



European Forest Types: toward an automated classification

Francesca Giannetti¹ · Anna Barbati² · Leone Davide Mancini² · Davide Travaglini¹ · Annemarie Bastrup-Birk³ · Roberto Canullo⁴ · Susanna Nocentini¹ · Gherardo Chirici¹

Received: 22 March 2017 / Accepted: 13 November 2017
© INRA and Springer-Verlag France SAS, part of Springer Nature 2017

Abstract

• **Key message** The outcome of the present study leads to the application of a spatially explicit rule-based expert system (RBES) algorithm aimed at automatically classifying forest areas according to the European Forest Types (EFT) system of nomenclature at pan-European scale level. With the RBES, the EFT system of nomenclature can be now easily implemented for objective, replicable, and automatic classification of field plots for forest inventories or spatial units (pixels or polygons) for thematic mapping.

• **Context** Forest Types classification systems are aimed at stratifying forest habitats. Since 2006, a common scheme for classifying European forests into 14 categories and 78 types (European Forest Types, EFT) exists.

• **Aims** This work presents an innovative method and automated classification system that, in an objective and replicable way, can accurately classify a given forest habitat according to the EFT system of nomenclature.

• **Methods** A rule-based expert system (RBES) was adopted as a transparent approach after comparison with the well-known Random Forest (RF) classification system. The experiment was carried out based on the information acquired in the field in 2010 ICP level I plots in 17 European countries. The accuracy of the automated classification is evaluated by comparison with an independent classification of the ICP plots into EFT carried out during the BioSoil project field survey. Finally, the RBES automated classifier was tested also for a pixel-based classification of a pan-European distribution map of beech-dominated forests.

This article is part of the Topical Collection on *ICP Forests*

Handling editor: Walter Seidling

Contribution of the co-authors Conceptualization: G. Chirici; methodology: G. Chirici, A. Barbati, R. Canullo, F. Giannetti; software: F. Giannetti; formal analysis: G. Chirici, A. Barbati, A. Bastrup-Birk; data curation: F. Giannetti, L.D. Mancini; writing—original draft: F. Giannetti; writing—review & editing: G. Chirici, A. Barbati, R. Canullo, F. Giannetti, A. Bastrup-Birk, S. Nocentini, D. Travaglini, S. Nocentini, D. Travaglini.

✉ Gherardo Chirici
gherardo.chirici@unifi.it

Francesca Giannetti
francesca.giannetti@unifi.it

Anna Barbati
barbati.sisfor@unitus.it

Leone Davide Mancini
leone.mancini@unitus.it

Davide Travaglini
davide.travaglini@unifi.it

Annemarie Bastrup-Birk
Annemarie.Bastrup-Birk@eea.europa.eu

Roberto Canullo
roberto.canullo@unicam.it

Susanna Nocentini
susanna.nocentini@unifi.it

¹ Department of Agricultural, Food and Forestry System, Università degli Studi di Firenze, Via San Bonaventura, 13-50145 Florence, Italy

² Department for Innovation in Biological, Agro-Food and Forest System, Università degli Studi della Tuscia, Viterbo, Italy

³ European Environmental Agency, Kongens Nytorv 6, 1050 Copenhagen, Denmark

⁴ Plant Diversity and Ecosystems Management Unit, School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy

- **Results** The RBES successfully classified 94% of the plots, against a 92% obtained with RF. When applied to the mapped domain, the accuracy obtained with the RBES for the beech forest map classification was equal to 95%.
- **Conclusion** The RBES algorithm successfully automatically classified field plots and map pixels on the basis of the EFT system of nomenclature. The EFT system of nomenclature can be now easily and objectively implemented in operative transnational European forest monitoring programs.

Keywords European Forest Type · Expert system · Classification · GIS · Vegetation · Algorithm · ICP forests

1 Introduction

Widespread challenges related to sustainable forestry and forest management, conservation and restoration of habitats, biodiversity maintenance, land management and monitoring, and climate change have generated a global need for classification of natural habitats and ecosystems (Vaz et al. 2015; Flanagan et al. 2015; Buffa and Villani 2012; Dengler et al. 2009; Song et al. 2005). Ecological or habitat classification is targeted primarily to the creation of a “common language” (Flanagan et al. 2015; Barbati et al. 2014, Mc Roberts et al. 2011; Shumchenia and King 2010) that can be used to describe, assess, monitor, and manage ecological communities.

A classification is needed to describe the diversity of forest ecosystems, to analyze their change over long time series, spanning over a full range of spatial and geographic scales, and providing knowledge of reference conditions and current states of ecosystems to make decisions about conservation and resource management (Faber-Langendoen et al. 2014). Vegetation classification systems are based on a set of criteria, including physiognomy (growth forms, structure) and floristics (compositional similarity and characteristic species combinations), in conjunction with ecological characteristics, including site factors, disturbance, bioclimate, and biogeography (Faber-Langendoen et al. 2014; Barbati et al. 2014; Chytrý 2012; EEA 2006).

More specifically, a “forest type classification scheme” is aimed at stratifying a large forest area (e.g., stocked forest land) into smaller and more homogeneous units to facilitate the analysis, interpretation, and reporting of forest data (Barbati et al. 2014; Mc Roberts et al. 2011).

In this regard, the increasing availability of large forest plot datasets all over the world (Ewald 2003) could be the basis for a well-founded supra-regional vegetation classification (Dengler et al. 2009). Several authors have proposed different approaches that allow the consistent classification of such data (Bruehlheide 2000; Chytrý et al. 2002; Dengler et al. 2009; Knollová et al. 2005; Illyés et al. 2007). The vegetation classification over large geographical areas and whole continents is nowadays a realistic aim (Dengler et al. 2009; Rodwell et al. 2002; Peet et al. 2001).

The revision of the European Nature Information System (EUNIS) terrestrial habitat classification, based on georeferenced vegetation plot data, has recently led to the identification of 37 distinct forest habitat types (EUNIS level 3),

considering stocked forest land only (EEA 2015a, 2015b). The EUNIS classification provides a pan-European reference for documenting, monitoring, and assessing the quality of habitats at the European level (<http://eunis.eea.europa.eu/>). On the other hand, the EUNIS forest types are difficult to use in monitoring or inventories as reference classification due to, e.g., a high number of classes and different spatial coverage.

The European Forest Types (EFT) classification system has proven to be operational and easier to use to facilitate understanding, interpretation, and communication of data on indicators describing the status and trends of forests, and forest management in Europe. The EFT were proposed (EEA 2006; Barbati et al. 2007) as a reference scheme for reporting on sustainable forest management (SFM) indicators, including those related to the assessment of biodiversity, according to ecologically homogeneous strata at pan-European level (EEA 2006; Forest Europe 2011; Mc Roberts et al. 2011). A recent revision of the EFT classification allows classifying stocked forest land of the pan-European region into 14 categories and 78 types (Barbati et al. 2016; Caudullo et al. 2016). The 14 categories represent groups of ecologically distinct forest communities dominated by specific assemblages of tree species, either native to Europe (categories 1–13) or introduced/alien (category 14). The exact definition of boundaries between classes is a common problem in nomenclature schemes adopted in vegetation classification (Mucina 1997). In the EFT scheme, a set of classification rules is provided with the nomenclature, in order to allow the user to assign in the field forest plots to categories and types (Barbati et al. 2006). The set of classification rules is based on (i) the main factors (biogeographical, climatic, elevation, edaphic conditions) that determine the latitudinal/altitudinal zonation of European forest vegetation or, otherwise, the appearance of azonal communities (hydrological regimes); (ii) the identification of dominant (single or group) of tree species, based on basal area/crown cover information.

The 78 types are designed to further describe the variety of forest communities covered by each category, and the breaking points between different types are defined in terms of species compositions. The European Atlas of Forest Tree Species recently documented by a tree species matrix the list of dominant or secondary trees species that characterize each class of the EFT nomenclature (Pividori et al. 2016).

We refer to Barbati et al. (2014) for a detailed description of the main characteristics of each category and type; nevertheless, key ecological and compositional features of the 14 categories are briefly summarized in [Annex 1](#).

EFT were extensively used in several research projects, for example, for documenting the presence of mixed-species stands in European forests (Bravo-Oviedo et al. 2014), quantifying fuel flammability of forest vegetation (Corona et al. 2014), and defining post-fire management measures in forest of burned areas (Barbati et al. 2010).

