



D2.1 – Pilot above ground dataset

***Compilation of readily available pre-processed
aboveground dataset and identification of data gaps***

Planned delivery date (as in DoA): M12

Due date of deliverable: 31/12/2024

Work Package: WP2

Work Package leader: FVA-BW

Deliverable leader: FVA-BW

Version: 1.0

Grant Agreement number: 101081177

WILDCARD — HORIZON-CL5-2022-D1-02 / HORIZON-CL5-2022-D1-02-05

Project Acronym: WILDCARD

Effects of rewilding in forests and agricultural lands on carbon sequestration and diversity

Duration of the project: January 1st, 2024 – December 31st, 2027

VERSION HISTORY~

Version	Date	Author /Reviewers	Partner	Description
0.1	20. 10.2024	Jan Zálešák Kamill Král Jan Holík	VUKOZ	Very First Outline
0.2	24. 10.2024	Jan Zálešák Kamil Král Martin Krůček Jan Holík	VUKOZ	First Draft
0.3	26. 10.2024	Lucia Seebach	FVA-BW	Review description (e.g. review of draft deliverable 0.1)
0.4	04.11.2024	Peter Meyer	NW-FVA	Data gap analysis
0.5	04.12.2024	Peter Meyer Elisabeth Klingberg	NW-FVA	Data gap analysis
0.6	06.12.2024	Peter Meyer Elisabeth Klingberg Jan Zálešák Kamil Král Jan Holík Tim Nairn Lucia Seebach	NW-FVA VUKOZ FVA-BW	Consolidated draft
0.7	07.12.2024	Giorgio Alberti	UNIUD	Revision
1.0	20.12.2024	Project Steering Board	All partners	Approval

CITATION:

Jan Holík, Yannek Käber, Elisabeth Klingberg, Kamil Král, Martin Krůček, Peter Meyer, Tomáš Přívětivý, Lucia Seebach, Jan Zálešák (2024). D2.1. - Pilot Aboveground Dataset and identification of gaps. Horizon Europe Project No. 101081177, European Commission, 34 pp.

ACKNOWLEDGEMENT



**Funded by
the European Union**

DISCLAIMER

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them



Table of contents

Executive Summary	3
Keywords	4
1. Introduction	5
2. Terms, Definitions and Data Requirements	6
2.1 Site	6
2.2 Time since abandonment (TSA)	6
2.3 Proforestation	6
2.4 Minimum dendrometric data requirements	7
3. Approach for dendrometric data collection, harmonization and integration	8
3.1. Data collection and harmonization	8
3.1.1 General approach	8
3.1.2 WILDCARD Site Metadata Table	8
3.1.3 Dendrometric Data Delivery	8
3.1.4 Metadata collection	9
3.2. Data Integration and Integrity	9
4. Database Structure and Description	11
4.1. Technical description	11
4.2. Individual tables of the database	12
5. Database Content	17
6. Gap analysis	18
6.1. Approach	18
6.2. Results of the gap analysis	20
6.3. Map of readily available sites	23
6.4. Gap filling steps	23
7. References	26
Annex	27
1.1. Abbreviations and long names of involved institutes	27
1.2. Gap overview for TSA 'Primeval'	28
1.3. Gap overview for TSA '>60'	29
1.4. Gap overview for TSA '30-60'	30
1.5. Gap overview for TSA '0-30'	31

Executive Summary

Background

The WILDCARD project investigates the effects of rewilding in European forests ('proforestation') and on agricultural lands ('afforestation') on carbon (C) sequestration and biodiversity. In Work Package 2 (WP2), this project specifically examines the effects of proforestation, utilizing a range of existing long-term dendrometric datasets across Europe. These datasets, provided initially by WILDCARD project partners at this pilot stage and later supplemented by external partners, require harmonization to enable consistent analyses across ecological gradients and temporal scales. Additionally, the representativeness of the available datasets must to be evaluated to ensure the validity and robustness of WP2's findings.

Deliverable's objectives

- Develop a unified database structure to integrate diverse dendrometric datasets.
- Provide a pilot dataset on aboveground forest dynamics, serving as a foundation for broader analyses of C and biodiversity trajectories.
- Identify and quantify gaps in the available data concerning ecologically relevant patterns in European forests, including length of Time Since Abandonment (TSA) and coverage of primeval forests, across both ecological and temporal gradients.

Methods

Data Collection and Harmonization: Pre-processed datasets from data providers were aligned with the WILDCARD's data schema, followed by central harmonization to ensure consistency across the data.

Database Development: A relational database (PostgreSQL with PostGIS) was developed to manage the data, integrating both spatial and ecological metadata.

Gap Analysis: A systematic comparison of the available datasets against defined ecological combinations (Forest Type Categories (FTC), ecoregions, soil properties) and TSA timelines was conducted to assess data coverage and identify major gaps.

Summary of Results

- A pilot database was successfully developed, including data from 136 sites, 8404 plots, and over 473,000 tree records.
- Initial gap analysis revealed missing data for the Boreal, Pannonian and Mediterranean ecoregions and limited representativity of the Atlantic Region. Alpine and Beech FTCs are well represented, while several other FTCs are still missing. The TSA gradient was rather well covered for the represented combinations of FTC/ecoregion/soil properties but exhibited major gaps across the whole length of ecological gradients.
- Data gaps highlight the need for additional external datasets and interpolation methods to fill underrepresented gradients.

The pilot dataset, stored on a PostgreSQL server, is currently accessible only to WILDCARD project partners via custom access credentials. No public access is planned at this stage. For access control purposes, credentials must be created. Please contact Jan Zalesak (zalesak@vukoz.cz) for further assistance.



Key Messages

- The pilot dataset provides a strong foundation for studying aboveground forest dynamics, supporting WILDCARD's objectives of assessing proforestation's role in climate and biodiversity strategies.
- Data gaps remain in underrepresented forest types and regions, requiring targeted data acquisition and methodological adjustments.
- Future work will focus on expanding dataset coverage, utilizing statistical models for interpolation, and validating findings through literature and additional data contributions.
- This deliverable advances the understanding of proforestation by providing tools and frameworks for comprehensive pan-European forest analyses.

Keywords

Rewilding, proforestation, data structure, dendrometric database, available data

1. Introduction

Within Work Package 2 (WP2), task 2.1 “Aboveground implications of proforestation” relies heavily on existing dendrometric and biodiversity datasets. These datasets are primarily provided by WILDCARD partners and members of the European Forest Reserve Initiative (EuFoRia) network, who are not WILDCARD partners. As a result, the datasets may vary in several aspects, including sampling design, measurement thresholds and plot sizes. Therefore, it is essential to establish a common database structure to integrate these diverse datasets into a cohesive and consistent framework, enabling comprehensive data analysis.

The database is supplemented by a site metadata table (“WILDCARD Metadata Table”), which provides an overview of the data coverage in relation to ecological conditions, forest types, time since abandonment (TSA), availability of biodiversity data and several other key attributes. At this stage, the WILDCARD Metadata Table serves as a foundation for a data gap analysis aimed at quantifying the extent to which the diversity of the study region and varying timescales of rewilding are represented.

This deliverable defines the minimum data requirements for analysing C and biodiversity trajectories in European forest reserves within the WILDCARD project. It outlines the process of dendrometric data collection and harmonization, the structure and the current content of the common dendrometric database and the approach to the data gap analysis.

At this stage of the project, the pilot aboveground dataset (D2.1) includes only dendrometric data from WILDCARD project partners, except for the gap analysis (Section 6), which incorporates data from all sites where metadata collection (TSA, forest type, soil type) is already complete. In the next phase, the dataset will gradually be expanded to include data from by external data-providers, primarily from the EuFoRia network. Additionally, preliminary results of the data gap analysis are presented and potential approaches for gap filling are discussed. As the aboveground dataset continues to grow beyond the deadline of this deliverable, the data gap analysis will be updated at regular intervals until the end of June 2025 (Milestone 5: “Collection and harmonisation of existing aboveground data from established networks completed”).

Beyond task 2.1 (“Aboveground implications of proforestation”), this deliverable is also closely linked to task 2.2 (“Belowground implications of proforestation”) and task 2.3 (“Resistance and resilience of carbon stocks and biodiversity to future climate change including disturbance”). In task 2.2, the aboveground provides the foundation for a gap analysis related to belowground C stocks and the selection of high intensity sampling sites (HIS-below). In task 2.3, the dataset will be used to validate a range of models on forest dynamics.

2. Terms, Definitions and Data Requirements

2.1 Site

In the context of WP2 within the WILDCARD project, a *site* (Figure 1) is defined as a relatively homogenous study area with respect to [EEA European Forest Type](#) categories, soil water and nutrient class, and time since abandonment (TSA). A site may consist of one or more plots. The definition of a (quasi-)homogenous site should be feasible using the existing dataset of data-providers. Any additional field mapping and refinement is not anticipated, and therefore, limited internal variability in site conditions is considered acceptable.

