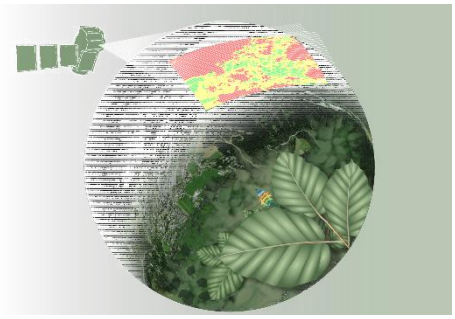
Gross Primary Production maps 2017-2023



Data Product Factsheet, Version 2026-01-12

General description

Data Product Title: Gross Primary Production maps from 2017 to 2023
Short description:
Gross Primary Production (GPP) is defined as the total amount of carbon fixed by an ecosystem in a given time. GPP maps are provided with a 10 m spatial resolution every 5-days for the period June 2017-December 2023. The period differs site by site depending upon in-situ data availability. The highest temporal resolution of 5-days is achieved in periods with 50% free-cloud conditions in the area of interest. In other cases, when cloud cover exceeds this threshold, the data is excluded, resulting in reduced frequency.
The GPP data products are calculated with a data-driven approach combining Earth Observations with in-situ carbon data. In specific, Sentinel-2 Multispectral data are combined with in situ GPP data derived from ICOS eddy covariance tower systems using XGBoost (Chen and Guestrin, 2016) machine learning algorithm (Spinosa et al., 2024).
The map extend corresponds to the boundaries of the selected long-term observation facilities as provided on the site registry DEIMS-SDR (e.g. https://deims.org/d4854af8-9d9f-42a2-af96-f1ed9cb25712). If the site boundaries cover an area smaller than 1 km² or the site boundaries are not registered in DEIMS-SDR with only point coordinates provided, the map extend is defined by a 1km x 1km bounding box to ensures consistency across all sites and balance computational efficiency with data availability. This extension has also been discussed with users engaged in the co-definition phase of the project.
The methodology integrates data from diverse ecosystems to estimate Gross Primary Production (GPP) using machine learning. Data pre-processing includes selecting sites based on data availability and completeness, extracting environmental data from ICOS, and estimating remote sensing indices from Sentinel 2 data. An XGBoost model is trained and evaluated using MAE, RMAE, and R^2 to predict time series of GPP. The model is then used for the computation of 5-day GPP maps within the ecosystem boundaries box. Annual error metrics are provided to ensure model accuracy.

Use cases:
The data product aims to provide input for an Essential Biodiversity Variable that can inform on the states and trends of ecosystems functioning. It also provides support to researchers working on ecosystem services valuing.

- **Potential user groups**:
  - Site and Platform Managers,
  - Protected Area Managers
  - Research community
- **Application scenarios/examples**:
  - Analysing trends and changes in ecosystem processes, conservation and restoration progresses assessments, scenarios for biodiversity and ecosystem services (e.g. for IPBES), improving carbon budget calculations.

Contact:

- **Contact person**: Anna Spinosa, Mario Alberto Fuentes Monjaraz, Valeria Mobilia

Data Product Specification:

- **Data Product Type:** Scientific
- **Data Level**: 3 (derived data product)
- **Data Product License:** CC-BY 4.0 Int.

- **Thematic scope**: Ecosystem processes and functions
- **Linked Standard eLTER Observation (SO)**: Gross primary production (SOBIO_090)
- **Linked Variable:** Gross Primary Production

- **Temporal extent**: 2017-2023, upon in-situ data availability
- **Spatial coverage:** Long-term observation sites; boundaries of the long-term observation facilities as provided on the site registry DEIMS-SDR (e.g. https://deims.org/d4854af8-9d9f-42a2-af96-f1ed9cb25712) or corresponding to a 1 km x 1km bounding box if the site boundaries cover an area smaller than 1 km² or the site boundaries are not registered in DEIMS-SDR with only point coordinates provided. Currently produced for 16 sites.
- **Spatial resolution:** 10x10m
- **Geographic projection:** EPSG:4326
- **Temporal resolution:** 5 days (restriction 50% cloud free images)
- **Temporal usability:** Archive with on-demand updates

- **Update frequency:** As needed, on-demand

Format and provision

Format specification

- **Standard**: Cloud optimized GeoTIFF (CoG)

Metadata and documentation

- **Standard**: Open Geospatial Consortium (OGC) following inter-alia ISO19139 encoding
- **URL**: https://catalog.elter.cerit-sc.cz/
- Metadata include minimum information on: Title, Abstract, Data Licence, Keywords (also used to include a reference to the site record on DEIMS.org as well as a DOI for the data record as a whole), Method description (including instrument, platform information, and measured attributes), Variable and Standard Observation, Geographic bounding box or polygon, Temporal extent, Responsible party name, Contact email address, and Citation.

