Spatiotemporal modeling of vegetation dynamics in a changing environment: combining Earth observation and machine learning

Carmelo Bonannella

Propositions

1. Machine learning requires accurate space and ground-based observations for monitoring of critical spatiotemporal forest changes and vegetation dynamics.
   (this thesis)

2. Integrated, large-area forest monitoring is restricted by limited access to quality reference datasets.
   (this thesis)

3. Scholars, striving for tenure, unintentionally fuel the dominance and profits of five major commercial publishers.

4. Active management of old-growth forests contributes more to climate change mitigation than afforestation and reforestation efforts.

5. The systemic lack of funding in academia reflects the lack of long-term planning in our current society.

6. Meritocracy in society is a myth masking deep-rooted inequalities.

7. Modern society has lost its environmental literacy.

Propositions belonging to the thesis, entitled

Spatiotemporal modeling of vegetation in a changing environment: combining Earth observation and machine learning

Carmelo Bonannella
Wageningen, 2 July 2024

Spatiotemporal modeling of vegetation dynamics in a changing environment: combining Earth observation and machine learning

Carmelo Bonannella

Thesis committee

Promotor:
Prof. Dr Martin Herold
Professor of Geo-information Science and Remote Sensing
Wageningen University & Research

Co-promotor:
Dr Sytze de Bruin
Associate Professor, Laboratory of Geo-information Science and Remote Sensing
Wageningen University & Research

Other members:
Prof. Dr Frank Sterck, Wageningen University & Research
Prof. Dr Mana Gharun, University of Münster
Dr Anton Vrieling, University of Twente
Dr Emiel van Loon, University of Amsterdam

This research was conducted under the auspices of the C.T. de Wit Graduate School of Production Ecology & Resource Conservation (PE&RC)

Spatiotemporal modeling of vegetation dynamics in a changing environment: combining Earth observation and machine learning

Carmelo Bonannella

Thesis
submitted in fulfilment of the requirements for the degree of doctor at
Wageningen University
by the authority of the Rector Magnificus
Prof. Dr C. Kroeze,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 2 July 2024
at 1:30 p.m. in the Omnia Auditorium.

Carmelo Bonannella
Spatiotemporal modeling of vegetation dynamics in a changing environment: combining Earth observation and machine learning
224 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2024)
With references, with summary in English

DOI 10.18174/655208

Summary

Vegetation, ranging from dense forests to open grasslands, plays a vital role in sustaining Earth's environmental health. These diverse ecosystems are crucial, offering services like carbon sequestration, soil stabilization, and water regulation, while providing habitats for countless species. Forests, in particular, stand as ecological powerhouses, encompassing nearly a third of the planet's land area and harboring around 80% of terrestrial biodiversity. They are instrumental in regulating freshwater flow and precipitation, crucial for agriculture, and are key players in atmospheric carbon absorption. Supporting the livelihoods of over 1.6 billion people globally, forests' ecological and socio-economic roles are intertwined. Consequently, their protection and judicious management are essential to ensure the ongoing provision of these vital ecosystem services and the sustenance of life on Earth.

Traditional forest monitoring methods, including field surveys and National Forest Inventories (NFI) campaigns, are pivotal for understanding forest composition, structure, and health. Established in the early 20th century, these approaches are crucial for forestry management and environmental assessments. NFIs in particular, provide comprehensive, country-specific data, supporting national forestry policies and international reporting. Yet, these traditional methods encounter limitations in scalability, frequency, and logistics, especially for monitoring large, remote, or inaccessible areas. With the accelerating pace of climate change, these methods fall short in monitoring dynamic ecological variables. They struggle to track shifts in species composition, alterations in species ranges, biodiversity or forest disturbances such as wildfires, pest outbreaks, droughts, and heatwaves. These limitations hinder their effectiveness in capturing the rapid ecological transformations and the evolving dynamics of forest ecosystems.

The limitations of traditional monitoring systems are highlighted in various policy frameworks, which emphasize the need for comprehensive and high-quality monitoring of all forested areas. For instance, the European Union (EU)'s new Forest Strategy for 2030, aligned with the European Green Deal and the EU Biodiversity Strategy for 2030, sets a vision to improve the quantity, quality, and resilience of European forests. It underscores the necessity of strategic forest monitoring, data collection, and coherent governance to help the transition towards forests that are more adapted and resilient to climate change and ensure their multifunctionality for the future decades.

The integration of Earth observation data has initiated a transformative era in forest and vegetation monitoring, significantly enhancing our capacity to assess and track vegetation dynamics. Among the different Earth observation tools, satellites, especially those from Landsat and Sentinel missions, offer a vast spatial coverage and a detailed enough spatial resolution that is crucial for observing large forested areas and detecting changes over time. The Global Forest Watch (GFW), an initiative utilizing mainly Landsat satellite imagery to monitor forests globally, exemplifies the application of Earth observation in large-scale forest monitoring. These satellite-based observations provide vital data on forest cover changes, deforestation rates, and reforestation efforts, contributing to a more nuanced understanding of global forest ecosystems and integrating traditional international reports on the state of the forests like the FAO's Forest Resource Assessment (FRA).

Machine learning technologies, particularly in the realm of image processing, have transformed the way we analyze data from remotely sensed sources. Machine learning applications have proven effective in processing complex datasets from satellite missions, including the intricacies of hyperspectral data, which present challenges due to their high dimensionality. These machine learning methods are adept at identifying patterns and changes in forest landscapes, such as variations in tree species composition, signs of forest degradation, and impacts of climate change. The ability of machine learning algorithms to process and analyze large collections of satellite imagery has opened new possibilities for comprehensive and dynamic forest monitoring.