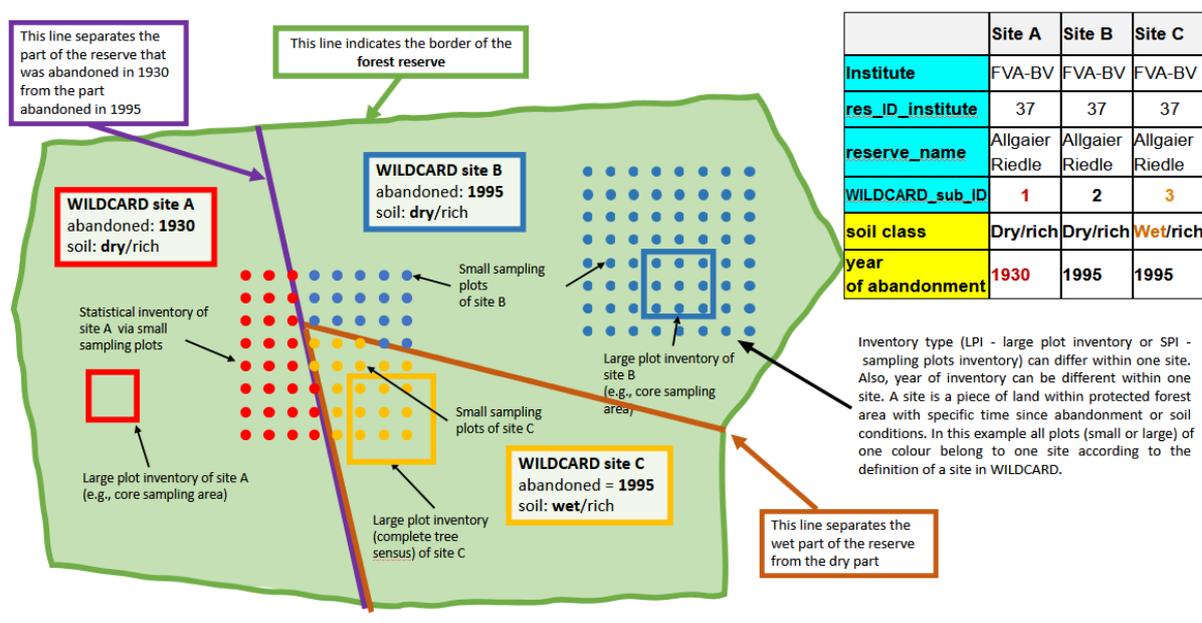


Figure 1. Definition of the SITE and its labelling in the WILDCARD Metadata table.

2.2 Time since abandonment (TSA)

Time Since Abandonment (TSA) refers to the number of years since the cessation of forest management at a specific site. In the context of the dendrometric data analysis, TSA is the duration (years) between the year of abandonment and the year of data collection. If the exact year of abandonment is unknown, the year of reserve designation can be used as a proxy.

2.3 Proforestation

Proforestation is defined as the spontaneous natural development of a forest after the cessation of forest management (Kun et al., 2020; Moomaw et al., 2019)



2.4 Minimum dendrometric data requirements

The following standards have been established for the use of dendrometric data in the WILDCARD project:

1. Dendrometric data have been collected from **permanent plots** in strictly unmanaged, old-growth or primary forests. At least one (but preferably more) tree census is available. For each site, either one or several large (≥ 0.5 ha) or multiple small sample plots ($n \geq 5$, each ≥ 500 m²) have been established. A single tree census/ inventory is acceptable, provided the plot location can be re-identified in the field for subsequent re-measured.
2. For each site, information on **time since abandonment** (TSA), [EEA European Forest Type](#) and soil water and nutrient class is available.
3. **Individual trees/stems and deadwood objects** have been measured. For each object, data on species (or genus; for deadwood: broadleaved/coniferous and decay class), diameter, length/height (with possible sub-sampling) and basic status (alive/dead) are available. If deadwood has not been measured at the level of individual objects (stems/pieces), plot-level estimates of coarse woody debris (CWD) volume/biomass are acceptable, provided the estimation method is documented and included as metadata with the data delivery. In exceptional cases, estimates of deadwood/CWD volume/biomass may be based on simplified site-level calculations.

3. Approach for dendrometric data collection, harmonization and integration

3.1. Data collection and harmonization

3.1.1 General approach

This European-wide dendrometric data harmonization is built on three pillars:

- i. basic pre-harmonization by the data-providers, following the work description and data template provided by the data-holding institution (VUKOZ);
- ii. internal data harmonization by the data-holding institution (VUKOZ) prior to uploading the data into the relational database and during related data treatment;
- iii. comprehensive data documentation (metadata) for each data delivery. Metadata are an integral part of both dendrometric data submission and the final relational database.

Gathering and harmonising data collected in different countries, using various methodological approaches and at different hierarchical levels of the inventory (site, plot, tree/ stem and coarse woody debris), and often maintained in different national languages, is an extremely challenging task. As a result, pre-processing must be carried out by the data-provider before data submission. This process includes translating data fields and values into English, as well as applying prescribed labelling and formatting for at least a set of basic attributes. To facilitate this, a data template has been created that all data-providers must follow.

3.1.2 WILDCARD Site Metadata Table

The collection of existing dendrometric data begins with the establishment, update and completion of the [WILDCARD Site Metadata](#) Table. This Table is based on the Metadata table of [EuFoRIA](#), the primary external source of WP2 dendrometric data not directly provided by WILDCARD project partners. However, there are some key conceptual differences between the two metadata tables. In the WILDCARD Site Metadata Table, each site has a single row dedicated solely to the site's environmental conditions without including methodological specifics of data collection. Each site record contains all necessary information, such as the year of abandonment, inventory years, soil water and soil nutrient class, EEA Forest Type, latitude/longitude and availability of additional data (e.g. LiDAR, or biodiversity data for selected guilds). Ultimately, every dendrometric data submission should include all sites listed for given data-provider in this table and vice-versa.

3.1.3 Dendrometric Data Delivery

A dendrometric data delivery follows the completion and updating of the WILDCARD Metadata Table. The details of dendrometric data delivery is outlined in the [Work Description document](#). The delivery distinguishes between mandatory “BASIC” variables, optional “EXTENDED” variables and metadata, all of which together form the core content of the database (see section 4.2).

3.1.4 Metadata collection

As mentioned earlier, extensive data documentation is the third and crucial pillar of the data collection and harmonization. Some dendrometric data attributes cannot be harmonized mechanically without substantial loss or bias of information. Additionally, the optimal approach to harmonization may vary depending on the specific purpose. For these reasons in some cases, the database retains the original values provided by each data-provider (e.g. deadwood decay classes), while also documenting the adopted definitions in the related metadata. It is therefore essential that the data be accompanied by all necessary metadata, which forms an integral part of the final database.

3.2. Data Integration and Integrity

Establishing a robust framework for dendrometric data management has been crucial for this part of the project. The proposed framework encompasses the design, development, documentation, dissemination and maintenance of relevant data standards, ensuring the availability, consistency and usability of all dendrometric datasets. The approach also involves providing support to data-providers during the data pre-processing phase by maintaining ongoing communication to clarify any uncertainties regarding the data and its required formatting. Before and after data submission, the data-holding institution ensures the correctness, completeness and consistency to verify their compliance with the prescribed standards. These standards are outlined in the following documents:

- Detailed Work Description;
- Dendrometric Data Template;
- WILDCARD Site Metadata Table.

Based on these standards, the data-holding institution also ensures adherence to geographical, spatial and methodological data frameworks prior to data submission and defines and addresses any necessary harmonization to account for variations across data-providers' data standards. Once the data is delivered and harmonized, the database implementation upholds data integrity through a multi-layered system that enforces validation at every stage of the data lifecycle.

To ensure consistent and comprehensive data integrity from early stages of data delivery, a **web-based application** (<https://databasechecking-euforia.streamlit.app/>) has been developed. This platform allows both the data-provider and the database product owner (VÚKOZ) to collaboratively perform the data submission validation process in one centralized location, with proper tracking and backlogging. The following section provides an overview of how data integrity is enforced at a high level.

First, **completeness and correctness checks** ensure that all required columns (data fields) are present and that the submitted data adheres to the defined data formats and value ranges. This validation policy prevents missing or incomplete data, as well as invalid entries.

Advanced **consistency and plausibility checks** are then conducted to identify anomalies, such as implausible diameter measurements or mismatched tree statuses. In cases of ambiguity, communication with the data-providers resolves any uncertainties, ensuring that the data meets high-quality standards before being entered into the central database.

After data is stored in the central database, full **access to this data is restricted** to the database product owner, ensuring that only justified and approved changes are made, following a strictly defined and documented process. **Data-providers are granted read only access**, allowing them to use the data for scientific purposes while ensuring the security and confidentiality of the entrusted data, preventing unauthorised people. For specific scientific needs, **access to externals can be provided**, but only upon written request (formal data agreement), with justification and approval from the project leadership, prioritizing data security.

Rules such as unique identifiers for trees, plots, and sites, strict transitions in tree statuses (e.g., alive to dead), and constraints on spatial and geometrical definitions ensure that the data accurately reflects real-world conditions.

Finally, periodic data integrity tests are conducted to identify errors in existing records. By systematically comparing historical (pre-database) and current data, the system flags inconsistencies, promoting accuracy and traceability. Together, these measures guarantee the reliability and scientific validity of the database, supporting accurate querying and analysis for ecological research.