Data provision

- **Data service**
  - **Standard:** Open Geospatial Consortium (OGC) using OGC WMS/WFS/CSW
  - **Content:** data layers for gross primary production (GPP) for the time period and a respective long-term observation facility
  - **URL:** https://spatialnode.elter.cerit-sc.cz/geoserver/web/wicket/bookmarkable/org.geoserver.web.demo.MapPreviewPage?2&filter=false

- **Knowledge package**
  - **Standard:** ZIP package for download
  - **Content:** Knowledge package containing
    - data layers for gross primary production (GPP) for the time period and a respective long-term observation facility (as zip archive)
    - data layers for valid areas to estimate error of gross primary production for the year and a respective long-term observation facility (as zip archive)
    - data error estimates per year and a respective long-term facility (as zip archive)
    - data product factsheet (as pdf file)
    - data license document (in B2SHARE metadata)
    - acknowledgement document (as csv file)
    - citation document (in B2SHARE metadata)
  - **URL:** https://b2share.eudat.eu/
  - **Identification:** DOI is issued as persistent identifier by B2SHARE

Data layers

Gross Primary Production (GPP)

- **Title: <site name> - Gross Primary Production (2017-2023)**
- **Abstract:** *Gross Primary Production (GPP) represents the total amount of carbon fixed by plants in an ecosystem over a specific period. GPP data products are derived using a data-driven approach that integrates Earth observation data with in-situ carbon flux measurements. Specifically, GPP estimations combine Sentinel-2 multispectral imagery with carbon flux data from eddy covariance towers, employing the XGBoost (Chen and Guestrin, 2016) machine learning algorithm for prediction. The resulting GPP maps are generated at a 10-meter spatial resolution with a temporal frequency of 5 days, covering the period from June 2017 to December 2023, upon in-situ data availability. The temporal resolution is contingent upon 50% free-cloud conditions in the area of interest, with lower frequencies occurring during periods of cloud cover above the threshold. The spatial extent of the GPP maps corresponds to the boundaries of long-term observation sites as recorded in the DEIMS-SDR registry (e.g., https://deims.org/d4854af8-9d9f-42a2-af96-f1ed9cb25712). For sites where the boundary area is smaller than 1 km², or if only point coordinates are available in DEIMS-SDR, the maps are constrained to a 1 km x 1 km bounding box.*
- **Method description:** *The methodology integrates multiple data sources via machine learning techniques to estimate Gross Primary Production (GPP) across different ecosystems. The process begins with data pre-processing, including the selection of sites based on criteria such available vegetation information, at least three full years of eddy covariance flux data. GPP and environmental data (e.g. air temperature, vapor pressure deficit, etc.) are extracted from the ICOS database across different ecosystem types. Then, different remote sensing (RS) indices (e.g NDVI, EVI, etc.) are estimated in GEE using Sentinel-2 MSI data as the mean value of the pixels found inside the climatological footprint 70 (an area were 70% of the GPP measurements are coming from). Both data coming from ICOS dataset and RS indices are used as predictors for the model.*

The data is split, with 70% used for training and 30% for testing. In the model setup, an XGBoost model is trained using the selected environmental and RS based indices. The model parameters are fine-tuned to improve accuracy. The remaining 30% of testing data is used to evaluate the model's performance by comparing its predictions against in-situ GPP data. Error metrics like Mean Absolute Error (MAE), Root Mean Absolute Error (RMAE), and R^2 are provided. The maps computation phase applies the trained model to ecosystem boundaries from the DEIMS website to generate 5-day GPP maps.

- **Data license:** *CC-BY 4.0 Int*[1]
- **Data level:** 3
- **Keywords:** *AGAME, GPP, Sentinel-2 MSI, RS, eLTER, primary, vegetation, plant, production, carbon, EBV, ecosystem function*
- **Variable:** *Gross Primary Production (url:* http://vocabs.lter-europe.net/EnvThes/21074)
- **Standard Observation:** *Gross Primary Production (SOBIO_090, url:* https://vocabs.lter-europe.net/so/090)
- **Site reference:** https://deims.org/<deims.id>
- **Temporal extent:** Generally, 2017-2023, but depending on in-situ data availability
- **Geographic bounding box:** <automatic based on geoTIFF>
- **Geographic projection:** EPSG:4326
- **Format specification:** Cloud Optimised geoTIFF (CoG)
- **Responsible party name**: Mario Fuentes, Valeria Mobilia, Anna Spinosa
- **Contact email address**:
    - Anna.Spinosa@deltares.nl
    - valeria.mobilia@deltares.nl
    - Mario.FuentesMonjaraz@deltares.nl
- **Citation:** *AGAME Project Team* (2024) <site name> - Gross Primary Production 2017-2023. Data product. DOI:xxx.xxx/xxx