The integration of Earth observation and machine learning technologies with traditional ground-based methods, such as NFIs, is pivotal in creating a robust forest monitoring system. This combination allows for the validation and enhancement of satellite-based observations with detailed ground-truth data. By merging the broad spatial coverage of satellite data with the accuracy and specificity of field data, we can achieve a more accurate, timely, and holistic view of forest ecosystems. This integrated approach is instrumental in addressing the limitations of traditional methods and fulfilling the need for rapid and responsive forest monitoring in the face of global environmental changes. It represents a significant step forward in our ability to manage and conserve forest resources effectively, ensuring their sustainability for future generations.

Thus, the objective of this thesis is to integrate Earth observation and machine learning technologies with field data to enhance our understanding of ongoing vegetation dynamics and the overall monitoring and management of forest ecosystems. This work mainly contributes to this goal by developing and applying novel methods at different spatial and temporal scales for vegetation modeling. More specifically, in order to accomplish this objective, four research questions have been formulated and addressed in this work: (1) What is the impact of climate change on potential biomes distribution based on Earth observation and machine learning methods, and what are the projected shifts in vegetation under various climate change scenarios? (2) What combination of Earth observation

and machine learning methods allows to map and analyze the distribution of forest tree species at high resolution? (3) How can these methods be applied to capture trends and disturbance impacts on forest tree species distributions and how do these reflect the ongoing changes in forest ecosystems? (4) What is the effect of coordinate precision of NFI data on the accuracy of high-resolution tree species classification models?

The thesis is composed of six chapters, with Chapters 1 and 6 framing the core content and Chapters 2 to 5 addressing the research questions presented above. **Chapter 1** serves as an introduction, delineating the current advancements in Earth observation and machine learning as they apply to vegetation modeling and monitoring. It also identifies existing research gaps and explores potential approaches for addressing them.

Chapter 2 investigates the projected shifts in vegetation under various climate scenarios by developing a data-driven approach based on machine learning techniques and pollen reconstruction datasets coming from different surveys that have been harmonized over the years. The chapter specifically focuses on modeling the potential distribution of global biomes, incorporating a range of climate variables to predict the current distribution and how different climate change scenarios could alter vegetation patterns and distributions in the future. Through this analysis, the chapter provides insights into the expected changes in biome locations and compositions, particularly the contraction of tropical rainforests and the expansion of higher latitude biomes like boreal forests, at polar biomes' expense. The chapter offers a picture of how climate change could reshape the global ecological landscape. Given the unpredictability of the future, the method is conservative in its conclusions: after analyzing every emission pathway scenario, it focuses on identifying key change areas, irrespective of specific emission pathways, providing a nuanced understanding of climate change's impact on the global ecological landscape.

Chapter 3 presents a more sophisticated data-driven approach than Chapter 2, integrating 305 environmental predictors, primarily from Earth observation data, employing advanced machine learning techniques alongside over 3 million tree occurrence points from different surveys or citizen science projects. This approach is used to model the potential and actual distribution of 16 key European forest tree species from 2000 to 2020 at high resolution. The chapter's core is the development and application of ensemble models for each species and distribution type, employing a series of machine learning algorithms such as random forest, gradient boosting, and neural networks, and utilizing spatial cross-validation to enhance predictive accuracy and avoid model overfitting. The influence of the spatial resolution of the environmental predictors on model performances was also tested. As expected, not only ensemble models outperformed individual models, but ensemble models using high spatial resolution data proved to be consistently better than the ones using coarse resolution data.

Chapter 4 delves into the changing dynamics of six key European forest tree species between 2000 and 2020, building upon the high resolution distribution maps generated

in Chapter 3. This chapter focuses on the response in the realized distribution of these species over the last twenty years by conducting a trend analysis on the distribution maps and then degrading the results to a coarser resolution to find the prevalent spatial patterns for each species. Findings indicate a predominant stability across species, with most areas exhibiting consistent distribution patterns. Nonetheless, the chapter also uncovers areas of concern, marked by declining species presence, especially at the latitudinal edges of their natural ranges. Additionally, the chapter examines the impact of forest disturbances on species distribution, noting a doubling in the range affected by disturbances such as wind and fire. The findings highlight the need for proactive conservation and long-term planning to ensure the resilience of European forests against climate-induced changes.

Chapter 5 investigates the impact of coordinate precision in National Forest Inventory (NFI) data on the accuracy of high-resolution tree species classification models. Focusing on seven European tree species, the chapter compares machine learning models trained on true, precise plot coordinates with those trained on coordinates derived from publicly available data. The chapter describes a procedure to estimate the precise plot coordinates location with the aid of high resolution forest masks and forest types layers for the country of the Netherlands. The analysis demonstrates that models utilizing true coordinates consistently outperform those based on estimated coordinates. This is evident in the higher scores reached across performance metrics used for both hard classification and probabilistic classification problems across most species classes for the model using precise data. The chapter highlights the critical importance of coordinate precision in ecological modeling and underscores the need for high-quality, accurate NFI data for effective tree species classification using Earth observation data and machine learning methods.

Lastly, **Chapter 6** summarises the main findings and results of the previous chapters and provides an overview on how they contribute towards filling the research gaps identified by this thesis. Following that, it offers a reflection and outlook, exploring potential pathways for future applications and new research prospects in the integration of Earth observation data with machine learning and field data for vegetation modeling.