In the State of European Forests 2011 (SOEF 2011) (FOREST EUROPE 2011), EFT were operationally implemented in a pilot project to report on some of the SFM quantitative indicators: forest area and its variation, growing stock volume, and deadwood (Barbati et al. 2011). Twenty-eight countries reported the indicators stratified according to the 14 EFT categories. These countries represent approximately 83% of the EU forest area (excluding the Russian Federation). The EFT allowed to report the pan-European SFM indicators according to ecologically sound units to frame indicators values. However, EFT-based reporting has not been implemented in the latest State of European Forests (SOEF 2015; FOREST EUROPE 2015). The SOEF 2015, in fact, applies a very simplified forest type classification and only for reporting on forest area (predominantly coniferous, predominantly broadleaved and mixed forests) and growing stock (broadleaves and conifers). By consequence, the aggregation of indicator values into broadly defined species groups imply reducing the information conveyed, especially in the countries characterized by a forest area with high variability in ecological conditions and level of naturalness. For instance, this can lead some countries to aggregate SFM indicators of non-native coniferous plantations (e.g., Sitka spruce, Douglas fir) with those of native conifer stands (e.g., Scot pine, fir, spruce). Likewise, in the countries with a wide bioclimatic variability (e.g., Italy, France, and Spain), the indicators of coniferous forests of the Alpine biogeographical region are averaged with those of the Mediterranean ones. In addition, the three forest type classes adopted by SOEF 2015 (FOREST EUROPE 2015) are not helpful in monitoring differential phenomena of gains and losses in forest area, e.g., climate change-related. That is, a stable value of broadleaved dominated forest area may actually hide gains and losses in the forest area covered by drought-sensitive and drought-tolerant forest types (e.g., a decline and substitution of beech forest by thermophilous deciduous species).

Though the motive for the decision of reporting SOEF indicators by these broadly defined forest types is not reported by any official document, it is reasonable to assume that the need of reducing the reporting burden for national correspondents represents a key factor. In fact, experts responsible for compiling country-level information for the SOEF process, in personal communications, complained

about the lack of clear quantitative criteria to unambiguously assign sampling units (e.g., National Forest Inventories plots) to the EFT classes. For this reason, it appears particularly useful to develop a straightforward method for the objective classification of European forests according to the EFT system of nomenclature.

Classification of a forest area according to a given forest type scheme should be usually accomplished by field crew during a field survey (Westfall 2009). However, this step is subjective and relies on the personal experience of the personnel involved (Hédél 2007; Westfall 2009). Consequently, when large monitoring programs are carried out by different field crews, the information collected in the field can be affected by inconsistency, due to the fact that different crews may classify the same forest area according to different forest types (Westfall 2009).

Remote sensing techniques and Geographic Information Systems (GIS) support the production and availability of digital maps of environmental variables (Ioannis et al. 2006) such as biogeographical regions (EEA 2016), climate data (Hijmans et al. 2005), global bioclimatic classification systems (Rivas-Martínez et al. 2004), orography, geology, soils (Panagos et al. 2011), water bodies (EEA 2013a), wetlands (EEA 2013b), and tree species distribution (San-Miguel-Ayanz et al. 2016). This information opens up for new possibilities to develop automated vegetation classification models (Millington et al. 2002), to map the presence and abundance of a given species (Duveneck et al. 2015), community presence (Ackers et al. 2015; Zimmermann and Kienast 1999), community structure and composition (Ohman and Gregory 2002; Adamo et al. 2015), or forest types (Beard et al. 2013).

Formalized classifications systems which attempt to derive vegetation maps at the continental scale level are evolving in recent years (Jiménez-Alfaro et al. 2014), since they are crucial for supporting nature conservation actions and forest planning in the European Union (Douda et al. 2016).

Several methods were applied to test for the automated classification of vegetation and forest habitats on the basis of data mining or artificial intelligence approaches in a GIS environment (Adamo et al. 2015; Openshaw and Openshaw 1997). Automated classification methods most commonly consist of induction rules, decision trees, random forest, rule-based expert system (RBES) (Pérez-Ortiz et al. 2016; Capelo et al. 2007; Cook et al. 1996; Andrew 1996), artificial neural networks (Wang et al. 2005), genetic algorithms (Chen and Yao 2008), and support vector machines (Torbick et al. 2007). Among these, RBES, also known as knowledge-based systems (Andrew 1996; Capelo et al. 2007; Pérez-Ortiz et al. 2016), is one of the most widely applied approaches for vegetation classifications since they solve the problems by rules derived from human expert knowledge (Grunwald 2009; Hayes-Roth 1985).

An expert system can be defined as “a hierarchy of rules, or a decision tree that describes the conditions under which a set of low level constituent information (user-defined variables, raster imagery, vector coverages, spatial models, external program and simple scalars) gets abstracted into a set of high level information classes” (Bingyuan et al. 2014). Conventional RBES uses human expert knowledge to solve real-world problems that normally would require human intelligence.

Several scientists demonstrated the potential of RBES applied to spatial data to improve the accuracy of landscape and ecological classifications (Gao et al. 2004). RBES offers a way to codify information and use it to make predictions (Czajkowski et al. 2007) based on basic axioms of logical set theory usually consisting of *if-then* rules (Robinove 1986). Further, it offers important advantages in terms of simplicity and computational efficiency (Robinove 1986).

The aim of this study was to develop and test a spatially explicit algorithm that automatically classifies a given forest habitat according to the EFT nomenclature (Barbati et al. 2014). Two different non-parametric classification approaches were tested: (i) a rule-based expert system (RBES) classification algorithm and (ii) a Random Forests (RF) classification tree.

The two systems were tested to classify by EFT categories 2010 level I plots from the Forest Focus/International Cooperative program on assessment and monitoring of air pollution effects on forest (ICP Forests), based on data collected in the field in the framework of the BioSoil-Biodiversity project (JRC 2011). After having analyzed the results achieved, the automated classifier that performed better (the RBES) was then tested for a pixel-based classification of a raster pan-European distribution map of beech-dominated forests into the two corresponding EFT categories (6—beech forest and 7—mountainous beech forest).

2 Materials and methods

2.1 Study area

The two tested classification algorithms, RBES and RF, have been developed to classify stocked forest land of the pan-European region according to the EFT classification. The available sample data to test the classifiers cover 17 EU countries where ICP Forests Level I data on forest structural variables have been collected in the framework of the BioSoil Forest Biodiversity Demonstration project (JRC 2011): Austria, Belgium, Czech Republic, Denmark, Estonia, France, Finland, Germany, Ireland, Italy, Latvia, Poland, Slovak Republic, Slovenia, Spain, Sweden, and UK. These countries on the basis of the SOEF 2015 (with the exclusion of the Russian Federation) represent 52% of the European area and 62% of the European forest area (Forest Europe 2015).

2.2 Data

2.2.1 ICP level I biosoil biodiversity plots

The two systems were tested on 2010 level I plots acquired in the field between 2005 and 2008 from the Forest Focus/International Cooperative program on assessment and monitoring of air pollution effects on the forest (ICP Forests) collected in the field in the framework of the BioSoil-Biodiversity project. As input data, we used the basal area per tree species and the best available wall-to-wall pan-European digital maps of several environmental variables.

Raw data were obtained and pre-elaborated in the framework of the official activities of the UNECE ICP Forests Network (Working Group on Forest Biodiversity 2007). The BioSoil plot is a circle with an area of 2000 m² (radius of 25.24 m) (Puletti et al. 2017). The database includes information on a number of structural and compositional variables (Bastrup-Birk et al. 2007; Hiederer and Durrant 2010). For testing the RBES algorithm, we used the following data: geographic position of the plots (x and y coordinates), diameter at breast height (1.30 m) (DBH), and species of each callipered tree. A total of 105 forest tree species was found in the field in the 2010 plots.

In the BioSoil project, each plot in the field was assigned to an EFT category according to the 14 categories documented in EEA (2006). For the purposes of this study, the information collected in the field is considered error free and is used as reference truth to quantify the accuracy of the classification algorithms.

2.2.2 Current beech forest distribution map

The current beech forest distribution map was created by the JRC at 1-km resolution in the framework of the European Atlas of Forest Tree Species (San-Miguel-Ayanz et al. 2016). The map was used to test the automated classification at pixel level. The map relies on statistical interpolation of field observations and remotely sensed information, performed with the Constrained Spatial Multi-Frequency Analysis (C-SMFA). Further details are provided by de Rigo et al. (2016). To test the classification system, we used the pixels where the relative probability of presence of beech is $\geq 50\%$ (i.e., pixels of the map with a value ≥ 0.5) in order to derive a map for beech-dominated forest areas.

2.2.3 Digital maps

Six digital thematic maps were used in this study as input geodatasets for the automated classifiers.

