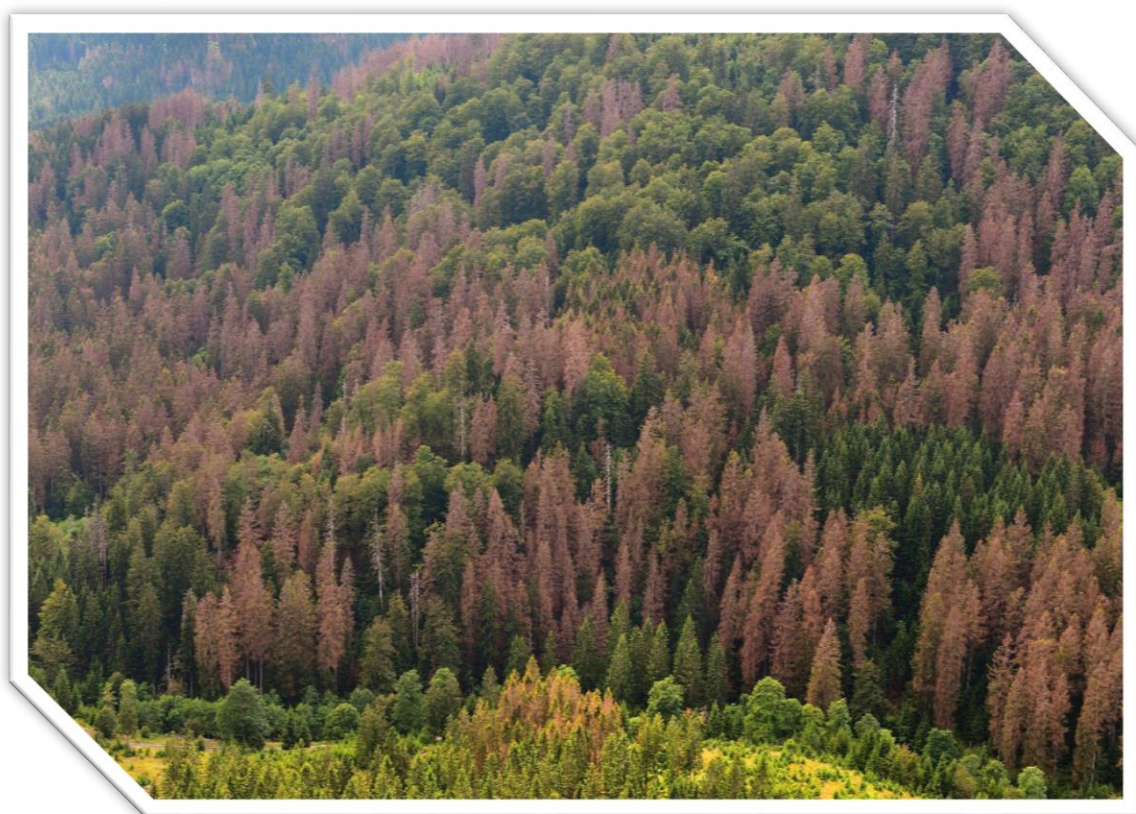


Copernicus Land Monitoring Service – High Resolution Layer - Vegetated Land Cover Characteristics

*SPECIFIC CONTRACT NO 3506/R0-COPERNCA/EEA.60009 IMPLEMENTING
FRAMEWORK SERVICE CONTRACT NO. EEA/DIS/R0/21/013*

D1.4 FEASIBILITY STUDY ON THE EXTENSION OF THE HRL TREE & COVER AND FOREST PORTFOLIO TOWARDS A EUROPEAN FOREST TYPE MAP



Date: 2025-09-08

Doc. Version : 1.1

Content ID: D1.4 EFT Feasibility

Table of contents

Table of contents	2
List of Figures	3
List of Tables.....	4
Document Control Information:.....	5
Document Approver(s) and Reviewer(s):	5
Document history:	5
1 Introduction	6
2 The European Forest Type nomenclature	6
2.1 Required tree-type information.....	8
2.2 Harmonization of current HRL Tree Cover & Forests nomenclature and the EFT nomenclature	12
3 EO-based high-resolution tree-type mapping at national and continental scale – state-of-the-art	12
3.1 Europe – Bonannella et al., 2022	12
3.2 Canada – Hermosilla et al., 2022, 2024	13
3.3 Germany - Blickensdörfer et al., 2024	14
3.4 Netherlands - Francini et al., 2024.....	17
3.5 Poland - Grabska-Szwagrzyk et al., 2024	18
3.6 France - Mouret et al., 2024	21
3.7 Summary	21
4 Potential reference data for training and validation	22
4.1 Map data	22
4.1.1 Austria	22
4.1.2 Czech Republic	22
4.1.3 France.....	22
4.1.4 Spain.....	25
4.1.5 Sweden	25
4.1.6 Others.....	25
4.2 Point and plot data.....	26
4.2.1 National Forest Inventories	26

4.2.2	ICP Forest	27
4.2.3	LUCAS	31
4.2.4	GBIF	33
4.2.5	Further	37
4.2.5.1	EU Forest	37
4.2.5.2	EUFGIS	37
4.2.5.3	Conifers of the World, Atlas Florae Europaeae	37
4.2.5.4	Curated benchmark datasets	38
5	Recommendations	38
	• Thematic content	38
	• Spatial resolution	39
	• Target accuracy	39
	• EO and auxiliary input data	39
	• Reference Input data	40
	• Update cycle	40
	List of Abbreviations & Acronyms	41
	References	43

List of Figures

Figure 2-1: Hierarchical attribution of species (Level 3) to potential tree type (Level 2) and leaf type classes (Level 1)	11
Figure 3-1: Comparison of the HRL Dominant Leaf Type 2021 with two species maps (Realized Distribution Probabilities - RDB) for 2018-2020 elaborated by Bonannella et al. 2022. The comparison illustrates the lack of spatial detail in the modelled RDB maps.	13
Figure 3-2: Relative predictor variable importance ranking across Canada's forested ecosystem tiles sorted by median rank value (Hermosilla et al., 2022).	14
Figure 3-3: Accuracies for pure-stands accuracy vs all-stands accuracy from (Blickensdörfer et al. 2024)	15
Figure 3-4: Visual comparison of the tree type map from Blickensdörfer et al., 2024 and the HRL DLT 2021 for an area in the Spessart (LAEA: 4280670m, 3006870m)	17
Figure 3-5: Tree type map for the Netherlands as presented by Francini et al., 2024.	18
Figure 3-6: F-11 score for 16 tree types using proportional and disproportional sample allocation.	19
Figure 3-7: Visual comparison of the tree type map from Grabska-Szwagrzyk et al., 2024 and the HRL DLT 2021 for an area near Kielce (LAEA: 5077235m, 3146295m)	20
Figure 4-1: Visual comparison of the forest outlines of the BD Forêt® V2 and the leaf type mapping in the VLCC DLT product. While there is a strong resemblance of main spatial structures the generalization of the BD Forêt® V2 yields many stands that are assigned to mix classes. In many cases the assignment of the class (e.g. yellow ellipse = closed spruce and fir forests) omits mixture components as well.	24

Figure 4-2: Distribution of ICP Forests Level 1 plots with different levels of available information.	29
Figure 4-3: Distribution of dominant genera in the ICP Forests Level 1 plots which have been updated since 2018 and are provided with second-level precision coordinates.	29
Figure 4-4: Distribution of ICP Forests Level 2 plots with different levels of available information.	30
Figure 4-5: Distribution of dominant genera in the ICP Forests Level 2 plots which have been updated since 2018 and are provided with second-level precision coordinates.	30
Figure 4-6: Spatial distribution of reference points for four genera extracted from LUCAS 2018.	31
Figure 4-7: Sample counts per genus for reference points extracted from LUCAS 2018.	32
Figure 4-8: Spatial distribution of reference points for two genera extracted from LUCAS 2022.	32
Figure 4-9: Sample counts per genus for reference points extracted from LUCAS 2022.	33
Figure 4-10: Histograms for the spatial uncertainty attribute and year of the observation for all records initially retrieved from the GBIF database.	34
Figure 4-11 Spatial distribution of reference points in the data extracted from GBIF after filtering and clipping to the extent of the EEA38 + UK.	35
Figure 4-12: Distribution of tree types in the data extracted from GBIF after filtering and clipping to the extent of the EEA38 + UK.	36
Figure 4-13 Example for spatial clustering GBIF records in urban areas. Berlin, Germany, (background Google Maps).	36
Figure 4-14: Spatial distribution of records and derived tree species richness from the EU-Forest dataset (figure from Mauri et al. 2017).	37
Figure 5-1 Working tree type legend adopted in this report.	39

List of Tables

Table 2-1: The 14 European Forest Type Categories and minimum required information on leaf and tree type information required as input.	7
Table 2-2: Overview of datasets used for the EFT modelling in (Giannetti et al. 2018).	8
Table 4-1: Overview of filtering steps applied to the tree species observations extracted from GBIF.	34

Document Control Information:

Document Title	D1.4 Feasibility study on the extension of the HRL Tree & Cover and Forest portfolio towards European Forest Type map
Project Title	Copernicus Land Monitoring Service – High Resolution Layer - Vegetated Land Cover Characteristics
Document Author	André Stumpf (GAF AG)
Project Owner	Luca Battistella (EEA)
Project Manager	Christian Siegert (GAF AG)
Document Code	D1.4 EFT Feasibility
Document Version	1.0
Distribution	Public
Date	2025-09-08

Document Approver(s) and Reviewer(s):

Name	Role	Action	Date
André Stumpf	Author	Initial version	2025-08-18
Stephanie Wegscheider	Reviewer	Review	2025-08-21

Document history:

Revision	Date	Created by	Short description of changes
1.0	2025-08-21	André Stumpf, Stephanie Wegscheider	Initial issue
1.1	2025-09-05	Luca Battistella, André Stumpf	Final review

1 Introduction

Since 2012 the Copernicus Land Monitoring Service has implemented regular updates of the maps on tree cover and forests for the extent EEA38 countries and the UK. This includes raster maps of the High-Resolution Layers - Dominant Leaf Type and the Tree Cover Density. Since the initial reference years, the methods and earth observation input data used for the generation of those products have been continuously improved. The focus thereby has been so far to improve accuracy and spatio-temporal resolution; as a result, the Dominant Leaf Type layer is today provided on an annual basis with 10-meter spatial resolution and omission / commission error rates that are generally below 10%. However, the currently available Dominant Leaf Type, provides still only a basic distinction between broadleaved and coniferous tree cover which is not sufficient for use cases that require more detailed thematic information on genus or even species levels.

In the frame of production project High-Resolution Layers Vegetated Land Cover Characteristics (Framework Service Contract EEA/DIS/RO/21/013) EEA has issued a consultancy request (Specific Contract 3506/RO-COPERNICA/EEA.60009) to assess the feasibility and requirements for expanding the nomenclature of the CLMS HRL Tree Cover and Forest with tree species information. The goal of providing additional tree species information should be to better support national and EU-level forest policy needs, enable the derivation of a forest typology that closely aligns with the European Forest Type classification (European Environment Agency 2007) and improve the thematic detail of the HRL Tree Cover & Forests.

2 The European Forest Type nomenclature

About 20 years ago an international consortium of experts has carried out a study to propose a user-friendly forest types classification to address shortcomings of existing forest classification systems for MCPFE (Ministerial Conference on the Protection of Forests in Europe) reporting and to improve the reporting on sustainable forest management in Europe. The results of this process are described in (European Environment Agency 2007) and includes the proposition of the European Forest Types (EFT) nomenclature. The EFT is a hierarchical nomenclature on two levels, with 14 so called categories that are further subdivided in 76 types. On the first level it contains the categories listed in Table 2-1 which are defined through a mix of variables in Appendix II — Classification keys in (European Environment Agency 2007) including:

- Leaf type (Coniferous, Broadleaved deciduous, Broadleaved evergreen, Mixed)
- Tree genus or species
- Hydrology (Wet, Mesic dry, Waterlogged, Riparian or alluvial, Dry or seasonally wet)
- Biogeographic region (Boreal, Boreo-nemoral, Atlantic or continental, Alpine, Mediterranean, Macaronesian or Anatolian, Vegetation belt)
- Thermophilous deciduous species = yes/no
- Oligotrophic = yes/no
- Land use / structure (i.e. Plantations)

Table 2-1: The 14 European Forest Type Categories and minimum required information on leaf and tree type information required as input.

European Forest Type Categories	Minimum required leaf type or tree type information
1. Boreal forest	Predominance of Coniferous
2. Hemiboreal forest and nemoral coniferous and mixed broadleaved-coniferous forest	Predominance of Coniferous or mix of Broadleaved deciduous and Coniferous
3. Alpine coniferous forest	Predominance of Coniferous
4. Acidophilous oak and oak-birch forest	Predominance of Quercus, presence of Betula
5. Mesophytic deciduous forest	Predominance of Broadleaved deciduous
6. Beech forest	Predominance of Fagus
7. Mountainous beech forest	Predominance of Fagus
8. Thermophilous deciduous forest	Predominance of Broadleaved deciduous
9. Broadleaved evergreen forest	Predominance of Broadleaved evergreen
10. Coniferous forests of the Mediterranean, Anatolian and Macaronesian regions	Predominance of Coniferous
11. Mire and swamp forest	Predominance of Broadleaved deciduous
12. Floodplain forest	Predominance of Broadleaved deciduous
13. Non-riverine alder, birch, or aspen forest	Predominance of Broadleaved deciduous, dominated by Alnus, Betula or Populus
14. Plantations and self-sown exotic forest	(Eucalyptus, Pinus, Picea, Picea, Pseudotsuga, Robinia, Prunus, Populus, Tsuga, Ailanthus)

The nomenclature is thus defined though a mix of environmental variables, not all of which can be assessed from EO data alone at the relevant accuracy and spatial resolution. Previous works (Giannetti et al. 2018) has laid out approaches to implement a spatially explicit mapping of the EFT nomenclature at 1km² spatial resolution by combining map data from different existing sources. An overview of the data sources used in (Giannetti et al. 2018) is provided in Table 2-2. However, most of the used input data sources are on rather coarse spatial scales, which can be suitable for variables that vary only over large distances (e.g. biogeographic regions), whereas in particular the usage of coarse bioclimatic and tree species maps still limits the capability to discriminate local environmental variabilities.

While some variables could potentially be derived already from more recent datasets with higher spatial resolution (e.g. HRL Water & Wetness 2018, EU-Hydro, HRL Water Cover Duration, Copernicus DEM 10m, E-OBS gridded dataset, HRL Dominant Leaf Type, see overview in Table 2-2) there is currently still a lack of a tree type¹ (genus or species-level) mapping with Pan-European coverage and high spatial resolution (10-100m). In this context it is also worth noting that The European Atlas of Forest Tree Species (European Commission: Joint Research Centre 2016) provides rather a probability of presence modelled based on occurrence records. This should be understood as a potential distribution of the species rather than an actual observed distribution which is not fully compatible with the HRL Tree Cover & Forests maps that are based on EO-data and updated on a yearly basis. A more recent publication (Giannetti and Zorzi 2023) proposed a refined approach to generate an EFT map at 100m spatial resolution, but still needed to rely on an up-sampled version of the European Atlas of Forest Tree Species.

Considering the lack of high-resolution geospatial information on tree species distribution the focus of this study has therefore been set on the **feasibility of a tree type map** that could extend the current HRL Tree Cover & Forests map portfolio and constitute an essential input dataset for

¹ In this report tree type is used as an umbrella term for nomenclatures that might often combine species, genus and leaf type level information.

the derivation of an EFT map. This includes in particular technical consideration being the state-of-the-art of EO-based tree type mapping (chapter 3) and the availability of suitable reference data (chapter 4). While user requirements have been assessed in a separate consultancy study² we still provide here an attempt to define a potential tree type nomenclature as a working hypothesis for the remainder of this study.

Table 2-2: Overview of datasets used for the EFT modelling in (Giannetti et al. 2018)

Dataset title	Source	Spatial resolution and used variable	Potential alternatives
European Atlas of Forest Tree Species	(European Commission: Joint Research Centre 2016)	1km ² , relative probability of presence for 39 tree species	Currently none at 10-100m resolution
Biogeographical regions dataset	(European Environment Agency 2016)	1:1 000 000 to 1: 10 000 000, Biogeographical boundaries obtained from the EU Member States and from the Emerald Network countries.	AgERA5 (Copernicus Climate Change Service 2020), E-OBS (Copernicus Climate Change Service 2020)
Bioclimatic Map of Europe	(Rivas-Martínez et al. 2004)	1:16 000 000, thermoclimatic belts of Europe in 5 regions, 9 subregions, 34 provinces, and 88 subprovinces	AgERA5 (Copernicus Climate Change Service 2020), E-OBS (Copernicus Climate Change Service 2020)
Natural Vegetation map of Europe	(Bohn 2000)	1:2 500 000 scale mapping of natural vegetation types	currently non
HRL Water and Wetness 2015	(European Environment Agency 2018)	20m raster map of wetness and water bodies	EU-Hydro Riparian Zones HRL Water Cover Duration
EU-DEM	(European Environment Agency 2011)	Digital Surface Model at 25m spatial resolution	Copernicus DEM

2.1 Required tree-type information

As a starting point we extracted a list of 104 common tree species from the chorological data compilation for the main European tree and shrub species (Caudullo et al. 2017). This list was complemented with further species listed in the matrix presented in (Pividori et al. 2016) if the

² <https://land.copernicus.eu/en/technical-library/hrl-vegetation-layers-comparative-analysis-forest-classification-and-temporal-consistency>

species was designated as dominant for at least one of the 76 EFTs. The compiled list of 113 species (Level 3) is presented in Figure 2-1.