3.2.1. Column Presence Tests

The on-line application verifies the presence of required columns and checks that their values fall within allowed ranges, based on predefined configurations. These configurations include validation rules such as predefined column names and data types (e.g., non-null, numeric). The system ensures that each column meets its specified requirements, confirming that the data adheres to the structure and constraints defined by the database schema. This validation process helps identify missing and incomplete data, incorrect formats, column mismatches and typographical errors.

3.2.2. Data Range and Value Checks

For each column in the dataset, the application checks that the values fall within the expected range and meet specific validation criteria. This ensures the logical correctness of data, such as verifying that measurements like diameter at breast height (DBH) or geographic coordinates fall within plausible and permissible ranges. These checks typically identify issues such as typographical errors, incorrect data units or systems (e.g. DBH in mm vs. cm; angles in degrees vs. radians), mismatched coordinate systems, or any inconsistent or incorrect data entries.

3.2.3. Consistency Checks Across Censuses

The application also performs checks across multiple inventory years (tree censuses) to track changes and ensure that records remain consistent and accurate over time. For example, it flags discrepancies such as unexpected changes in tree status (e.g. a tree previously recorded as dead is marked as alive, or a fragmented tree is classified as complete), significant reductions in tree diameter, or appearance of missing or reappearing tree data in subsequent inventories. Additionally, it identifies values that do not comply with internal thresholds.

3.2.4. Automation of Data Upload

This step ensures that the database is populated efficiently and error-free. A custom-developed application automates the data upload process for the database product owner, streamlining the integration of data from the providers. By minimizing manual intervention, the

system reduces the risk of human error during data entry, ensuring adherence to the defined schema and maintaining data integrity.

3.2.5. Running of Testing Queries

The final step involves running a series of database queries to generate basic stand statistics, including Mean/Min/Max site DBH (mm), Stand basal area ($\text{m}^2 \text{ha}^{-1}$) and site tree/stem density (stems ha^{-1}) for standing trees/stems. For lying stems (deadwood), the queries assess Mean/Min/Max site lower and upper log diameters (mm), the number of lying logs per hectare (logs ha^{-1}) and lying site deadwood volume ($\text{m}^3 \text{ha}^{-1}$), provided these data are included in the delivery. Similarly, for CWD, the total site CWD volume ($\text{m}^3 \text{ha}^{-1}$) is quantified, if available in the dataset. A critical aspect of this step is comparing the values derived from the database with the respective stand statistics provided by the data-providers from their original dataset (prior to database integration).

Overall, this framework not only ensures that incoming data is valid and reliable, but also that it is correctly implemented within the complex relational database environment.

4. Database Structure and Description

To store and query project data, we use the PostgreSQL relational database management system. The WILDCARD WP2 dendrometric database houses harmonized, quality checked data on standing living trees, snags, and lying deadwood, all recorded on permanent plots across unmanaged forest sites in Europe.

The core database entities include forest sites, sampling plots and dendrometric data for both living and dead trees (standing and lying), each described by relevant quantitative and qualitative attributes. These entities are interlinked with one another and with related metadata using foreign key relations.

The database was designed to integrate data from the two main types of forest inventories: **large plot inventories (LPI)** and **sampling plot inventories (SPI)**. The LPI approach involves sampling trees on large-scale plots, such as those used by the ForestGEO global network of permanent forest plots (<https://forestgeo.si.edu/>). The SPI approach, in contrast, samples trees on multiple small-scale plots, as seen in national forest inventory programs (e.g. <https://research.fs.usda.gov/programs/nfi#overview>).

By integrating all dendrometric data into a single relational database, the system facilitates the efficient, harmonized use and querying of data from various sources, serving the needs of WP2 within the WILDCARD project.

4.1. Technical description

The database has been built within the technical framework of the Czech Meta Centrum (metacentrum.cz), which operates as a part of the [European Grid Initiative \(EGI.eu\)](http://EuropeanGridInitiative.org). Meta Centrum provides researchers with access to computational and data-based resources and

facilities. Within Meta Centrum's Kubernetes cluster, we run the precompiled PostgreSQL package (ver.14) on its servers, leveraging the center's infrastructure for its reliability, robustness and security. Additionally, we take advantage of the integrated PostGIS extension (version 3.2) of PostgreSQL, enabling efficient handling of geographical data (PostGIS 3.2).

The database schema consists of six interlinked main tables: `sites`, `site_design`, `plot`, `tree`, `CWD` and `species_list`, which together integrate the various levels of dendrometric data. Each of these tables serves a specific purpose. Below is a summary of the key types of data collected in each table:

- **Inventory Data:** Fields such as `inventory_year` and `inventory_type` track the timing and method of data collection, which is essential for ensuring comparability across time and sites.
- **Tree quantitative and qualitative Data:** Fields like `DBH`, `height`, `life` and `integrity` provide crucial information on the physical condition and growth patterns of trees over time.
- **Species Data:** The `species_list` table stores species-level information, supporting biodiversity studies and the analysis of forest composition.
- **Geospatial Data:** Spatial records are stored in Well Known Text ([WKT](#)) format. The presence of spatial data (`geom` fields) in the `site_design`, `plot`, and `tree` tables means that the database integrates spatial information across all data levels. This, along with the `epsg_code` field in the `site_design` table that specifies the Coordinate Reference System (CRS) used, ensures that all spatial data are interpreted consistently according to the designated CRS, guaranteeing compatibility with GIS tools for mapping and spatial analyses.
- **Extended Attributes:** Several tables include `extended_attributes`, stored as JSON fields, allowing for flexible and scalable storage of additional data that may not fit into the predefined schema.
- **Metadata:** Any metadata necessary to explain measurement design and conditions, which cannot be fully standardized or harmonised, is included in this section.

4.2. Individual tables of the database

This section provides the list of all tables along with descriptions of their main columns (data fields).

4.2.1 Sites Table

The Sites table contains environmental information and data availability details for the forest sites, along with key identification parameters, as described in the following.

Identification data fields:

- **institute:** The institute providing and owning the raw dendrometric data.
- **site_name:** institutional name of the reserve/study area associated with the site.
- **site_id:** institutional ID of the reserve/study area associated with the site.
- **wildcard_sub_id:** a site sub-identifier given by the institute who provided the data when splitting a reserve/study area into more (quasi-)homogenous sites following the WILDCARD criteria.
- **composed_site_id:** uniquely identifies a given (quasi-)homogenous site accommodating this format: `institute__site_id__site_name__wildcard_sub_id`. As a



unique site identifier, `composed_site_id` also links multiple tables (e.g., Site Metadata Table, `sites`, `site_design`, `plots`, `cwd` and `tree`), ensuring universal site data accessibility.

- **record_id**: this field is a unique, auto-incrementing numeric identifier for each record in the site table. In this table, the `record_id` represents a unique immutable identifier of each site. It will be further referred to and used in the Site Metadata Table as **wildcard_db_id**: it is a numeric equivalent of the `composed_site_id`.

Environmental data fields:

- **forest_type**: The classification of the [EEA European Forest Type](#).
- **year_abandonment**: The year when forest has been abandoned, i.e. no management has been applied since this year.
- **soil_water**: The classification of soil water content on the site by WILDCARD specifications.
- **soil_nutrient**: The classification of soil nutrient status on the site by WILDCARD specifications.

Data availability/research intensity data fields:

- **his**: A boolean field marking whether a site is a high-intensity sampled site.
- **lidar**: A boolean field marking whether WILDCARD relevant LiDAR data has been/ about to be acquired on the site.
- **biodiversity**: A boolean field marking whether any biodiversity has been sampled on the site.

4.2.2 Site Design Table

The Site Design table describes the data collection design for each site in a given year, including plot arrangements and data collection parameters such as target woody structures (e.g., tally trees or sampled population) and size thresholds. It also groups all plots associated with a specific sampling design for each site and inventory year combination.

Identification data fields:

- **record_id**: A unique, auto-incrementing identifier for each record.
- **site identifiers**: `composed_site_id`, `site_name`, `site_id`, `wildcard_sub_id`.

Sampling design fields:

- **inventory_year**: The year the tree record was collected.
- **inventory_type**: Indication whether this is a LPI or SPI plot.
- **no_plots**: number of plots of that design (their list is in extended attributes column)
- **circle_no**: The number of circles with that SPI design.
- **circle_radius**: radius of the (sub-)circle.
- **standing_alive_threshold**: DBH from which the tree is measured.
- **standing_dead_threshold**: DBH from which the tree is measured.
- **lying_alive_threshold**: Diameter from which the tree is measured.
- **lying_dead_threshold**: Diameter from which the tree is measured.
- **species_pool**: A definition of the target species pool for the site and inventory year.
- **lis_cwd**: A boolean field indicating if coarse woody debris (CWD) sampling was conducted.

Geometry data fields:

- **geom**: The geometry field that holds spatial data including the location, shape and size of the site.
- **stem_coord_ref_point**: A reference point on the tree used for tree stem coordinates (stem base or stem centre at breast height).
- **epsg_code**: Coordinate reference system (CRS) used for spatial data.

Additional fields:

- **other**: Additional information not captured by other fields.
- **extended_attributes**: A JSONB field for storing extra attributes or metadata related to the site design, like *plot_list* for the site design (in JSON).

4.2.3. Plots Table

The Plots table contains information on individual plots within each site, including their geospatial characteristics and sampling details. It outlines the plot-site relationship, plot ID, coordinates and potentially several topographic features such as slope, aspect and elevation. Unlike the Site Design table where all LPI or SPI plots of the same data collection design in a given inventory year are described by a single record, the Plots table provides a detailed description for each plot.