The *map of biogeographic regions of Europe*, prepared in scale 1:10,000,000 by the EEA (EEA 2015a, 2015b), which contains the official delineations of nine regions. The map was developed for the Habitats Directive (92/43/EEC) including the design of the Natura 2000 and EMERALD networks related to the implementation of the Convention on the Conservation of European Wildlife and Natural Habitats (Bern Convention) (Council of Europe 1979). The *Bioclimatic Map of Europe* (Rivas-Martínez et al. 2004), at a 1:16,000,000 scale, identifies the thermoclimatic belts of Europe in 5 regions, 9 subregions, 34 provinces, and 88 subprovinces (Rivas-Martínez et al. 2004). The *Natural Vegetation map of Europe*, in a scale of 1:10,000,000 (Bohn et al 1994), was used to map the potential range of distribution of oligotrophic soils where acidophilous oakwoods can occur (class F1, see rule 5 hereinafter described).

From the Copernicus Land Monitoring Services project at the European Environment Agency (EEA) (<http://land.copernicus.eu/pan-european/high-resolution-layers>), we used two high-resolution layers, both available as raster boolean maps having a geometric resolution of 20 m: (a) *wetlands* and (b) *permanent waterbodies*. Finally, we used the *pan-European Digital Elevation Model* (DEM) with an original spatial resolution of 1 arc sec available at EEA (2010).

2.2.4 Non-native tree species database

We used the list of tree species identified as alien at country level by the project DAISIE—Delivering Alien Invasive Species Inventories for Europe (<http://www.europe-aliens.org/>). The 105 tree species surveyed in ICP plots were accordingly reclassified as non-native (if included in the DAISIE list) or native (if not included in the DAISIE list) in the 17 investigated countries.

2.2.5 Methodology

For each plot, we calculated the percent basal area of each tree species surveyed out of the total basal area of the plot and, on the basis of the geographic location of the center of the plots, we extracted the information from the six digital maps.

2.2.6 Rule-based expert system EFT algorithm

According to Barbati et al. (2014), a species is considered dominant when its basal area per plot represented at least 50% of the total plot basal area. Species that represented less than 5% of the total plot basal area were not considered. The RBES algorithm is a system of conditional (*if-then*) rules which recursively partitions the data set into smaller subdivisions on the basis of a set of rules defined at each branch of the decision tree.

- Rule 1: we classified in forest category 14 (introduced and alien tree species forest) all the plots where the dominant species was not native at the country level, according to the DAISIE database.
- Rule 2: forest categories 11–12 correspond to azonal forest communities characterized by specific hydrological regimes. We assigned to category 12 (floodplain forests) the plots within a 1-km buffer created from the COPERNICUS permanent waterbodies high-resolution layer (EEA 2013a, 2013b), and in category 11 (mire and swamp forest), the plots falling within areas mapped by the COPERNICUS wetland high-resolution layer (EEA 2013a, 2013b).
- Rule 3: forest categories 1, 2, and 3 are defined on the basis of the biogeographical regions and species composition. The geographic boundaries of the Boreal and of the Alpine biogeographical regions, required for the classification of categories 1 (boreal forest) and 3 (Alpine coniferous forest) respectively are taken from the map of biogeographic regions of Europe (EEA 2015a, 2015b). The limit of the hemiboreal area where category 2 (nemoral coniferous and mixed broadleaved coniferous forest) is distributed is instead delineated from the unit “temperate hemiboreal bioclimate” of the Bioclimatic Map of Europe (Rivas-Martínez et al. 2004).
- Rule 4: forest categories 6 and 7 are beech-dominated communities with a different altitudinal zonation. In Mediterranean and Alpine biogeographic regions, the plots with altitude less than 650 m a.s.l. were assigned to category 6, the others to category 7; in Continental, Atlantic, Pannonian, Hemiboreal, and Boreal regions, the altitude threshold for category 6 was set to 350 m.
- Rule 5: plots were assigned to category 4 (acidophilous oak and oak-birch forest) on the basis of species composition and spatially congruence with geographical distribution of acidophilous oakwoods (F1 class) delineated in the Natural Vegetation map of Europe (Bohn 1994).
- Rule 6: to classify the remaining categories 5 (mesophytic deciduous forest), 8 (thermophilous deciduous forest), 9 (broadleaved evergreen forest), and 13 (non-riverine alder, birch, or aspen forest), we used the information on the tree species composition only.

The information available in the matrix of European Forest Types’ dominant tree species published in the European Atlas of Forest Tree Species (Pividori et al. 2016) was used for identifying dominant trees to classify the plots. The system was implemented in an R code, using the packages *raster* (Hijmans 2015), *shape* (Soetaert 2014), and *spatial.tools*

(Greenberg 2014). The overall decision tree used in the RBES is available as [Annex 2](#) to this contribution.

2.2.7 Random Forest classification tree

RF classification tree is a non-parametric “ensemble learning” algorithm based on a machine learning approach (Breiman 2001). Classification trees create a set of rules for binary partitioning input dataset in regions (called nodes) which are as much homogeneous as possible on the basis of the available predicting variables (Breiman et al. 1984). RF is an enhancement of traditional decision trees because it consists of a large number of trees, and the final classification is determined as the majority of the results obtained by the trees. In this study, we used RF as it is implemented in R-cran RandomForest package on the basis of 500 trees (Liaw and Wiener 2002). The code permits to assess the total accuracy of the classification with a bootstrapping approach and the importance of each specific predictor through the Mean Decrease Gini coefficient (Liaw and Wiener 2002).

2.2.8 Accuracy assessment

The accuracy of the two classification approaches was evaluated against the EFT classification at category level assigned in the field. To do so, we first created a confusion matrix comparing the 2010 plots in terms of field-based and

algorithm attributions to EFT categories. On the basis of the confusion matrix, we calculated the overall accuracy as the percent of plots, where the two classifications agree out of a total of 2010 plots and the kappa index of agreement (KIA) (Cohen 1960); for each category, we also calculated omission (EO) and commission (EC) errors and the user and producer accuracy.

We followed the same approach also to evaluate the accuracy of the beech map produced for the two EFT categories (6 and 7). In this later case, we compared the result of the classification for the pixels relative to the position of the 2010 ICP plots.

Data availability The R code of the algorithm is available from the corresponding author upon request.

3 Results

3.1 ICP classification

The overall accuracy of the EFT classifications was high, 0.94 and 0.90, and the KIA 0.64 and 0.60 for RBES and RF, respectively (Tables 1 and 2, Fig. 1).

The errors of omission of the RBES algorithm ranged between 0.09 (category 10) and 0.219 (category 11), while the errors of commission ranged between 0 (categories 9) and 0.34 (category 14) (Table 1).

Table 1 The confusion matrix is calculated by comparing the EFT category assigned to forest plots in the field with the EFT category returned by the RBES EFT algorithm. Overall accuracy = 0.94, KIA = 0.64 (see [Annex 1](#) for categories description) In Italics the number of plot where the EFTs classification obtained by RF classification is congruent with the ones attributed in the Field by crew.

RBES	Field classification														Total	User accuracy
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
1	<i>574</i>			1		1								3	579	0.99
2		<i>195</i>		1	1	1					1	1		1	201	0.97
3			<i>212</i>		4	1	1	1			1			2	222	0.95
4				<i>16</i>	1			2						1	20	0.80
5	2		2		<i>54</i>	1	4	1				1			65	0.83
6	4	1		5		<i>140</i>		1						2	153	0.92
7	1		4		1	6	<i>326</i>	5						9	352	0.93
8	1					2		<i>128</i>				2		12	145	0.88
9									<i>55</i>					1	56	0.98
10										<i>10</i>					10	1.00
11				1	4	1					32	2	1		41	0.78
12					1						1		<i>30</i>		32	0.94
13					1		1	1				1	1	22	27	0.81
14		4	1			3		4				2	1	92	107	0.86
Total	582	200	219	24	67	156	332	143	55	11	35	39	24	123	<i>2010</i>	
Producer accuracy	0.99	0.98	0.97	0.67	0.81	0.90	0.98	0.90	1.00	0.91	0.91	0.77	0.92	0.75		

In Italics the number of plot where the EFTs classification obtained by RF classification is congruent with the ones attributed in the Field by crew

Table 2 The confusion matrix is calculated by comparing the EFT category assigned to forest plots in the field with the EFT category returned by the RF classification tree. Overall accuracy = 0.92, KIA = 0.60 (see Annex 1 for category description)

RF	Field classification														Total	User accuracy
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
1	<i>579</i>										26	23	7	4	639	0.91
2		<i>196</i>		2	1	2	7	1			1	3	1	4	218	0.90
3			<i>210</i>		2	3	3			2	1			3	224	0.94
4				<i>17</i>	1									1	19	0.89
5					<i>58</i>			1				1			60	0.97
6	3	1		5		<i>145</i>	10	1							165	0.88
7		3	8		1	6	<i>312</i>	1						3	334	0.93
8			1					<i>139</i>					1	1	142	0.98
9									<i>55</i>	1				1	57	0.96
10										<i>8</i>					8	1.00
11											<i>7</i>				7	1.00
12												<i>8</i>			8	1.00
13					3							1	<i>15</i>		19	0.79
14					1								3	<i>106</i>	110	0.96
Total	582	200	219	24	67	156	332	143	55	11	35	39	24	123	2010	
Producer accuracy	0.99	0.98	0.96	0.71	0.87	0.93	0.94	0.97	1.00	0.73	0.20	0.21	0.63	0.86		

In Italics the number of plot where the EFTs classification obtained by RF classification is congruent with the ones attributed in the Field by crew

The errors of omission of RF classification tree ranged between 0 (category 10, 11, and 12) and 0.138 (category 6), while the error of commission ranged between 0 (category 9) and 4.0 (category 11) (Table 2).

