NDEPOSITION_EMEP data were extracted for for each forest site from the corresponding gridcell in the ISIMIP data described in Frieler K., R. Betts, E. Burke, P. Ciais, S. Denvil, D. Deryng, K. Ebi, T. Eddy, K. Emanuel, J. Elliott, E. Galbraith, S.N. Gosling, K. Halladay, F. Hattermann, T. Hickler, J. Hinkel, V. Huber, C. Jones, V. Krysanova, S. Lange, H.K. Lotze, H. Lotze-Campen, M. Mengel, I. Mouratiadou, H. Müller Schmied, S. Ostberg, F. Piontek, A. Popp, C.P.O. Reyer, J. Schewe, M. Stevanovic, T. Suzuki, K. Thonicke, H. Tian, D.P. Tittensor, R. Vautard, M. van Vliet, L. Warszawski, F. Zhao (accepted pending revisions) Assessing the impacts of 1.5°C global warming - simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). Geoscientific Model Development (https://www.isimip.org/gettingstarted/#input-data-bias-correction). The dataset contains additionally the variable *forcingCondition*.

Table 13: Description of NDEPOSITION_ISIMIP2B addtional variables included in the database.

variable	type	units	description	source
forcingCondition	TEXT	adimensional	This category refers to the conditions underlying the climatic forcing, e.g. following historical CO2 time series, preindustrial picontrol runs or representative concentration pathways (rcp).	ISIMIP

Dataset CO2_ISIMIP

The CO2 dataset contains annual global concentrations of atmospheric CO2 for several forcing conditions.

Table 14: Description of CO2_ISIMIP variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site	TEXT	adimensional	Site name
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
forcingCondition	TEXT	adimensional	This category refers to the conditions underlying the climatic forcing, e.g. following historical CO2 time series or representative concentration pathways (rcp).
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)
co2_ppm	REAL	ppm	CO2 mean global concentrations for the different different forcing conditions: RCP and historical values (1975-2013)

Dataset ATMOSPHERICHEATCONDUCTION

The ATMOSPHERICHEATCONDUCTION data contains half-hourly measurements of the variables listed in the table below and was obtained from FLUXNET2015 data.

Table 15: Description of ATMOSPHERICHEATCONDUCTION variables included in the database.

variable	type	Units	description
record_id	INTEGER	Adimensional	Record ID as decimal number
site_id	INTEGER	Adimensional	Site code as decimal number (01-99)
date	TEXT	Adimensional	Date in format YYYY-MM-DD hh:mm:ss. Derived from TIMESTAMP_START
year	INTEGER	YYYY	Year with century as decimal number (0000-9999). Derived from TIMESTAMP_START
mo	INTEGER	MM	Month as decimal number (01-12). Derived from TIMESTAMP_START
day	INTEGER	DD	Day of the month as decimal number (01-31). Derived from TIMESTAMP_START
hCORRJOINTUNC_Wm2	REAL	W m-2	Joint uncertainty estimation for h as sqrt(hRANDUNC^{2} + ((hCORR75 - hCORR25) / 1.349)^{2})
hCORR_Wm2	REAL	W m-2	Sensible heat flux, corrected hFMDS by energy balance closure correction factor
hFMDS_Wm2	REAL	W m-2	Sensible heat flux, gapfilled using MDS method
hFMDS_qc	INTEGER	adimensional	Quality flag for hCORR. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
leCORRJOINTUNC_Wm2	REAL	W m-2	Joint uncertainty estimation for le
leCORR_Wm2	REAL	W m-2	Latent heat flux, corrected le_FMDS by energy balance closure correction factor
leFMDS_Wm2	REAL	W m-2	Latent heat flux, gapfilled using MDS method
leFMDS_qc	INTEGER	adimensional	Quality flag for leCORR. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
timestampEnd	TEXT	YYYYMMDDHHMM	ISO timestamp end of averaging period - short format
timestampStart	TEXT	YYYYMMDDHHMM	ISO timestamp start of averaging period - short format

Dataset FLUX

The FLUX data contains half-hourly measurements of the variables listed in the table below and was obtained from FLUXNET2015 data.