Identification data fields:

- **unique_site_design_id**: A foreign key linking to the *site_design* record_id.
- **record_id** (*unique_plot_id*): Unique auto-incrementing identifier for each plot.
- **composed_site_id**: identifier of site the plot belongs to.
- **spi_id**: Unique plot identifier for a sampling plot inventory.
- **lpi_id**: Unique plot identifier for a large plot inventory.

Sampling design fields:

- **inventory_year**: The year in which the plot was surveyed.
- **circle_no**: The number of the sampling circle associated with the plot.
- **circle_radius**: The radius of the sampling circle.
- **plot_sampled**: Boolean indicating whether plot was sampled during the inventory.
- **sampled_area**: The area of the plot that was sampled, in hectares.
- **consistent_id_standing**: Boolean to indicate if IDs of standing trees are kept across inventories.
- **consistent_id_lying**: Boolean to indicate if IDs of lying trees are kept across inventories.
- **geom**: A geometry field holding plot coordinates

Additional fields:

- **extended_attributes**: A JSONB field for extra plot-specific attributes not covered by other columns.

4.2.4. Tree Table

The Tree table provides detailed information about individual trees within the plots, including their size, position, taxonomy and status.

Identification data fields:

- **institute:** The institute providing and owning the raw dendrometric data.
- **composed_site_id:** Unique site identifier.
- **record_id:** A unique, auto-incrementing identifier for each tree record.
- **tree_id:** An internal identifier for the tree individual within the plot.
- **stem_id:** An internal identifier for each stem of one tree, if a tree has multiple stems.
- **piece_id:** An identifier of pieces of one stem.

Inventory data fields:

- **inventory_year:** The year the tree record was collected.
- **consistent_id:** A boolean, True if the tree ID is consistent across inventories.
- **date:** The date of the tree measurement.

Status data fields:

- **life:** The life status of the tree, e.g., alive ('A'), dead ('D').
- **position:** The position of the tree *sensu* standing ('S') or lying ('L').
- **integrity:** The integrity of the tree: complete ('C'), fragmented ('F'), or other ("NA").
- **decay:** A classification of tree decay class.
- **species_code:** A code representing the species of the tree.

Measurement data fields:

- **dbh:** Diameter at breast height (mm).
- **diameter_1:** The diameter measured at 'lower' end of the tree (mm).
- **diameter_2:** The diameter measured at 'upper' end of the tree (mm).
- **height:** The height of the standing tree (m).
- **length:** The length of the lying tree (m).
- **geom:** A geometry field representing the tree's spatial location.

Additional data fields:

- **extended_attributes:** A JSONB field for any extra attributes related to the tree.

4.2.5. CWD Table

The coarse woody debris (CWD) data table is designed to store information on plot-level estimates of CWD volume (calculated using line intersect sampling), aggregated by species and decay classes. It contains the following data fields:

- **composed_site_id:** Unique site identifier.
- **spi_id:** Unique plot identifier in a sampling plot inventory.
- **lpi_id:** Unique plot identifier for a large plot inventory.
- **inventory_year:** The year the tree record was collected.
- **species_code:** A code representing the species of the tree.
- **decay:** A classification of wood decay class.
- **volume:** CWD wood volume on the plot (m³)

4.2.6. Metadata Table

This table evolves based on the metadata collected from the data-providers. It stores descriptive metadata for the various tables and fields in the database.

- **record_id**: A unique, auto-incrementing identifier for each metadata record.
- **institute**: The institution or organization providing the metadata.
- **table_name**: The name of the table to which the metadata applies.
- **field_name**: The specific field in the table that this metadata refers to.
- **field_description**: A description of the field provided by the data-provider.
- **extended_attributes**: any extra information provided that does not fit in any of the predefined columns.
- **value_code**: Any code used by the provider that needs further explanation and metadata collection.
- **value_description**: Meaning of the code including all parameters needed to use the data.
- **other**: Additional metadata information, if required.

4.2.7. Species List Table

The Species List table contains taxonomic information on the botanical species surveyed within the forest plots. It includes the following data fields:

- **species_code**: A unique, 6-character code for each species.
- **genus**: The genus of the tree/woody individual.
- **species**: The species of the tree/woody individual.
- **variety**: The variety or subspecies, if applicable.
- **full_scientific**: The full scientific name of the tree species.
- **other**: Additional information or classification for the species, if needed.

5. Database Content

This section provides an overview of the current structure and volume of collected and harmonized data successfully implemented in the database, including basic summary statistics on participating institutes, countries, number of sites, plots and tree records (Table 1). As data collection continues, this section will remain dynamic, with updates expected weekly as the database is further populated.

Table 1. Summary table showing how many data entries are there for each data table as well as summary for number of participating countries and institutions.

Timestamp: 5.12.2024

Table	Number of records
Countries	5
Institutes/WILDCARD partners	5
Sites	136
Site_designs	310
Plots	8 404
Trees	473 521

As previously mentioned, the pilot dataset (D2.1), stored on a PostgreSQL server, is currently accessible only to WILDCARD project partners via **custom access credentials** (sensitive dataset). No public access is planned at this stage. For access control purposes, credentials must be created. Please contact Jan Zalesak (zalesak@vukoz.cz) for further assistance.



6. Gap analysis

6.1. Approach

6.1.1 Context

The data gap analysis evaluates the extent to which the available dataset covers the ecological gradients represented within the WILDCARD project (see [Milestone 1](#) “Delineation of relevant ecoregions and ecological gradients, connection to existing databases”). Ecological gradients are represented by combinations of categorical variables such as forest type categories, ecoregions and other relevant factors, which together encompass the full range of ecological constellations in European forests. These ecological constellations form the basis for WILDCARD WP2, which aims to assess the overall potential of proforestation for climate and biodiversity protection in Europe. A “data gap” is defined as the “missing representation of a relevant ecological constellation within the European forests”. The WILDCARD Metadata Table serves as input for a script-based gap-analysis (R 4.4.1), which calculates the extent to which the available dataset covers the full range of ecological constellations (see attachments). Based on the results, strategies for gap filling are identified and discussed.

6.1.2 Methodology

The ecological gradients are represented by a nested system of categorical variables. This approach was chosen because it simplifies the process of defining, identifying and communicating data gaps, making it more transparent. However, when fitting statistical models, continuous variables will be used wherever possible and appropriate. To capture the diversity of ecological constellations, we apply the following surrogate variables:

- The six selected biogeographical regions of Europe (EEA, 2016), also called **ecoregions** (see [Milestone 1](#)).
- The twelve EEA **forest type categories** (FTC) (EEA, 2007).
- The combinations of **soil water** regime (3 classes) and **soil nutrient** status (2 classes), which are ecologically relevant for each FTC. For example, dry conditions in mire and swamp forests, or eutric conditions in acidophilous oak/oak birch forests, are excluded (Table 2). The distinct classes are defined as follows:
 - soil nutrient status: **dystric** with $\text{pH} \leq 5$, **eutric** with $\text{pH} > 5$
 - soil water regime: **dry, mesic, wet/waterlogged**
- The timeframe for rewilding is represented by four TSA-classes (0-30 years, 30-60 years, >60 years and primeval forests). This classification is aligned with Buchwald et al. (2005), and is consistent with the most recent and comprehensible evaluation of primary forests in Europe (Sabatini et al., 2018, 2021).

Not all combinations of classes exist within European forests: certain forest types are restricted to specific ecoregions, and most forest types occur only under particular soil water and nutrient regimes (as mentioned above). Therefore, a reference matrix was created to reflect all potentially relevant combinations, against which the representativity of the dataset is measured (Table 2).



Table 2. Forest Type Categories represented in the different ecoregions, their soil nutrient supply (dy = dystrophic, e = eutric) and soil water regime (d = dry, m = mesic, w = wet/waterlogged). If a class combination is ecologically not relevant, such as 'dry mires' or 'nutrient-rich acidophilous oak forest', then the corresponding soil property is not mentioned.

Forest Type Category	Ecoregions												Σ combinations	
	Alpine		Atlantic		Boreal		Continental		Mediterranean		Panonian			
	nutrient	water	nutrient	water	nutrient	water	nutrient	water	nutrient	water	nutrient	water		
<i>Boreal</i>	dy	d, m, w			dy	d, m, w								6
<i>Hemiboreal</i>			dy	d, m, w	dy	d, m, w	dy	d, m, w						9
<i>Alpine coniferous</i>	dy, e	d, m, w												6
<i>Acidophilous oak/oak-birch</i>			dy	d, m, w	dy	d, m, w	dy	d, m, w	dy	d, m, w	dy	d, m, w		15
<i>Mesophytic deciduous</i>	dy, e	d, m	dy, e	d, m	dy, e	d, m	dy, e	d, m	dy, e	d, m	dy, e	d, m		24
<i>Beech</i>			dy, e	d, m	dy, e	d, m	dy, e	d, m	dy, e	d, m	dy, e	d, m		20
<i>Mountainous beech</i>	dy, e	d, m					dy, e	d, m	dy, e	d, m				12
<i>Thermophilous deciduous</i>	dy, e	d, m	dy, e	d, m			dy, e	d, m	dy, e	d, m	dy, e	d, m		20
<i>Broadleaved evergreen</i>									dy, e	d, m, w				6
<i>Coniferous forest Mediterranean, ...</i>									dy, e	d, m, w				6
<i>Mire and swamp</i>	dy, e	w	dy, e	w	dy, e	w	dy, e	w	dy, e	w	dy, e	w		12
<i>Floodplain</i>			e	w	e	w	e	w	dy, e	w	dy, e	w		7
Σ combinations	23		21		20		25		35		19			143

Table 2 shows 143 combinations of ecoregion, forest type, and soil type that are considered ecologically relevant. When combined with the four classes of TSA, this results in **572 relevant combinations in total**. However, the WILDCARD Metadata Table includes combinations that do not match the predefined ecologically relevant categories. For instance, spruce-dominated stands classified as boreal forests in the Continental or Atlantic ecoregion could be reclassified as “Hemiboreal forest and nemoral coniferous and mixed broadleaved-coniferous forest”, which are also present in the nemoral zone of Europe (Atlantic and Continental ecoregion). If any combinations persist that are not included in the reference matrix and it is confirmed that the classification is correct, these datasets are still incorporated into the statistical models.