The classification key for the 14 EFT categories (European Environment Agency 2007) has been analysed regarding the minimum required leaf type or tree type information. The result is presented in Table 2-1 and shows that a minimum of three leaf types (Broadleaved deciduous, Broadleaved evergreen and Coniferous) and five genera (Fagus, Quercus, Alnus, Betula and Populus) would be sufficient as input. This should, however, not only include information on the dominant leaf type and genus but also on mixture components.

The list of species in Figure 2-1 can subsequently be mapped to the minimum required leaf types (Level 1) and genera (Level 2). In addition, we suggest here to distinguish among Pinus and Picea at Level 2. Though they are not strictly required for the classification key of the EFT categories, they include the most common species in Europe and their distinction would certainly be interesting for other applications beyond the EFT. In addition, we also added Eucalyptus and Olea Europea; though the definition EFT category "Plantations and self-sown exotic forest" depends mainly on the forest management intensity, the two species account for largest part of plantations, especially in the Southern Europe.

Level 1	Level 2	Level 3
Broadleaved evergreen	Quercus (evergreen spp.)	Quercus ilex
		Quercus coccifera
		Quercus suber
		Quercus alnifolia
		Quercus rotundifolia
	Olea europaea	Olea europaea ssp. Sylvestris
		Olea europaea ssp. Cerasiformis
		Olea europaea L. subsp. europaea var. Europaea ³
	Eucalyptus	Eucalyptus sp.
		Ceratonia siliqua
		Pistacia atlantica
		Phoenix theophrasti
		Laurus nobilis
		Ceratonia siliqua
		Chamaerops humilis
		Ilex aquifolium
Coniferous	Picea	Picea abies
		Picea omorika
		Picea orientalis
	Pinus	Pinus brutia
		Pinus cembra
		Pinus halepensis
		Pinus heldreichii
		Pinus mugo
		Pinus nigra
		Pinus peuce
		Pinus pinaster
		Pinus pinea
		Pinus sylvestris
	Other Coniferous	Larix decidua
		Pseudotsuga menziesii
		Abies alba
		Abies borisiiregis
		Abies cephalonica
		Abies cilicica
		Abies nebrodensis
		Abies nordmanniana
		Abies numidica
		Abies pinsapo
		Cupressus dupreziana
		Cupressus sempervirens
		Cedrus atlantica
		Cedrus libani
		Tetraclinis articulata
		Taxus baccata
		Juniperus communis
		Juniperus drupacea
		Juniperus excelsa
		Juniperus foetidissima
		Juniperus oxycedrus
		Juniperus phoenicea
		Juniperus thurifera

Legend

Minimum requirement needed for EFT Level 1

among the 5 most common species

dominant in at least one EFT subclass

to be consider because of area coverage

generic class for none-further defined genus

1 semi-evergreen

2 except Acer sempervirens = evergreen

3 cultivated form

Level 1	Level 2	Level 3
Broadleaved deciduous	Fagus	Fagus sylvatica
		Fagus orientalis
		Fagus moesiaca
	Quercus (deciduous spp.)	Quercus canariensis
		Quercus cerris
		Quercus cerrioides
		Quercus petraea
		Quercus robur
		Quercus frainetto
		Quercus pubescens
		Quercus pyrenaica
		Quercus ithaburensis subsp. Macrolepis
		Quercus trojana ¹
		Quercus faginea ¹
	Betula	Betula pendula
		Betula pubescens
		Buxus balearica
	Alnus	Alnus cordata
		Alnus glutinosa
		Alnus incana
		Alnus viridis
	Populus	Populus alba
		Populus nigra
		Populus tremula
	Other Broadleaved deciduous	Acer campestre
		Acer granatense
		Acer heldreichii
		Acer hyrcanum
		Acer monspessulanum
		Acer opalus
		Acer platanoides
		Acer pseudoplatanus
		Fraxinus angustifolia
		Fraxinus excelsior
		Fraxinus ornus
		Carpinus betulus
		Carpinus orientalis
		Salix alba
		Salix caprea
		Salix eleagnos
		Tilia cordata
		Tilia platyphyllos
		Tilia tomentosa
		Ostrya carpinifolia
		Castanea sativa
		Ulmus glabra
		Ulmus laevis
		Ulmus minor
		Aesculus hippocastanum
		Arbutus unedo
		Celtis australis
		Liquidambar orientalis
		Juglans regia
		Platanus orientalis
		Prunus avium
		Prunus mahaleb
		Prunus padus
		Sorbus aria
		Sorbus aucuparia
		Sorbus domestica
		Sorbus torminalis

Figure 2-1: Hierarchical attribution of species (Level 3) to potential tree type (Level 2) and leaf type classes (Level 1).

2.2 Harmonization of current HRL Tree Cover & Forests nomenclature and the EFT nomenclature

Beside requirements resulting from the EFT nomenclature and other potential use cases the compatibility of a new map product with the existing HRL Tree Cover & Forest products is an important consideration. During an analysis of EFT nomenclature described in (European Environment Agency 2007) two potential conflicts have been identified for an integration of the EFT definitions with the existing HRL Tree Cover & Forests portfolio:

- The dominance of a species in EFT terms is generally defined through the dominance of the basal area. Since the basal-area cannot be easily quantified through satellite remote sensing methods the dominance in crown cover fraction is typically considered in the calibration of the HRL Dominant Leaf Type product. While both quantities are closely related they are not the same.
- The category “Alpine coniferous forest” in the EFT explicitly includes *Pinus mugo* (dwarf Pines) which are explicitly excluded from the HRL Dominant Leaf Type product to enable the derivation of forest type product that is aligned with the FAO definition (FAO 2000)

Further characteristics of existing HRL Tree Cover & Forests need to be considered to enable an alignment of a new tree type mapping with the readily available layers.

- Spatial resolution of 10m
- Extent of the tree cover and forest should be aligned with the extent of the HRL Dominant Leaf Type, Tree Cover Density and Forest Type maps, respectively.
- Tree type should be aligned with the leaf type as depicted in the Dominant Leaf Type map.
- The dominant observable species / tree type should be depicted; in analogy to the DLT depicting the dominant leaf type.
- Update cycle of either 1 year to align with status layers (DLT and TCD) or 3 years to align with FTY and TCPC
- The minimum mapping unit should be aligned either with DLT and TCD (pixel-based) or the FTY layer (MMU of 0.5ha).

3 EO-based high-resolution tree-type mapping at national and continental scale – state-of-the-art

In this chapter a review of the current state-of-the-art for tree type mapping is provided. To this end we focus on recent studies that go beyond a leaf type classification only (i.e. aiming at least at genus level information) and target mappings at national and /or continental scale.

3.1 Europe – Bonannella et al., 2022

(Bonannella et al. 2022) used a total of 305 environmental variables including 108 spectral seasonal metrics derived from time-series of Landsat ARD, reference data derived from GBIF (GBIF, The Global Biodiversity Information Facility 2025), LUCAS (Eurostat 2025) and EU-Forest (Mauri et al. 2017) and ensemble models to model realized and potential distribution for six reference periods (1) 2000–2002, (2) 2002–2006, (3) 2006–2010, (4) 2010–2014, (5) 2014–2018

and (6) 2018–2020 and 15 species. The resulting maps are produced at 30m resolution and provide reasonable accuracies for the presence / absence of most of the 16 species. It is, however, important to note that the realized distribution provides rather little spatial detail when compared to the currently available HRL (Figure 3-1) and does not allow to infer the dominance of one species over the other. This is certainly related to the use of numerous predictors with rather coarse native spatial resolution and could be amplified by the fact that the EU-Forest (1km² grid) is part of the training data whereas the chorological maps are also used as predictor variables.

The resulting maps are available at <https://zenodo.org/records/6516590>.

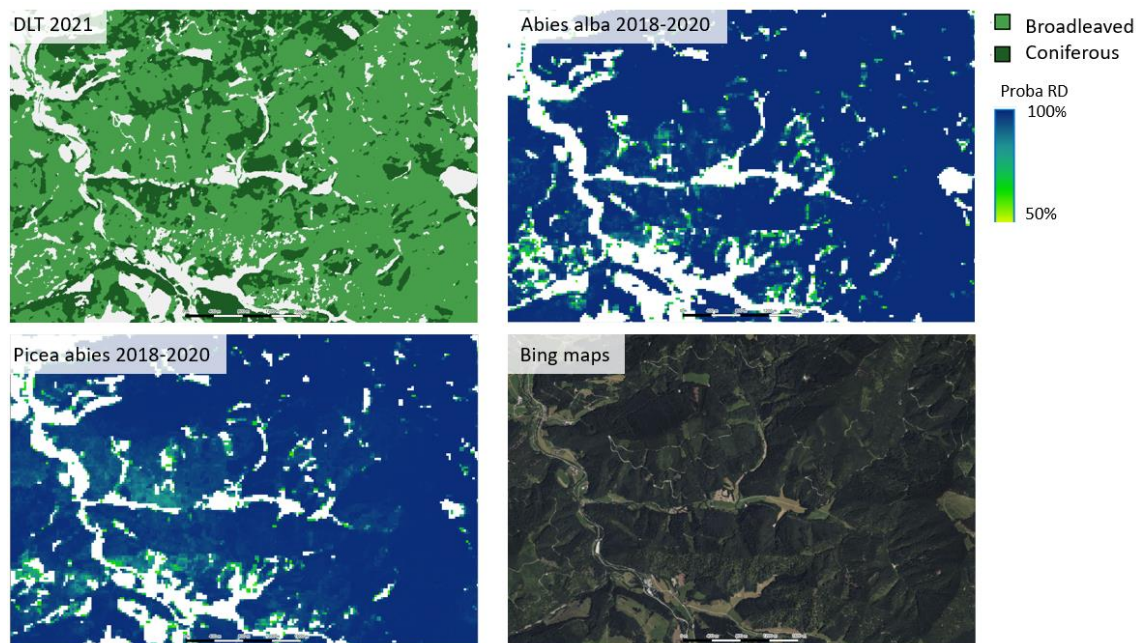


Figure 3-1: Comparison of the HRL Dominant Leaf Type 2021 with two species maps (Realized Distribution Probabilities - RDB) for 2018-2020 elaborated by Bonannella et al. 2022. The comparison illustrates the lack of spatial detail in the modelled RDB maps.

3.2 Canada – Hermosilla et al., 2022, 2024

(Hermosilla et al. 2022) used Landsat surface reflectance image composites for 1984 - 2019, topographic data, climate data, Multi-Source Land Surface Phenology (MS-LSP, based on L-8 and S-2), geographic coordinates (lat, lon) and NFI polygons to train a Random Forest classifier to map the dominant of 37 tree species across the Canadian territory for the reference year 2019. The produced maps were evaluated favourably with an overall accuracy of 93.1%, whereas the high overall accuracy was largely driven by the dominance and high accuracy of *Picea mariana* which constitutes more than 57.30% of the stands in Canada. User's accuracy values ranged from 72.7% to 100%, with an average of 91.4% whereas Producer's accuracies were generally lower ranging from 0% to 100% with an average value of 65.9%. The lowest accuracies were generally achieved for species with low prevalence. For both training and testing only stands in which one species reached 50% dominance were used and samples at the border of stands were excluded.

The authors attribute the high-accuracies also to the fact that model training was performed per regions which typically limits the number of species that have to be dealt with per model. Interestingly, the most important variables for the mapping are variables which are related to location and environmental conditions rather than the EO data, such that geographic coordinates, precipitation, minimum and maximum temperature and elevation consistently

rank with the highest importance. The authors comment that the strong predictive power of such variables might be closely related to the dense and regular samples available through the regular sampling grid of the Canadian NFI and might be less impactful if used with training data that does not follow a regular grid. The accuracies were assessed on an independent test set representing 30% of the available reference data; a minimum distance of 45m was assured among the samples to reduce the impact of spatial autocorrelation. From the paper it is not obvious through, if this separation is indeed sufficient to make the assessed accuracies fully independent of the training data.

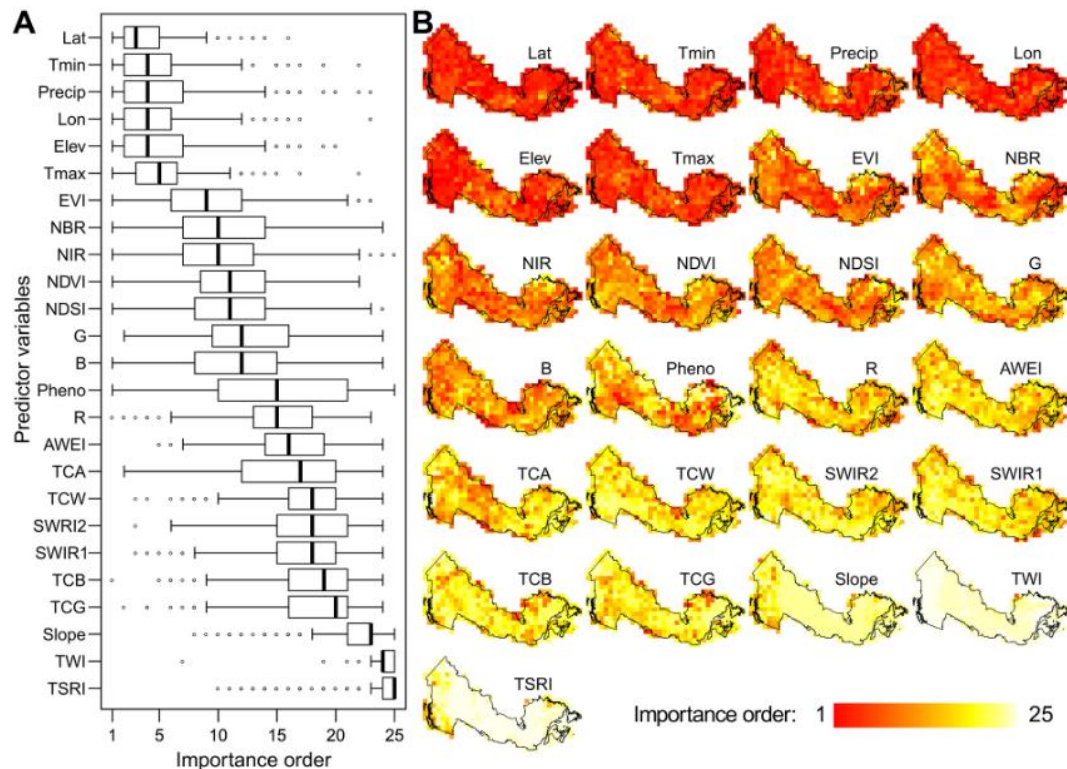


Figure 3-2: Relative predictor variable importance ranking across Canada's forested ecosystem tiles sorted by median rank value (Hermosilla et al., 2022).

(Hermosilla et al. 2024) later extended this study produce annual maps for the years 1984 to 2022. In this work, the methodology was modified to use climate variables such as annual minimum temperature, maximum temperature, and precipitation averaged over a five-year window to reduce variability in the output. Phenological variable were disregarded since they were not available for the full period at sufficient spatial resolution. The time-series of probabilities was smoothed using a Hidden Markov Model. The overall accuracies of the annual maps were at 86.1 %, and hence somewhat lower than the 2019 map from the previous study, the F1-score decreased as well and was 9% lower in average. The study provides an extensive analysis of the evolution of the species composition; the accuracy of the changes is not assessed. The annual tree species maps as well as several related datasets are available at https://opendata.nfis.org/mapserver/nfis-change_eng.html.

3.3 Germany - Blickensdörfer et al., 2024

(Blickensdörfer et al. 2024) have conducted a national scale mapping of 11 tree species classes for Germany using Sentinel-1 and Sentinel-2 time series for 2017 and 2018 in combination with

reference data extracted from NFI data for the years 2011/2012. The satellite image time series were interpolated to harmonized 5-day intervals (Sentinel-2) or aggregated to monthly composites (Sentinel-1) and complemented by further environmental variables on topography, climate, weather and soil moisture. A Random Forest classifier was used at the underlying machine learning model. To address the time-lag between the reference data from the NFI (2012) and the satellite data all plots that were detected with severe disturbance (Senf and Seidl 2020) were excluded.