Variable	Unit	Value range	Definition
GPP	g C/m²/day	nan: masked areas 0 - 20	Total amount of carbon fixed by an ecosystem in a given period of time, expressed as g C/m²/day

[1] See https://creativecommons.org/licenses/by/4.0/legalcode.en

Figure 1.- Hohes Holz – Gross Primary Production (2021-06-18) product. Mock-up of an AGAME Gross Primary Production data product for TERENO Harz - Central German Lowland - Hohes Holz site (https://deims.org/ddd2e8d2-44db-420e-8fa4-6b4fe1b00c78) in 2021-06-18 within GEOSS.

Gross Primary Production Error Estimate

- **Title: <site name> - Gross Primary Production (GPP) Valid Areas <year>**
- **Abstract:** *Annual comparisons between modelled and in-situ measurements derived for Gross Primary Production.*
- **Method description:** *Annual performance comparisons are conducted between model predictions and in-situ measurements at site's station. Specifically, the yearly average modelled GPP within the climatological footprint 70, is compared with the annual mean value of GPP derived from in-situ data. Validation evaluation metrics, as MAE and R^2 are provided. Next, K-means clustering is used to generate a classification map, identifying areas with similar vegetation cover to the eddy covariance tower fetch. Pixels are classified as either 1: similar vegetation to the climatological footprint or 0: different vegetation type. Assuming stable or similar environmental conditions across the area, the same error metrics calculated for the climatological footprint 70, can be used to assess the model's accuracy.*
- **Data license:** *CC-BY 4.0 Int*
- **Data level:** *3*
- **Keywords:** *AGAME, GPP, Sentinel-2 MSI, RS, eLTER, primary, vegetation, plant, production, error estimate, uncertainty*
- **Variable:** *Uncertainty*
- **Standard Observation:** *Gross Primary Production (SOBIO_090, url:* https://vocabs.lter-europe.net/so/090)
- **Site reference:** *https://deims.org/<deims.id>*
- **Temporal extent:** *2017 – 2023,* upon in-situ data availability
- **Geographic bounding box:** *<automatic based on geoTIFF>*

- **Geographic projection:** EPSG:4326
- **Format specification:** Cloud Optimised geoTIFF (CoG) and CSV (for the whole period)
- **Responsible party name:** *Anna Spinosa, Valeria Mobilia, Mario Fuentes*
- **Contact email address**:
    - Anna.spinosa@deltares.nl
    - valeria.mobilia@deltares.nl
    - Mario.FuentesMonjaraz@deltares.nl
- **Citation:** *AGAME Project Team* (2024) <site name> - Gross Primary Production Error Estimate <year>. Data product. DOI:xxx.xxx/xxx

Variable	Unit	Value range	Definition
GPP Valid Areas	boolean	0-1	Indicate the areas that have similar vegetation to the climatological footprint fetch. For those areas, the metrics of performance calculated for the climatological footprint 70 are provided both in the maps metadata as well as separate csv file.

Figure 2.- Hohes Holz – Gross Primary Production Valid Areas Map (2021) product. Mock-up of an AGAME Gross Primary Production data of an uncertainty data product for TERENO Harz - central german lowland - Hohes Holz site (https://deims.org/ddd2e8d2-44db-420e-8fa4-6b4fe1b00c78) in 2021 within GEOSS.

Gross Primary Production and model's input variables

- **Title: <site name> - Gross Primary Production and model's input variables (2017-2023)**

- **Abstract:** This dataset provides a multilayer data set in NetCDF format including the <site name> - Gross Primary Production (2017-2023), and the input variables used in the algorithm to estimate GPP. The layers included in this dataset are described in Annexes in the Remote sensing indices tables and the Environmental data from ERA5 models. This last dataset is resampled to 10 m resolution, but their values assumed constant and equal to the eddy covariance fetch estimations. Additionally, this dataset provides the SCL and QA60, used to estimate the cloud coverage and mask the clouds in the GPP product.
- **Method description:** See description on Gross Primary Production (GPP) data layer method description.
- **Data license:** *CC-BY 4.0 Int*
- **Data level:** *3*
- **Keywords:** *AGAME, GPP, EBV, Sentinel-2 MSI, RS, eLTER, primary production, ecosystem functioning, vegetation, plant*
- **Variable:** *Gross Primary Production (url:* http://vocabs.lter-europe.net/EnvThes/21074)
- **Standard Observation:** *Gross Primary Production (SOBIO_090, url:* https://vocabs.lter-europe.net/so/090)
- **Site reference:** *https://deims.org/<deims.id>*
- **Temporal extent:** 01-06-2017 – 31.12.2023
- **Geographic bounding box:** *<site extent>*
- **Format specification:** netCDF
- **Responsible party name**: *Anna Spinosa, Valeria Mobilia, Mario Fuentes*
- **Contact email address**:
    - Anna.Spinosa@deltares.nl
    - valeria.mobilia@deltares.nl
    - Mario.FuentesMonjaraz@deltares.nl
- **Citation:** *AGAME Project Team* (2024) <site name> - Gross Primary Production and model's input variables 2017-2023. Data product. DOI:xxx.xxx/xxx