Using the species composition only in the RF algorithm, the overall accuracy dropped to 0.62, thus indicating that information from the geospatial dataset is critically relevant for the EFT classification. The ranking order of these predictors evaluated on the basis of the indicator Mean Decrease gini (MDG) was the following: latitude (MDG = 253.73); bioregions (MDG = 212.51); elevation (MDG = 115.42); longitude (MDG = 97.13); soils (MDG = 14.00); proximity to river (MDG = 9.10); and presence of wetland (MDG = 1.34)

Per-class user's accuracies (UA), values higher than 0.85 were reached in most EFT; the lowest UA value was 0.78 for category 11. Producer's reliabilities are significantly higher in RBES than in RF (ranging between 0.2 and 0.4) for some EFT (10 to 13 classes). Per-class producer's accuracy ranges from good (~ 0.7 for class 4) to perfect (class 9) in RBES, and from very low (~ 0.2 for class 11 and 12) to perfect (class 9) in RF.

3.2 Beech map classification

Since the RBES algorithm produced slightly better results, we used it to produce a classification of the current distribution of beech forests in the EFT categories 6 - beech forest and 7 - mountain beech forest (Fig. 2).

The overall accuracy of the map evaluated on the basis of the ICP forest plots classified in the field was equal to 0.93 (Table 3). In (Fig. 1) it is reported the EFTs classification obtained by RBES of the ICP Forest plots.

4 Discussion

Both RBES and RF classification algorithms produced very satisfactory results in the automated classification of the European Forest Types categories. The RBES, compared to the RF approach, is preferable for two main reasons: (1) because RBES is a totally transparent approach that can be easily replicated with different input data from field surveys or thematic maps; (2) because the RBES in our test demonstrated a better performance, in terms of producer accuracy.

The two classification approaches produced similar results for categories 1, 2, 3, 4, and 9. RF produced better results for categories 4, 5, 6, 8, and 14 while the RBES performed better for categories 7, 10, 11, 12, and 13. Analyzing in detail the source of errors in the misclassification of categories 4, 5, 6, and 8, we found that in 63% of these plots the classification in the field was affected by some inconsistency between dominant species, as quantified by basal area data, and the EFT category identified in the field. Two possible explanations can be hypothesized. First, clear rules linking tree species composition to forest categories, as presented in Pividori et al. (2016), were not available at the time when the BioSoil project

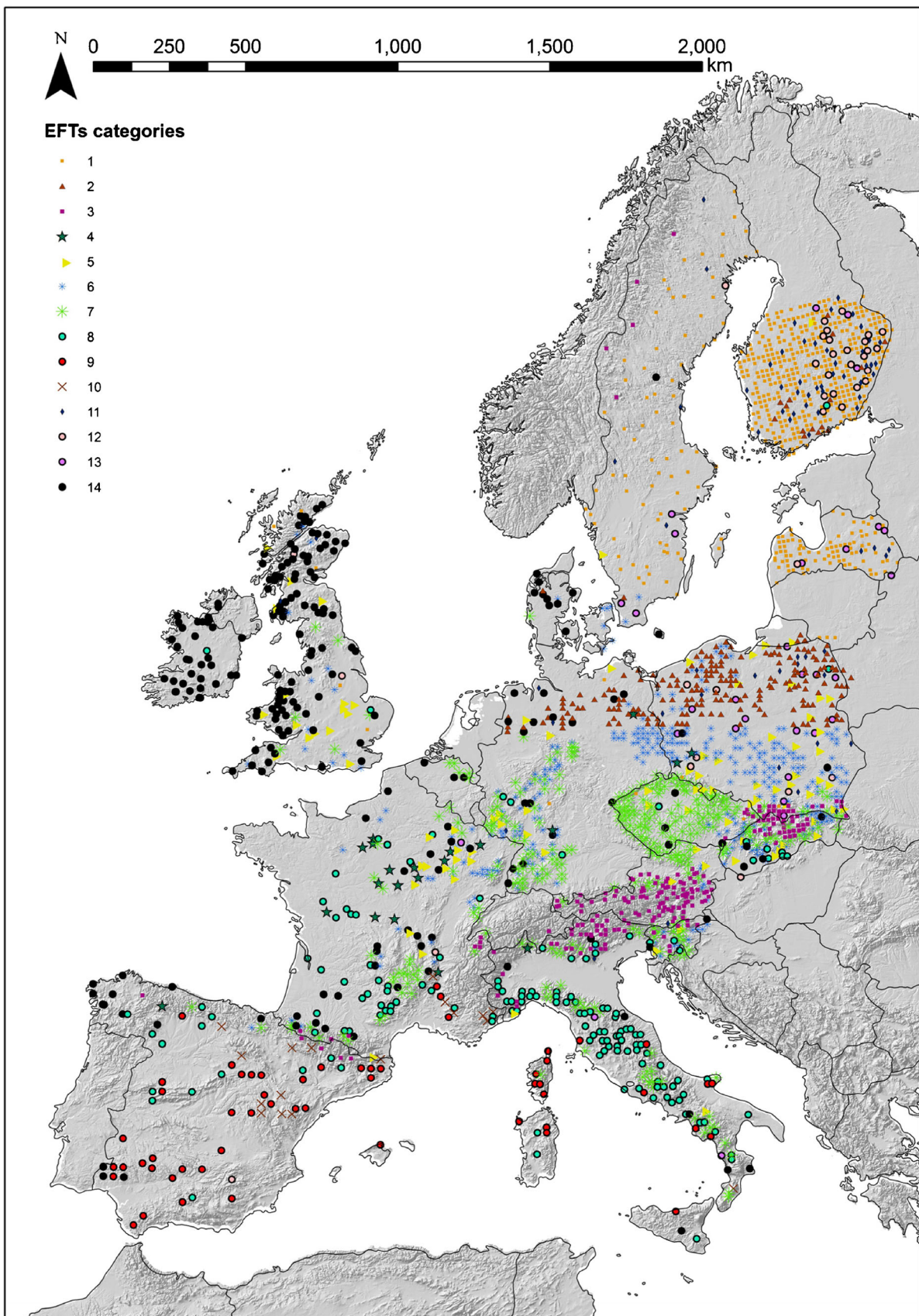


Fig. 1 ICP plots classified into the categories of the European Forest Types in the 17 considered countries by the RBES algorithm (see [Annex 1](#) for the description of EFT categories)