Contents

Chapter 1

Introduction

1.1 Background

1.1.1 Forests and climate change

Forests and other vegetation ecosystems are extremely valuable for humanity and the planet, playing an important role in environmental sustainability and ecological balance. Their significance extends beyond simply providing resources; they are crucial for maintaining biodiversity, regulating water cycles, conserving soil and offering spaces for recreation and spiritual connection. Forests are particularly important in this sense, covering nearly a third of the planet's land area (FAO, 2022), representing the most widespread vegetation type globally (Banskota et al., 2014) and hosting about 80% of terrestrial biodiversity (Cazzolla Gatti et al., 2022). These ecosystems not only support over 1.6 billion people globally but also function as major carbon sinks, absorbing and storing carbon dioxide (CO_2) from the atmosphere, crucial in mitigating climate change (IPCC, 2021).

The health and stability of these ecosystems face increasing threats from climate change. Changes in temperature and precipitation patterns severely impact forest health and biodiversity (Dale et al., 2001; Forzieri et al., 2021; Keenan, 2015; Lindner et al., 2010; Rounsevell et al., 2018). The increasing frequency and intensity of extreme weather events such as wildfires, storms, and droughts compound these threats, leading to widespread destruction and long-term impacts on forest recovery and resilience (Allen et al., 2010; Patacca et al., 2023; Seidl et al., 2017). Furthermore, climate change has exacerbated the incidence of pest infestations and diseases, which pose a significant risk to tree populations and forest compositions (Ayres and Lombardero, 2000; Hlásny et al., 2021a; Hlásny et al., 2021b). The intricate balance within forest ecosystems is further disrupted, as phenology and growth rates are affected: phenological shifts, such as earlier spring leaf-out or later autumn senescence, can disrupt the synchronization between forests and the species that depend on them. These changes in timing can have cascading effects on ecosystems, altering food webs and habitats (Cleland et al., 2007).

Understanding the intricacies of forest ecosystems and their fundamental processes is vital for the multifaceted role they play; a series of measurements, analysis and experiments is needed to understand, for instance, the potential forests may have as nature-based climate solutions (Baldocchi and Penuelas, 2019; Seddon, 2022), their responses to climate change (Anderegg et al., 2022; Ruiz-Benito et al., 2020) and how new forest management strategies can be implemented to build climate-smart forests (Verkerk et al., 2020).

1.1.2 Current status of vegetation monitoring

In response to climate change pressures, the United Nations Framework Convention on Climate Change (UNFCCC) has been instrumental in global climate action for forest protection. First in 2008, the UNFCCC established the REDD+ (or REDD-plus) framework. REDD+ (Reducing Emissions from Deforestation and Forest Degradation) aims to

create a financial value for the carbon stored in forests. It encourages developing countries to reduce emissions from forested lands and invest in sustainable forest management. REDD+ extends beyond just curbing deforestation – it encompasses a broader spectrum of strategies including forest conservation, sustainable management, and enhancing forest carbon stocks. The framework has been pivotal in shaping national forest monitoring systems, aligning them with the goal of climate change mitigation (De Sy, 2016). More recently, the COP28 conference in Dubai in 2023 marked a significant milestone in this effort. COP28 concluded with an agreement signaling the "beginning of the end" of the fossil fuel era. This transition away from fossil fuels is supported by the establishment of a new fund under the UNFCCC, with commitments totaling over USD 600 million, aimed at addressing loss and damage, including those affecting forests (Wise, 2023).

After the significant advancements made at COP28, global initiatives like the Bonn Challenge gain renewed importance. Launched in 2011 by by the German government and the International Union for Conservation of Nature (IUCN), the Bonn Challenge aims to restore 350 million hectares of the world's deforested and degraded lands by 2030. To track restoration commitments, a systematic framework called the Bonn Challenge Barometer was established (Dave et al., 2018). The Bonn Challenge serves as a platform for countries and organizations to contribute towards the global forest restoration goal, aligning with other international efforts like the Aichi Biodiversity Targets (CBD, 2011) and the Paris Agreement (UNFCCC, 2018).

Europe's regional actions in forest conservation and climate change mitigation are exemplified by the European Union's new forest strategy, a part of the European Green Deal (European Commission, 2021a). The European Green Deal mirrors the global commitments of COP28 and the Bonn Challenge, adapting them to the continental context. Key elements include promoting sustainable forest bioeconomy, ensuring the sustainable use of wood resources, protecting primary and old-growth forests, and encouraging re- and afforestation (European Commission, 2021b). This strategy, promoting sustainable forest bioeconomy and reforestation, is supported by the European Commission's proposed regulation on a forest monitoring framework, integrating remote sensing with traditional monitoring practices (European Commission, 2023).