Variable	Unit	Value Range	Definition
GPP	g C/m²/day	nan: masked areas; 0 - 20	Total amount of carbon fixed through ecosystem photosynthesis in a given area and period, expressed as g C/m²/day
Clr	-	0 - 5	Chlorophyll Index Red (ClR) – Sensitive to chlorophyll content in vegetation
EVI	-	0 - 1	Enhanced Vegetation Index (EVI) – Optimized for vegetation monitoring
EVI2	-	0 - 1	Two-band Enhanced Vegetation Index (EVI2) – Similar to EVI, but uses only two bands
NDVI	-	-1 to 1	Normalized Difference Vegetation Index (NDVI) – Measures vegetation health and density
MNDVI	-	-1 to 1	Modified Normalized Difference Vegetation Index (MNDVI) – Variation of NDVI, often used to reduce atmospheric noise
MNDWI	-	-1 to 1	Modified Normalized Difference Water Index (MNDWI) – Often used to delineate open water features
NDII	-	-1 to 1	Normalized Difference Infrared Index (NDII) – Sensitive to plant water content

LSWI	-	-1 to 1	Land Surface Water Index (LSWI) – Sensitive to moisture content in vegetation and soil
SCL	-	0 - 11	Scene Classification Layer (SCL) – Indicates the type of land cover and quality (e.g., clouds, water, vegetation)
QA60	-	0 or 1	Quality Assurance (QA60) – Bit mask indicating cloud presence: Bits 0-9: Unused Bit 10: Opaque clouds (0: None, 1: Present) Bit 11: Cirrus clouds (0: None, 1: Present)
LW_IN_ERA	W/m²	0 - 500	Downward longwave radiation flux (ERA5)
LW_IN_JSB_ERA	W/m²	0 - 500	Adjusted downward longwave radiation flux (JSBACH model in ERA5)
PA_ERA	hPa	900 - 1100	Surface pressure (ERA5)
P_ERA	mm/day	0 - variable (depends on region)	Total precipitation (ERA5)
SW_IN_ERA	W/m²	0 - 1000	Downward shortwave radiation flux (ERA5)
TA_ERA	°C	-50 to 50	Air temperature at 2 meters (ERA5)
VPD_ERA	hPa	0 - 50	Vapour pressure deficit (ERA5)
WS_ERA	m/s	0 - 30	Wind speed at 10 meters (ERA5)

Production

Method description

The methodology integrates multiple data from different sources via machine learning techniques to estimate Gross Primary Production (GPP) across different ecosystem types. The process starts with the selection of the sites, based on criteria such data availability. Sites with at least three full years of eddy covariance flux data are selected across different ecosystem types. GPP and environmental data (e.g., air temperature, vapour pressure deficit) are then extracted from the ICOS database, level 2 data products are used as they form the final, quality-checked ICOS RI dataset. The daily data provided from the ARCHIVE data product, which has already been gap-filled is used (ICOS Community Portal, https://www.icos-cp.eu/). Then, remote sensing indices (e.g. NDVI, EVI, etc.) are estimated after applying the cloud filter, in Google Earth Engine (GEE) using Sentinel-2 MSI data. Indices are computed for each pixel within the selected bounding area and averaged to obtain a daily value per image. An interpolation followed to get a daily time series for each variable derived from Sentinel-2 data. Both the in situ and remote sensing variables are used as predictors for the model. Additional variables are derived, those include elevation, latitude and longitude, ecosystem type. Those are incorporated in the model as categorical explanatory variables.

In the model setup, an XGBoost model is trained using 70% of the data, the remaining 30% are used for testing purposes. The model parameters are fine-tuned. The output of this trained model is a time series of GPP prediction. The 30% training data is used to evaluate the model against in-situ GPP data, using metrics like Mean Absolute Error (MAE), Root Mean Absolute Error (RMAE), and R^2.