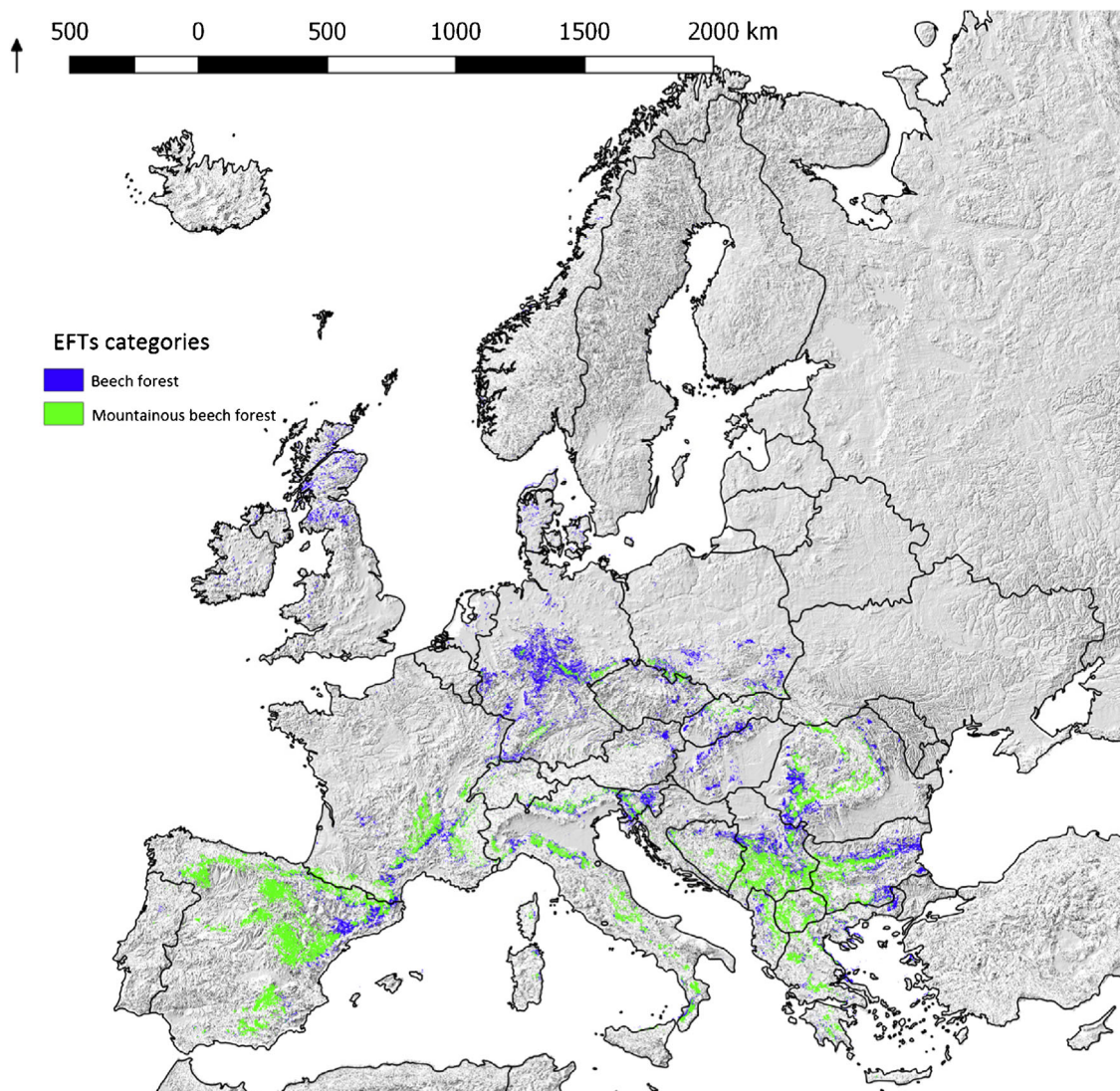


Fig. 2 Beech map classified in two EFT categories (6—beech forest and 7—mountainous beech forest)

was carried out. Thus, in these cases, the inconsistency can be attributed to the subjective evaluation of the field crew. Secondly, if it happens that the field plot is located in a small stand, surrounded by a wider forest type with a different composition, the plot is attributed to the largely represented forest category dominated by forest trees other than those observed in

the plot. In this case, the source of error derives from an ambiguity between field survey and forest classification scales.

The RBES classification errors in category 14 are instead to be attributed to a change in the definition of this specific category, occurring as a consequence of the EFT modifications introduced by Barbati et al. (2014). More specifically, plantations of native

Table 3 The confusion matrix is calculated by comparing the EFT category of the plot assigned in the field with the EFT category returned by the RBES algorithm

Beech map	PLOT			Total	User accuracy	Omission error
	6	7	Other categories			
6	124	12	32	168	0.73	0.27
7	4	289	28	321	0.90	0.10
Outside beech map	28	31	1462	1521	0.96	0.04
Total	156	332	1522	2010		
Producer accuracy	0.79	0.87	0.96	Overall accuracy		0.93
Commission error	0.21	0.13	0.04			

species were included in category 14 during the BioSoil project, while the revised EFT (Barbati et al. 2014) includes in this category forest stands dominated by non-native species.

The rest of the errors introduced by the RBES algorithm are most probably due to the limited spatial resolution of thematic maps used, which cannot capture the small-scale variability of forest habitats that instead were detected during the field survey. This is especially true for category 12. The buffer around waterbodies of 1 km is in fact most probably too wide, but it was the minimum possible value to be consistent with the pixel size of the analysis which is imposed by the trees species map resolution (1 km × 1 km) currently available at European level (San-Miguel-Ayanz et al. 2016).

When the RBES system is applied to stratify into EFT other forest type-based maps, as in the case of the beech map provided by the JRC (de Rigo et al. 2016), we found an excellent accuracy (OA 0.93) when comparing pixel-based classification with the ground survey carried out in the ICP level I plots. In this case, classification errors are mainly due to the coarse resolution of the beech map if compared to the field survey unit dimension.

Our findings ultimately demonstrate that the RBES system can be easily applied to objectively classify forest inventory plots by EFT. The information needed on stand variables is in fact routinely collected in European forest inventories: (i) basal area (derived by callipered DBH), (ii) trees species, and (iii) geographic location of the plot, and based on plot position, information on other environmental variables can be automatically extracted from thematic maps.

Such an automated RBES system could be routinely applied in the classification of all European forest inventory plots which is the operation needed to obtain European forest statistics based on EFT categories, overcoming the limited informative power of the current SOEF 2015 report which is instead based on the adoption of only three broad species groups (conifers, broad-leaves, mixed forests) (FOREST EUROPE 2015).

5 Conclusions

The outcome of the present study leads to the application of a spatially explicit RBES algorithm which is able to automatically classify a stocked forest area according to the EFT system of nomenclature at pan-European scale level. This system is a decision tree based on two main information sources: (1) the basal area per tree species and (2) the geographic position which is used to query several different thematic pan-European digital thematic maps. In order to better understand the potentiality of the RBES approach, we carried out a comparison with the well-known data mining RF algorithm.

The comparison was based on the data collected in 2010 ICP Forests Level I plots surveyed in the framework of the BioSoil Forest Biodiversity Demonstration project. The results obtained with the RBES approach in terms of overall

accuracy (94%) and KIA (0.62) were very satisfactory and better than those obtained with RF. All the variables we used in the classifications are relevant. On the basis of the MDG analysis carried out in the RF approach, we found that without the use of geographic variables (those variables coming from maps and not from the field work), the overall accuracy of the classification decreases by 30%. The maps currently available at pan-European level have limited resolution, for the discrimination of local site environmental variability; it is thus highly recommended that in the future high-resolution layers also for bioclimatic variables could become available.

The RBES system was able to classify all the plots in a way that was consistent, to a large extent, with the field-based classification performed by the crews. The automated classification could be easily applicable to the data collected in the NFIs in Europe, for example, to stratify in a harmonized way NFI datasets into EFT and process-associated forest type-based SFM indicators to be reported in the framework of Forest Europe reporting (SOEF).

The RBES system, available in a complete form as the classification diagram in Annex 1 to this contribution, can be potentially applied across different spatial scales. As we demonstrated for the beech forests, the algorithm can be used to derive EFT maps, at least when digital raster maps with the percent presence of the forest species (in terms of basal area or forest cover) are available. Moreover, multipurpose large-scale inventories would benefit from the integration of the European forest types, in terms of increased capability to assess mechanisms that influence changes of the forest systems and the provision of forest ecosystems (Corona 2016).

The rule-based expert system here presented can significantly contribute to the large-scale deployment of the EFT classification in forest monitoring initiatives in Europe. We also hope that in the future more accurate and detailed pan-European digital maps related to forest and environmental variables would become available to improve the quality of the classification and, therefore, the perspectives of future operational use of this system.

Acknowledgements This research was carried out in the context of ICP Forest group “Upscaling & Spatially explicit estimation of biophysical variables with remote sensing”, coordinated by Prof. Gherardo Chirici. The authors wish to thank the guest editor Walter Seidling and the two anonymous reviewers for their positive contribution to improving the manuscript.