The gap analysis is based on all sites (records) for which the necessary metadata – such as ecoregion, forest type category, soil type, and TSA - are complete. This includes all sites provided by the WILDCARD partners as well as a number of additional sites from external partners, such as members of the EuFoRla network (who are not WILDCARD partners).

The gap analysis is conducted at the level of distinct ecological combinations, which are defined by a combination of forest type category, soil type, ecoregion, and TSA class. For example, 'mountainous beech forest with mesic soil water regime and nutrient rich soil in the continental region with TSA 30-60 years'. The analysis quantifies the **proportion of ecological combinations for which data is available from at least one site, compared to the total number of ecologically relevant combinations** as shown in Table 2.

This analysis can be conducted at any hierarchical level within the dataset, such as for the entire dataset, at individual class levels (e.g. forest type categories, ecoregions, or TSA classes), or for combinations of these levels (e.g. Beech forests in the Continental ecoregion).

An example is provided in Figure 2. The resulting value represents the percentage of relevant ecological combinations covered by our dataset – hence the term ‘Coverage of ecological gradient’.

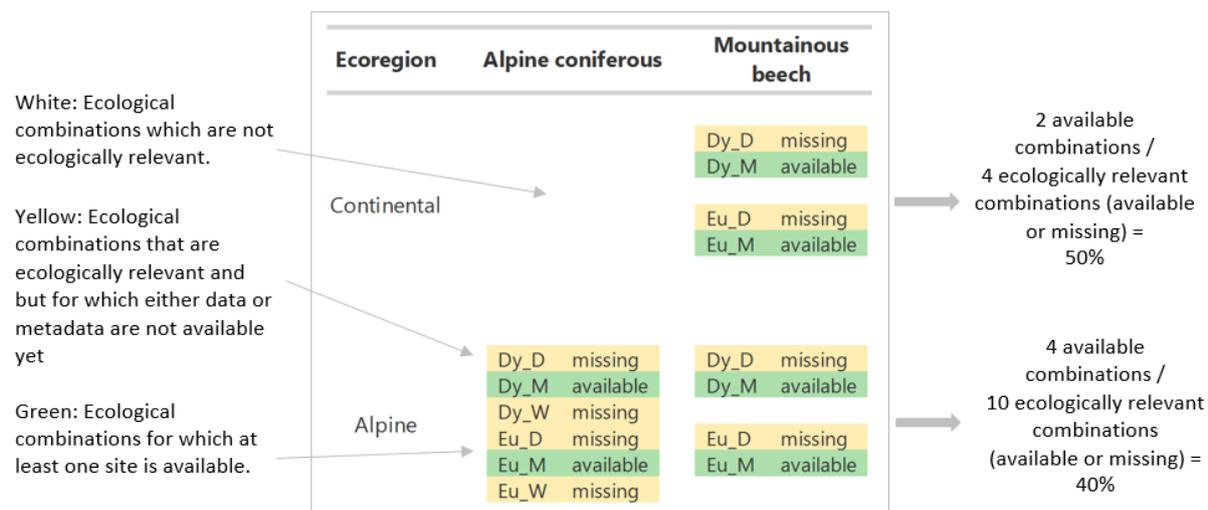


Figure 2. Example - calculating the coverage of the ecological gradient, here within the Continental and the Alpine ecoregion with two forest type categories each, with TSA 'Primeval'. Letter codes s. Table 2 (e. g. Dy_W = dystric waterlogged, ...)

6.2. Results of the gap analysis

In this section, we present the findings of the gap analysis for the three categories: TSA, forest type category, and ecoregion.

TSA is crucial for our research question. Overall, the data coverage is limited, with only a few of the ecologically relevant combinations represented (Figure 3). However, it is important to highlight that for certain common combinations of forest type x ecoregion, the TSA gradient is relatively well covered. This provides a strong foundation for statistical modelling. For more detailed information, please refer to Tables 4-7 in the Annex.

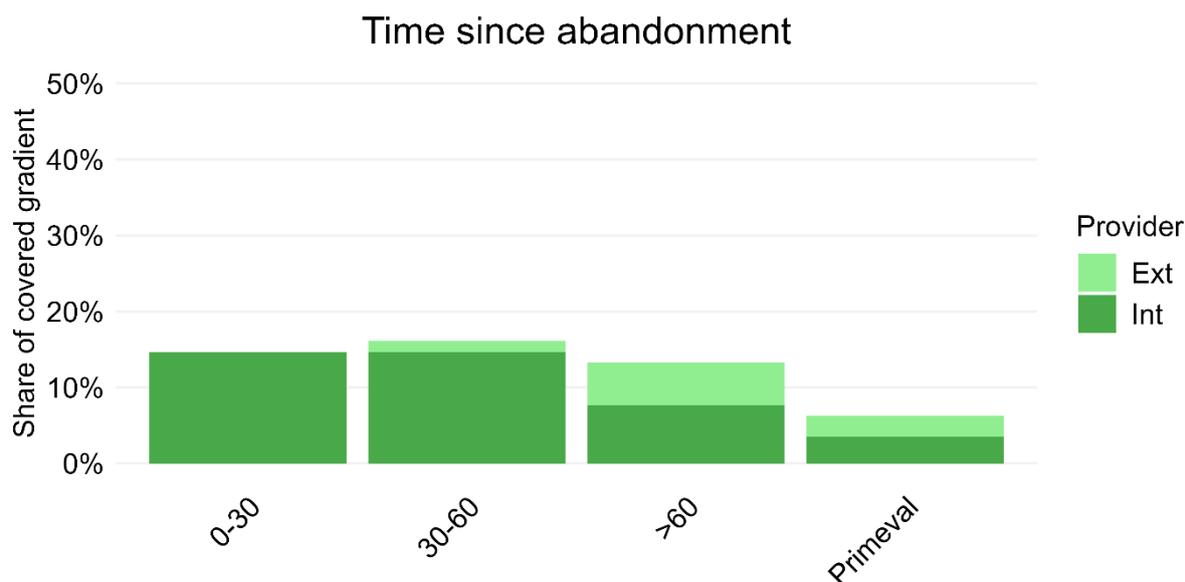


Figure 3. Share of covered ecological constellations for different 'time since abandonment (TSA)' classes (status as of 3 December 2024). The coverage is calculated as the 'proportion of distinct ecological combinations for which data is available from at least one site, compared to all ecologically relevant combinations'. The maximum possible coverage value is 100%. TSA is specified in years. Data-providers include both WILDCARD partners and external ones, primarily other EuFoRla members.

The coverage of different forest type categories and ecoregions is less uniform than the TSA classes (Figure 4 and Figure 5). For alpine coniferous forests and mountainous beech forests, more than a quarter of the ecologically relevant combinations are represented by at least one available site. However, the Boreal and Mediterranean forest type categories are scarcely represented in the dataset (Figure 4).

In terms of ecoregions, the dataset shows relatively good coverage of the Continental ecoregion, with just under half of relevant ecological combinations represented. However, there are significant gaps in the Alpine and Atlantic ecoregions. Mediterranean, Boreal and Pannonian ecoregions remain largely underrepresented in the current dataset (Figure 5).