Once the time series model is established, the model is used for the computation of 5-day GPP maps. The maps computation phase applies the trained model to ecosystem boundaries. The spatial extent of the GPP maps corresponds to the boundaries of long-term observation sites as recorded in the DEIMS-SDR registry (e.g., https://deims.org/d4854af8-9d9f-42a2-af96-f1ed9cb25712). For sites

where the boundary area is smaller than 1 km², or if only point coordinates are available in DEIMS-SDR, the maps are constrained to a 1 km x 1 km bounding box.
Annual performance comparisons are conducted between model predictions and in-situ measurements. Specifically, the yearly average modelled GPP within the climatological footprint 70, is compared with the annual mean value of GPP derived from in-situ data. Validation evaluation metrics, as MAE and R^2 are provided. Next, K-means clustering, an unsupervised classification algorithm, is used to generate classification maps, identifying areas with similar vegetation cover to the eddy covariance tower fetch. Pixels are classified as either 1: similar vegetation to the climatological footprint or 0: different vegetation type. Assuming stable or similar environmental conditions across the area, the same error metrics calculated for the climatological footprint 70, can be used to assess the model's accuracy. This process provides spatial and temporal insights into ecosystem productivity. Annual performance comparisons are conducted between model predictions and in-situ measurements for the extend of the climatological footprint 70. Additional in-situ data are however required for the provision of a fully spatially accurate product.

Producer: Deltares

Input data sources (provenance):

- ***In-situ data:*** ICOS Level 2 data[2],[3]
    - In-situ data for the model calibration and validation were downloaded from the ICOS Carbon portal (https://doi.org/10.18160/S6HM-CP8Q; see Annex).
- ***EO data:*** Sentinel-2 L2A images
- ***Site information:*** DEIMS-SDR[4]
- ***Climate model:*** ERA5 provided through ICOS infrastructure[5],[6]

- **Input variables**:
    - Daily Gross Primary Production derived from in-situ eddy covariance measurements (GPP_DT_VUT_REF, ICOS)
    - ERA5 environmental data provided through ICOS infrastructure:
        - Downward longwave radiation flux (LW_IN_ERA)
        - Adjusted downward longwave radiation flux (LW_IN_JSB_ERA, JSBACH model in ERA5)
        - Surface pressure (PA_ERA)
        - Total precipitation (P_ERA)2
        - Downward shortwave radiation flux (SW_IN_ERA)
        - Air temperature at 2 meters (TA_ERA)
        - Vapour pressure deficit (VPD_ERA)
        - Wind speed at 10 meters (WS_ERA)
    - Remote sensing indices derived from Sentinel-2 imagery
        - Chlorophyll Index Red (CLr),

[2] See https://data.icos-cp.eu/portal/
[3] See https://data.icos-cp.eu/portal/#%7B%22filterCategories%22%3A%7B%22project%22%3A%5B%22icos%22%5D%2C%22theme%22%3A%5B%22ecosystem%22%5D%2C%22type%22%3A%5B%22etcArchiveProduct%22%5D%7D%7D
[4] See https://deims.org/
[5] See https://data.icos-cp.eu/portal/
[6] See https://data.icos-cp.eu/portal/#%7B%22filterCategories%22%3A%7B%22project%22%3A%5B%22icos%22%5D%2C%22theme%22%3A%5B%22ecosystem%22%5D%2C%22type%22%3A%5B%22etcArchiveProduct%22%5D%7D%7D

- Enhanced Vegetation Index (EVI),
- Two-bands Enhanced Vegetation Index (EVI2),
- Land Surface Water Index (LSWI),
- Modified Normalized Difference Vegetation Index (MNDVI),
- Modified Normalized Difference Water Index (MNDWI),
- Normalized Difference Infrared Index (NDII),

- **Minimum requirements for temporal and spatial resolution**:
  - Daily estimations in the eddy covariance tower fetch for the Gross Primary Production and environmental data. Daily estimations at spatial resolution of 10 meters for the remote sensing indices.
- **Minimum number of sites**: any
- **Habitat requirements**: any

Data and technical requirements

- Cloud computing infrastructure providing Sentinel-2 imagery: Google Earth Engine (GEE) python API to filter and access the data.
- API providing in-situ data with ICOS standards

Processing steps including aspects of integration of data using different methods: (e.g. harmonisation of methods, calculation of accuracy and uncertainty)

- Data pre-processing:
  - Sites election based on data availability and completeness, a minimum of 4 years of continuous Gross Primary Production, environmental data, and site site-specific data is required. Estimation of remote sensing indices for the selected sites.
  - Extraction of Gross Primary Production and environmental data from ICOS.
  - Data combination and split in training and testing data (70% and 30% of the selected sites respectively)
- Model Setup:
  - Training of a XGBoost model with training dataset for time series prediction
  - Model tuning and model performance evaluation calculating MAE, RMAE, MAE, and R^2 between model estimations and in-situ data from the testing dataset.
- Maps computation:
  - Estimation of Gross Primary Production using the derived XGBoost model for each Sentinel-2 acquisition over the area of interest with 50% or less cloud coverage.
  - Computation of error metrics (MAE, RMAE, R^2) for the yearly average modelled GPP within the climatological footprint 70, which is compared the annual mean GPP from in-situ data.
  - Identification of areas with the same vegetation cover of the eddy covariance tower fetch within the bounding box. Classification maps are obtained using an unsupervised algorithm (K-means algorithm).