Table 16: Description of FLUX variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
date	TEXT	adimensional	Date in format YYYY-MM-DD hh:mm:ss. Derived from TIMESTAMP_START
year	INTEGER	YYYY	Year with century as decimal number (0000-9999). Derived from TIMESTAMP_START
mo	INTEGER	MM	Month as decimal number (01-12). Derived from TIMESTAMP_START
day	INTEGER	DD	Day of the month as decimal number (01-31). Derived from TIMESTAMP_START
gppDtCutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Gross Primary Production, from Daytime partitioning method, reference selected from GPP versions using a model efficiency approach. Based on corresponding NEE_CUT_XX version
gppDtCutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Gross Primary Production, calculated as stdev(gppDtCut_XX) / sqrt(40). SE from 40 half-hourly gppDtCut_XX
gppDtVutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Gross Primary Production, from Daytime partitioning method, reference version selected from GPP versions using a model efficiency approach. Based on corresponding neeVut_XX version
gppDtVutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Gross Primary Production, calculated as stdev(gppDtVut_XX) / sqrt(40. SE from 40 half-hourly gppDtVut_XX
gppNtCutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Gross Primary Production, from Nighttime partitioning method, reference selected from GPP versions using a model efficiency approach. Based on corresponding NEE_CUT_XX version
gppNtCutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Gross Primary Production, calculated as stdev(gppNtCut_XX) / sqrt(40). SE from 40 half-hourly gppNtCut_XX
gppNtVutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Gross Primary Production, from Nighttime partitioning method, reference version selected from GPP versions using a model efficiency approach. Based on corresponding neeVut_XX version
gppNtVutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Gross Primary Production, calculated as (stdev(gppNtVut_XX) / sqrt(40)). SE from 40 half-hourly gppNtVut_XX
neeCutRefJointunc_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Joint uncertainty estimation for neeCutRef, including random uncertainty and USTAR filtering uncertainty [sqrt(neeCutRef_RANDUNC$^{2 + ((NEE_CUT_84 - NEE_CUT_16) / 2)}$2)] for each half-hour
neeCutRef_qc	INTEGER	adimensional	Quality flag for neeCutRef. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.

neeCutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Net Ecosystem Exchange, using Constant Ustar Threshold (CUT) across years, reference selected on the basis of the model efficiency
neeVutRefJointunc_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Joint uncertainty estimation for neeVutRef, including random uncertainty and USTAR filtering uncertainty [sqrt(neeVutRef_RANDUNC$^{2 + ((neeVut_84 - neeVut_16) / 2)}$2))] for each half-hour
neeVutRef_qc	INTEGER	adimensional	Quality flag for neeVutRef. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
neeVutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Net Ecosystem Exchange, using Variable Ustar Threshold (VUT) for each year, reference selected on the basis of the model efficiency
recoDtCutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Ecosystem Respiration, from Daytime partitioning method, reference selected from RECO versions using a model efficiency approach. Based on corresponding NEE_CUT_XX version
recoDtCutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Ecosystem Respiration, calculated as stdev(recoDtCut_XX) / sqrt(40). SE from 40 half-hourly recoDtCut_XX
recoDtVutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Ecosystem Respiration, from Daytime partitioning method, reference selected from RECO versions using a model efficiency approach. Based on corresponding neeVut_XX version
recoDtVutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Ecosystem Respiration, calculated as stdev(recoDtVut_XX) / sqrt(40). SE from 40 half-hourly recoDtCut_XX
recoNtCutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Ecosystem Respiration, from Nighttime partitioning method, reference selected from RECO versions using a model efficiency approach. Based on corresponding NEE_CUT_XX version
recoNtCutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Ecosystem Respiration, calculated as stdev(recoNtCut_XX) / sqrt(40. SE from 40 half-hourly recoNtCut_XX
recoNtVutRef_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Ecosystem Respiration, from Nighttime partitioning method, reference selected from RECO versions using a model efficiency approach. Based on corresponding neeVut_XX version
recoNtVutSe_umolCO2m2s1	REAL	umolCO2 m-2 s-1	Standard Error for Ecosystem Respiration, calculated as stdev(recoNtVut_XX) / sqrt(40. SE from 40 half-hourly recoNtCut_XX
timestampEnd	TEXT	YYYYMMDDHHMM	ISO timestamp end of averaging period - short format
timestampStart	TEXT	YYYYMMDDHHMM	ISO timestamp start of averaging period - short format

Dataset METEOROLOGICAL

The METEOROLOGICAL data contains half-hourly measurements of the variables listed in the table below and was obtained from FLUXNET2015 data.