This research was cofunded by the Accademia Italiana di Scienze Forestali and by geoLAB—Laboratory of Forest Geomatics at the Department of Agriculture, Alimentary and Forest Systems of the Università degli Studi di Firenze in the framework of a PhD Student Scholarship to Francesca Giannetti.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Annex 1

Table 4 The most recent version of the system of nomenclature of the European Forest Types at category level (14 categories) from Barbati et al. (2014)

EFT category	Name	Definition
1	Boreal forest	Extensive boreal, species-poor forests, dominated by <i>Picea abies</i> and <i>Pinus sylvestris</i> . Deciduous trees including birches (<i>Betula</i> spp.), aspen (<i>Populus tremula</i>), rowan (<i>Sorbus aucuparia</i>), and willows (<i>Salix</i> spp.) tend to occur as early colonizers
2	Hemiboreal and nemoral coniferous and mixed broadleaved-coniferous forest	Latitudinal mixed forests located in between the boreal and nemoral (or temperate) forest zones with similar characteristics to EFT 1, but a slightly higher tree species diversity, including also temperate deciduous trees like <i>Tilia cordata</i> , <i>Fraxinus excelsior</i> , <i>Ulmus glabra</i> , and <i>Quercus robur</i> . Includes also pure and mixed forests in the nemoral forest zone dominated by coniferous species native within the borders of individual FOREST EUROPE member states like <i>Pinus sylvestris</i> , pines of the <i>Pinus nigra</i> group, <i>Pinus pinaster</i> , <i>Picea abies</i> , and <i>Abies alba</i>
3	Alpine forest	High-altitude forest belts of central and southern European mountain ranges, covered by <i>Picea abies</i> , <i>Abies alba</i> , <i>Pinus sylvestris</i> , <i>Pinus nigra</i> , <i>Larix decidua</i> , <i>Pinus cembra</i> , and <i>Pinus mugo</i> . Includes also the mountain forest dominated by birch of the boreal region
4	Acidophilous oak and oak-birch forest	Scattered occurrence associated with less fertile oligotrophic soils of the nemoral forest zone; the tree species composition is poor and dominated by acidophilous oaks (<i>Q. robur</i> , <i>Q. petraea</i>) and birch (<i>Betula pendula</i>).
5	Mesophytic deciduous forest	Related to medium rich soils of the nemoral forest zone; forest composition is mixed and made up of a relatively large number of broadleaved deciduous trees: <i>Carpinus betulus</i> , <i>Quercus petraea</i> , <i>Quercus robur</i> , <i>Fraxinus</i> , <i>Acer</i> and <i>Tilia cordata</i> .
6	Beech forest	Widely distributed lowland to submountainous beech forest. Beech, <i>Fagus sylvatica</i> , and <i>F. orientalis</i> (Balkan) dominate, locally important is <i>Betula pendula</i> .
7	Mountainous beech forest	Mixed broadleaved deciduous and coniferous vegetation belt in the main European mountain ranges. Species composition differs from EFT 6, including <i>Picea abies</i> , <i>Abies alba</i> , <i>Betula pendula</i> , and mesophytic deciduous tree species. Includes also mountain fir dominated stands.
8	Thermophilous deciduous forest	Deciduous and semi-deciduous forests mainly of the Mediterranean region dominated by thermophilous species, mainly of <i>Quercus</i> ; <i>Acer</i> , <i>Ostrya</i> , <i>Fraxinus</i> , and <i>Carpinus</i> species are frequent as associated secondary trees. Includes also <i>Castanea sativa</i> -dominated forest
9	Broadleaved evergreen forest	Broadleaved evergreen forests of the Mediterranean and Macaronesian regions dominated by sclerophyllous or lauriphyllous trees, mainly <i>Quercus</i> species.
10	Coniferous forests of the Mediterranean	Anatolian and Macaronesian regions. Varied group of coniferous forests in Mediterranean, Anatolian, and Macaronesian regions, from the coast to high mountains. Dry and often poorly developed soils limit tree growth. Several tree species, including a number of endemics of <i>Pinus</i> , <i>Abies</i> , and <i>Juniperus</i> species.
11	Mire and swamp forest	Wetland forests on peaty soils widely distributed in the boreal region. Water and nutrient regimes determine the dominant tree species: <i>Pinus sylvestris</i> , <i>Picea abies</i> , or <i>Alnus glutinosa</i>
12	Floodplain forest	Riparian and riverine species-rich forests characterized by different assemblages of species of <i>Alnus</i> , <i>Betula</i> , <i>Populus</i> , <i>Salix</i> , <i>Fraxinus</i> , and <i>Ulmus</i> .
13	Non-riverine alder, birch or aspen forest	Pioneer forests dominated by <i>Alnus</i> , <i>Betula</i> , or <i>Populus</i> .
14	Introduced tree species forest	Forests dominated by introduced tree species (sensu Pan-European indicator 4.4), occurring on a wide range of site conditions which otherwise would develop forests of above categories. Introduced tree species can be identified at regional (recommended) or national level and comprise tree species that are not native to Europe (e.g., <i>Eucalyptus</i> spp., <i>Robinia pseudoacacia</i> , <i>Acacia dealbata</i> , <i>Ailanthus altissima</i> , <i>Prunus serotina</i> , <i>Quercus rubra</i> , <i>Fraxinus alba</i> , <i>Picea sitchensis</i> , <i>Pinus contorta</i> , <i>Pinus banksiana</i> , <i>Pseudotsuga menziesii</i> , <i>Tsuga heterophylla</i>); tree species native to Europe, but not naturally occurring within the borders of individual FOREST EUROPE member states; tree species native only in some regions of an individual FOREST EUROPE country

Annex 2

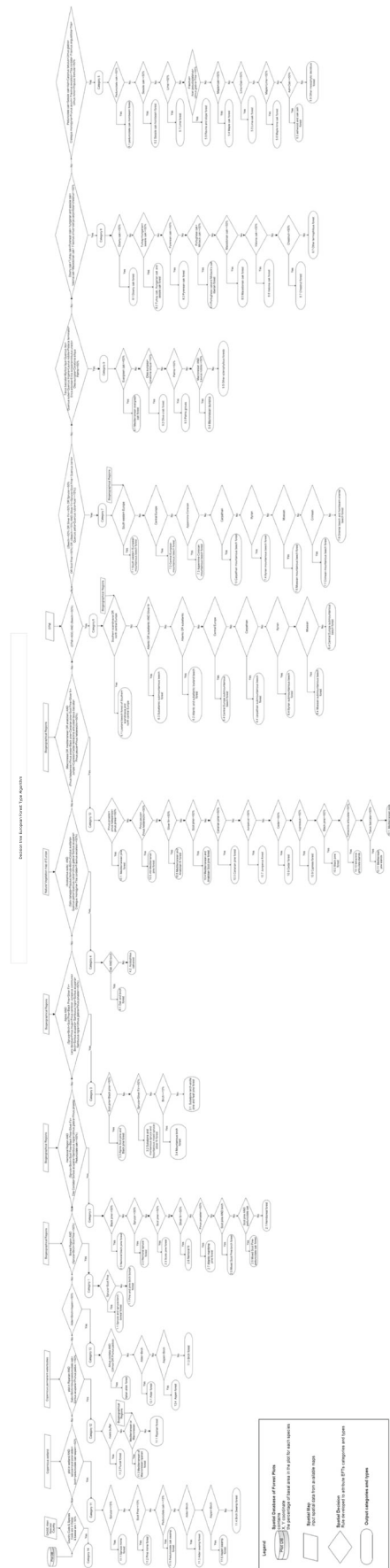


Fig. 3 The image in full resolution is available on line at <http://150.217.83.201/owncloud/index.php/s/Eb2DBopgJT6LLg>

References

- Ackers SH, Davis RJ, Olsen KA, Dugger KM (2015) The evolution of mapping habitat for northern spotted owls (*Strix occidentalis caurina*): a comparison of photo-interpreted, Landsat-based, and lidar-based habitat maps. *Remote Sens Environ* 156:361–373. <https://doi.org/10.1016/j.rse.2014.09.025>
- Adamo M, Tarantino C, Lucas RM, Tomaselli V, Sigismondi A, Mairota P, Blonda P (2015) Combined Use of Expert Knowledge and Earth Observation Data for the Land Cover Mapping of an Italian Grassland Area: An EODHaM System Application. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). N.p., 3065–3068. doi:10.1109/IGARSS.2015.7326463
- Andrew M (1996) Information Systems project redefinition in New Zealand : will we ever learn?. *Aus. Comput J* 28:27–40
- Barbati A, Corona P, Marchetti M (2006) European forest types. Categories and types for sustainable forest management and reporting. European Environment Agency, EEA Technical report No. 9/2006, ISSN 1725-2237. Available at: http://www.foresteurope.org/docs/other_meetings/2006/wfc/WFC_4_eea_technical_report_92006.pdf. Accessed 20 November 2017. Accessed 20 Nov 2017
- Barbati A, Corona P, Marchetti M (2007) A Forest typology for monitoring sustainable forest management: the case of European Forest Types. *Plant Biosyst.* 141(1):93–103. <https://doi.org/10.1080/11263500601153842>
- Barbati A, Arianoutsou M, Corona P, de las Heras J, Fernandes P, Moreira F, Papageorgiou K, Vallejo R, Xanthopoulos G (2010) Post-fire forest management in southern Europe: a COST action for gathering and disseminating scientific knowledge. *IForest* 3(1):5–7. <https://doi.org/10.3832/ifer0523-003>
- Barbati A, Corona P, Marchetti M (2011) Annex 1: Pilot Application of the European Forest Types. In Michalak R. (eds.), *FOREST EUROPE, UNECE and FAO 2011. State of Europe's Forests 2011. Status and trends in Sustainable Forest Management in Europe*. Oslo: 259–273
- Barbati A, Marchetti M, Chirici G, Corona P (2014) European Forest Type and Forest Europe SFM indicators: tools for monitoring progress on forest biodiversity conservation. *Forest Ecol Manag* 321: 145–151. <https://doi.org/10.1016/j.foreco.2013.07.004>
- Bastrup-Birk A, Neville P, Chirici G, Houston T (2007) *The BioSoil Forest biodiversity field manual*. ICP Forests, Hamburg
- Beard JS, Beeston GR, Harvey JM, Hopkins AJM, Shepherd DP (2013) *The vegetation of Western Australia at the 1:3,000,000 scale. Explanatory memoir. Second edition*. Conserv Sci Western Aust 9: 1–152
- Bingyuan C, Ma S, Cao H (2014) *Ecosystem assessment and fuzzy systems management*. Springer (ISBN 978-3-319-03448-5), New York/London, p 529
- Bohn U, Neuhäusl R, unter Mitarbeit von / with contributions by Gollub G, Hettwer C, Neuhäuslová Z, Raus, Th, Schlüter H, Weber H (2000/2003): *Karte der natürlichen Vegetation Europas / Map of the Natural Vegetation of Europe*. Maßstab / Scale 1 : 2 500 000. Münster (Landwirtschaftsverlag), Bundesamt für Naturschutz (BfN)/Federal Agency for Nature Conservation Konstantinstr. 110, 53179 Bonn, Germany. Avilable on line https://is.muni.cz/el/1431/podzim2012/Bi9420/um/Bohn_etal2004_Map-Nat-Veg-Europe.pdf Accessed 20 November 2011
- Bravo-Oviedo A, Pretzsch H, Ammer C, Andenmatten E, Barbati A, Barreiro S, Brang P, Bravo F, Coll L, Corona P, Den Ouden J, Ducey MJ, Forrester DI, Giergiczny M, Jacobsen JB, Lesinski J, Löff M, Mason B, Matovic B, Metslaid M, Morneau F, Motiejunaite J, O'Reilly C, Pach M, Ponette Q, Del Rio M, Short I, Skovsgaard JP, Soliño M, Spathelf P, Sterba H, Stojanovic D, Strelcova K, Svoboda M, Verheyen K, Von Lüpke N, Zlatanov T