Link to the model

- **GitHub (https://github.com/eLTER-RI/agame)**
- **Validation and quality assurance procedures**
  - Comparison with in-situ observations for valid ecosystems.
  - Performance indices (MAE, RMAE, MAE and R^2).

Production effort estimation

- Data pre-processing: 2 months
- Model setup: 2 months.
- Maps computation: 2 months

Production frequency

- On demand

Citation

AGAME Project Team (2024) Gross Primary Production maps 2017-2023. Data Product Factsheet. https://doi.org/10.23728/b2share.14832ed336a44b3c8e284996ffa3202c

Acknowledgement

This work on the AGAME Gross Primary Production data product is funded by the European Space Agency (ESA, contract no. 4000143740/24/I-AG) in the frame of the GEOSS Platform Plus project (Horizon Europe, GA No. GA.Nr. 101039118). The work done is based on the requirements from eLTER contributing in addition to the eLTER Site Information Cluster. In-situ data for model calibration and validation has been derived from the ICOS Carbon Portal.

We would like to thank the ICOS Research Infrastructure for providing the flux tower data used in this study. We thank the teams at the ICOS stations of San Rossore 2, Hyltemossa, Torgnon, Mejusseaume, Hartheim, Hainich, Davos, Svartberget, Bilos, Hyytiala, Hohes Holz, Vielsalm, Brasschaat, Norunda, Fontainebleau-Barbeau and Tharandt for their effort in collecting data.

References

Spinosa, A., Eleveld, M., Mallast, U., Peterseil, J., Mobilia, V., Karisma, K., Fuentes-Monjarez, M., El Serafy, G. (2024) Automated Gross Primary Production Application for Monitoring Ecosystem Health within GEOSS. IGARSS Conference Proceedings 2024, Athens, Greece. Pp. doi: 10.1109/IGARSS53475.2024.10642481.

Chen T., and Guestrin C. (2016), "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

Annex - Input variables

Site information

Name	Site ID (ICOS)	DEIMS.iD	Country	Ecosystem Type	Years of data	ICOS Data Citation
San Rossore 2	IT-SR2	https://deims.org/29d0ad65-2274-45d5-9f64-b600f9faf983	Italy	evergreen needleleaf forests	4	Arriga, N., Goded, I., Dell'Acqua, A., Matteucci, M. (2025). ETC L2 ARCHIVE from San Rossore 2, 2019-01-01–2025-10-01, ICOS RI, https://hdl.handle.net/11676/BwK-6iDNKSg7642Ll55MOO69
Hyltemossa	SE-Htm	https://deims.org/b27549de-1e6f-4f3e-9e22-4d4f3dec61e3	Sweden	evergreen needleleaf forests	6	Heliasz, M., Kljun, N., Biermann, T., Holst, J., Holst, T., Kornacher, P., Linderson, M., Mölder, M., Rinne, J. (2025). ETC L2 ARCHIVE from Hyltemossa, 2018-01-01–2025-10-01, ICOS RI, https://hdl.handle.net/11676/TZ238Cd2oU7G-uKzHGjZLqJu
Torgnon	IT-Tor	https://deims.org/a03ef869-aa6f-49cf-8e86-f791ee482ca9	Italy	grasslands	7	E. Cremonese, M. Galvagno, U. Morra di Cella, ETC L2 Fluxes from Torgnon, 2015-12-31–2024-12-31, Ecosystem Thematic Centre (2025). URL: https://hdl.handle.net/11676/oWX79D_bh4t41D7IWxkmEmsF
Mejusseaume	FR-Mej	https://deims.org/a5496211-d63f-494b-8da7-64e9abf8898b	France	grasslands	4	Flechard, C., Flechard, C., Fauvel, Y., Hamon, Y., Jacotot, A. (2025). ETC L2 Fluxes from Mejusseaume, 2018-12-31–2024-12-31, ICOS RI, https://hdl.handle.net/11676/hocQhquWY1rcGj_qPXZM3fds
Hartheim	DE-Har	https://deims.org/213a9ea7-0155-408d-8763-0a9fabf4a666	Germany	mixed forests	4	Christen, A., Baab, F., Kolbe, S., Plapp, T., Redepenning, D., Schindler, D. (2025). ETC L2 Fluxes from Hartheim, 2019-12-31–2024-12-31, ICOS RI, https://hdl.handle.net/11676/tuVVgJ68UZO_kT2qfBP_YAIY
Hainich	DE-Hai	https://deims.org/5b409a72-2a45-4238-a501-e24f1a2900db	Germany	deciduous broadleaf forests	5	Knohl, A., Tiedemann, F., Klosterhalfen, A., Markwitz, C., Siebicke, L. (2024). ETC L2 Fluxes from Hainich, 2018-12-31–2023-12-31, ICOS RI, https://hdl.handle.net/11676/DOFlHcQWEUNCUk9TqSZEj4SL
Davos	CH-Dav	https://deims.org/a547dab2-859a-414c-b148-0e7df8de5773	Switzerland	evergreen needleleaf forests	4	Feigenwinter, I., Etzold, S., Gharun, M., Hortnagl, L., Meier, P., Liechti, K., Stutz, T., Burri, S., Zweifel, R., Buchmann, N., Baur, T., Eugster, W., Gessler, A., Hug, C., Hummel, D., Häni, M., Hüglin, C., Kumar, S., Marty, M., Schmitt Oehler, M., Staudinger, M., Sutter, F., Thimonier Rickenmann, A., Trotsiuk, V., Waldner, P., Wilhelm, M., Zhu, J., Zimmermann, S. (2025). ETC L2 Fluxes from Davos, 2018-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/VCcKtEYQO468-OhYpZcf78GC