The evolution of forest monitoring, influenced by various policies, has been a journey from traditional methods to more modern approaches. Initially focused on managing and conserving forest resources, as seen with the early forest inventories of the Republic of Venice, these methods have long been a cornerstone of forest management (Appuhn, 2010; Susmel, 1994). To provide another European example, countries like Germany have traditionally relied on detailed field inventories, conducted every decade, to guide forest management plans. Traditional methods primarily involved ground-based surveys and manual inventories, each characterized by direct field observations and measurements. Foresters and ecologists would physically traverse forested areas to gather data on various

parameters, including tree species identification, diameter at breast height (DBH), tree height, canopy cover, and understory vegetation types. These methods offer high-quality, localized information, vital for comprehending forest dynamics, health, and biodiversity and allowing for decisions about resource allocation, conservation, and exploitation. A crucial aspect of traditional forest monitoring is the implementation of National Forest Inventories (NFIs). NFIs provide systematic and comprehensive data on forest resources at a national scale and they are instrumental in guiding national forestry policies and resource management strategies, often serving as a country's primary source of forestry data (Fig. 1.1). While NFIs have different histories in different countries, the first examples of NFI date back to the late nineteenth century and the beginning of the twentieth century (Tomppo et al., 2010; United States Department of Agriculture (USDA), 2007; Von Berg, 1995), with some countries like Norway recently celebrating a century old NFI program (Breidenbach et al., 2020). In nations where the forest industry is a major sector, NFIs play a critical role within the national forest cluster. NFIs serve not only forestry purposes but also provide essential strategic information for national industrial policies (Tomppo et al., 2010). Due to their importance, NFIs must be regularly updated to reflect both management activities like silvicultural treatments and natural factors such as infestations and wildfires; the usual time between revisits of the same field plot ranges between 5 and 10 years (Gschwantner et al., 2022).

However, the rapid changes in forest dynamics driven by climate change (Popkin, 2021) have brought into question the adequacy of traditional monitoring methods such as the NFIs. These concerns stem from the rapidly evolving forest conditions which these inventories may not accurately reflect. This has led to a realization of the growing need for more timely and dynamic monitoring approaches. Additionally, many required forest attributes for developing adaptive strategies are either not commonly recorded or difficult to derive from in traditional sources (Goodbody et al., 2021). Simultaneously, the evolving scope of forestry and environmental policies, especially in response to global challenges like climate change and biodiversity loss, has increasingly exposed the limitations of traditional monitoring methods. Frameworks such as the REDD+ and the Paris Agreement have significantly influenced the direction and methods of forest monitoring. The shift has led to the development of more comprehensive and sophisticated forest monitoring practices, as traditional methods often fall short of the scale and frequency of data required by modern policy needs (Wulder et al., 2024). Ground surveys, while detailed, are limited in scope and impractical for extensive assessments, especially in remote or inaccessible regions. This concern is further compounded by the resource-intensive nature of traditional methods, which demand considerable manpower and logistical planning, leading to delays in information availability (Fassnacht et al., 2024; Zweifel et al., 2023)

In addressing the limitations of traditional forest monitoring methods, forestry and environmental monitoring have increasingly adopted advanced technologies such as Earth Observation (EO) data, Geographic Information Systems (GIS), and Machine Learning