The study paid particular attention to the underrepresentation of mixed stands which is common issue for tree type mapping and points out that accuracies might often be overestimated when simply excluding mixed plots from training and test data. The underrepresentation of mixed plots in the training data was partially addressed through the generation of pseudo-labels for plots with a mix of one deciduous species and one evergreen species. Similarly, the accuracy assessment was conducted to evaluate the contrast between pure and mixed stands (Figure 3-3).

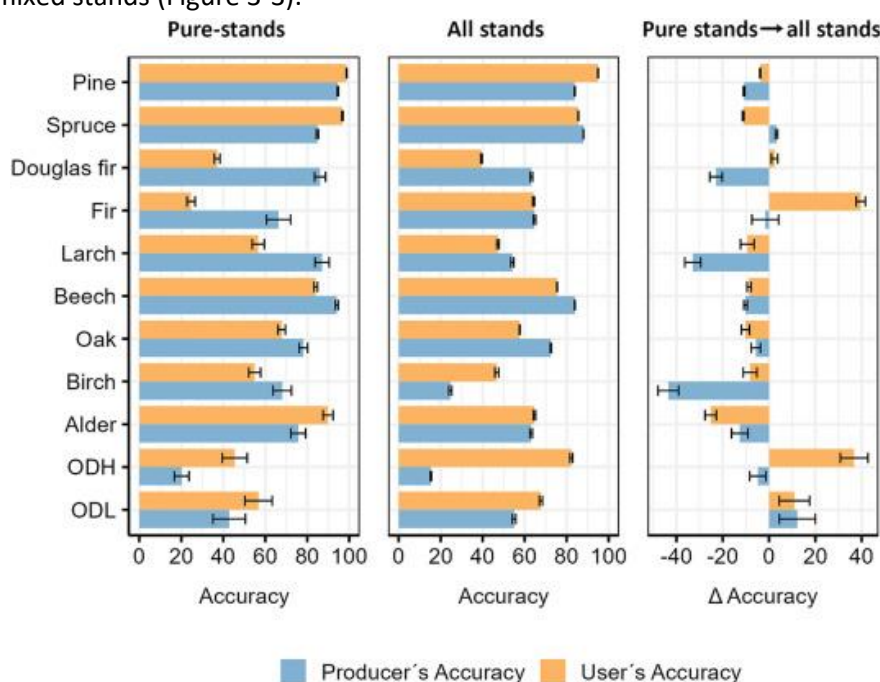


Figure 3-3: Accuracies for pure-stands accuracy vs all-stands accuracy from (Blickensdörfer et al. 2024)

The classification was applied within a forest mask based on the HRL TCD 2018 (>50% tree cover density), the Forest Additional Support Layer (FADSL, which excludes tree cover occurring in agricultural and urban areas) and German digital landscape model (DLM) for 2018 (which considers only areas that are forest, bog or swamp) and an MMU of 0.25 ha.

In summary the study shows promising results (F1-score, 72% and 97%) for 5 most common species in pure stands (spruce and pine, beech, oak and alder). For the other classes, however the F1-scores were only 28-69%. The accuracies dropped in average by 11.5% if mixed stands were included in the evaluation.

The authors conclude that their approach could serve as a steppingstone for tree species mapping efforts at the European scale, while further research is still needed to address some major challenges:

- Positional accuracy of NFI plot data which is not optimized for remote sensing applications
- Systematic sampling of the NFI plot data leading to small sample sizes for rare species

- Difficulties to map minor species which predominantly occur as minor admixture alongside dominant species

While the study states that large amount of NFI data are available for many countries it does not go into further details regarding general obstacles to access such data.

A visual comparison with the existing HRL Dominant Leaf Type (Figure 3-4) shows similar spatial patterns suggesting a general compatibility between the mapping approaches. Differences in the mapping of smaller landscape elements are due to the forest definition adopted by (Blickensdörfer et al. 2024) whereas the HRL Dominant Leaf Type maps tree cover.

The resulting raster map at 10m spatial resolution can be downloaded from https://www.openagrar.de/receive/openagrar_mods_00084346.

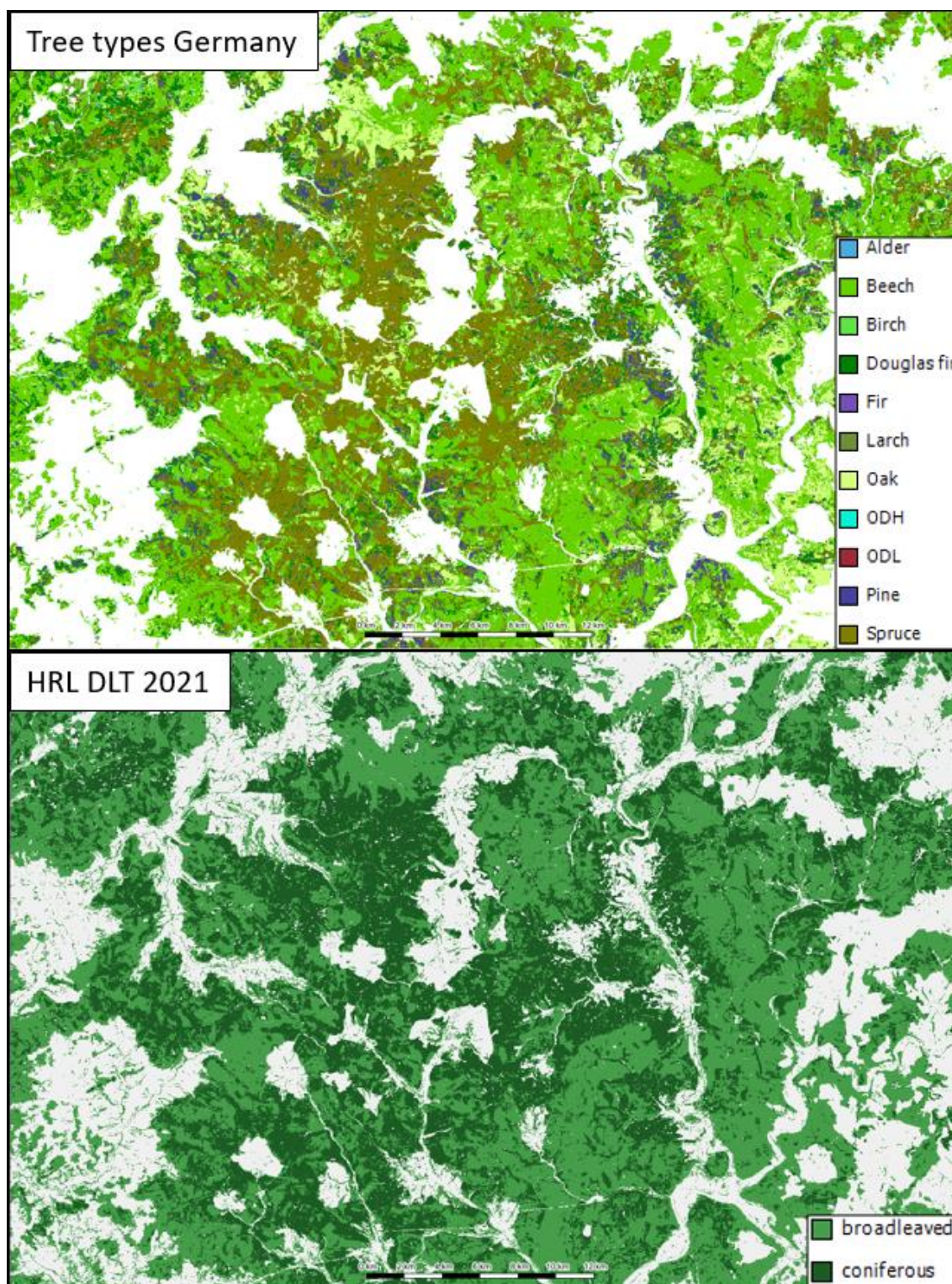


Figure 3-4: Visual comparison of the tree type map from Blickensdörfer et al., 2024 and the HRL DLT 2021 for an area in the Spessart (LAEA: 4280670m, 3006870m).

3.4 Netherlands - Francini et al., 2024

(Francini et al. 2024) presented a study on mapping seven tree type classes in the Netherlands combining Sentinel-2 time-series and NFI data to train a Random Forest classifier for the reference years 2019-2020. Samples were filtered to consider only plots for which the dominant species reached a basal area proportion of at least 80%. With the best set of features the overall

accuracy of the resulting map was assessed at 80.2%. Area proportions of the mapped tree types were compared with NFI-based estimates at a 10-km providing mixed results especially for rarer classes.

According to the published paper the map data is available upon request only.

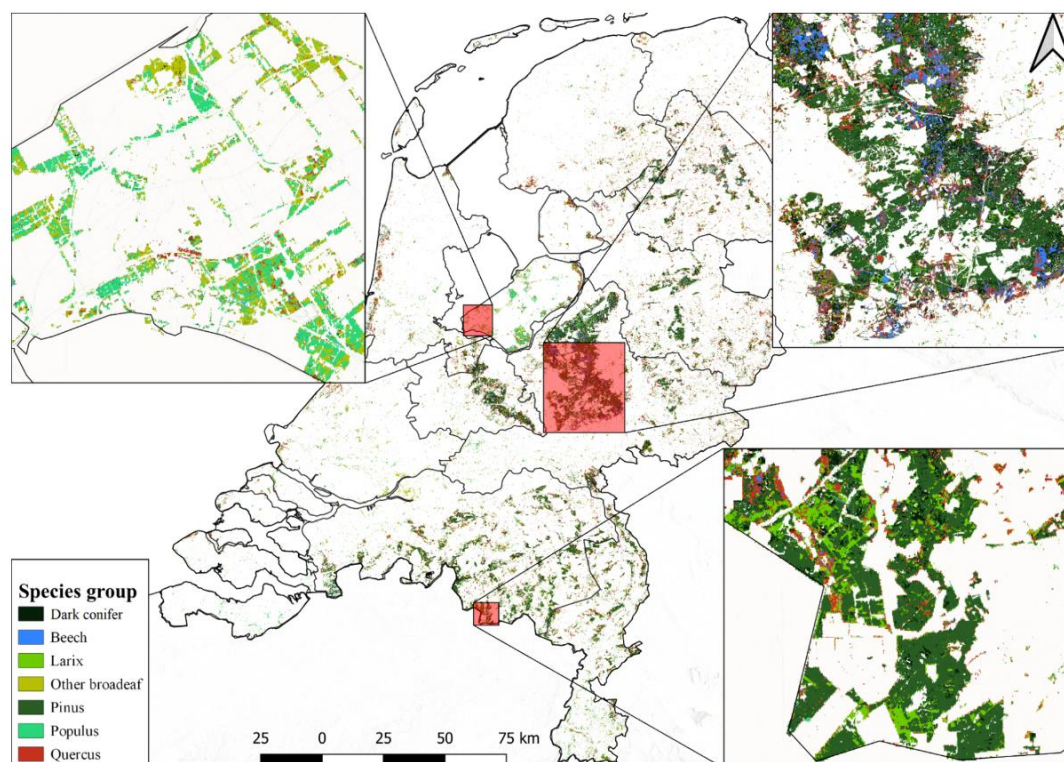


Figure 3-5: Tree type map for the Netherlands as presented by Francini et al., 2024.

3.5 Poland - Grabska-Szwagrzyk et al., 2024

(Grabska-Szwagrzyk et al. 2024) present a study to map the 16 dominant tree types in Poland using seasonal means extracted from Sentinel-2 data for the reference years 2018–2021.

Reference data were obtained from the Polish Forest Data Bank (FDB)³ and pre-selected with a focus on polygons representing pure stands with a single-species dominance and with trees older than 10 years. The dominance threshold varied among species from 100% for Scots pine to 60 %–80 % for less common species such as *Populus* spp. or *Pseudotsuga menziesii* where lower thresholds had to be set to obtain a sufficient sample size. A combination of ESA Worldcover 2021 and Google Dynamic World was used to generate a forest mask; samples outside this mask were removed. From the paper it is not fully clear which forest definition was adopted. To match the FDB polygons with homogeneous areas in the available imagery (i.e. areas which are likely to represent the dominant species) an image segmentation was performed on a seasonal S-2 composite for summer 2021. Only segments larger than 0.5 ha and with at least 60% overlap to the FDB stands were selected. The resulting 4500 polygons were divided into 2999 for training (~400k pixels) and 1501 for testing. Pixels with a mean NDVI below 0.6 in summer 2021 were excluded from the analysis since most of them corresponded to clear cuts.

³ From the paper we were not able to determine the reference years of the FDB that has been used.

The seasonal S-2 means were used together with six ancillary environmental variables on elevation, climate, precipitation, temperature and soils to train a Random Forest classifier on the extracted training data. Oversampling of smaller classes was test to improve their accuracies.

The conducted accuracy assessment indicates values of approximately 80% overall accuracy. Oversampling of smaller classes improved the result for most smaller classes but degraded the overall accuracy. The good overall accuracy is to a large degree due to the high accuracy for Pinus with an F-1 score > 90% (Figure 3-6) and representing 47.5% of the forests in Poland.

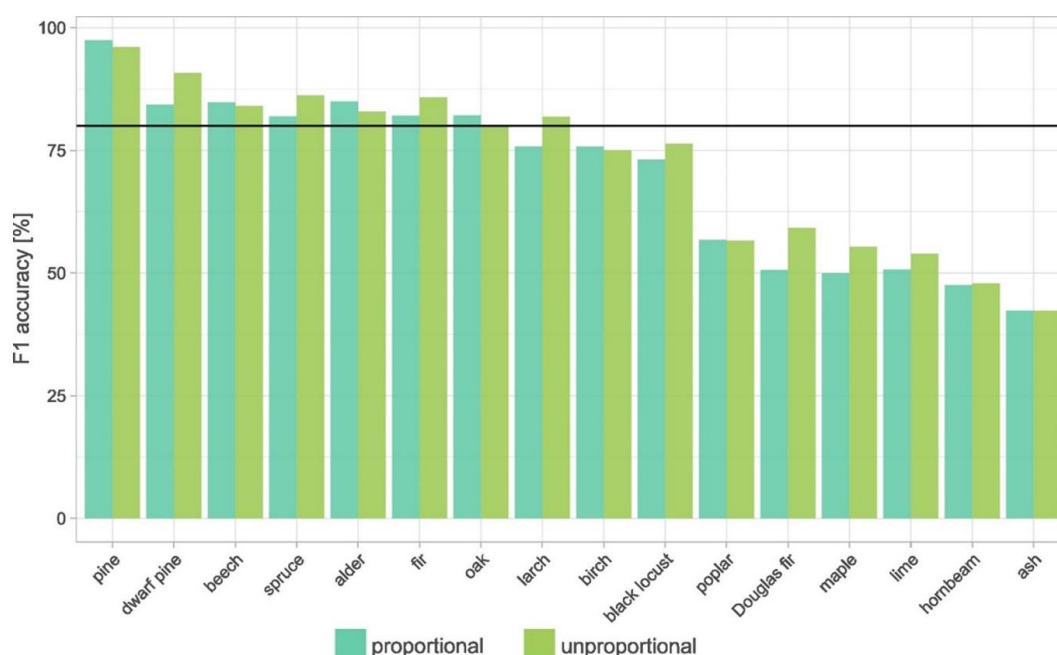


Figure 3-6: F-11 score for 16 tree types using proportional and disproportional sample allocation.

A visual comparison with the existing HRL Dominant Leaf Type (Figure 3-7) shows similar spatial patterns suggesting a general compatibility of the mapping approaches. Smaller landscape elements such as tree rows and small tree cover patches, as well as fruit tree plantations are mostly omitted in the map of (Grabska-Szwagrzyk et al. 2024) which is to be expected given the focus on forest trees.