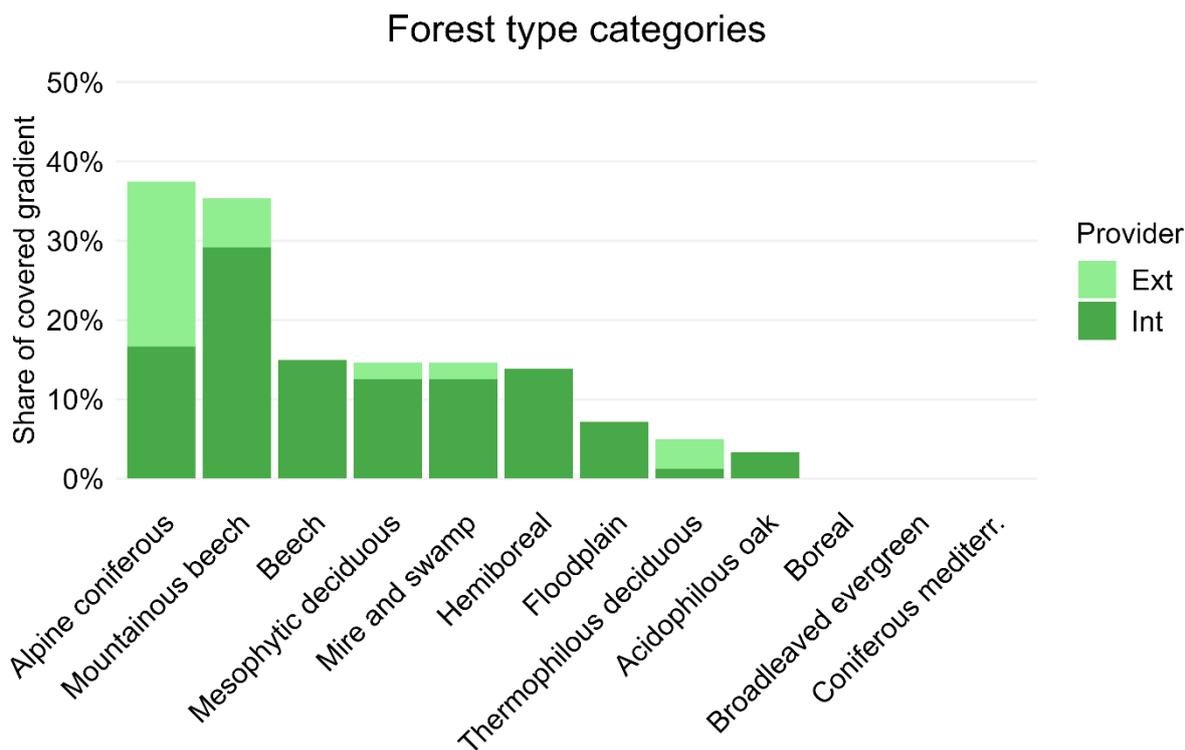


Figure 4. Share of covered ecological gradients for different forest type categories (Status as of 3 December 2024). The covered ecological gradient is calculated as the 'proportion of distinct ecological combinations for which data is available from at least one site, compared to all ecologically relevant combinations'. The maximum value is 100%. Forest type categories are based on the definitions provided by the European Environment Agency (2007). Data-providers are either internal (i.e. WILDCARD partners) or external partners (mostly other EuFoRla members).

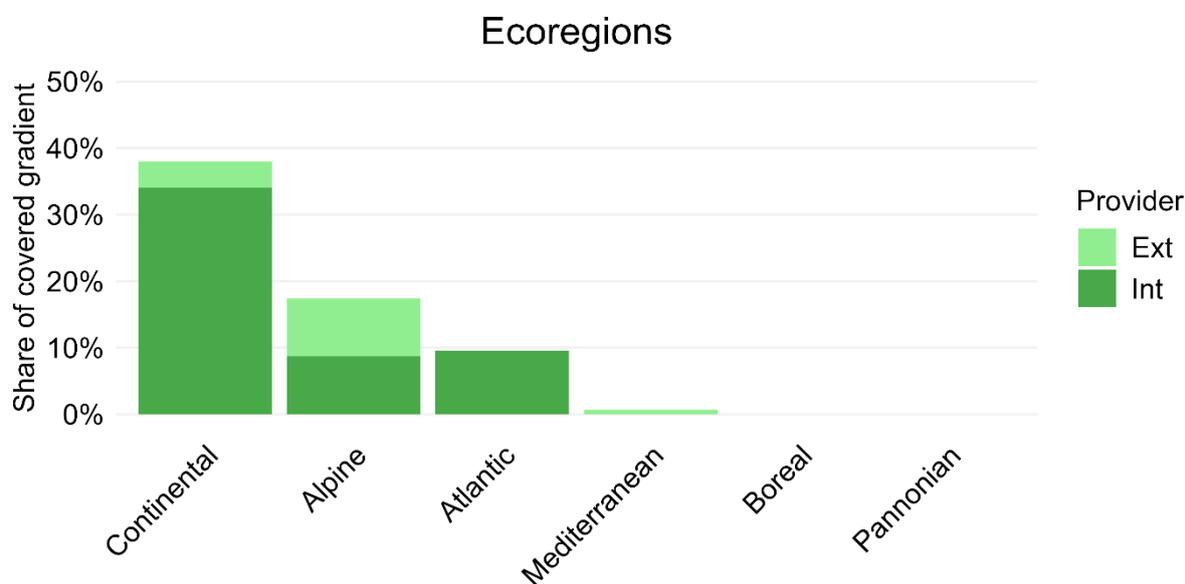


Figure 5. Share of covered ecological gradients for different selected ecoregions (status as of 3 December 2024). Covered ecological gradient is calculated as the 'proportion of distinct ecological combinations for which data is available from at least one site, compared to all ecologically relevant combinations'. The maximum value is 100%. Ecoregions follow the definition provided by the European Environment Agency (2016), as outlined in Milestone 1. Data-providers are either internal (i.e. WILDCARD partner), or external partners (mostly other EuFoRla members).

6.3. Map of readily available sites

Most sites are located in Central Europe, which explains the uneven coverage of ecoregions and forest type categories (Figure 6).

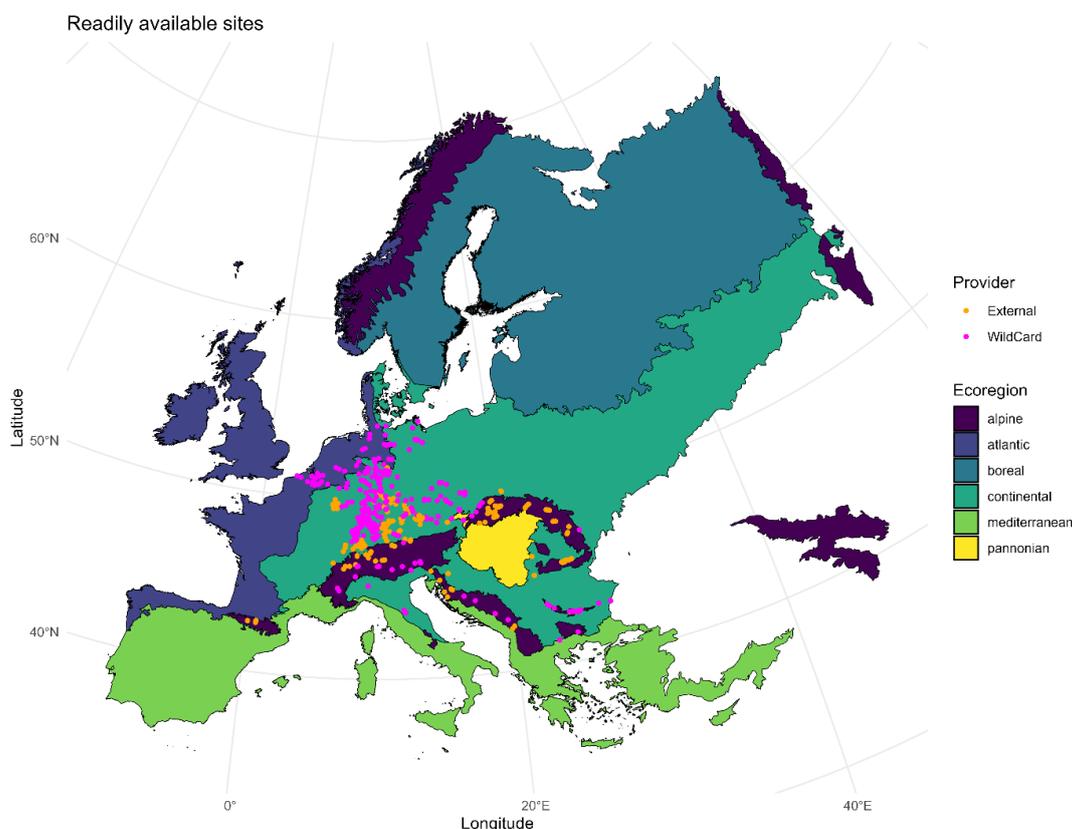


Figure 6. Locations of currently available sites, differentiated between those provided by WILDCARD partners (pink) and those provided by external partners (orange).

6.4. Gap filling steps

Currently, the dataset includes all the sites provided by the WILDCARD project partners, along with some additional sites contributed by external organisations that are members of EuFoRIA. Furthermore, WILDCARD WP2 has established contacts with colleagues in Sweden who have expressed interest in providing additional data. As a result, more external sites from other institutions will be incorporated before June 2025 (the due date for Milestone 5 “Collection and harmonisation of existing aboveground data from established networks completed”). We anticipate that this will significantly help in addressing the identified gaps (Table 3). However, we recognize that statistical modelling will still need to account for the persistence of certain gaps in the data.

Table 3. Progress in completing the WILDCARD Metadata Table. We distinguish sites from potentially interested external partner organisations (for full names, see Annex 1.1), sites which are currently being added ('in progress') and sites that are already available, mostly provided by WILDCARD partners.