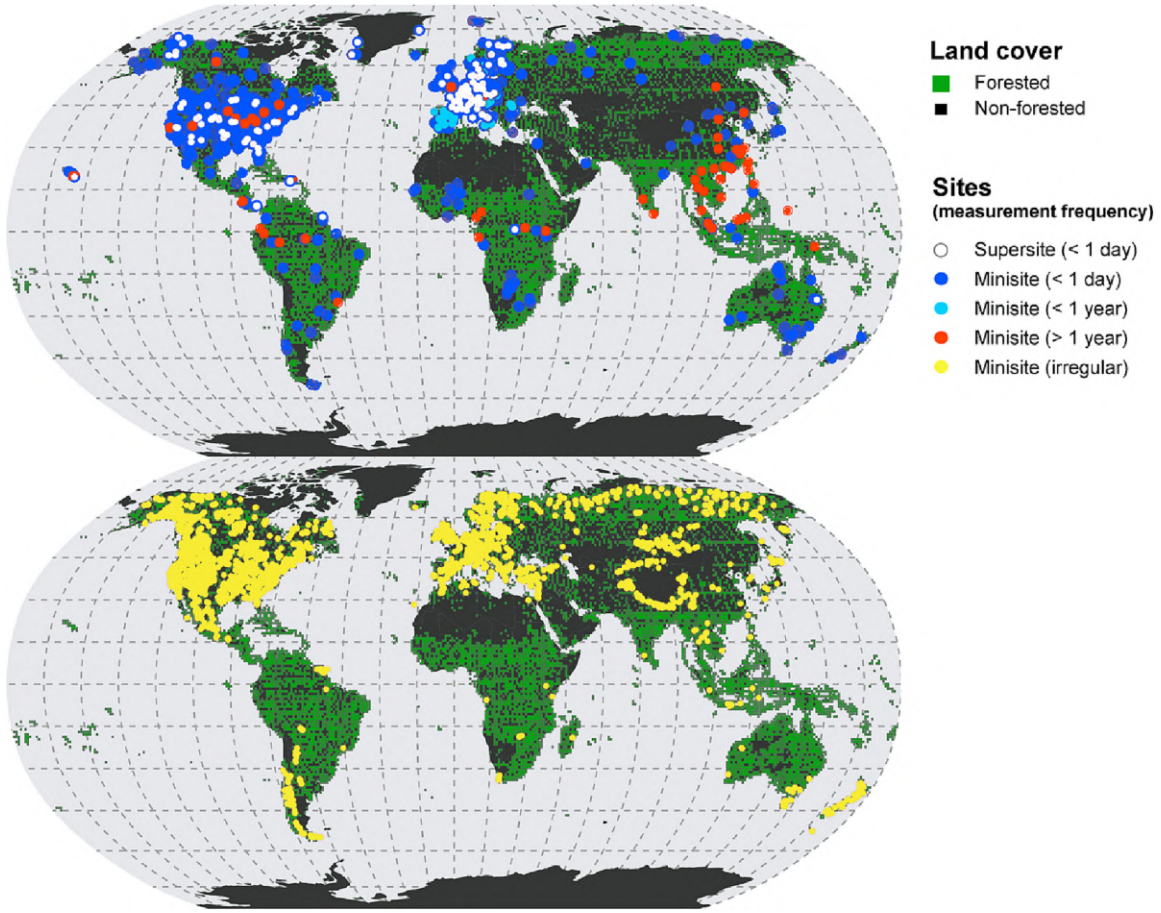

Figure 1.1: Distribution of global vegetation research infrastructure, adapted from Zweifel et al. (2023). Green and black squares indicate forested and non-forested land cover. Circles indicate the location of existing forest observation infrastructure. White circles indicate forest research with a high density of measurement devices and a high measurement frequency (i.e. ICP forest (ICP Forests, 2010)). Other-colored circles refer to smaller infrastructures with a lower density of devices and various database update frequencies of less than one day, less than one year and more than one year. Lower map includes locations with databases not regularly updated, like the TRY plant trait database (Kattge et al., 2020) or NFIs. Note that only NFIs whose precise location is public are included in the figure. Refer to the original publication for more detailed information.

(ML) techniques. These technologies offer several advantages over traditional methods, including broader area coverage, higher frequency, and reduced manpower requirements, facilitating near real-time monitoring. This is crucial for adaptive management and timely policy responses to environmental changes and emergencies (Goodbody et al., 2021; Wulder et al., 2024). The integration of these advanced techniques with traditional methods represents a critical advancement in forest monitoring and assessing spatiotemporal changes in forest resources, combining the detailed, local insights of traditional surveys with the expansive coverage and efficiency of modern technologies. What began as an exploration

of the use of black and white aerial photographs has evolved substantially. Today, these technologies have become indispensable in the scientific domain of forestry, transitioning from a novel concept to operational tools.

1.1.3 Earth observation and machine learning for spatiotemporal modeling of vegetation

Earth observation systems harness spectral, spatial, and temporal information over a local to global scale through a variety of sensors aboard satellites. The latest advances and investments in Earth Observation Programmes (EOP) for global environmental data acquisitions enable a comprehensive assessment of the status of vegetation: optical sensors, like those on the Landsat series satellites from the NASA and the U.S. Geological survey, and the Sentinel-2 satellite from the European Copernicus program, offer high resolution (respectively, 30 m and 20-10 m) imagery that is crucial for mapping forest cover (Hansen et al., 2013) and forest changes (Potapov et al., 2020; Potapov et al., 2022), tree species (Breidenbach et al., 2021; Grabska et al., 2020; Hermosilla et al., 2022) or forest disturbances and forest recovery (**Bonannella** et al., 2022a; Senf and Seidl, 2021a; White et al., 2017). Radar sensors, such as those on the Sentinel-1 satellite, penetrate cloud cover and provide data regardless of weather conditions, making them invaluable for year-round monitoring of forest biomass and structure (Babiy et al., 2022) or detecting deforestation (Reiche et al., 2021; Welsink et al., 2023). LiDAR (Light Detection And Ranging) technology, although less commonly available from space, provides detailed three-dimensional information about forest canopy structure when deployed on airborne platforms (Næsset, 2002; White et al., 2013); the GEDI (Global Ecosystem Dynamics Investigation) mission exemplifies the usage of spaceborn LiDAR, delivering unparalleled detail on canopy vertical structure on global scale (Dubayah et al., 2020). Hyperspectral data enables precise applications in vegetation type discrimination (Fricker et al., 2019; Zhang et al., 2020) or stress detection (Einzmann et al., 2021) due to its ability to capture a wide spectrum across hundreds of narrow bands and discern subtle spectral differences indicative of various plant properties.

Most of the continuous EOP such as MODIS, Landsat or Sentinel have also implemented a free data policy (Woodcock et al., 2008), which has been a pivotal factor in the widespread adoption of EO data for vegetation modeling applications. Furthermore, improvements in the spatiotemporal resolution of EO data have significantly enhanced our ability to detect and monitor minor changes on the land surface with greater level of detail (Berner and Goetz, 2022; Higgins et al., 2023; Lang et al., 2023). This democratization of data access has not only spurred a surge in remote sensing's application in forestry but has also underscored the essential role of freely available data in driving research and practical interventions in forest management and conservation (Fig. 1.2).