Table 17: Description of METEOROLOGICAL variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
date	TEXT	adimensional	Date in format YYYY-MM-DD hh:mm:ss. Derived from TIMESTAMP_START
year	INTEGER	YYYY	Year with century as decimal number (0000-9999). Derived from TIMESTAMP_START
mo	INTEGER	MM	Month as decimal number (01-12). Derived from TIMESTAMP_START
day	INTEGER	DD	Day of the month as decimal number (01-31). Derived from TIMESTAMP_START
timestampEnd	TEXT	YYYYMMDDHHMM	ISO timestamp end of averaging period - short format
timestampStart	TEXT	YYYYMMDDHHMM	ISO timestamp start of averaging period - short format
lwInFMDS_Wm2	REAL	W m-2	Longwave radiation, incoming, gapfilled using MDS
lwInFMDS_qc	REAL	adimensional	Quality flag for lwInFMDS. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
lwInF_Wm2	REAL	W m-2	Longwave radiation, incoming, consolidated from lwInFMDS and lwInERA. lwInFMDS used if lwInFMDS_qc is 0 or 1.
lwInF_qc	INTEGER	adimensional	Quality flag for lwInF. 0 = measured; 1 = good quality gapfill; 2 = downscaled from ERA
pF_mm	REAL	mm	Precipitation consolidated from p and pERA
pF_qc	INTEGER	adimensional	Quality flag for pF. 0 = measured; 2 = downscaled from ERA
p_mm	REAL	mm	Precipitation.
paF_kPa	REAL	kPa	Atmospheric pressure consolidated from pa and paERA
paF_qc	INTEGER	adimensional	Quality flag for paF. 0 = measured; 2 = downscaled from ERA.
pa_kPa	REAL	kPa	Atmospheric pressure
swInFMDS_Wm2	REAL	W m-2	Shortwave radiation, incoming, gapfilled using MDS (negative values set to zero, e.g., negative values from instrumentation noise).
swInFMDS_qc	REAL	adimensional	Quality flag for swInFMDS. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
swInF_Wm2	REAL	W m-2	Shortwave radiation, incoming consolidated from swInFMDS and swInERA (negative values set to zero). swInFMDS used if swInFMDS_QC is 0 or 1
swInF_qc	INTEGER	adimensional	Quality flag for swInF. 0 = measured; 1 = good quality gapfill; 2 = downscaled from ERA

taFMDS_degC	REAL	degree Celsius	Air temperature, gapfilled using MDS method
taFMDS_qc	INTEGER	adimensional	Quality flag for taFMDS. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
taF_degC	REAL	degree Celsius	Air temperature, consolidated from taFMDS and taERA
taF_qc	INTEGER	adimensional	Quality flag for taF. 0 = measured; 1 = good quality gapfill; 2 = downscaled from ERA.
vpdFMDS_hPa	REAL	hPa	Vapor Pressure Deficit, gapfilled using MDS.
vpdFMDS_qc	INTEGER	adimensional	Quality flag for vpdFMDS. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
vpdF_hPa	REAL	hPa	Vapor Pressure Deficit consolidated from vpdFMDS and vpdERA. vpdFMDS used if vpdFMDS_qc is 0 or 1.
vpdF_qc	INTEGER	adimensional	Quality flag for vpdF. 0 = measured; 1 = good quality gapfill; 2 = downscaled from ERA
wsF_ms1	REAL	m s-1	Wind speed, consolidated from ws and wsERA. ws used if measured.
wsF_qc	INTEGER	adimensional	Quality flag of wsF.0 = measured; 2 = downscaled from ERA.
ws_ms1	REAL	m s-1	Wind speed

Dataset SOILTS

The soil time series data contains half-hourly measurements of the variables listed in the table below and was obtained from FLUXNET2015 data.