Provider	Institute	interested	in progress	available
<i>WILDCARD Partner</i>	<i>NWFVA</i>	0	0	142
	<i>FVA-BW</i>	0	0	110
	<i>VUKOZ</i>	0	0	41
	<i>INBO</i>	0	0	25
	<i>DISAFA - UNITO</i>	0	1	19
	<i>IBER-BAS</i>	0	0	17
	<i>UNIUD</i>	0	0	4
	<i>USV</i>	0	0	1
	<i>WR</i>	0	123	
<i>External partner organisations</i>	<i>CULS</i>	0	2	71
	<i>LWF</i>	0	0	71
	<i>WSL</i>	0	8	48
	<i>BFNP</i>	0	0	26
	<i>ANF</i>	0	0	10
	<i>UNITBV</i>	0	0	3
	<i>INCDS M Dracea</i>	0	0	2
	<i>UL</i>	0	14	0
	<i>URK</i>	7	3	0
	<i>WuH NRW</i>	54	0	0
	<i>TUZVO</i>	25	0	0
	<i>INRAE</i>	24	0	0
	<i>UNITUS</i>	11	0	0
	<i>UNIPD</i>	8	0	0
	<i>BGD-NP</i>	5	0	0
	<i>UEF</i>	5	0	0
	<i>WULS</i>	4	0	0
	<i>ICIFOR, INIA-CSIC</i>	3	0	0
<i>NPS</i>	2	0	0	
<i>SLU</i>	1	0	0	

The impact of data gaps would be less significant if they were distributed randomly (MCAR = missing completely at random). However, this is not the case: the Pannonian, Boreal and Mediterranean ecoregions are significantly underrepresented in our dataset, which especially limits the coverage of climate gradient and forest type categories. As a result, for half of the forest type categories, there is little to no data available. These include thermophilous deciduous, acidophilous oak/oak-birch, broadleaved evergreen, coniferous forest of Mediterranean regions, floodplain, and mire and swamp forests.

The Pannonian, Boreal and Mediterranean ecoregions are significantly underrepresented in our dataset, which particularly limits the coverage of climate gradient and forest type categories. The only two forest types that are reasonably well covered in terms of TSA gradients are alpine coniferous forests and mountainous beech forests in the Alpine and

Continental ecoregions. The data available for these two regions and forest types allow us to make progress towards the central research question: the relationship between TSA and aboveground carbon and biodiversity. To make reliable pan-European projections, which include the full range of forest types across all European ecoregions, we will need to adopt a hybrid approach that includes incorporating additional data (see Table 3), applying creative statistical solutions and validating results with relevant literature.

To predict above-ground C storage and sequestration, we plan to fit generalised additive mixed models (Ravindra et. al, 2019). In these models, climate, soil nutrient and soil water content will be represented as continuous variables rather than categorical ones. This approach will enable cautious extrapolation to underrepresented ranges, particularly in climate variables. As long as the range of data is sufficiently broad, we can fill the gaps in between using interpolations. However, extrapolations will need to be interpreted with caution. To assess the robustness of models, we intend to perform a sensitivity analysis e.g., (leverage analysis, leaving out available constellations) to evaluate how the model's predictions vary based on the data gaps.

In addition, we will consult the extensive body of literature on biomass storage in forest ecosystems with three key objectives:

1. Identifying additional data sources: we will look for further datasets covering underrepresented forest types and ecoregions.
2. Improving understanding of ecological drivers: we aim to deepen our understanding of the key ecological drivers influencing aboveground C storage in forests. This will help us develop additional methods to address data gaps in a more ecologically meaningful way.
3. Validating our findings: by comparing our dataset with findings from the literature, we will be able to assess the plausibility of our results and ensure that they align with existing knowledge in the field.

For forest type categories as a categorical variable, we may need to adopt a different approach. One possible solution is to transform FTCs into continuous variables representing the proportions of the main tree species within each forest type. These variables could then be integrated into the modelling process.

In addition to our dataset of abandoned forests, we also have access to a dataset from 'ICP Forests' (Galluzzi et. al, 2019), which contains data on managed forests. We are currently exploring the potential of this dataset as a baseline for comparing C and biodiversity trajectories following the cessation of forest management.

7. References

- Buchwald, E. (2005). A hierarchical terminology for more or less natural forests in relation to sustainable management and biodiversity conservation. In *Proceedings: Third expert meeting on harmonizing forest-related definitions for use by various stakeholders* (pp. 17-19).
- EEA, 2007. European forest types - Categories and types for sustainable forest management reporting and policy. In: European Environment Agency (Ed.), Technical Report 9, p. 111.
- EEA, 2016. Biogeographic regions, Europe 2016, data set version 1, available from European Environmental Agency. <https://doi.org/10.2909/960998c1-1870-4e82-8051-6485205ebbac>
- Galluzzi, M., Giannetti, F., Puletti, N., Canullo, R., Rocchini, D., Bastrup-Birk, A., & Chirici, G. (2019). A plot-level exploratory analysis of European forest based on the results from the BioSoil Forest Biodiversity project. *European Journal of Forest Research*, 138, 831-845.
- Kun, Z., DellaSala, D., Keith, H., Kormos, C., Mercer, B., Moomaw, W.R., Wiezik, M., 2020. Recognizing the importance of unmanaged forests to mitigate climate change. *GCB Bioenergy* 12, 1034–1035. <https://doi.org/10.1111/gcbb.12714>
- Moomaw, W.R., Masino, S.A., Faison, E.K., 2019. Intact Forests in the United States: Proforestation Mitigates Climate Change and Serves the Greatest Good. *Front. For. Glob. Chang.* 2, 1–10. <https://doi.org/10.3389/ffgc.2019.00027>
- Ravindra, K., Rattan, P., Mor, S., & Aggarwal, A. N. (2019). Generalized additive models: Building evidence of air pollution, climate change and human health. *Environment international*, 132, 104987.
- Sabatini, F. M., Burrascano, S., Keeton, W. S., Levers, C., Lindner, M., Pötzschner, F., ... & Kuemmerle, T. (2018). Where are Europe's last primary forests?. *Diversity and distributions*, 24(10), 1426-1439.
- Sabatini, F. M., Bluhm, H., Kun, Z., Aksenov, D., Atauri, J. A., Buchwald, E., ... & Kuemmerle, T. (2021). European primary forest database v2. 0. *Scientific data*, 8(1), 220.

Annex

1.1. Abbreviations and long names of involved institutes

Status	Institute	Name
<i>WILDCARD Partner</i>	<i>WR</i>	Wageningen University and Research
	<i>NW-FVA</i>	Northwest German Forest Research Institute
	<i>FVA-BW</i>	Forest Research Institute Baden-Württemberg, Germany
	<i>VUKOZ</i>	Silva Tarouca Research Institute, Pruhonice, Czech Republic
	<i>EVINBO</i>	Research Institute for Nature and Forest, Belgium
	<i>DISAFA - UNITO</i>	Università degli Studi di Torino, Italy
	<i>IBER-BAS</i>	Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Bulgaria
	<i>UNIUD</i>	University of Udine, Italy
	<i>USV</i>	University of Suceava, Romania
<i>External partner organisations</i>	<i>CULS</i>	Czech University of Life Sciences Prague, Czech Republic
	<i>LWF</i>	Bavarian State Institute of Forestry, Germany
	<i>WSL</i>	Swiss Federal Institute for Forest, Snow and Landscape Research, Switzerland
	<i>BFNP</i>	Bavarian Forest National Park, Germany
	<i>ANF</i>	Administration de la nature et des forests, Luxembourg
	<i>UNITBV</i>	Transilvania University of Braşov, Romania
	<i>INCDS M Dracea</i>	National Institute for Research and Development in Forestry "Marin Dracea", Romania
	<i>UL</i>	Romania University of Ljubljana, Slovenia
	<i>URK</i>	University of Agriculture in Kraków, Poland
	<i>WUR</i>	Wageningen University & Research, Netherlands
	<i>WuH NRW</i>	Wald und Holz NRW (State Forest Enterprise, Northrhine-Westfalia), Germany
	<i>TUZVO</i>	Technical University in Zvolen, Slovakia
	<i>INRAE</i>	National Research Institute for Agriculture, Food and Environment, France
	<i>UNITUS</i>	University of Tuscia, Italy
	<i>UNIPD</i>	University of Padua, Italy
	<i>BGD-NP</i>	Berchtesgaden Forest National Park
	<i>UEF</i>	University of Eastern Finland
	<i>WULS</i>	Warsaw University of Life Sciences, Poland
	<i>ICIFOR, INIA-CSIC</i>	Institute of Forest Science (ICIFOR) INIA-CSIC, Spain
	<i>NPS</i>	National Park Sumava, Czech Republic
<i>SLU</i>	Swedish University of Agricultural Sciences, Sweden	



1.2. Gap overview for TSA ‘Primeval’

Available and missing ecological combinations in primeval forests, based on the readily available dataset.