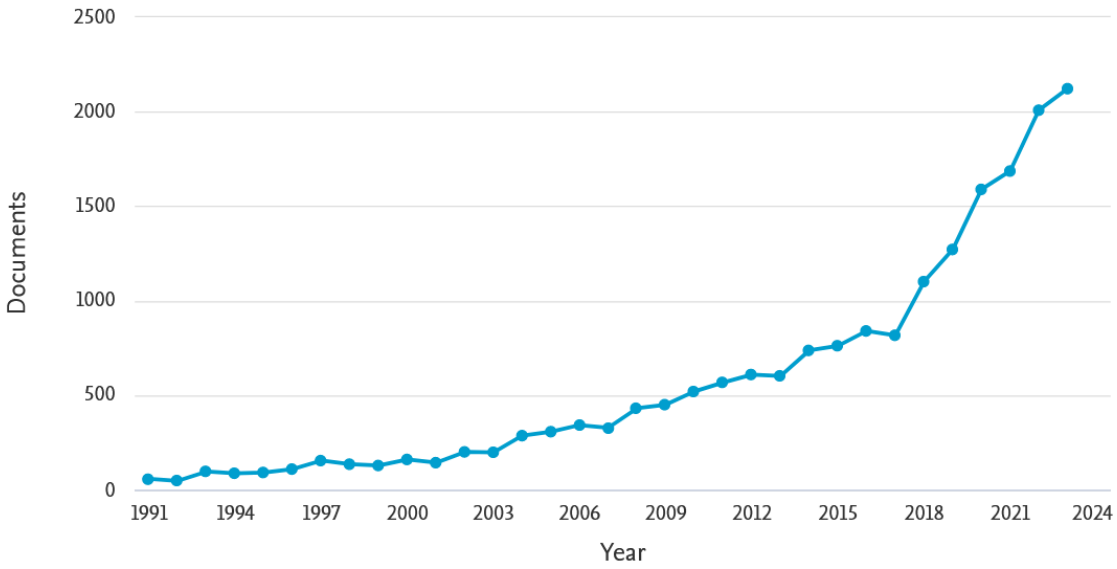

Figure 1.2: Yearly publications indexed by Scopus from 1991 to 2023. The queries included were: "remote sensing" AND "modeling" AND "vegetation" OR "forest". A total of 18979 documents were found; an increasing trend can be seen from 2008 first (Landsat open archive), with a sudden spike in 2017 (Cloud computing geospatial analysis with Google Earth Engine).

For this reason, and thanks to the exponential increase in computing power of the last decade (Gorelick et al., 2017; Yang et al., 2017), solutions such as Machine Learning algorithms have recently become increasingly important in interpreting the vast and complex datasets provided by EO technologies. ML tries to learn the relationship between the response and the predictors through the observation of dominant patterns (Breiman, 2001b). Contrary to traditional statistical models, no kind of ecological assumptions are explicitly embedded in ML algorithms: ML can be especially useful when dealing with data gathered without a specific and rigorous sampling design (Ij, 2018). ML algorithms have great potential to analyze the large amount of EO data, enabling the mapping and monitoring of changes on multiple geographical scales in a timely manner through reproducible research (Gobeyn et al., 2019). Liu et al. (2018) provide a complete review of forest ecology applications of three commonly used ML methods: decision-tree learning (DT), artificial neural network (ANN) and support vector machines (SVMs). In their review, they show how ML algorithms have extensively been used for species distribution models (SDM), carbon cycles, hazard (avalanches, wildfire and windstorm) assessment and prediction and other more specific tasks like predicting aboveground biomass, growing stock volume or tree height; Mountrakis et al. (2011) provide a review focused only on SVMs using remotely sensed data for different vegetation modeling applications, such as monitoring chlorophyll concentration, gross primary productivity and evapotranspiration, or for classification of hyperspectral data. Random Forests, an ensemble learning method that utilizes multiple decision trees, has frequently been used for tree species classification

(Breidenbach et al., 2021; Grabska et al., 2020; Waser et al., 2021). Hamedianfar et al. (2022) provide a review on the usage of Deep Learning methods, a branch of ML, particularly on estimating forestry characteristics and time-series dynamics (i.e. DBH, height, aboveground biomass, height increment etc), reporting more than a hundred of studies for the period 2017–2020.

In this sense, the Global Forest Change product from Hansen et al. (2013) stands as a pioneering effort in global-scale, high resolution forest monitoring. Leveraging the information provided by the Landsat mission image archive, this initiative successfully mapped global tree cover, delineates areas of forest loss and gain for the period 2000–2012, all at 30 m pixel level (Fig. 1.3). While not free from limitations especially for local analysis (O'Lear, 2020), given the importance of the product the maps have been updated yearly and included as one of the main component of the Global Forest Watch (GFW) initiative.