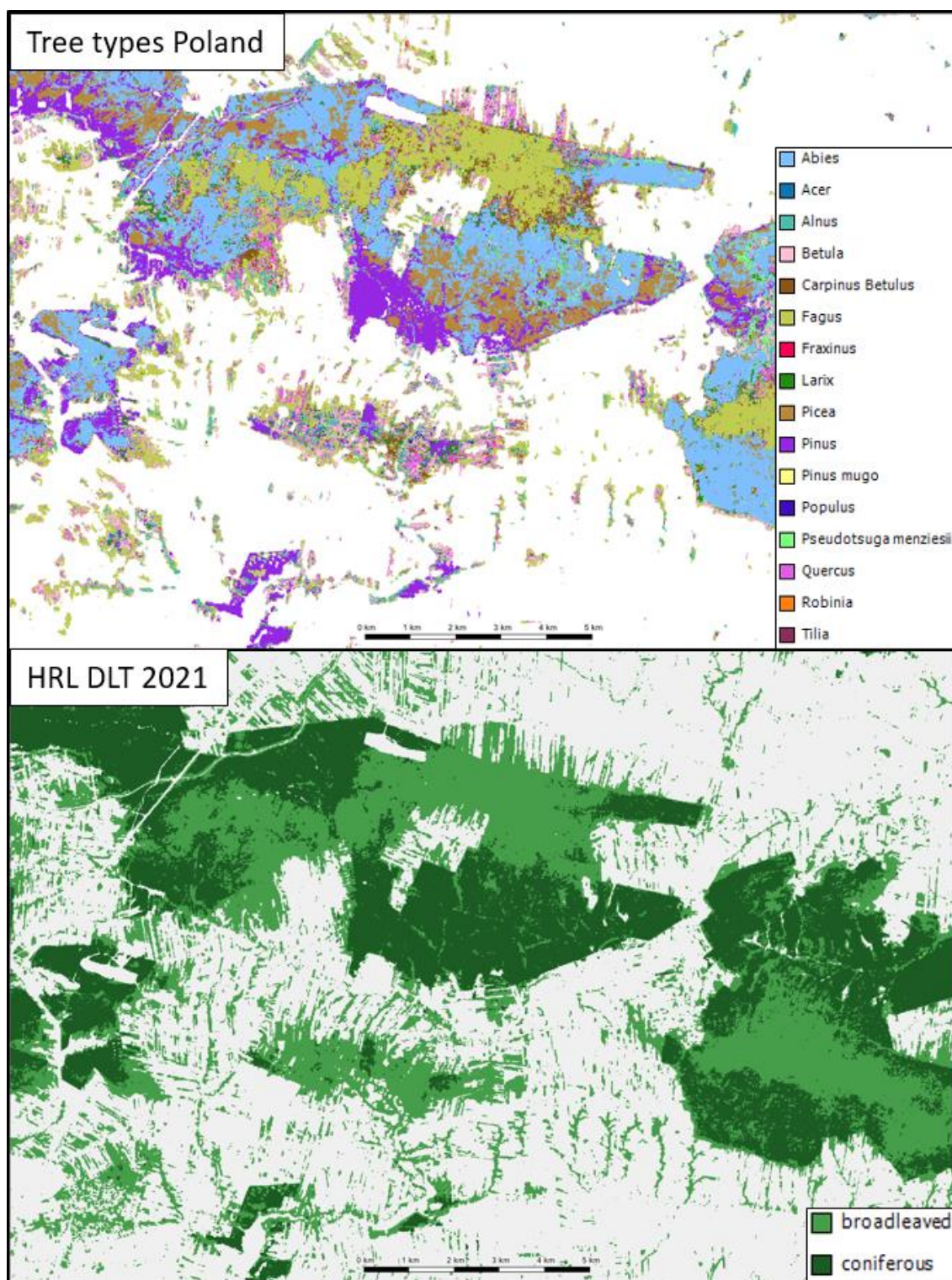


Figure 3-7: Visual comparison of the tree type map from Grabska-Szwagrzyk et al., 2024 and the HRL DLT 2021 for an area near Kielce (LAEA: 5077235m, 3146295m).

The tree type map as well as training and test data are available at <https://doi.org/10.5281/zenodo.10180469>. A view service is available at <https://ee-aweaksbarg.projects.earthengine.app/view/speciesmappl>.

3.6 France - Mouret et al., 2024

(Mouret et al. 2024) present a study on tree type mapping in France. Although the study does not cover the full national territory it still seems relevant in the context of this report. The study targets the classification of the 10 most common species in the Centre Region of France based on time-series of Sentinel-2 for 2 years (2019-2020) and reference data extracted from the NFI. From the later around 4400 plots were extracted using only plots where the dominant species reached 75% coverage. The S-2 time-series are harmonized to 10-day intervals.

The study compares three different deep learning architectures with a classical Random Forest approach and combined with different strategies to address class imbalance. All in all, the deep learning architectures outperform the classic Random Forest algorithm, differences among the DL techniques are minor. In terms of F-1 score a relatively simple deep MLP architecture performed slightly better (0.81), whereas Lightweight Temporal Attention Encoder (LTAE) led to better balanced accuracy metrics (0.83). The tested techniques to tackle class imbalance generally helped to improve the balanced accuracy while having no significant impact on the overall accuracies. The paper is focused in particular on the comparison of the different approaches and does not go into further detail regarding the analysis of the resulting maps and the separation of tree covered areas and other land cover types.

The authors also report a comparison of using LTAE on raw data with cloud masks (i.e. not application of cloud masks and not interpolation) as proposed in (Bellet et al. 2024). The obtained results, however, had slightly lower accuracies and showed borders among neighbouring S-2 tiles that point to overfitting of specific observations within individual S-2 tiles. The code is available at https://framagit.org/fl.mouret/tree_species_classification_iota2 and <https://framagit.org/iota2-project/iota2>. The training data and resulting maps are not public.

3.7 Summary

A series of recent publications on tree type mapping on national and continental scale show promising results for the 5-6 leading tree species and homogeneous stands with accuracies from 70% to more than 90% depending on the study and the design of the validation. Further research is, however, still needed to achieve similar results for less common species / taxa and to correctly determine the dominant species in mixed stands which are currently mapped with significantly lower accuracies.

Most studies rely on a combination of optical time-series from Landsat or Sentinel-2 and samples extracted from National Forest Inventories (NFI). Only (Blickensdörfer et al. 2024) also consider Sentinel-1 time-series.

In general, the authors agree, that the availability of the accurate NFI reference is a key ingredient to obtain reliable training and validation data. An exception is the work presented by (Bonannella et al. 2022) which uses a type of crowd-sourced training data and a large set of environmental variables; the results should, however, be considered rather as a coarse scale modelling of occurrences rather than an actual mapping of the dominant tree type.

In this regard it is important to also recall that the reported accuracy metrics should not be considered in isolation. The accuracies reported by (Hermosilla et al. 2022) and (Hermosilla et al. 2024) for tree type mapping in Canada are for example significantly higher than comparable studies in Europe. This might be at least partially due to the large homogenous forests stands in Canada which allow that a quasi-interpolation from regular-spaced NFI sample plots would already yield rather high accuracies. This is also corroborated by the fact that environmental positional variables such as precipitation, temperature or lat/lon coordinates rank highest in the model presented by (Hermosilla et al. 2022). It is unlikely that the similar accuracies could be achieved for more heterogenous landscape such as in Central and Southern Europe.

All presented studies use external tree cover / forest masks which are applied on top of the tree type classification. On the one hand, the reported accuracies typically focus on the confusion among tree type and do not include errors from the applied masks. On the other hand, this highlights the importance of existing tree cover mappings such as included in the HRL Tree Cover & Forests as a first step towards tree type mapping.

4 Potential reference data for training and validation

The chapter attempts to provide an overview of available reference datasets that could be considered as input for the generation of training and validation datasets. The chapter includes two sub sections to present map data with full spatial coverage (Section 4.1) and point or polygon data that can be characterized as samples (Section 4.2).

4.1 Map data

4.1.1 Austria

The Austrian authority for the national forest inventory Bundesforschungszentrum für Wald (BFW) provides a web viewer for a tree type map which has been derived through a combination of the NFI data 2016-2021 and Sentinel-2: <https://www.waldinventur.at/?x=1474545.38136&y=6059660&z=7.75968&r=0&l=1111#/map/1/mBaumartenkarte/Bundesland/erg9>

The map distinguishes 13 tree type and tree type mixture (e.g. Picea – Pinus). We were not able to find further details on the methodology used or the accuracy of the provided map.

4.1.2 Czech Republic

The Czech Forestry Institute (previously UHUL, renamed to NLI as of 2025-01-01: <https://nli.gov.cz/en/o-uhul/>) provides a web viewer for tree type map which has been derived through a combination of the NFI data and Sentinel-2 series: <https://geoportal.uhul.cz/mapy/MapyDpz.html>

There are three time-stamps of the tree type mapping (Lesní dřeviny) for the reference years 2017, 2019 and 2022 combining NFI data and Sentinel-2 time-series. The map distinguishes Pinus, Picea, Fagus, Quercus, other coniferous and other broadleaved and has been assessed with an overall accuracy of 85%⁴. Further details on the methodology or plans to continue this mapping in the future were not accessible at the time of writing of this report.

4.1.3 France

The BD FORÊT is an openly available (<https://geoservices.ign.fr/bdforet#telechargementv2>) vector database for continental France which maps 32 vegetation classes (mainly Forest types) with a minimum mapping area of 0.5 ha and a minimum mapping width of 20 meter. The database has been elaborated by IGN France for the reference period 2007-2018 based on aerial imagery and reference data from the national forest inventory. The database has recently been complemented by a beta version of a forest mask (Masque Forêt 2019-2022) which adheres to

⁴ https://www.uhul.cz/wp-content/uploads/08_Mlcousek-FMI_DPZ_AJ.pdf

the forest definition of the FAO. An improved version of this forest mask is planned for publication in 2025.

While the provided database is generally of high quality, the spatial and thematic resolution poses some limitations for direct use of the map as a reference of an EO-based mapping at 10m resolution. Figure 4-1 illustrates the partial mismatch between the outlines in the BD FORÊT and the heterogeneity of different dominant leaf types within the delineated stands. The nomenclature of the BD FORÊT, however, still includes classes for pure stands which could still be leveraged for calibration and validation. Considering the working legend presented in section 2.1 would be applicable for classes referring to the genera *Quercus*, *Fagus*, *Pinus* and *Populus* (plantations) only.

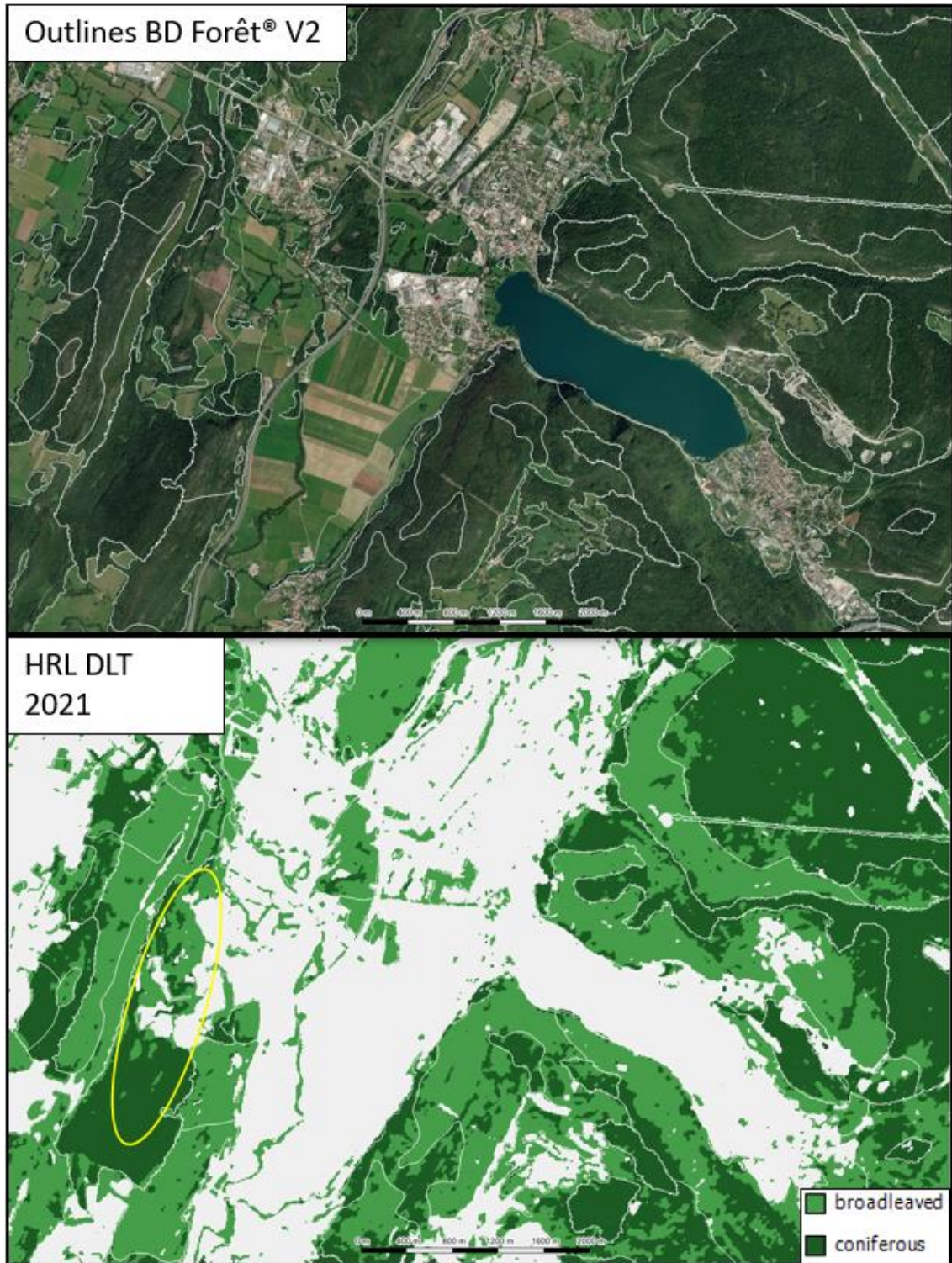


Figure 4-1: Visual comparison of the forest outlines of the BD Forêt® V2 and the leaf type mapping in the VLCC DLT product. While there is a strong resemblance of main spatial structures the generalization of the BD Forêt® V2 yields many stands that are assigned to mix classes. In many cases the assignment of the class (e.g. yellow ellipse = closed spruce and fir forests) omits mixture components as well.

4.1.4 Spain

The Mapa Forestal de España (MFE25 / MFE50) is publicly available (<https://www.miteco.gob.es/es/cartografia-y-sig/ide/descargas/biodiversidad/mfe.html>) at a scale of 1:25.000 as a vector-format database for each of the 17 autonomous communities of Spain. Exceptions are Valencia and Andalusia which are still only available at a scale of 1:50.000. The database provides rich information on the forest type and distinguishes among 178 species. For each polygon up to three of the most common species are named and a percentage cover is provided. The reference years for the regions vary between 2005 and 2023, a usage of the data for more recent reference years will thus in some cases require careful filtering to avoid or correct for areas with significant changes in the last 20 years.

4.1.5 Sweden

The Swedish University of Agricultural Sciences hosts (<https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/foreststatistics/slu-forest-map/>) a time-series of EO-based forest maps covering the years 2000, 2005 and 2010 (25 x 25 m) and 2015 (12.5 x 12.5m). Up until 2010, the maps were based solely on satellite images from Landsat and SPOT, whereas according to the documentation the 2015 map was based on temporary samples from the NFI's temporary samples plots from 2017-2019, normalized surface models derived aerial stereo images and Sentinel-2 imagery. Given that the latter is rather based on data from 2017 onward the suggested nominal reference year 2015 is somewhat confusing. The latest map edition also provides only partial coverage of focused on Southern Sweden and the western coast.

The product includes 12 map layers that describe volume per tree species, basal area, basal area-weighted average height, basal area-weighted average diameter and biomass. The volume maps provide an idea of the timber resources categorised by pine, spruce, beech, oak, birch, and other deciduous trees. An evaluation, against NFI data suggest a an RMSE of the total volume is approx. 85 m³sk/ha or 41 percent, the corresponding value for tree biomass is 44 tons TS/ha or 40 percent. The values generally decrease with increasing latitude, which is largely explained by the fact that the timber stocks also decrease with latitude.

4.1.6 Others

Further maps which are based on data from national and international authorities exist. One of the first attempts to a pan-European map of dominant tree species was conducted by (Brus et al. 2012) using a combination of ICP and NFI reference plots and a combination of co-kriging and multivariate regression. While, the resulting map of dominant species was evaluated with an overall accuracy of only 43% it still provides a useful indicator for the general distribution tree species across Europe.

A large effort to gather and combine data from ICP and NFI has also been conducted by (European Commission: Joint Research Centre 2016) and lead to the publication of the European Atlas of Forest Tree Species. A more recent technical report, details the maps produced for the atlas as well as for other project by JRC (Beck et al. 2023). This includes:

- **Chorological maps** a set of vector files indicating, in broad terms, where the species is native, and where it has been introduced and naturalized. Details on the methodology are provided in (Caudullo et al. 2017). The latest version of such maps are currently hosted here: <https://data.mendeley.com/datasets/hr5h2hcg4/18>
- **Relative Probability of Presence (RPP) maps** at 1 km resolution, which quantify the probability of finding at least one individual of the tree species in a plot placed randomly within the 1km grid cell. Available online at <https://forest.jrc.ec.europa.eu/en/european-atlas/atlas-data-and-metadata/>

- **Maximum Habitat Suitability (MHS) maps** range between 0 (low suitability) and 1 (high suitability). Where the species has actually been observed, MHS is 1. MHS values below 0.55 indicate a low likelihood for the survival of the species. Available online at <https://forest.jrc.ec.europa.eu/en/european-atlas/atlas-data-and-metadata/>

While the produced maps are unique invaluable indicators for the tree species distribution in Europe the spatial resolution and other characteristics are, however, hindering a straightforward usage in combination with EO data at spatial resolution of 10m. The RPP maps used in a prototype of European Forest Type map (Giannetti et al. 2018), for example, carries in fact rather little information on the dominant species in a given cell. Similar issue applies for an earlier attempt for statistical modelling of the distribution of 20 tree species groups in Europe (Brus et al. 2012).