Ecoregion	Alpine coniferous	Mountainous beech	Beech	Mesophytic deciduous	Mire and swamp	Hemiboreal	Floodplain	Thermophilous deciduous	Acidophilous oak	Broadleaved evergreen	Coniferous mediterr.	Boreal
Continental		Dy_D missing Dy_M available	Dy_D missing Dy_M available	Dy_D missing Dy_M missing		Dy_D missing Dy_M missing Dy_W missing		Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing			
		Eu_D missing Eu_M available	Eu_D missing Eu_M missing	Eu_D missing Eu_M available		Eu_W missing	Eu_W missing	Eu_D missing Eu_M missing				
Alpine	Dy_D missing Dy_M available Dy_W missing Eu_D missing Eu_M available Eu_W missing	Dy_D missing Dy_M available		Dy_D missing Dy_M missing		Dy_W missing		Dy_D missing Dy_M missing				Dy_D missing Dy_M missing Dy_W missing
		Eu_D missing Eu_M available		Eu_D missing Eu_M missing		Eu_W missing		Eu_D missing Eu_M missing				
Atlantic			Dy_D missing Dy_M missing	Dy_D missing Dy_M missing		Dy_D missing Dy_M missing Dy_W missing		Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing			
			Eu_D missing Eu_M missing	Eu_D missing Eu_M missing		Eu_W missing	Eu_W missing	Eu_D missing Eu_M missing				
Mediterranean		Dy_D missing Dy_M missing	Dy_D missing Dy_M missing	Dy_D missing Dy_M missing				Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing	
		Eu_D missing Eu_M available	Eu_D missing Eu_M missing	Eu_D missing Eu_M missing		Dy_W missing	Dy_W missing	Eu_D missing Eu_M missing		Eu_D missing Eu_M missing Eu_W missing	Eu_D missing Eu_M missing Eu_W missing	
Boreal			Dy_D missing Dy_M missing	Dy_D missing Dy_M missing		Dy_D missing Dy_M missing Dy_W missing			Dy_D missing Dy_M missing Dy_W missing			Dy_D missing Dy_M missing Dy_W missing
			Eu_D missing Eu_M missing	Eu_D missing Eu_M missing		Eu_W missing	Eu_W missing					
Pannonian			Dy_D missing Dy_M missing	Dy_D missing Dy_M missing				Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing			
			Eu_D missing Eu_M missing	Eu_D missing Eu_M missing		Dy_W missing	Dy_W missing	Eu_D missing Eu_M missing				

Available combinations are highlighted in green, meaning that at least one site is available for that ecological combination. Missing combinations are highlighted in yellow. Ecologically irrelevant combinations are not shown (white).

Abbreviations for the soil types: Dy_D = dystric/dry; Dy_M = dystric/mesic; Dy_W = dystric/wet; Eu_D = eutric/dry; Eu_M = eutric/mesic; Eu_W = eutric/wet.

See Figure 2 above for an explanation of how the coverage of ecological gradients is calculated on this basis.

1.3. Gap overview for TSA '>60'

Available and missing ecological combinations in forests that have been abandoned over 60 years ago, based on the readily available dataset.

Ecoregion	Alpine coniferous	Mountainous beech	Beech	Mesophytic deciduous	Mire and swamp	Hemiboreal	Floodplain	Thermophilous deciduous	Acidophilous oak	Broadleaved evergreen	Coniferous mediterr.	Boreal
Continental		Dy_D missing Dy_M available	Dy_D missing Dy_M available	Dy_D available Dy_M missing		Dy_D missing Dy_M missing Dy_W missing		Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing			
		Eu_D missing Eu_M available	Eu_D missing Eu_M available	Eu_D available Eu_M available	Dy_W available Eu_W available		Eu_W available		Eu_D available Eu_M available			
Alpine	Dy_D missing Dy_M available Dy_W missing Eu_D available Eu_M available Eu_W available	Dy_D missing Dy_M available Eu_D missing Eu_M available		Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D available Eu_M missing				Dy_D missing Dy_M missing Dy_W missing
			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing			
Atlantic			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_D missing Dy_M missing Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing			
Mediterranean		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing Eu_D missing Eu_M missing Eu_W missing	Dy_D missing Dy_M missing Dy_W missing Eu_D missing Eu_M missing Eu_W missing	
Boreal			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_D missing Dy_M missing Dy_W missing Eu_W missing			Dy_D missing Dy_M missing Dy_W missing			Dy_D missing Dy_M missing Dy_W missing
Pannonian			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing			

Available combinations are highlighted in green, meaning that at least one site is available for that ecological combination. Missing combinations are highlighted in yellow. Ecologically irrelevant combinations are not shown (white).

Abbreviations for the soil types: Dy_D = dystric/dry; Dy_M = dystric/mesic; Dy_W = dystric/wet; Eu_D = eutric/dry; Eu_M = eutric/mesic; Eu_W = eutric/wet.

See Figure 2 above for an explanation of how the coverage of ecological gradients is calculated on this basis.

1.4. Gap overview for TSA '30-60'

Available and missing ecological combinations in forests that have been abandoned between 30 and 60 years ago, based on the readily available dataset.

Ecoregion	Alpine coniferous	Mountainous beech	Beech	Mesophytic deciduous	Mire and swamp	Hemiboreal	Floodplain	Thermophilous deciduous	Acidophilous oak	Broadleaved evergreen	Coniferous mediterr.	Boreal
Continental		Dy_D missing Dy_M available Eu_D available Eu_M available	Dy_D missing Dy_M available Eu_D available Eu_M available	Dy_D available Dy_M available Eu_D available Eu_M available	 Dy_W available Eu_W missing	Dy_D available Dy_M available Dy_W missing	 Eu_D available Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing			
Alpine	Dy_D missing Dy_M missing Dy_W missing Eu_D missing Eu_M available Eu_W available	Dy_D missing Dy_M missing Eu_D missing Eu_M available		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	 Dy_W missing Eu_W missing			Dy_D missing Dy_M missing Eu_D missing Eu_M missing				Dy_D missing Dy_M missing Dy_W missing
Atlantic			Dy_D missing Dy_M available Eu_D missing Eu_M available	Dy_D missing Dy_M missing Eu_D missing Eu_M available	 Dy_W missing Eu_W available	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing				
Mediterranean		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	 Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing	
Boreal			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	 Dy_W missing Eu_W missing	Dy_D missing Dy_M missing Dy_W missing		Dy_D missing Dy_M missing Dy_W missing				Dy_D missing Dy_M missing Dy_W missing
Pannonian			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	 Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Dy_W missing				

Available combinations are highlighted in green, meaning that at least one site is available for that ecological combination. Missing combinations are highlighted in yellow. Ecologically irrelevant combinations are not shown (white).

Abbreviations for the soil types: Dy_D = dystic/dry; Dy_M = dystic/mesic; Dy_W = dystic/wet; Eu_D = eutric/dry; Eu_M = eutric/mesic; Eu_W = eutric/wet.

See Figure 2 above for an explanation of how the coverage of ecological gradients is calculated on this basis.

1.5. Gap overview for TSA '0-30'

Available and missing ecological combinations in forests that have been abandoned less than 30 years ago, based on the readily available dataset.

Ecoregion	Alpine coniferous	Mountainous beech	Beech	Mesophytic deciduous	Mire and swamp	Hemiboreal	Floodplain	Thermophilous deciduous	Acidophilous oak	Broadleaved evergreen	Coniferous mediterr.	Boreal
Continental		Dy_D available Dy_M available	Dy_D missing Dy_M available	Dy_D available Dy_M available		Dy_D available Dy_M available Dy_W available		Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing			
		Eu_D missing Eu_M available	Eu_D missing Eu_M available	Eu_D available Eu_M available			Eu_W missing		Eu_D missing Eu_M missing			
Alpine	Dy_D missing Dy_M missing Dy_W missing Eu_D missing Eu_M available Eu_W missing	Dy_D missing Dy_M missing Eu_D missing Eu_M available		Dy_D missing Dy_M missing Eu_D missing Eu_M missing			Dy_W missing Eu_W missing		Dy_D missing Dy_M missing Eu_D missing Eu_M missing			Dy_D missing Dy_M missing Dy_W missing
Atlantic			Dy_D missing Dy_M available Eu_D missing Eu_M available	Dy_D missing Dy_M missing Eu_D missing Eu_M available		Dy_D missing Dy_M missing Dy_W missing	Dy_W available	Dy_D missing Dy_M missing	Dy_D missing Dy_M available Dy_W missing			
					Eu_W available		Eu_W missing					
Mediterranean		Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing			Dy_W missing	Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing	Dy_D missing Dy_M missing Dy_W missing Eu_D missing Eu_M missing Eu_W missing	Dy_D missing Dy_M missing Dy_W missing Eu_D missing Eu_M missing Eu_W missing	
					Eu_W missing		Eu_W missing					
Boreal			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing		Dy_D missing Dy_M missing Dy_W missing			Dy_D missing Dy_M missing Dy_W missing			Dy_D missing Dy_M missing Dy_W missing
					Eu_W missing		Eu_W missing					
Pannonian			Dy_D missing Dy_M missing Eu_D missing Eu_M missing	Dy_D missing Dy_M missing Eu_D missing Eu_M missing				Dy_D missing Dy_M missing	Dy_D missing Dy_M missing Dy_W missing			
					Eu_W missing		Eu_W missing		Eu_D missing Eu_M missing			

Available combinations are highlighted in green, meaning that at least one site is available for that ecological combination. Missing combinations are highlighted in yellow. Ecologically irrelevant combinations are not shown (white).

Abbreviations for the soil types: Dy_D = dystric/dry; Dy_M = dystric/mesic; Dy_W = dystric/wet; Eu_D = eutric/dry; Eu_M = eutric/mesic; Eu_W = eutric/wet.

See Figure 2 above for an explanation of how the coverage of ecological gradients is calculated on this basis.