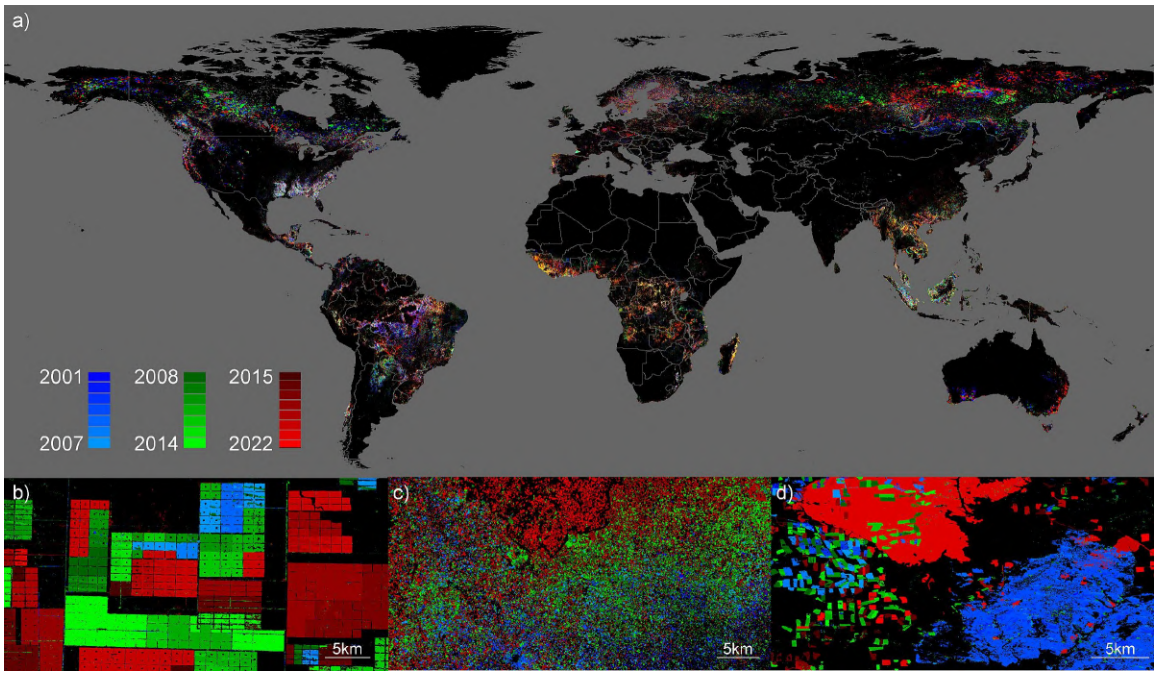

Figure 1.3: (a) Global forest cover loss, 2001-2022, shown averaged at a 0.05°resolution in red, green and blue by epoch. Insets are full resolution zooms of (b) Chaco woodland clearing in Paraguay, (c) shifting cultivation expansion in humid tropical forest of the Democratic Republic of the Congo and (d) boreal forest logging and fire in Siberia, Russia. Adapted from Hansen et al. (2013).

Building upon the methodology introduced by Hansen et al. (2013), the development of high resolution forest attribute products began to emerge. Potapov et al. (2021) produced the first 30 m resolution global canopy height map, using Landsat imagery and GEDI-derived measurements. Shortly thereafter, Lang et al. (2023) improved this work by employing Sentinel-2 data instead of Landsat to create a 10 m resolution canopy height product,

leveraging the same GEDI data. After many previous attempts done at coarser resolutions (Araza et al., 2023), Harris et al. (2021) mapped annual forest-related greenhouse gas emissions and removals at 30 m for the period 2001—2019, effectively adding aboveground biomass to the list of main forest attributes mapped at high resolution and global scale. An excellent example of integrating traditional methods with EO data and ML tools is provided by the National Terrestrial Ecosystem Monitoring System (NTEMS) developed by the Canadian Forest Service to provide national-scale baseline information on Canada's forested ecosystems (Hermosilla et al., 2016); based largely on Landsat data, the system provides time series high resolution maps of most of the main traditional forest attributes included in NFI surveys (on top of the ones previously mentioned, tree species composition was added as well) and additional forest attributes which have been recognized as essential information in the last decade, such as data on forest disturbances intensity, drivers and recover (White et al., 2014; White et al., 2017). Wulder et al. (2024) pushed the bar even higher by using this data collection to propose a stand-level satellite-based forest inventory (SBFI) system.

1.2 Research gaps

Despite the increase in availability of EO data and in opportunities provided by advanced ML solutions, having comprehensive and accurate information on vegetation across diverse ecosystems and different spatial and temporal scales remains a significant challenge. Due to their inherent complexity, the responses to climate change of vegetation ecosystems are highly variable, depending on specific spatial patterns and sensitivities to global climate shifts. This variability introduces a feedback mechanism in vegetation-climate interactions, raising concerns about the potential of climate change to significantly impede vegetation activities (Mennis, 2006) or causing a variety of unexpected changes in vegetation intensity (Chen et al., 2022; Reichstein et al., 2019). While traditionally process-based models such as Dynamic Global Vegetation Models (DGVMs) have been employed to study vegetation changes under various climate change scenarios (Hickler et al., 2012), these models, while invaluable, often face challenges in capturing the full spectrum of vegetation responses to changing climates due to their generalized assumptions and the complexity of ecological interactions (Lasslop et al., 2020). It is true that these models often incorporate complex interactions between climate, land use, and vegetation, but they also often struggle with uncertainties in climate projections, limited spatial resolution, the challenge of integrating diverse data sources for accurate predictions or the unavailability of input data (Bao et al., 2021; Beigaitė et al., 2022; Zaehle et al., 2005). Conversely, data-driven models using ML algorithms offer computational efficiency and adaptability by leveraging observational data and EO technologies, albeit with interpretability and extrapolation challenges (Meyer and Pebesma, 2021); one of the most common method using EO data for future forecasting involves employing ML models to create a relationship between the NDVI (Normalized

Difference Vegetation Index), used as a vegetation proxy, and climatic factors or Net Primary Productivity (NPP) (Bao et al., 2021). This method however presents limitations as it relies on a proxy rather than incorporating direct ground observations of vegetation within the targeted area: Huang et al. (2021) highlight in their review the limitations of the NDVI for accurately depicting vegetation status, particularly noting its saturation effects. These effects could lead to significant overestimations or underestimations when forecasting vegetation changes on a global scale especially in the tropical regions (Zeng et al., 2023). Another approach would be to train ML models using ground observations labeled according to a biome classification scheme to then extrapolate these models into the future using climatic projections, thus avoiding using proxies; examples in this sense are however few and limited in either spatial resolution or scale.