4.2 Point and plot data

This section focusses on point and plot data which have been reviewed in this study.

4.2.1 National Forest Inventories

All of the studies presented in chapter 3 rely on the availability of NFI data for reliable reference data for training and validation. While it would be desirable to use NFI data for all EEA38 countries and the UK, in many cases the access to such data is restricted in particular for permanent plots since a publication of the exact coordinates could increase the risk that samples could be manipulated and reduce the representativeness of the inventories (Schadauer et al. 2024). Hence it must be expected that the access to NFI data will be limited to specific countries and subsets in the near future. A complete survey of the availability of NFI data in all EEA38 countries and the UK was beyond the scope of this study, nevertheless we highlight here a few examples where NFI data is readily accessible. A good overview of the authorities responsible for the NFI data in the EU27 countries is given by <https://www.inventarioforestale.org/en/inventari-forestali-europei/>.

The countries for which, to the best of our knowledge, access to precise locations is currently in principle possible are listed below. In this context it is worth mentioning that for Germany there is a benchmark dataset available which combines Sentinel-2 time-series and labels from the German NFI via a publication by (Freudenberg et al. 2025). However, this does not include exact locations and since the Sentinel-2 time-series have been pre-processed with a custom pre-processing workflow the data cannot easily be combined with other datasets.

Sweden

A subset of field data collected for temporary sample plots is available via the Swedish University for Agricultural Sciences under <https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/listor/sample-plot-data/>. The database currently holds more than 40,000 records the years 2007 till 2023.

Spain

The 4th Spanish National Forest Inventory was conducted starting 2008 and includes an assessment of the species composition for parcels from the MFE25. The data is publicly available under <https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario->

forestal-nacional/cuarto_inventario.html and includes centre coordinates and identifiers of the MFE25 parcel.

Netherlands

The Nederlandse Bosinventarisatie Inventory (NBI, National Forest Inventory for the Netherlands) has been conducted since 1938; the latest 7th survey has been concluded in 2022. The 8th NBI is currently ongoing and will be concluded in 2026. The raw field data is available via <https://www.probos.nl/publicaties/overige/1094-mfv-2006-nbi-2012>. Exact coordinates to relate the provided species information to EO data can in principle be provided upon request and legal agreement to keep the provided data confidential (Mart-Jan Schelhaas 2025, personal communication, February 2025).

Finland

In March 2018 Finland launched an open forest data system which provides access to information on private, family-owned forests (about 60-70% of Finland's forest)⁵. The service is accessible for forest owners and forest professionals at: <https://www.metsakeskus.fi/sv/arenden/minskogfi/produktbeskrivning-av-minskogfi-for-proffs>

Details on the registration process and data access are available mainly in Finnish language and could not be followed up in the scope of this study.

Italy

The Italian National Forestry Information System (SINFor) provides cartographic and statistical information on the Italian forests. One publicly available reference dataset is the Carta Forestale d'Italia (CFI2020) which has been generated at a scale of 1:10000. The map can be visualized through a web-viewer; requests for direct access can be made for up to two of the 21 regions of Italy according to the official government website: <https://cfi-sinfor.crea.gov.it/Home/CartaForestale>

4.2.2 ICP Forest

The International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) was established in 1985 with the main purpose to monitor air pollutants, climate change and other stressors on the condition of forest ecosystems. According to the website⁶ the initiative combines a network of 40 European countries with around 6,000 observation plots distributed on a systematic 16 x 16 km transnational grid across Europe (Level 1) and 600 more intensively monitored plots for selected forest ecosystem at Level 2. The data is available upon request and subject to strict usage conditions⁷ and made available to GAF AG for the use in this study.

The ICP Forest database has a rather complex data structure which is documented at <https://icp-forests.org/documentation/Introduction/index.html>. This includes numerous attributes on biodiversity, stand characteristics, tree conditions, soils and so forth most of which could not be checked in depth within the scope of this study. The focus was thus put on assessing the possibility to extract tree species information at a spatial, temporal and thematic detail that

⁵<https://forest.fi/article/five-applications-that-use-open-forest-data-over-three-million-downloads-of-geographic-forest-data-maps-within-a-year/>

⁶ <https://www.icp-forests.net/>

⁷ <https://www.icp-forests.net/data-maps/data-requests>

would allow to retrieve reliable reference data from the database that could be matched to EO data such as Sentinel-1/2.

The obtained Level 1 plot files contained initially 21689 records, 1119 of which could not be matched with a unique ID (combination of plot_id and code_country) in the file that contains the surveys of the crown conditions (including tree species information per individual tree). Of the remaining 20570 plots only 11354 were considered further since the latest survey was in 2018 or more recent. However, most of the plots are provided with coordinates rounded to decimal minutes, leaving only 1817 plots (Figure 4-2) as a potential source for reference data. The coordinates of those plots are provided with a geolocation accuracy of 15-25m. The dominant tree species can be determined by parsing species information per tree. though no further information on the exact location of the tree is available. The distribution of dominant species in the remaining plots is presented in Figure 4-3. The 7 most common genera in the remaining Level 1 plots are Pinus, Picea, Betula, Quercus, Fagus, Alnus and Populus which are all considered in the second level of the hierarchical legend (Figure 2-1).

Similarly, the obtained Level 2 plot files contain 1177 records; 945 of which could be matched with a unique ID to a corresponding record of the crown condition surveys. Of those only 554 have been revisited since 2018 and finally only 498 which are provided with non-rounded coordinates Figure 4-5 . The dominant genera in this set are Pinus, Picea, Quercus and Fagus (Figure 4-5).

Based on this brief analysis it appears that the ICP Forest database could be valuable source of authoritative reference data to support an EO-based tree type mapping for the EEA38. However, the heterogenous spatio-temporal coverage is most likely insufficient and a combination with other reference datasets will be indispensable. Several open questions remain in particular regarding the impact of the following points:

- Uncertainties of the geolocation of 15-25m
- Limited information on the plot design; the plot size variable seems typically filled but it is not obvious if/how an exact spatial footprint of the plot can be derived.
- Other information on the tree height and crown conditions might be helpful to determine which tree type dominate the signal of the plot from a remote sensing perspective but could not be investigated in detail in this study.
- There is apparently no information on the location of every tree and thus it is not obvious if/ how the tree species distribution in mixed plots could be linked to 10m-resolution Sentinel-2 observations. It might therefore be necessary to further limit the number of usable plots to those which are dominated by one tree type.

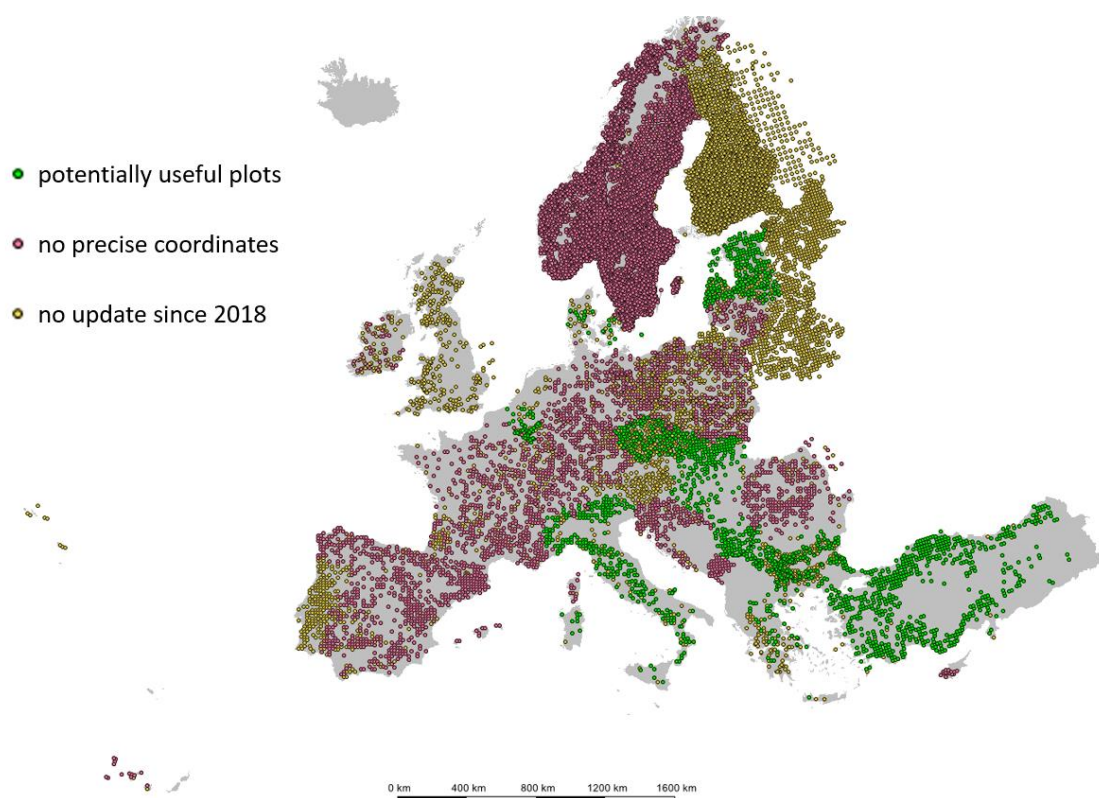


Figure 4-2: Distribution of ICP Forests Level 1 plots with different levels of available information.

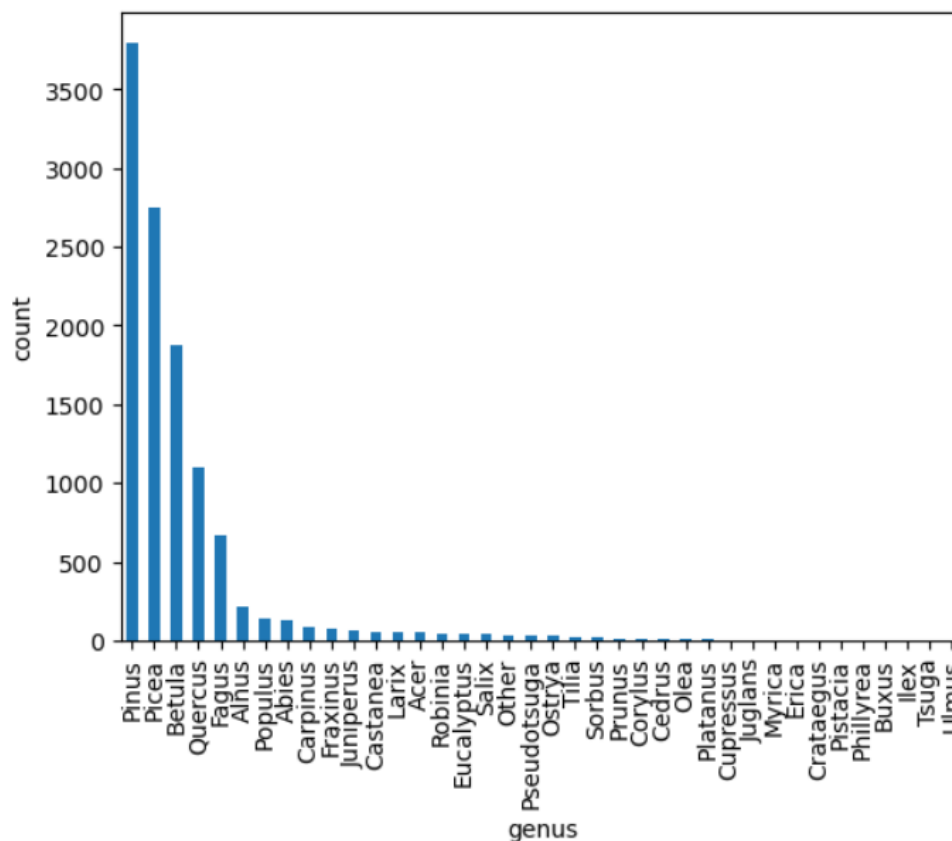


Figure 4-3: Distribution of dominant genera in the ICP Forests Level 1 plots which have been updated since 2018 and are provided with second-level precision coordinates.

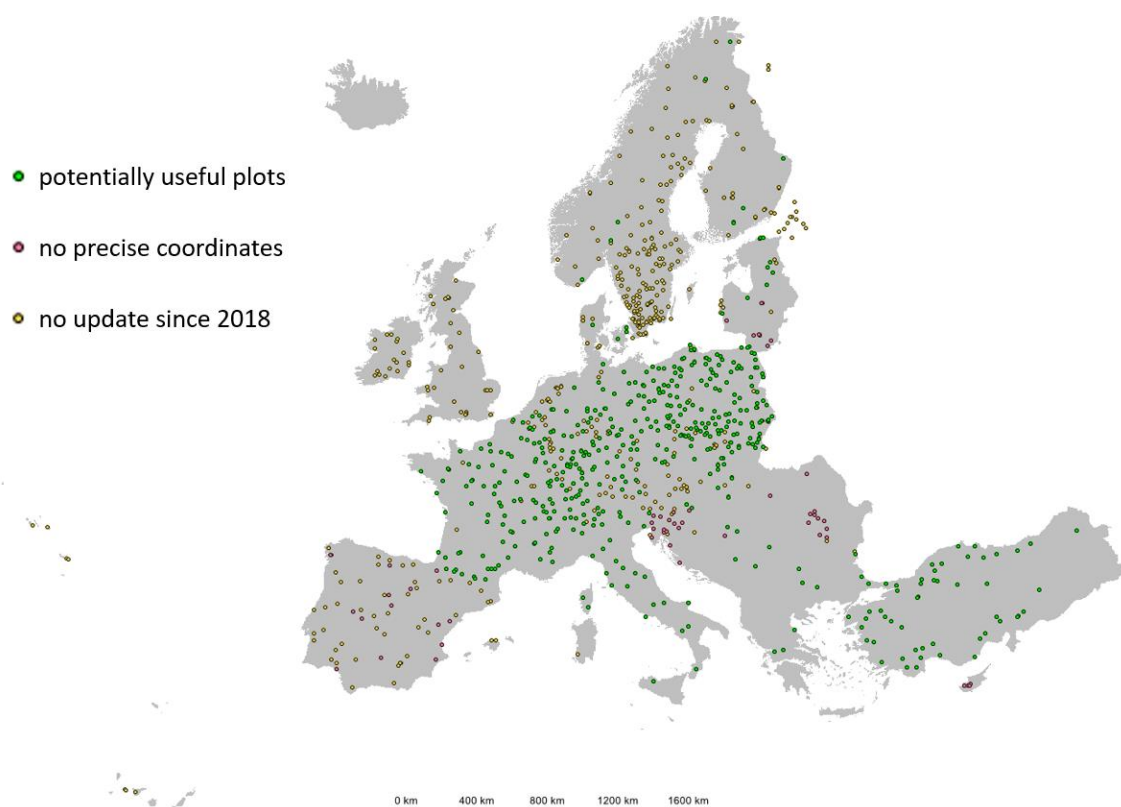


Figure 4-4: Distribution of ICP Forests Level 2 plots with different levels of available information.

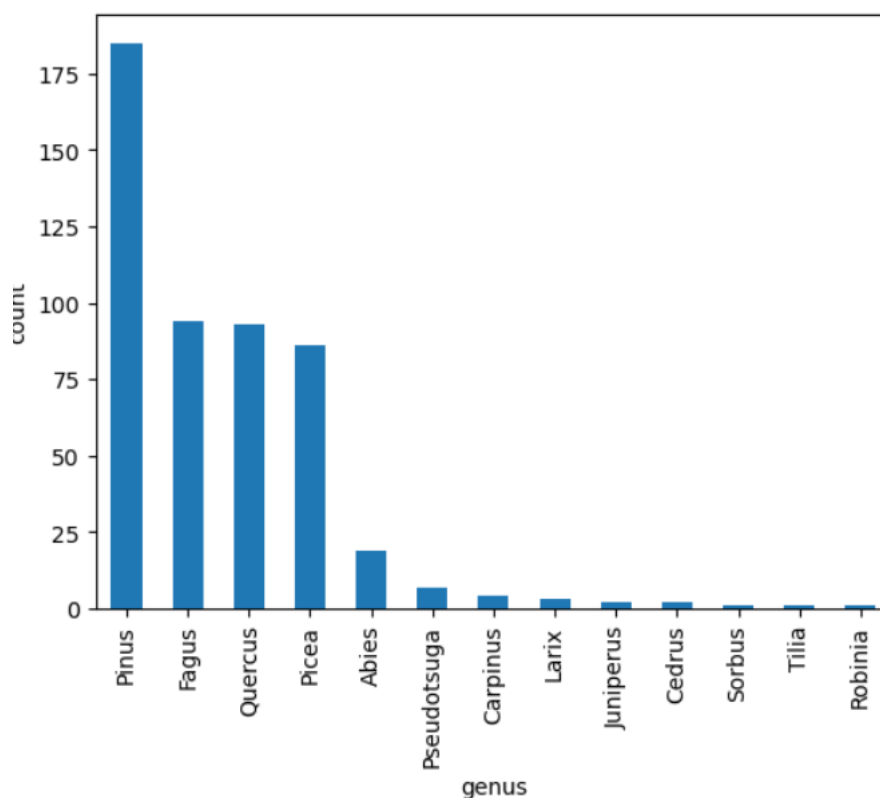


Figure 4-5: Distribution of dominant genera in the ICP Forests Level 2 plots which have been updated since 2018 and are provided with second-level precision coordinates.