Table 18: Description of SOILTS variables included in the database.

variable	type	units	description
record_id	INTEGER	adimensional	Record ID as decimal number
site_id	INTEGER	adimensional	Site code as decimal number (01-99)
date	TEXT	adimensional	Date in format YYYY-MM-DD hh:mm:ss. Derived from TIMESTAMP_START
year	INTEGER	YYYY	Year with century as decimal number (0000-9999). Derived from TIMESTAMP_START
mo	INTEGER	MM	Month as decimal number (01-12). Derived from TIMESTAMP_START
day	INTEGER	DD	Day of the month as decimal number (01-31). Derived from TIMESTAMP_START
timestampEnd	TEXT	YYYYMMDDHHMM	ISO timestamp end of averaging period - short format
timestampStart	TEXT	YYYYMMDDHHMM	ISO timestamp start of averaging period - short format
swcFMDS1_degC	REAL	percent	Soil water content, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
swcFMDS1_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
swcFMDS2_degC	REAL	percent	Soil water content, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
swcFMDS2_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
swcFMDS3_degC	REAL	percent	Soil water content, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
swcFMDS3_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
swcFMDS4_degC	REAL	percent	Soil water content, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
swcFMDS4_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
swcFMDS5_degC	REAL	percent	Soil water content, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
swcFMDS5_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
tsFMDS1_degC	REAL	degree Celsius	Soil temperature, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
tsFMDS1_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
tsFMDS2_degC	REAL	degree Celsius	Soil temperature, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
tsFMDS2_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
tsFMDS3_degC	REAL	degree Celsius	Soil temperature, gapfilled with MDS (numeric index "#" increases with the depth, 1 is shallowest)
tsFMDS3_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.

tsFMDS4_degC	REAL	degree Celsius	Soil temperature, gapfilled with MDS (numeric index “#” increases with the depth, 1 is shallowest)
tsFMDS4_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.
tsFMDS5_degC	REAL	degree Celsius	Soil temperature, gapfilled with MDS (numeric index “#” increases with the depth, 1 is shallowest)
tsFMDS5_qc	INTEGER	adimensional	Quality flag for tsFMDS#. 0 = measured; 1 = good quality gapfill; 2 = medium; 3 = poor.

Dataset MODIS

The original MODIS time series are available at the NASA Land Processes Distributed Archive Center (LP DAAC). The data were downloaded from the Land Product Subset web service of the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC). Five different datasets are included in the database:

- **MODIS_MOD09A1**
- **MODIS_MOD15A2**
- **MODIS_MOD11A2**
- **MODIS_MOD13Q1**
- **MODIS_MOD17A2**

The data comprise surface reflectance, land surface temperature, vegetation indexes, Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR), GPP and Net Photosynthesis. The MODIS variables are listed the table below.

Table 19: Description of MODIS variables included in the database

variable	type	units	description	source
record_id	INTEGER	adimensional	Record ID as decimal number	MOD09A1
site	TEXT	adimensional	Site name	MOD09A1
site_id	INTEGER	adimensional	Site code as decimal number (01-99)	MOD09A1
date	TEXT	adimensional	Date in format YYYY-MM-DD	MOD09A1
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)	MOD09A1
mo	INTEGER	MM	Month as decimal number (01-12)	MOD09A1
day	INTEGER	DD	Day of the month as decimal number (01-31)	MOD09A1
aB01_rad	REAL	radian	Angle in red. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
aB02_rad	REAL	radian	Angle in near infrared. Calculated with MOD09A1. Spatial resolution 0.5 km	MOD09A1
aB05_rad	REAL	radian	Angle in SWIR 1.Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
aB06_rad	REAL	radian	Angle in SWIR 2. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
evi8	REAL	adimensional	Enhance Vegetation Index. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite	MOD09A1
ndvi8	REAL	adimensional	Normalized Difference Vegetation Index. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite	MOD09A1
ndwi	REAL	adimensional	Normalized Difference Water Index. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite	MOD09A1
reflB01_percent	REAL	percent reflectance	Surface Reflectance Band 1 (620–670 nm) Red. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
reflB02_percent	REAL	percent reflectance	Surface Reflectance Band 2 (841–876 nm) NIR. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
reflB03_percent	REAL	percent reflectance	Surface Reflectance Band 3 (459–479 nm) Blue. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
reflB04_percent	REAL	percent reflectance	Surface Reflectance Band 4 (545–565 nm) Green. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
reflB05_percent	REAL	percent reflectance	Surface Reflectance Band 5 (1230–1250 nm) SWIR1. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1