Name	Site ID (ICOS)	DEIMS.iD	Country	Ecosystem Type	Years of data	ICOS Data Citation
Svartberget	SE-Svb	https://deims.org/c0705d0f-92c1-4964-a345-38c0be3113e1	Sweden	evergreen needleleaf forests	5	Peichl, M., Nilsson, M., Larmanou, E., Smith, P., Marklund, P., De Simon, G., Lofvenius, P., Dignam, R., Holst, J., Molder, M., Andersson, T., Boschetti, F., Kozii, N., Linderson, M., Ottosson-Löfvenius, M. (2025). ETC L2 Fluxnet (half-hourly) from Svartberget, 2018-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/qR-2VNMiMXAsYwMLzwucjOsR
Bilos	FR-Bil	https://deims.org/6f716444-c0bd-4a04-b72b-add3e302eef1	France	evergreen needleleaf forests	4	Domec, J., Loustau, D., Aluome, C., Chipeaux, C., Denou, J., DEPUYDT, J., Garrigou, C., Kruszewski, A., Lafont, S. (2025). ETC L2 Fluxes from Bilos, 2018-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/snw9tOxm4IuJClv8EUKmKmAJ
Hyytiala	FI-Hyy	https://deims.org/663dac80-211d-4c19-a356-04ee0da0f0eb	Finland	evergreen needleleaf forests	5	Mammarella, I., Aalto, J., Back, J., Kolari, P., Laakso, H., Levula, J., Matilainen, T., Pihlatie, M., Pumpanen, J., Taipale, R., Vesala, T. (2023). ETC L2 Fluxes from Hyytiala, 2017-12-31–2023-10-31, ICOS RI, https://hdl.handle.net/11676/Ke4Y-W71Kj_Tc8l2Q--ZZpMZ
Hohes Holz	DE-HoH	https://deims.org/ddd2e8d2-44db-420e-8fa4-6b4fe1b00c78	Germany	deciduous broadleaf forests	4	Rebmann, C., Dienstbach, L., Schmidt, P., Wiesen, R., Meis, J., Feldmann, I., Bastos Campos, F., Dejosez, S., Garcia Quiros, I., Gimper, S., Hautmann, D., Hildebrandt, A., Kempka, P., Paasch, S. (2025). ETC L2 Fluxes from Hohes Holz, 2018-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/2L1kNpUSC5efdHkzpoNCpyDT
Vielsalm	BE-Vie	https://deims.org/c4c1c0ca-5a43-4d19-ab50-34dad0af44e8	Belgium	mixed forests	3	Vincke, C., Bogaerts, G., Chebbi, W., Chebbi, W., Chopin, H., Demoulin, L., Douxfils, B., Engelmann, T., Faurès, A., Heinesch, B., Manise, T., Orgun, A., Piret, A., Thyrion, T. (2025). ETC L2 Fluxes from Vielsalm, 2019-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/ekoaXg7X-lvgLhR_LAj_2fqw
Brasschaat	BE-Bra	https://deims.org/68e6a8e5-d6d2-4c8c-91c4-10e7f87ac556	Belgium	evergreen needleleaf forests	3	Janssens, I., Bresseleers, S., De Meulder, T., Lefevre, L., Naseer, M., Roland, M., Segers, J., Van Look, J., Vande Sompele, A. (2025). ETC L2 Fluxes from Brasschaat, 2019-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/o4-UZZQCQbRq4S_kgyXYGAhs
Norunda	SE-Nor	https://deims.org/459284f2-c7cd-4bb1-96c3-2b38c338ff4a	Sweden	evergreen needleleaf forests	5	Kljun, N., Molder, M., Lehner, I., Bergström, G., Båth, A., Holst, J., Linderson, M. (2025). ETC L2 Fluxes from Norunda, 2017-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/GjBSkT2tl8tfZDyHv17REzmU