4.2.3 LUCAS

The Land Use / Cover Area frame Survey (LUCAS) is probably the most comprehensive in-situ survey of land cover for the EU.

LUCAS 2018 contains attributes on primary land cover (LC1) and secondary land cover (LC2) for which the nomenclature contains the following relevant woodland classes:

- C10 BROADLEAVED WOODLAND
- C20 CONIFEROUS WOODLAND
 - C21 Spruce dominated coniferous woodland
 - C22 Pine dominated coniferous woodland
 - C23 Other coniferous woodland
- C30 MIXED WOODLAND
 - C31 Spruce dominated mixed woodland
 - C32 Pine dominated mixed woodland
 - C33 Other mixed woodland

According to the technical documentation⁸ *“For all points in CXX the surveyor has to indicate the forest cover code in the respective “LC plant species” field, according to the forest type classification of the European Environment Agency.”* The 2018 survey thus includes information on the 14 EFT classes (Table 2-1).

An analysis of those attributes shows that LUCAS 2018 includes 120078 points with primary LC being Woodland, plus additional 58 points where secondary LC is Woodland. Applying a further filter to retain only direct observations with an observation distance of < 100m leaves 52331. Based on the class annotation for LC1, LC2 and LC1_SPEC it is possible to distinguish 4 different genera being *Pinus*, *Picea*, *Quercus* and *Fagus* at 30693. The spatial distribution and distribution among the 4 genera are displayed in Figure 4-6 and Figure 4-7, respectively.

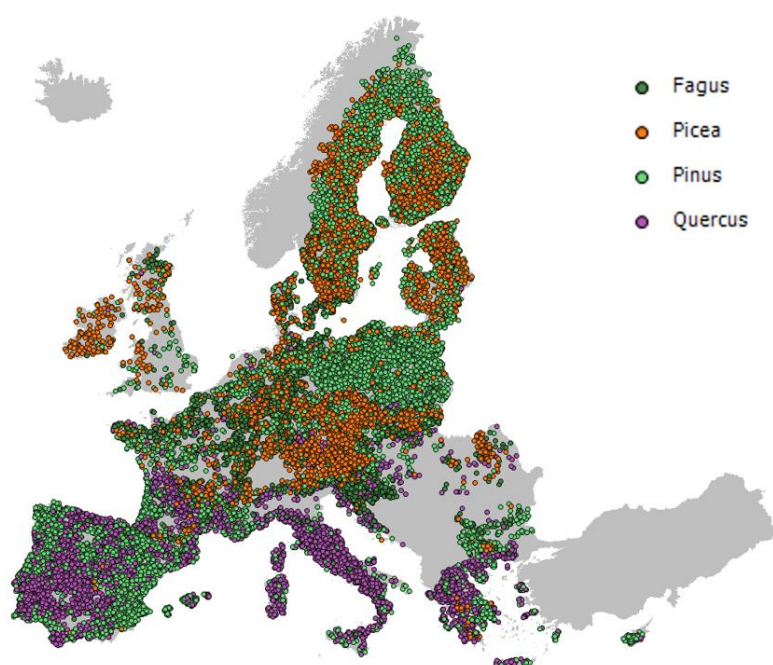


Figure 4-6: Spatial distribution of reference points for four genera extracted from LUCAS 2018.

⁸<https://ec.europa.eu/eurostat/documents/205002/8072634/LUCAS2018-C3-Classification.pdf>

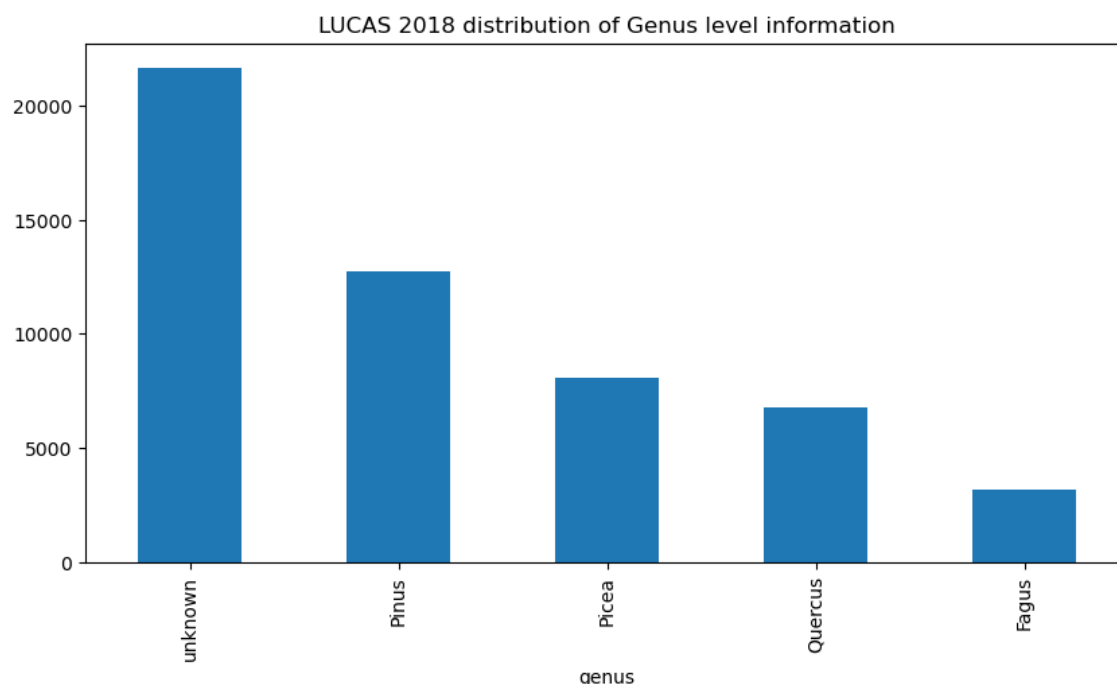


Figure 4-7: Sample counts per genus for reference points extracted from LUCAS 2018.

The same analysis has also been conducted with the LUCAS 2022 survey but revealed that that the usage of the EFT nomenclature has apparently been skipped in the 2022 survey. This change is not explicitly mentioned in the official documentation⁹ but the nomenclature cannot be found anymore in the documentation or the LC1_SPEC field. As only the two genera Pinus and Picea can be distinguished based on the information in LC1. Applying the same filtering for woodland and direct observations at distances < 100m yields thus 15346 points for which genus level information can be inferred Figure 4-8, Figure 4-9).

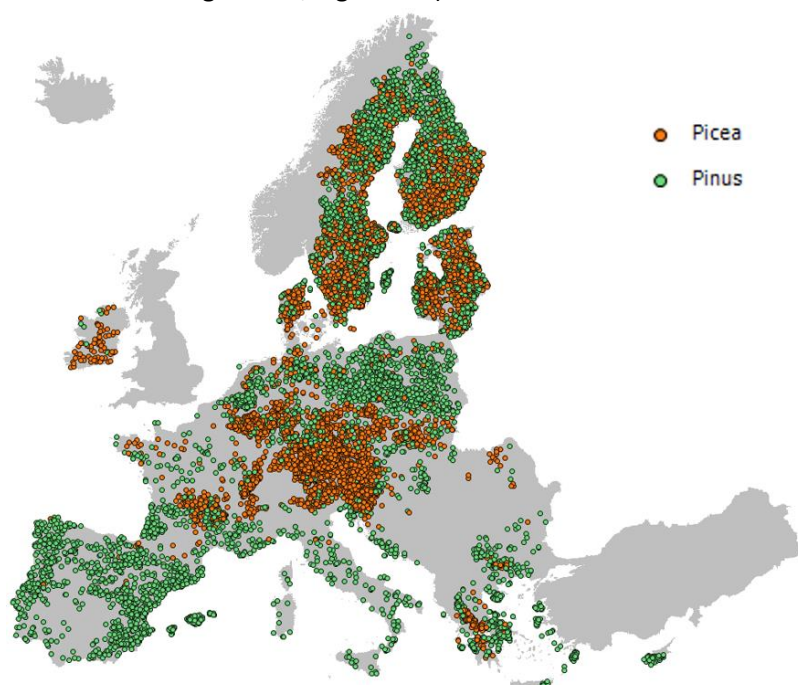


Figure 4-8: Spatial distribution of reference points for two genera extracted from LUCAS 2022.

⁹ <https://ec.europa.eu/eurostat/documents/205002/13686460/C3-LUCAS-2022.pdf>

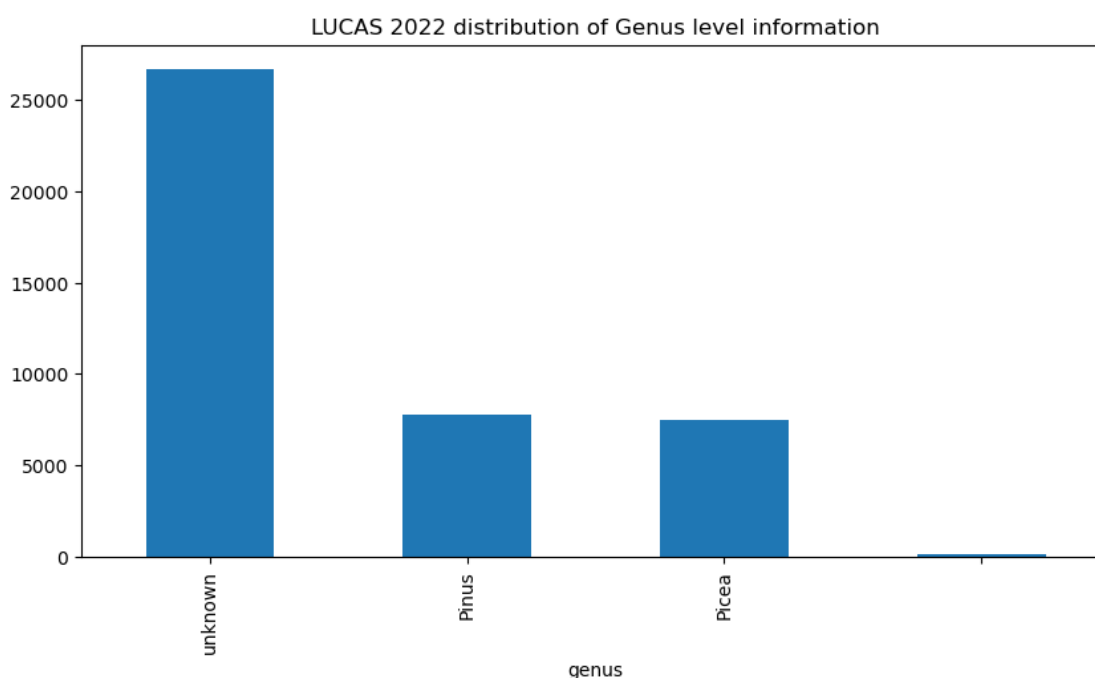


Figure 4-9: Sample counts per genus for reference points extracted from LUCAS 2022.

Based on this brief analysis it can be concluded that while essential in terms of coverage and homogeneity, the LUCAS data alone is currently rather limited in terms of thematic depth for tree type mapping. This is particularly true for the reference year 2022 where the EFT nomenclature has been removed. The year 2018 on the other hand could be in particular valuable to support and validate the derivation of an EFT map at least at the highest level (categories) of the EFT nomenclature.

4.2.4 GBIF

The Global Biodiversity Information Facility (GBIF) is “an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth”¹⁰. It provides extensive documentation, tooling and a community around a vast collection of occurrences datasets which can be accessed by anyone through the browser or with an API.

For this study we used the provided API to query all records for 110 tree species within the bounding box of the EEA38 countries and UK (DOMs have been left aside). At the time of the query (2024-10-31) this yielded a total ~19 Mio records. Given that GBIF is an open source database which includes records from various sources and crowd sourcing initiatives such as iNaturalist¹¹ the data quality, spatial precision and reference dates are heterogeneous and not directly suitable for remote sensing applications. Several metadata tags of the records such as coordinate uncertainties and observation year (Figure 4-10) can be used to filter out observations which are most likely less suitable for further processing already. An overview of tested / implemented filtering steps and the respective number of remaining samples is provided in Table 4-1. Discarding older and spatially uncertain observations drastically reduces the number of observations. While most filtering steps are based on GBIF's metadata other steps require a close look and more customized filters. This includes for example the removal of data ingested from the French NFI for which the coordinates are rounded to the nearest 1km grid but

¹⁰ <https://www.gbif.org/what-is-gbif>

¹¹ <https://www.inaturalist.org/>

the metadata refers to the precision of the original data. Further quality control effort and explorative data analysis seems required since it would probably reveal other such type of issues in the GBIF data. For multiple species records at the same location there is little information available to infer which one is dominant from a remote sensing perspective and records with duplicates have thus been disregarded in a final step.

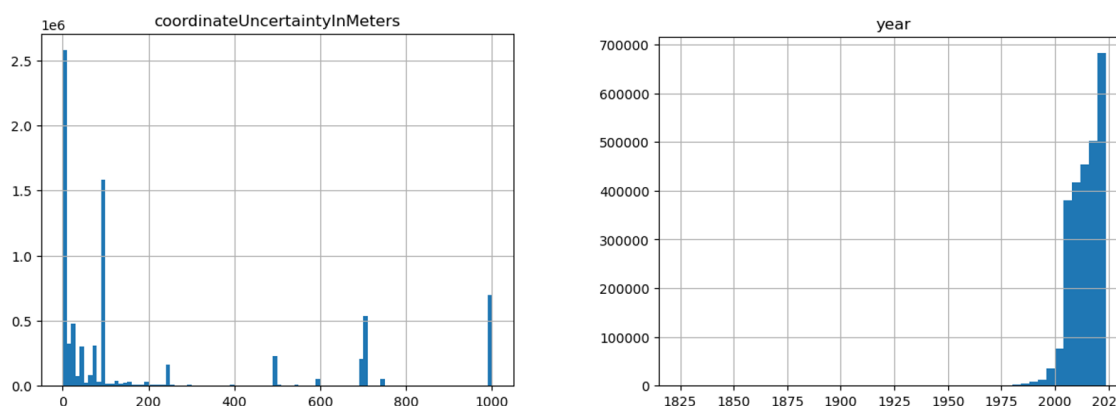


Figure 4-10: Histograms for the spatial uncertainty attribute and year of the observation for all records initially retrieved from the GBIF database.

Table 4-1: Overview of filtering steps applied to the tree species observations extracted from GBIF.

ID	Processing step	n records after filtering
1	Initial download	~19 Mio
2	Remove all with coordinateUncertaintyInMeters \leq 10m	2,575,859
3	Remove all observations older than the year 2000	2,522,083
4	Remove all observations with the following issues: 'IDENTIFIED_DATE_UNLIKELY', 'AMBIGUOUS_INSTITUTION', 'INDIVIDUAL_COUNT_CONFLICTS_WITH_OCCURRENCE_STATUS', 'COUNTRY_INVALID', 'DEPTH_UNLIKELY', 'CONTINENT_INVALID', 'MODIFIED_DATE_UNLIKELY', 'CONTINENT_COUNTRY_MISMATCH', 'FOOTPRINT_SRS_INVALID', 'TAXON_MATCH_SCIENTIFIC_NAME_ID_IGNORED', 'COORDINATE_REPROJECTED', 'FOOTPRINT_WKT_INVALID', 'BASIS_OF_RECORD_INVALID', 'COUNTRY_MISMATCH', 'CONTINENT_COORDINATE_MISMATCH', 'RECORDED_DATE_MISMATCH', 'GEODETIC_DATUM_INVALID', 'OCCURRENCE_STATUS_UNPARSABLE', occurrenceStatus = ABSENT	2,441,550
5	Remove all observations ingested from NFI France (actual precision only 1km)	1,974,838
6	Remove duplicates (same genus at the location)	1,644,309
7	Remove duplicates (different genera at the location)	1,130,935
8	Clip to extent of EEA38 + UK	1,073,195

After applying multiple filtering steps, approximately 1 Mio records remain for the EEA38 + UK. While a more advanced quality control of the remainder of the data was not possible within the frame of this study a few further observations are worth mentioning. The spatial density of samples is relatively high across Europe gets sparser towards South-Eastern Europe and Turkey (Figure 4-11). Mapping the species to the level 2 of tree types legend shows a class distribution which is probably not representative for the actual frequency of the different types (Figure 4-12). Picea, for example, is represented with relatively few samples in comparison with its

prevalence in statistically representative datasets such as LUCAS (cf, Figure 4-7). A possible explanation is that a typical plantation species such as *Picea* is perceived by GBIF contributors as a less relevant record in a database that is primarily dedicated to biodiversity. Another issue that might be related is the general clustering of observations in and around urban areas (Figure 4-13) which is at first-order probably a consequence of the density of potential / actual observers that contribute to GBIF. Further filtering steps seem likely necessary to at least partially correct for such biases without disregarding the numerous valuable records for European forests contained in GBIF database.