reflB06_percent	REAL	percent reflectance	Surface Reflectance Band 6 (1628–1652 nm) SWIR2. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
reflB07_percent	REAL	percent reflectance	Surface Reflectance Band 7 (2105–2155 nm) SWIR3. Fill value: NA. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
refl_qc	INTEGER	adimensional	Indicates the level of quality correction of the product (the seven bands) that is classified as follows: 0 = Highest quality, corrected product produced at ideal quality all bands; 2 = corrected product produced at less than ideal quality some or all bands, some bands could not be completely correct; 3 = interpolated, when corrected product has not been produced in one or some bands and they have been interpolated with the value Rt = (Rt-1 + Rt+1)/2; 4 = corrected product not produced, when product has not been completely corrected in one or some bands and could not be interpolated. Data may be wrong or filled with NA; 5 = Missing data, indicates that the product was not available for that date. Some of them correspond to specific continuous periods. All the columns filled with NA.	MOD09A1
sani_rad	REAL	radian	Shortwave Angle Normalized Index. Valid range: -3.14 - 3.14. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite.	MOD09A1
sasi_rad	REAL	radian	Shortwave Angle Slope Index. Spatial resolution: 0.5 km. Temporal resolution: 8-day composite	MOD09A1
record_id	INTEGER	adimensional	Record ID as decimal number	MOD15A2
site	TEXT	adimensional	Site name	MOD15A2
site_id	INTEGER	adimensional	Site code as decimal number (01-99)	MOD15A2
date	TEXT	adimensional	Date in format YYYY-MM-DD	MOD15A2
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)	MOD15A2
mo	INTEGER	MM	Month as decimal number (01-12)	MOD15A2
day	INTEGER	DD	Day of the month as decimal number (01-31)	MOD15A2
fpar	REAL	adimensional	Proportion of available radiation in the photosynthetically active wavelengths. Valid range: 0 - 1. Fill value: NA. Spatial resolution: 1 km. Temporal resolution: 8-day composite.	MOD15A2
fpar_qc	INTEGER	adimensional	Indicates the level of the product quality that is classified as follows: 0 = Good quality (main algorithm with or without saturation); 2 = Other quality (back-up algorithm or fill values)	MOD15A2
lai	REAL	adimensional	Leaf area index. Valid range: 0 - 10. Fill value: NA. Spatial resolution: 1 km. Temporal resolution: 8-day composite.	MOD15A2
lai_qc	INTEGER	adimensional	Indicates the level of the product quality that is classified as follows: 0 = Good quality (main algorithm with or without saturation); 2 = Other quality (back-up algorithm or fill values)	MOD15A2