- (2014) European mixed forests: definition and research perspectives. *Forest Systems* 23(3):518–533. <https://doi.org/10.5424/fs/2014233-06256>
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Bruelheide H (2000) A new measure of fidelity and its application to defining species groups. *J Veg Sci* 11(2):167–178. <https://doi.org/10.2307/3236796>
- Brus DJ, Hengeveld GM, Walvoort DJJ, Goedhart PW, Heidema AH, Nabuurs GJ, Gunia K (2011) Statistical mapping of tree species over Europe. *Eur J For Res* 131(1):145–157
- Buffa G, Villani M (2012) Are the ancient forests of the Eastern Po Plain large enough for a long term conservation of herbaceous nemoral species? *Plant Biosyst* 146(4):970–984. <https://doi.org/10.1080/11263504.2012.704887>
- Capelo J, Masquita S, Costa JC, Ribeiro S, Arsénio P, Neto C, Monteiro-Henriques T, Aguiar C, Honrado J, Espírito-Santo D, Lousã M (2007) A methodological approach to potential vegetation modeling using GIS techniques and phytosociological expert-knowledge: application to mainland Portugal. *Phytocoenologia* 37(3–4):399–415. <https://doi.org/10.1127/0340-269X/2007/0037-0399>
- Caudullo G, Pasta S, Giannetti F, Barbati A, Chirici G (2016) European forest classifications in San-Miguel. In: Ayanz J, de Rigo D, Caudullo G, Houston Durrant T, Mauri A (eds) *European Atlas of Forest Tree Species*. Publication Office of the European Union, Luxembourg, pp 32–33
- Chen M, Yao Z (2008) Classification techniques of neural networks using improved genetic algorithms. In: *Proceedings of 2nd International Conference on Genetic and Evolutionary Computing*, Washington. Article number 4637407, 115–119
- Chytrý M (2012) Vegetation of the Czech Republic: diversity, ecology, history and dynamics. *Preslia* 84(3):427–504
- Chytrý M, Tichý L, Holt J, Botta-Dukát Z (2002) Determination of diagnostic species with statistical fidelity measures. *J Veg Sci* 13(1):79–90. <https://doi.org/10.1111/j.1654-1103.2002.tb02025.x>
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46
- Cook SE, Corner R, Grealish GJ, Gessler PE, Chartres CJ (1996) A rule-based system to map soil properties. *Soil Sci Soc Am J* 60(1996):1893–1900. <https://doi.org/10.2136/sssaj1996.03615995006000060039x>
- Corona P (2016) Consolidating new paradigms in large-scale monitoring and assessment of forest ecosystems. *Environ Res* 144(Pt B):8–14. <https://doi.org/10.1016/j.envres.2015.10.017>
- Corona P, Ferrari B, Cartisano R, Barbati A (2014) Calibration assessment of forest flammability potential in Italy. *IForest* 7(5):300–305. <https://doi.org/10.3832/ifor1123-007>
- COUNCIL OF EUROPE (1979)- ETS 104 – Conservation of Wildlife and Natural Habitats, Bern, 19.IX.1979. available at <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680078aff>. Accessed 20 November 2017
- de Rigo D, Caudullo G, San-Miguel-Ayanz J (2016) European tree species distribution with Constrained Spatial Multi-Frequency Analysis. EUR - Scientific and Technical Research. Publications Office of the European Union. Available on line at http://forest.jrc.ec.europa.eu/media/atlas/Modelling_data_and_information_on_forest_tree_species.pdf
- Dengler J, Löbel S, Dolnik C (2009) Species constancy depends on plot size—a problem for vegetation classification and how it can be solved. *J Veg Sci* 20(4):754–766. <https://doi.org/10.1111/j.1654-1103.2009.01073.x>
- Douda J, Boulík K, Slezák M, Biurrun I, Nociar J, Havrdová A, Doudová J, Acíc S, Brisse H, Brunet J, Chytrý M, Claessens H, Csiky J, Didukh Y, Dimopoulos P, Dullinger S, FitzPatrick Ú, Guisan A, Horschler PJ, Hrivnák R, Jandt U, Kaçki Z, Kevey B, Landucci F, Lecomte H, Lenoir J, Paal J, Paternoster D, Pauli H, Pielech R, Rodwell JS, Roelandt B, Svenning JC, Šibík J, Šilc U, Škvorec Ž, Tsiripidis I, Tzonev RT, Wohlgemuth T, Zimmermann NE (2016) Vegetation classification and biogeography of European floodplain forests and alder cars. *Appl Veg Sci* 19(1):147–163. <https://doi.org/10.1111/avsc.12201>
- Duveneck MJ, Thompson JR, Wilson BT (2015) An imputed forest composition map for New England screened by species range boundaries. *Forest Ecol Manag* 347:107–115. <https://doi.org/10.1016/j.foreco.2015.03.016>
- EEA (2006) European forest types. Categories and types for sustainable forest management and reporting. European Environment Agency, EEA Technical report No. 9/2006. ISSN 1725-2237 Available at: http://www.foresteurope.org/docs/other_meetings/2006/wfc/WFC_4_eea_technical_report_92006.pdf. Accessed 20 November 2017
- EEA (2013) Digital Elevation Model over Europe (EU-DEM). European Environment Agency Available on-line at <https://www.eea.europa.eu/data-and-maps/data/eu-dem#tab-metadata>
- EEA (2013a) Copernicus Initial Operations 2011–2013 Land Monitoring Service pan-European Component, High Resolution Layer Permanent Water Bodies (PWB). Available at <http://land.copernicus.eu/pan-european/high-resolution-layers/permanent-water-bodies/view> Accessed 30 January 2016. In: European Environment Agency (2013) GIO land (GMES/Copernicus initial operations land) High Resolution Layers (HRLs) – summary of product specifications, European Environment Agency, Copenhagen K, Denmark
- EEA (2013b) Copernicus Initial Operations 2011–2013 Land Monitoring Service pan-European Component, High Resolution Layer Permanent Wetlands (WET) Available at <http://land.copernicus.eu/pan-european/high-resolution-layers/wetlands/view> Accessed 30 January 2016. In: European Environment Agency (2013). GIO land (GMES/Copernicus initial operations land) High Resolution Layers (HRLs) – summary of product specifications, European Environment Agency, Copenhagen K, Denmark
- EEA (2015a) Linking in situ vegetation data to the EUNIS habitat classification: results for forest habitats. European Environment Agency, EEA Report No. 18/2015. ISSN 1725-2237 Available on-line <<https://www.eea.europa.eu/publications/eunis-habitat-classification/download> Accessed 20 November 2017>
- EEA (2015b) The biogeographical regions dataset. Europe. European Environment Agency Available on-line at: <http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3#tab-metadata> : Accessed 20 November 2017
- Ewald J (2003) A critique for phytosociology. *J Veg Sci* 14(2):291–296. <https://doi.org/10.1111/j.1654-1103.2003.tb02154.x>
- Faber-Langendoen D, Keeler-Wolf T, Meidinger D, Tart D, Hoagland B, Josse C, Navarro G, Ponomarenko S, Saucier J-P, Weakley A, Comer P (2014) EcoVeg: a new approach to vegetation description and classification. *Ecol Monogr* 84(4):533–561. <https://doi.org/10.1890/13-2334.1>
- Flanagan NE, Richardson CJ, Ho M (2015) Connecting differential responses of native and invasive riparian plants to climate change and environmental alteration. *Ecol Appl* 25(3):753–767. <https://doi.org/10.1890/14-0767.1>
- FOREST EUROPE (2011) State of Europe’s Forests 2011. Forest Europe Liaison Unit/UNECE Timber Section/FAO, Oslo
- FOREST EUROPE (2015). State of Europe’s Forests 2015. Ministerial Conference on the Protection of Forests in Europe 2015. Available on-line <http://www.foresteurope.org/docs/fullsoef2015.pdf> . Accessed 20 November 2017
- Gao J, Chen H, Zhang Y, Zha Y (2004) Knowledge-based approaches to accurate mapping of mangroves from satellite data. *Photogramm Eng Remote Sens* 70(11):1241–1248. [10.14358/PERS.70.11.1241](https://doi.org/10.14358/PERS.70.11.1241)
- Greenberg JA (2014) Spatial functions meant to enhance the core functionality of the package “raster”, including a parallel processing engine for use with rasters. R-CRAN package, Available on-line