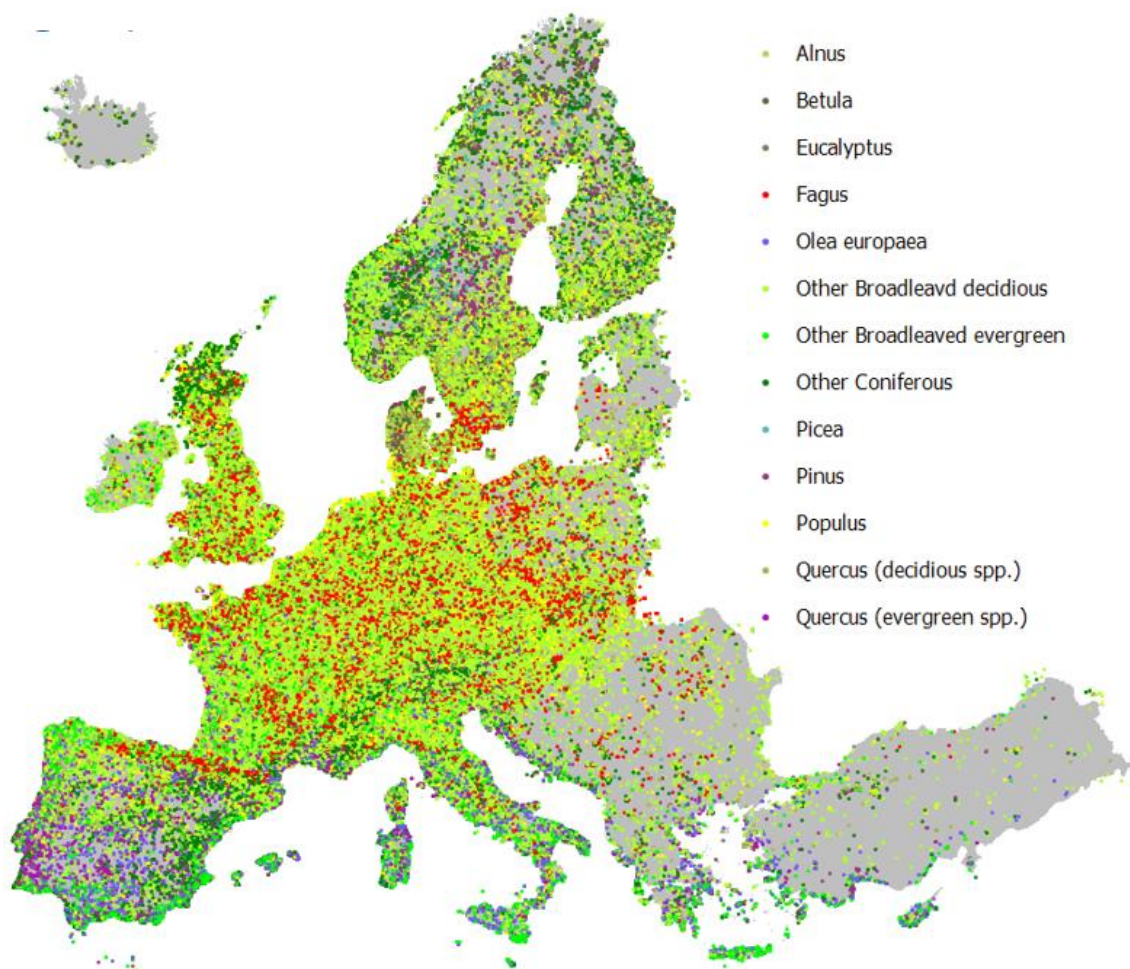


Figure 4-11 Spatial distribution of reference points in the data extracted from GBIF after filtering and clipping to the extent of the EEA38 + UK.

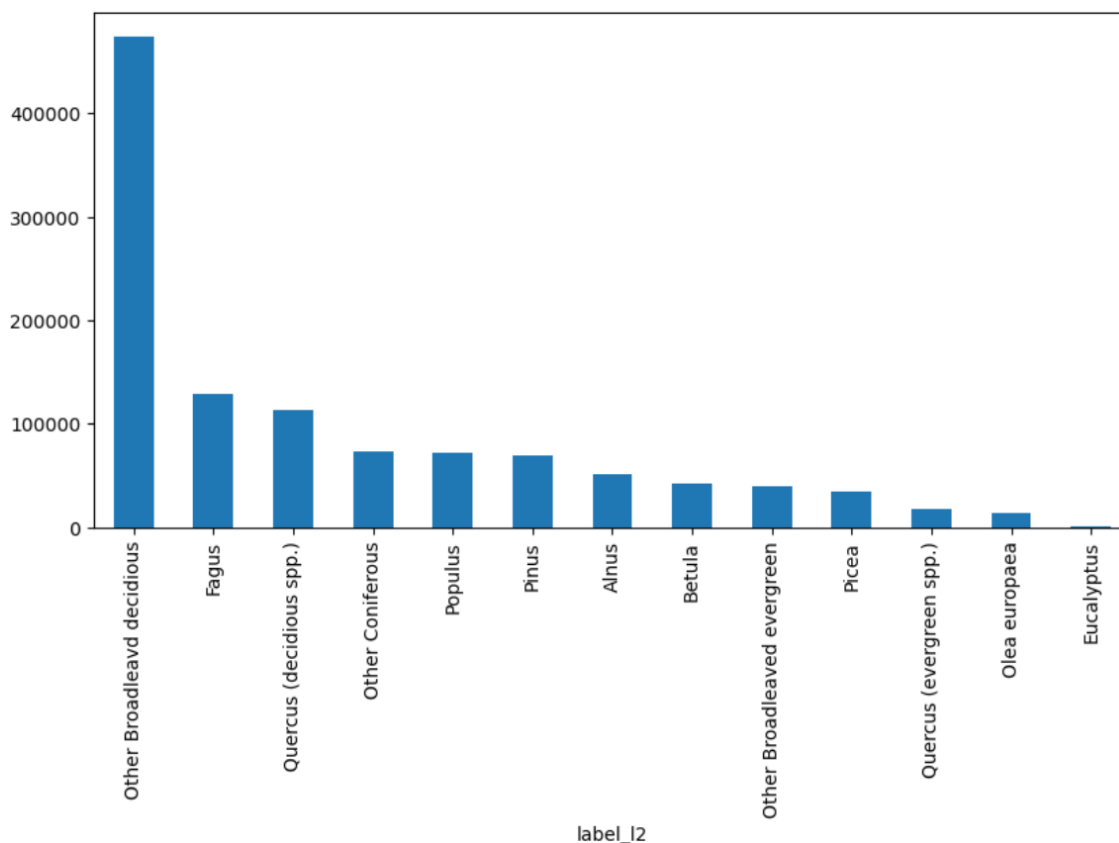


Figure 4-12: Distribution of tree types in the data extracted from GBIF after filtering and clipping to the extent of the EEA38 + UK.

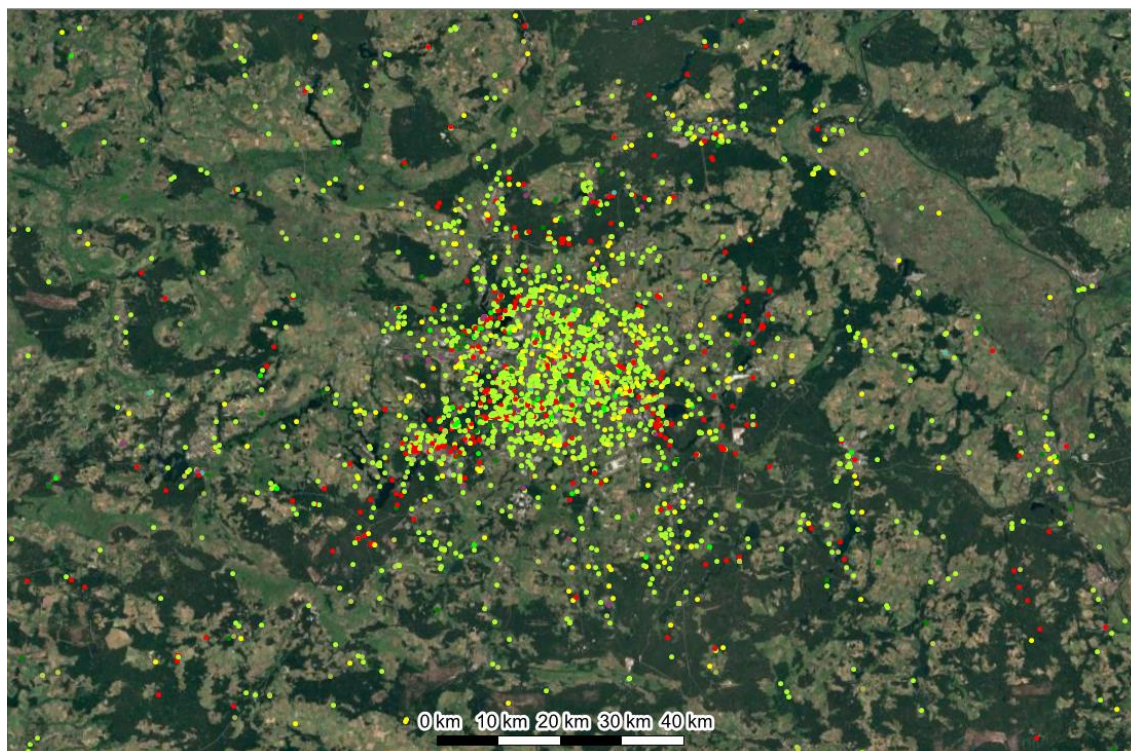


Figure 4-13 Example for spatial clustering GBIF records in urban areas. Berlin, Germany, (background Google Maps).

4.2.5 Further

Several previous works have already analysed the availability of reference data on tree species distribution for Europe (e.g. Beck et al. 2023) and a number of further datasets exist and are briefly mentioned in this section.

4.2.5.1 EU Forest

The EU-Forest dataset compiled by (Mauri et al. 2017) represents a unique database of 1,000,525 occurrence records compiled from NFI datasets (96%) and complemented by data from BioSoil and Forest Focus (4%). Since all records are georeferenced to the 1km INSPIRE grid and the temporal reference for each record is somewhat uncertain they cannot easily be linked to EO data at 10m spatial resolution. Nevertheless, the EU-Forest constitutes an important baseline dataset with broad spatial coverage and very detailed tree species information. The original Forest Focus and BioSoil data dating back to the early 2000s can currently not be made available to the public and are probably not fully suitable for EO analysis due to limitations in spatial precision (pers. comm. Giovanni Caudullo, JRC December 2024).

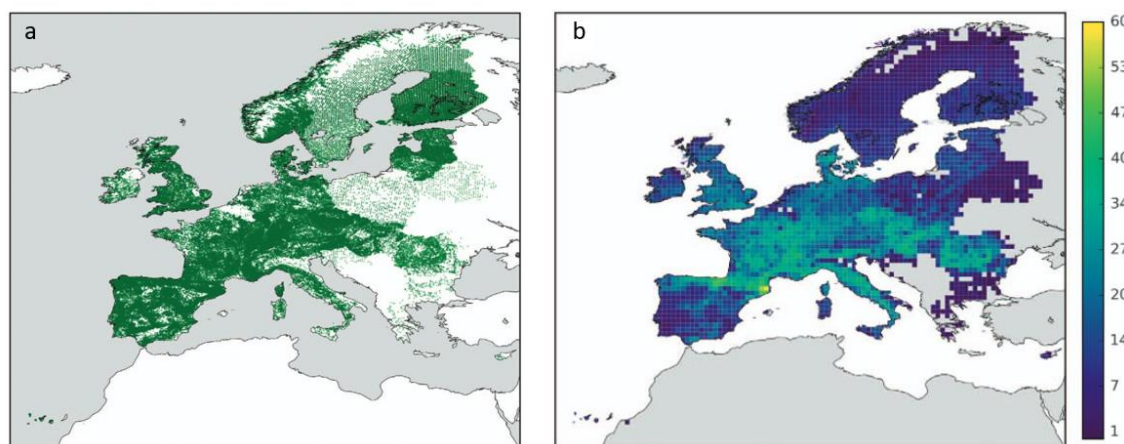


Figure 4-14: Spatial distribution of records and derived tree species richness from the EU-Forest dataset (figure from Mauri et al. 2017).

4.2.5.2 EUFGIS

The website of the European Information System on Forest Genetic Resources (EUFGIS) includes a database containing 3231 units and 113 tree species in 35 countries. By navigating the site, it is possible to extract tree species information for certain units. Typically, the available information includes the coordinates (variable precision), the size of the unit (in some cases) and the single or multiple species in the unit.

4.2.5.3 Conifers of the World, Atlas Florae Europaeae

Conifers of the World is a database accessible via <https://herbaria.plants.ox.ac.uk/bol/conifers> and contains 37,000 conifer herbarium records gathered from all continents. The data can be explored and exported from the website; however, the records are primarily of historical value, often dating back to first half of the 20th century, and have limited spatial precision. As a reference for mapping contemporary species distribution at high spatial resolution the data is probably of limited value. Similarly, the **Atlas Florae Europaeae** is a valuable database for the biodiversity of vascular plants but provides records only on a 50x50km grid.

4.2.5.4 Curated benchmark datasets

Beyond the datasets listed above with global, Pan-European and national coverages there are probably numerous further datasets with national and regional coverage that were not discovered or could not be investigated in greater depth in the frame of this study. In this context it is worth mentioning that recent benchmark datasets generated based on regional and national datasets are becoming an increasingly valuable source to fill gaps with high-quality curated reference datasets. This includes, for example, recent publication such as the TreeSatAI dataset based on forest inventory data from Lower Saxony (Ahlsweide et al. 2023), PureForest providing curated patch-based annotation based on the French Forest Inventory and the BD Forêt V2 (Gaydon and Roche 2024), the dataset described in (Weiser et al. 2022) which includes tree species labels for individual trees in 12 forest plots in the German state of Baden-Württemberg, or VHRTreeSpecies described in (Sertel and Topgul 2025) which connects VHR data with tree species labels derived from forest stand maps from the General Directorate of Forestry of Türkiye.

While the effort required to collect and harmonize numerous reference data with only regional data coverage can become prohibitively high at a Pan-European scale, the high quality of the reference data and/or coverage of areas not well represented in other datasets may still make them worthwhile in many cases.

5 Recommendations

Based on the available literature and results from previous works it can be stated that the production of a European Forest Type map following the nomenclature in (European Environment Agency 2007) is generally feasible and has already been showcased in (Giannetti and Zorzi 2023). However, a major obstacle for the derivation of an EFT map whose spatio-temporal detail is compatible with the CLMS HRL Tree Cover & Forest products, is the lack of high-resolution (10-100m) information on the observed distribution of tree species with Pan-European coverage. This report has, therefore focused on the potential EO-based tree species mapping at European and continental scale and the accessibility of reference data for training and validation.

Based on the reviewed literature and reference datasets a series of recommendations can be provided on the path forward towards an extension of the HRL Tree Cover & Forest with tree species information:

- **Thematic content**

As a working hypothesis we have adopted a classification scheme (Figure 5-1) which combines leaf type, phenology, species level and genera level information. This legend should be seen as a preliminary compromise between the requirements for an EFT map, the most common species and genera and technical feasibility. In general, it could still be considered to further extend the legend at level 2 by common genera such as *Abies*, *Pseudotsuga*, *Carpinus*, *Larix* or *Acer*. Considering the national scope of most available studies, further work is still need to determine a good trade off between high-thematic detail and high accuracy at a pan-European scale. The outcome of ongoing research projects which include prototyping of tree species maps such as EvoLand¹² and ForestPaths¹³ could provide additional relevant insights in this context.

¹² <https://www.evo-land.eu/clms-prototypes/#prototypes>

¹³ <https://forestpaths.eu/maps>

Regardless of the nomenclature selected, care should be taken to ensure that the tree type mapping is harmonized with the leaf type information already presented in the HRL DLT, avoiding contradictions between layers.