record_id	INTEGER	adimensional	Record ID as decimal number	MOD11A2
site	TEXT	adimensional	Site name	MOD11A2
site_id	INTEGER	adimensional	Site code as decimal number (01-99)	MOD11A2
date	TEXT	adimensional	Date in format YYYY-MM-DD	MOD11A2
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)	MOD11A2
mo	INTEGER	MM	Month as decimal number (01-12)	MOD11A2
day	INTEGER	DD	Day of the month as decimal number (01-31)	MOD11A2
lstDay_degK	REAL	degree Kelvin	Daytime land surface temperature. Valid range: 150 – 1310.7. Fill value: NA. Spatial resolution: 1 km. Temporal resolution: 8-day composite.	MOD11A2
lstDay_qc	INTEGER	adimensional	Indicates the level of quality of the product that is classified as follows: 0 = good quality; 2 = other quality; 3 = interpolated, 4 = pixel not produced (NA)	MOD11A2
lstNight_degK	REAL	degree Kelvin	Nighttime land surface temperature. Valid range: 150 – 1310.7. Fill value: NA. Spatial resolution: 1 km. Temporal resolution: 8-day composite.	MOD11A2
lstNight_qc	INTEGER	adimensional	Indicates the level of quality of the product that is classified as follows: 0 = good quality; 2 = other quality; 3 = interpolated, 4 = pixel not produced (NA)	MOD11A2
record_id	INTEGER	adimensional	Record ID as decimal number	MOD13Q1
site	TEXT	adimensional	Site name	MOD13Q1
site_id	INTEGER	adimensional	Site code as decimal number (01-99)	MOD13Q1
date	TEXT	adimensional	Date in format YYYY-MM-DD	MOD13Q1
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)	MOD13Q1
mo	INTEGER	MM	Month as decimal number (01-12)	MOD13Q1
day	INTEGER	DD	Day of the month as decimal number (01-31)	MOD13Q1
evi16	REAL	adimensional	Enhanced Vegetation Index. Valid range: -0.2 - 1. Fill value: NA. Spatial resolution: 250 meters. Temporal resolution: 16-day composite.	MOD13Q1
evi16_qc	REAL	adimensional	Indicates the level of the product quality that is classified as follows: 0 = good quality, index produced; 2 = other quality, index produced, but check other qc and index produced, but most probably cloudy; 3 = index not produced due to other reasons than cloud, thus fill values were substituted by an interpolated values when the previous and the following values were available index = (indext-1 + indext+1)/2.	MOD13Q1
ndvi16	REAL	adimensional	Normalized Difference Vegetation Index. Valid range: -0.2 - 1. Fill value: NA. Spatial resolution:	MOD13Q1

			250 meters. Temporal resolution: 16-day composite.	
ndvi16_qc	REAL	adimensional	Indicates the level of the product quality that is classified as follows: 0 = good quality, index produced; 2 = other quality, index produced, but check other qc and index produced, but most probably cloudy; 3 = index not produced due to other reasons than cloud, thus fill values were substituted by an interpolated values when the previous and the following values were available index = (indext-1 + indext+1)/2.	MOD13Q1
record_id	INTEGER	adimensional	Record ID as decimal number	MOD17A2
site	TEXT	adimensional	Site name	MOD17A2
site_id	INTEGER	adimensional	Site code as decimal number (01-99)	MOD17A2
date	TEXT	adimensional	Date in format YYYY-MM-DD	MOD17A2
year	INTEGER	YYYY	Year with century as decimal number (0000-9999)	MOD17A2
mo	INTEGER	MM	Month as decimal number (01-12)	MOD17A2
day	INTEGER	DD	Day of the month as decimal number (01-31)	MOD17A2
gpp_gCm2d	REAL	gC m-2 d-1	Gross Primary Production. Valid range: -375 – 375. Fill value: NA. Spatial resolution 1 km. Temporal resolution: 8-day accumulation.	MOD17A2
gpp_qc	INTEGER	adimensional	Indicates the level of the product quality that is classified as follows: 0 = good quality, the estimates were done using the main algorithm with or without saturation; 2 = other quality, the estimates were done using back-up algorithm.	MOD17A2
psNet_gCm2d	REAL	gC m-2 d-1	Net Photosynthesis (GPP – maintenance respiration). Valid range: -375 – 375. Fill value: NA. Spatial resolution 1 km. Temporal resolution: 8-day accumulation.	MOD17A2
psNet_qc	INTEGER	adimensional	Indicates the level of the product quality that is classified as follows: 0 = good quality, the estimates were done using the main algorithm with or without saturation; 2 = other quality, the estimates were done using back-up algorithm.	MOD17A2