Name	Site ID (ICOS)	DEIMS.iD	Country	Ecosystem Type	Years of data	ICOS Data Citation
Fontainebleau-Barbeau	FR-Fon	https://deims.org/dcd5ef29-9000-4c25-ac5c-8bf068e0c264	France	deciduous broadleaf forests	4	Berveiller, D., Dufrêne, E., Delpierre, N., Morfin, A., Francois, C., Vincent, G., Bazot, S., Soudani, K., Girardin, C., Guillot, T., Perot-Guillaume, C. (2025). ETC L2 Fluxes from Fontainebleau-Barbeau, 2018-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/WgPSFQlVYui2ANucafunTOY9
Tharandt	DE-Tha	https://deims.org/639cfe33-17b2-49fc-8e8d-3a2cb80231fc	Germany	evergreen needleleaf forests	3	Bernhofer, C., Eichelmann, U., Grünwald, T., Hehn, M., Mauder, M., Moderow, U., Prasse, H. (2025). ETC L2 Fluxes from Tharandt, 2019-12-31–2025-09-30, ICOS RI, https://hdl.handle.net/11676/CiyUpM0rxPnZcy3ekLJnfCaL

Site-specific information

Variable	Description
lat	Latitude – Geographic coordinate specifying the north-south position
lon	Longitude – Geographic coordinate specifying the east-west position
elevation	Elevation – Height above sea level
canopy_height	Canopy Height – Average height of the forest canopy
biom_evergreen needleleaf forests	Flag for Evergreen Needleleaf Forests – Indicates whether the location is classified as Evergreen Needleleaf Forests biome
biom_grasslands	Flag for Grasslands – Indicates whether the location is classified as grassland biome
biom_deciduous broadleaf forests	Flag for Deciduous Broadleaf Forests – Indicates whether the location is classified as deciduous broadleaf forests biome
biom_mixed forests	Flag for Mixed Forests – Indicates whether the location is classified as mixed forest biome
day	Day – Day of the year (1-365)
month	Month – Month of the year (1-12)
winter	Winter – Indicator variable or flag representing the winter season
spring	Spring – Indicator variable or flag representing the spring season
summer	Summer – Indicator variable or flag representing the summer season
fall	Fall – Indicator variable or flag representing the fall season

Environmental data downscaled from ERA5 models and in-situ data

Variables	Description
LW_IN_ERA	Downward longwave radiation flux (ERA5)
LW_IN_JSB_ERA	Adjusted downward longwave radiation flux (JSBACH model in ERA5)
PA_ERA	Surface pressure (ERA5)
P_ERA	Total precipitation (ERA5)
SW_IN_ERA	Downward shortwave radiation flux (ERA5)
TA_ERA	Air temperature at 2 meters (ERA5)
VPD_ERA	Vapor pressure deficit (ERA5)
WS_ERA	Wind speed at 10 meters (ERA5)
GPP_DT_VUT_REF	Gross Primary Production (GPP) - daytime (VUT_REF model) (ICOS)

Remote sensing indices

Vegetation Index	Description
Clr	Chlorophyll Index Red (ClR) – Sensitive to chlorophyll content in vegetation
EVI	Enhanced Vegetation Index (EVI) – Optimized for vegetation monitoring

Vegetation Index	Description
EVI2	Two-band Enhanced Vegetation Index (EVI2) – Similar to EVI, but uses only two bands
NDVI	Normalized Difference Vegetation Index (NDVI) – Measures vegetation health and density
MNDVI	Modified Normalized Difference Vegetation Index (MNDVI) – Variation of NDVI, often used to reduce atmospheric noise
MNDWI	Modified Normalized Difference Water Index (MNDWI) – Often used to delineate open water features
NDII	Normalized Difference Infrared Index (NDII) – Sensitive to plant water content
LSWI	Land Surface Water Index (LSWI) – Sensitive to moisture content in vegetation and soil