- <http://www.geog.illinois.edu/~jgrn/software-and-datasets/rasterengine-tutorial/>
- Grunwald S (2009) Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152(3–4):195–207. <https://doi.org/10.1016/j.geoderma.2009.06.003>
- Hayes-Roth F (1985) Rule-based systems. *Commun ACM* 28(9):921–932. <https://doi.org/10.1145/4284.4286>
- Hédél R (2007) Is sampling subjectivity a distorting factor in surveys for vegetation diversity? *Folia Geobot. Phytotaxon* 42:191–198
- Hiederer R, Durrant T (2010) Evaluation of BioSoil Demonstration Project—preliminary data analysis. Joint Research Centre, report no JRC56739. ISSN 1018–5593
- Hijmans RJ (2015) raster: Geographic data analysis and modeling. R-CRAN package, Available on line at <http://CRAN.R-project.org/package=raster>
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25(15):1965–1978
- Illyés E, Chytrý M, Botta-Dukát Z, Jandt U, Škodová I, Janišová M, Willner W, Hájek O (2007) Semi-dry grasslands along a climatic gradient across Central Europe: vegetation classification with validation. *J Veg Sci* 18(6):835–846. <https://doi.org/10.1111/j.1654-1103.2007.tb02600.x>
- Ioannis N, Vogiatzakis, Griffiths GH (2006) A GIS-based empirical model for vegetation prediction in Lefka Ori. *Crete Plant Ecol* 184(2): 311–323
- Jiménez-Alfaro B, Chytrý M, Rejmánek M, Mucina L (2014) The number of vegetation types in European countries: major determinants and extrapolation to other regions. *J Veg Sci* 25(3):863–872. <https://doi.org/10.1111/jvs.12145>
- JRC (2011) Biosoil biodiversity executive report. Report number: Joint Research Centre report 64509. Durrant, T., San-Miguel-Ayanz, J., Schulte, E., & Suarez-Meyer, A. (Eds). Publications Office of the European Union. ISSN 1018–5593 <http://publications.jrc.ec.europa.eu/repository/handle/JRC64509> Available on-line 20 November 2017
- Knollová I, Chytrý M, Tichý L, Hájek O (2005) Stratified resampling of phytosociological databases: some strategies for obtaining more representative data sets for classification studies. *J Veg Sci* 16(4):479–486. <https://doi.org/10.1111/j.1654-1103.2005.tb02388.x>
- Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2(3):18–22
- Mc Roberts RE, Chirici G, Winter S, Barbati A, Corona P, Marchetti M, Hauk E, Brändli U-B, Beranova J, Rondeaux J, Sanchez C, Bertini R, Barsoum N, Alberdi Asencio I, Condéz S, Saura S, Neagu S, Cluzeau C, Hamza N (2011) Prospects for harmonized biodiversity assessments using national forest inventory data. In: Chirici G, Winter S, McRoberts RE (eds) *National Forest Inventories: contributions to Forest biodiversity assessments*. Springer, Heidelberg, pp 41–97. https://doi.org/10.1007/978-94-007-0482-4_3
- Millington A, Walsh S, Osborne PE (2002) GIS and remote sensing applications in biogeography and ecology. Kluwer Academic Publishers
- Mucina L (1997) Classification of vegetation: past, present and future. *J Veg Sci* 8(6):751–760. <https://doi.org/10.2307/3237019>
- Ohman J, Gregory M (2002) Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbour imputation in coastal Oregon, USA. *Can J For Res* 32(4):725–741. <https://doi.org/10.1139/x02-011>
- Openshaw S, Openshaw C (1997) *Artificial intelligence in geography*. Wiley, London
- Panagos P, Jones A, Bosco C, Senthil Kumar PS (2011) European digital archive on soil maps (EuDASM): preserving important soil data for public free access. You can download the article in press for your documentation. *Int J Digital Earth* 4(5):434–443. <https://doi.org/10.1080/17538947.2011.5965803>
- Peet R, Harris J, Grossman D, Jennings M, Walker MD (2001) An information infrastructure for vegetation science: project overview and progress report. Available on-line at <http://labs.bio.unc.edu/Peet/Proposals/plotsummary.htm> Accessed 20 November 2017
- Pérez-Ortiz M, Peña JM, Gutiérrez PA, Torres-Sánchez J, Hervás-Martínez C, López-Granados F (2016) Selecting patterns and features for between- and within- crop-row weed mapping using UAV-imagery. *Expert Syst Appl* 47:85–94. <https://doi.org/10.1016/j.eswa.2015.10.043>
- Pividori M, Giannetti F, Barbati A, Chirici G (2016) European Forest Types: tree species matrix in San-Miguel. In: Ayanz J, de Rigo D, Caudullo G, Houston Durrant T, Mauri A (eds) *European Atlas of Forest Tree Species*. Publication Office of the European Union, Luxembourg, pp 34–35
- Puletti N, Giannetti F, Chirici G, Canullo R (2017) Deadwood distribution in European forests. *Journal of Maps* 13(2):733–736. <https://doi.org/10.1080/17445647.2017.1369184>
- Robinove CJ (1986) *Principles of logic and the use of digital geographic systems*. Department of Interior, US Geological Survey, Reston, VA
- Rodwell JS, Schaminée JHJ, Mucina L, Pignatti S, Dring J, Moss D (2002) The diversity of European vegetation—an overview of phytosociological alliances and their relationships to EUNIS habitats. *Rapport EC-LNV* 2002(054):1–168
- San-Miguel-Ayanz J, de Rigo D, Caudullo G, Houston DT, Mauri A (2016) *European Atlas of Forest Tree Species*. Publication Office of the European Union, Luxembourg, pp 32–33. <https://doi.org/10.2788/038466>
- Shumchenia EJ, King JW (2010) Comparison of methods for integrating biological land physical data for marine habitat classification. *Cont Shelf Res* 30(16):1717–1729. <https://doi.org/10.1016/j.csr.2010.07.007>
- Song M, Zhou C, Ouyang H (2005) Simulated distribution of vegetation types in response to climate change on the Tibetan Plateau. *J Veg Sci* 16(3):341–350. <https://doi.org/10.1111/j.1654-1103.2005.tb02372.x>
- Vaz AS, Marcos B, Gonçalves J, Monteiro A, Alves P, Civantos E et al (2015) Can we predict habitat quality from space? A multi-indicator assessment based on an automated knowledge-driven system. *Int J Appl Earth Obs Geoinf* 37:106–113. <https://doi.org/10.1016/j.jag.2014.10.014>
- Wang ZY, Leung KS, Klir GJ (2005) Applying fuzzy measures and non-linear integrals in data mining. *Fuzzy Sets Syst* 156(2005):371–380. <https://doi.org/10.1016/j.fss.2005.05.034>
- Working Group on Forest Biodiversity (2007) *The BioSoil Forest Biodiversity field manual*. In: JRC 2011, Evaluation of BioSoil Demonstration Project: Forest Biodiversity. Publications Office of the European Union, Luxembourg. <https://doi.org/10.2788/84823>
- Zimmermann NE, Kienast F (1999) Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *J Veg Sci* 10(4):469–482. <https://doi.org/10.2307/3237182>