Level 1	Level 2
Broadleaved evergreen	Quercus (evergreen spp.)
	Olea europaea
	Eucalyptus
	Other Broadleaved evergreen
Broadleaved deciduous	Fagus
	Quercus (deciduous spp.)
	Betula
	Alnus
	Populus
	Other Broadleaved deciduous
Coniferous	Picea
	Pinus
	Other Coniferous

Figure 5-1 Working tree type legend adopted in this report.

• *Spatial resolution*

Given the spatial resolution of Sentinel-2 as the primary data source and the currently available HRL DLT and TCD a spatial resolution of 10m is recommended. The additional implementation of an MMU of 0.25 ha or 0.5 ha could be considered to allow a focus on homogenous stands. However, it should be considered that neither the HRL FTY nor the HRL DLT include an MMU on the leaf type datasets. An MMU on the actual tree type being depicted would therefore inevitably lead to inconsistencies with the HRL FTY and HRL DLT.

• *Target accuracy*

A series of recent publications on tree type mapping on national and continental scale show promising results for the 5-6 leading tree species and homogenous stands with accuracies from 70% to more than 90% depending on the study and the design of the validation. Further research is, however, still needed to achieve similar results for less common species / taxa and to correctly determine the dominant species in mixed stands which are currently mapped with significantly lower accuracies. In this context, overall accuracies at pan-European scale of 75- 80% might be attainable. However, the feasibility of a given target accuracy will strongly depend on the class depth and the amount of accessible high-quality training data (see point Reference Input data) below.

As stated above, determining a satisfactory trade-off between class-depth, accuracy and the effort to collect and curate training data at pan-European scale will still require considerable experimentation. While currently ongoing research projects such as EvoLand and ForestPaths are addressing some of the open questions, a ramp up phase would certainly be beneficial to iterate and test different settings before an operational roll out.

• *EO and auxiliary input data*

Sentinel-2 time-series are undoubtedly a primary input data source for a tree species classification at pan-European scale. Furthermore, most recent studies also include sets of environmental variables such as topography, latitude/longitude, temperature or precipitation which can be beneficial to consider the natural habitat boundaries of certain species. However, care must be taken when including such features to avoid degradation of spatial details or artefacts caused by native spatial resolutions significantly coarser than 10m. Some studies also

suggest the additional use of Sentinel-1 data time-series. Given the additional data volume that would need to be processed the cost-benefit of Sentinel-1 time-series should be carefully evaluated before proceeding to an operational roll out. Similarly, VHR imagery could be considered as an additional input data source, keeping in mind that this would not only increase the computational footprint of a pan-European mapping but also the complexity of the model architecture required to ingest both Sentinel time-series and VHR imagery.

- **Reference Input data**

Most studies and maps cited above rely on reference data derived from National Forest Inventories and emphasize the high importance of such high-quality data. However, as shown in chapter 4, there are currently relatively few regional and national institutions publicly accessible with precise coordinates. It therefore seems indispensable to complement available NFI data with further accessible sources such as LUCAS, ICP Forest, GBIF and other available source to gather a reference dataset which is sufficiently representative for European countries currently covered by the CLMS. Indeed, the accessibility, collection, harmonization and quality control of reference data currently constitute the most important obstacle for the calibration and validation of a high-quality pan-European tree species map. This study has enabled important first steps in mapping out and partially harmonizing various sources; further strategies to curate and quality control combined datasets are currently being explored in the frame of the EvoLand project. While more effort for data cleaning could also be undertaken during a ramp up phase (i.e. before an operational roll out of a tree type map), there are currently still some hard limitations on the shareability of NFI data that would need to be solved at a different institutional level. Ongoing projects such as monifun¹⁴ will hopefully pave the way forward for to facilitate exchange and harmonization of forest information in Europe.

- **Update cycle**

The HRL Tree Cover & Forests portfolio is currently following a yearly update cycle for the status layers and a three-yearly update cycle for change layers. Given that species composition is typically fairly stable over time an update of a tree type mapping every three years is probably sufficient.

¹⁴ <https://www.monifun.eu>

List of Abbreviations & Acronyms

Abbreviation	Name
AgERA5	Daily surface meteorological data set for agronomic use, based on ERA5
ARD	Analysis Ready Data
BD FORÊT	Reference vector database for forests of France
BFW	Bundesforschungszentrum für Wald
Biosoil	ICP Forests demonstration project
CFI2020	Carta Forestale d'Italia
CLMS	Copernicus Land Monitoring Service
DEM	Digital Elevation Model
DLM	Digital Landscape Model
DLT	Dominant Leaf Type
DOM	Départements d'outre-mer
EEA	European Environment Agency
EEA38	The 32 member and 6 cooperating countries of the EEA
EFT	European Forest Type
EO	Earth Observation
E-OBS	ENSEMBLES daily gridded observational dataset
ERA5	ECMWF Re-Analysis
ESA	European Space Agency
EU	European Union
EUFGIS	European Information System on Forest Genetic Resources (EUFGIS)
EU-Forest	high-resolution tree occurrence dataset for Europe
EU-Hydro	pan-European hydrographic dataset
FADSL	Forest Additional Support Layer
FAO	Food and Agriculture Organization
FDB	Polish Forest Data Bank
Forest Focus	ICP Forests demonstration project
FTY	Forest Type
GBIF	Global Biodiversity Information Facility
HRL	High Resolution Layer
ICP Forest	International Co-operative Programme on Forests
IGN	Institut national de l'information géographique et forestière
JRC	Joint Research Centre of the European Commission
lat	geographic latitude
LC1, LC2 and LC1_SPEC	Primary and secondary land cover attributes in the LUCAS survey
lon	geographic longitude
LTAE	Temporal Attention Encoder
LUCAS	Land Use and Cover Area Frame Statistical Survey

MCPFE	Ministerial Conference on the Protection of Forests in Europe
MFE25 / MFE50	Mapa Forestal de España at scale of 1:25000 and 1:50000
MHS	Maximum Habitat Suitability
MMU	Minimum Mapping Unit
MS-LSP	Multi-Source Land Surface Phenology
NBI	Nederlandse Bosinventarisatie, National Forest Inventory for the Netherlands
NDVI	Normalized Difference Vegetation Index
NFI	Natioanl Forest Inventory
NLI	Czech Forestry Institute (previously UHUL)
RDB	Realized Distribution Probabilities
RMSE	Root Mean Square Error
RPP	Relative Probability of Presence
S-1	Sentinel-1
S-2	Sentinel-2
SINFor	Sistema Informativo Nazionale Forestale
SPOT	Satellite Pour l'Observation de la Terre
TCD	Tree Cover Density
UK	United Kingdom

References

- Ahlswede, Steve, Christian Schulz, Christiano Gava, et al. 2023. 'TreeSatAI Benchmark Archive: A Multi-Sensor, Multi-Label Dataset for Tree Species Classification in Remote Sensing'. *Earth System Science Data* 15 (2): 681–95. <https://doi.org/10.5194/essd-15-681-2023>.
- Beck, P. S., G. Caudullo, G. Forzieri, et al. 2023. *Tree Species Distribution Data and Maps for Europe*. Publications Office of the European Union. <https://doi.org/doi/10.2760/602977>.
- Bellet, Valentine, Mathieu Fauvel, Jordi Inglada, and Julien Michel. 2024. 'End-to-End Learning for Land Cover Classification Using Irregular and Unaligned SITS by Combining Attention-Based Interpolation With Sparse Variational Gaussian Processes'. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17: 2980–94. <https://doi.org/10.1109/JSTARS.2023.3343921>.
- Blickensdörfer, Lukas, Katja Oehmichen, Dirk Pflugmacher, Birgit Kleinschmit, and Patrick Hostert. 2024. 'National Tree Species Mapping Using Sentinel-1/2 Time Series and German National Forest Inventory Data'. *Remote Sensing of Environment* 304 (April): 114069. <https://doi.org/10.1016/j.rse.2024.114069>.
- Bohn, Udo. 2000. *Karte Der Natürlichen Vegetation Europas = Map of the Natural Vegetation of Europe*. 1:2 500 000. Bundesamt für Naturschutz. <https://www.synbiosys.alterra.nl/eurovegmap/>. https://www.researchgate.net/profile/Nugzar_Zazanashvili/publication/320299343_Subarktische_boreale_nemoral-montane_Lichtwalder_sowie_subalpine_und_oromediterrane_Vegetation_Subarctic_boreal_and_nemoral-montane_open_woodlands_as_well_as_subalpine_and_oro-Mediterranean_vegetation/links/59ddf204a6fdccc2e0f38633/Subarktische-boreale-nemoral-montane-Lichtwaelder-sowie-subalpine-und-oromediterrane-Vegetation-Subarctic-boreal-and-nemoral-montane-open-woodlands-as-well-as-subalpine-and-or-Mediterranean-vege.pdf?__cf_chl_tk=Q3xvJ7_3Bcja_42y8Je0bXF7JqOODyFh.8znE9SfFv4-1752593464-1.0.1.1-5BfTNP1nLIUt5SNALHBn.S77y03sGH.S3UGmejKg9kc.
- Bonannella, Carmelo, Tomislav Hengl, Johannes Heisig, et al. 2022. 'Forest Tree Species Distribution for Europe 2000–2020: Mapping Potential and Realized Distributions Using Spatiotemporal Machine Learning'. *PeerJ* 10 (July): e13728. <https://doi.org/10.7717/peerj.13728>.
- Brus, D. J., G. M. Hengeveld, D. J. J. Walvoort, et al. 2012. 'Statistical Mapping of Tree Species over Europe'. *European Journal of Forest Research* 131 (1): 145–57. <https://doi.org/10.1007/s10342-011-0513-5>.
- Caudullo, Giovanni, Erik Welk, and Jesús San-Miguel-Ayanz. 2017. 'Chorological Maps for the Main European Woody Species'. *Data in Brief* 12 (June): 662–66. <https://doi.org/10.1016/j.dib.2017.05.007>.
- Copernicus Climate Change Service. 2020. 'E-OBS Daily Gridded Meteorological Data for Europe from 1950 to Present Derived from in-Situ Observations'. ECMWF. <https://doi.org/10.24381/cds.151d3ec6>.

- European Commission: Joint Research Centre. 2016. *European Atlas of Forest Tree Species*. Edited by G Caudullo, D De Rigo, A Mauri, T Houston Durrant, and J San-Miguel-Ayanz. Publications Office of the European Union. <https://doi.org/10.2760/776635>.
- European Environment Agency. 2007. *European Forest Types Categories and Types for Sustainable Forest Management Reporting and Policy*. 2nd edition, May 2007. EEA Technical Report No 9/2006.
- European Environment Agency. 2011. 'European Digital Elevation Model (EU-DEM)'. Prod-ID: DAT-193-En. Version EU-DEM (raster) - version 1.1. <https://sdi.eea.europa.eu/catalogue/datahub/api/records/3473589f-0854-4601-919e-2e7dd172ff50/formatters/xsl-view?output=pdf&language=eng&approved=true>.
- European Environment Agency. 2016. 'Biogeographical Regions'. Prod-ID: DAT-85-En. Version Europe 2016, ver. 1. <https://www.eea.europa.eu/en/datahub/datahubitem-view/11db8d14-f167-4cd5-9205-95638dfd9618?activeAccordion=334349>.
- European Environment Agency. 2018. 'Water and Wetness 2015 (Raster 20 m), Europe, 3-Yearly, Nov. 2020'. Version 07.00. European Environment Agency. GeoTIFF. <https://doi.org/10.2909/25ED7A97-E3A6-42ED-A403-D636D6880E6D>.
- Eurostat. 2025. 'LUCAS'. <https://ec.europa.eu/eurostat/de/web/lucas/overview>.
- FAO. 2000. *FRA 2000 ON DEFINITIONS OF FOREST AND FOREST CHANGE*. Forest Resources Assessment WP 33. The Forest Resources Assessment Programme. Food and Agricultural Organisation. <https://www.fao.org/4/ad665e/ad665e03.htm>.
- Francini, Saverio, Mart-Jan Schelhaas, Elia Vangi, et al. 2024. 'Forest Species Mapping and Area Proportion Estimation Combining Sentinel-2 Harmonic Predictors and National Forest Inventory Data'. *International Journal of Applied Earth Observation and Geoinformation* 131 (July): 103935. <https://doi.org/10.1016/j.jag.2024.103935>.
- Freudenberg, Maximilian, Sebastian Schnell, and Paul Magdon. 2025. 'A Sentinel-2 Machine Learning Dataset for Tree Species Classification in Germany'. *Earth System Science Data* 17 (2): 351–67. <https://doi.org/10.5194/essd-17-351-2025>.
- Gaydon, Charles, and Floryne Roche. 2024. 'PureForest: A Large-Scale Aerial Lidar and Aerial Imagery Dataset for Tree Species Classification in Monospecific Forests'. Version 2. Preprint, arXiv. <https://doi.org/10.48550/ARXIV.2404.12064>.
- GBIF, The Global Biodiversity Information Facility. 2025. 'What Is GBIF?' <https://www.gbif.org/what-is-gbif>.
- Giannetti, Francesca, Anna Barbati, Leone Davide Mancini, et al. 2018. 'European Forest Types: Toward an Automated Classification'. *Annals of Forest Science* 75 (1): 6. <https://doi.org/10.1007/s13595-017-0674-6>.
- Giannetti, Francesca, and Ilaria Zorzi. 2023. *Gridded Dataset of European Forest Types*. OptFor-EU D1.1. https://optforeu.eu/wp-content/uploads/2023/07/OptFor-EU_D1.1_Gridded-Dataset-of-EFT_rda.pdf.

- Grabska-Szwagrzyk, Ewa, Dirk Tiede, Martin Sudmanns, and Jacek Kozak. 2024. 'Map of Forest Tree Species for Poland Based on Sentinel-2 Data'. *Earth System Science Data* 16 (6): 2877–91. <https://doi.org/10.5194/essd-16-2877-2024>.
- Hermosilla, Txomin, Alex Bastyr, Nicholas C. Coops, Joanne C. White, and Michael A. Wulder. 2022. 'Mapping the Presence and Distribution of Tree Species in Canada's Forested Ecosystems'. *Remote Sensing of Environment* 282 (December): 113276. <https://doi.org/10.1016/j.rse.2022.113276>.
- Hermosilla, Txomin, Michael A. Wulder, Joanne C. White, Nicholas C. Coops, Christopher W. Bater, and Geordie W. Hobart. 2024. 'Characterizing Long-Term Tree Species Dynamics in Canada's Forested Ecosystems Using Annual Time Series Remote Sensing Data'. *Forest Ecology and Management* 572 (November): 122313. <https://doi.org/10.1016/j.foreco.2024.122313>.
- Mauri, Achille, Giovanni Strona, and Jesús San-Miguel-Ayanz. 2017. 'EU-Forest, a High-Resolution Tree Occurrence Dataset for Europe'. *Scientific Data* 4 (1). <https://doi.org/10.1038/sdata.2016.123>.
- Mouret, Florian, David Morin, Milena Planells, and Cécile Vincent-Barbaroux. 2024. 'Tree Species Classification at the Pixel-Level Using Deep Learning and Multispectral Time Series in an Imbalanced Context'. Version 2. Preprint, arXiv. <https://doi.org/10.48550/ARXIV.2408.08887>.
- Pividori, M, F Giannetti, A Barbati, and G Chirici. 2016. 'European Forest Types: Tree Species Matrix'. *European Atlas of Forest Tree Species*, e01f162.
- Rivas-Martínez, S., A. Penas, and T.E. Díaz. 2004. 'Bioclimatic Map of Europe'. Cartographic Service. University of León, Spain. <https://web.archive.org/web/20061014163920/http://www.globalbioclimatics.org/form/maps.htm>.
- Schadauer, Klemens, Rasmus Astrup, Johannes Breidenbach, et al. 2024. 'Access to Exact NATIONAL FOREST INVENTORY Plot Locations Must Be Carefully Evaluated'. *New Phytologist* 242 (2): 347–50. <https://doi.org/10.1111/nph.19564>.
- Senf, Cornelius, and Rupert Seidl. 2020. 'Mapping the Forest Disturbance Regimes of Europe'. *Nature Sustainability* 4 (1): 63–70. <https://doi.org/10.1038/s41893-020-00609-y>.
- Sertel, Elif, and Sule Nur Topgul. 2025. 'Comparative Analysis of Deep Learning Approaches for Forest Stand Type Classification: Insights from the New VHRTreeSpecies Benchmark Dataset'. *International Journal of Digital Earth* 18 (1). <https://doi.org/10.1080/17538947.2025.2522394>.
- Weiser, Hannah, Jannika Schäfer, Lukas Winiwarter, Nina Krašovec, Fabian E. Fassnacht, and Bernhard Höfle. 2022. 'Individual Tree Point Clouds and Tree Measurements from Multi-Platform Laser Scanning in German Forests'. *Earth System Science Data* 14 (7): 2989–3012. <https://doi.org/10.5194/essd-14-2989-2022>.