Understanding the distribution and dynamics of vegetation at the ecosystem level sets the foundation for examining more granular ecological units. Given the importance of forests in the overall balance of vegetation ecosystems, tree species, being the fundamental building blocks of forests, play a crucial role in determining overall ecosystem stability. Identifying the range, constraints and drivers of species distribution is the domain topic of species distribution modeling; while being a relatively recent area of focus thanks to GIS tools, delineating species distributions has always been a primary goal of ecology (Andrewartha and Birch, 1954) and is today even more important in the context of climate change (Afuye et al., 2021). Commonly, SDMs rely on climatic or bioclimatic factors at a coarse spatial resolution ($\geq$1 km), a choice mostly dictated by data availability rather then species ecology (Guisan and Thuiller, 2005; Mayer and Cameron, 2003; Porfirio et al., 2014), while in the temporal dimension long time averages (30-–50 years) are often used (Iturbide et al., 2018a). Due to the prohibitive costs of field campaigns solely for species data collection, most SDM studies depend on pre-existing, often outdated, datasets or the Global Biodiversity Information Facility (GBIF) database, rather than incorporating diverse, up-to-date observational data from varied sources. Furthermore, combining EO data, and specifically the use of high spatial resolution, with ML algorithms for SDM has only recently started to be explored (Gelfand and Shirota, 2021; Hefley and Hooten, 2016; Pérez Chaves et al., 2018). High spatial resolution (<100 m) maps, despite usually having slightly lower accuracy, are more beneficial for conservation purposes and management decisions than coarser but more accurate maps (Gottschalk et al., 2011; Guisan et al., 2013; Manzoor et al., 2018; Prates-Clark et al., 2008). Similarly, there is a growing need for mapping products with finer temporal resolution to accurately document the rapid ongoing ecological shifts caused by climate change. Given the focus of SDM research on long term future forecasting (2080–2100) (Williams et al., 2007), these studies tend to assess future species distributions by extrapolating the relationship between a species ecological niche and its environment under equilibrium assumption. In the case of tree species, there is general consensus that the species range would either shift towards the poles (Berner and Goetz, 2022; Zhu et al., 2012; Zhu et al., 2014) or at higher elevations (Feeley

et al., 2011; Maharjan et al., 2023). However, this approach risks either underestimating or overestimating potential shifts in species distribution, as it does not account for the ongoing changes that could be the result of not only climate but of interactions between the environment, local conditions and other species. No studies have fully utilized dense time series of realized species distribution maps produced with EO data to assess ongoing shifts, which could allow for the early detection of areas at risk or the establishment of high value conservation areas (i.e. ecological *refugia*). Furthermore, the identification of an ongoing trend could allow for the development of data-driven approaches to future forecasting.

NFI data has traditionally been used to collect and assess information on tree species (Ewald, 2012; Rigling et al., 2013; Scherrer et al., 2022), yet they lack comprehensive wall-to-wall coverage and their decadal frequency limits data collection to build statistically significant trends. Integrating NFI data with the complementary information provided by EO data and ML models presents a tested, successful strategy for creating detailed, high-resolution products with higher temporal frequency (Hermosilla et al., 2022; Strickland et al., 2020; Wulder et al., 2024). These results are possible only when using the precise geographical coordinates of the NFI plots. However, the common practice of degrading NFI plot geographic coordinates' precision of public or open-access datasets, primarily for privacy concerns, poses a significant barrier to remote sensing practitioners. This degradation limits the integration of location-specific NFI observations with the dynamic mapping capabilities afforded by EO data and ML algorithms, affects the spatial accuracy of the derived maps and limits the scope of ecological insights that can be derived from these products (Ceccherini et al., 2022; Fassnacht et al., 2024).

1.3 Research objectives

The overall objective of this PhD thesis is to integrate Earth observation data, machine learning technologies and field plot data to enhance our understanding of ongoing vegetation dynamics and the overall monitoring and management of forest ecosystems. More specifically, this thesis aims to answer the following research questions:

1. What is the impact of climate change on potential biomes distribution based on Earth observation and machine learning methods? And what are the projected shifts in vegetation under various climate change scenarios?

2. What combination of Earth observation and machine learning methods allows to map and analyze the distribution of forest tree species at high resolution?

3. How can these methods be applied to capture trends and disturbance impacts on forest tree species distributions and how do these reflect the ongoing changes in forest ecosystems?

4. What is the effect of coordinate precision in NFI data on the accuracy of high-resolution tree species classification models?

1.4 Thesis overview

The thesis is composed of six main chapters, with Chapters 1 and 6 framing the core content and Chapters 2 to 5 addressing each one of the research questions presented in Section 1.3. **Chapter 1** serves as an introduction, delineating the current advancements in Earth observation and machine learning as they apply to vegetation modeling and monitoring. It also identifies existing research gaps and explores potential approaches for addressing them.

Chapter 2 investigates the projected shifts in vegetation under various climate scenarios by developing a data-driven approach based on machine learning techniques and pollen reconstruction datasets coming from different surveys that have been harmonized over the years. The chapter specifically focuses on modeling the potential distribution of global biomes, incorporating a range of climate variables to predict the current distribution and how different climate change scenarios could alter vegetation patterns and distributions in the future. Through this analysis, the chapter provides insights into the expected changes in biome locations and compositions, particularly the contraction of tropical rainforests and the expansion of higher latitude biomes like boreal forests, at polar biomes' expense. The chapter offers a clearer picture of how climate change could reshape the global ecological landscape; given the unpredictability of the future, the method is rather conservative in its conclusions: after analyzing every emission pathway scenario, it focuses on identifying key change areas, irrespective of specific emission pathways, providing a nuanced understanding of climate change's impact on the global ecological landscape.

Chapter 3 presents a more sophisticated data-driven approach than Chapter 2, integrating 305 environmental predictors, primarily from Earth observation data, employing advanced machine learning techniques alongside over 3 million tree occurrence points from different surveys or citizen science projects. This approach is used to model the potential and actual distribution of 16 key European forest tree species from 2000 to 2020 at high resolution. The chapter's core is the development and application of ensemble models for each species and distribution type, employing a series of machine learning algorithms such as random forest, gradient boosting, and neural networks, and utilizing spatial cross-validation to enhance predictive accuracy and avoid model overfitting. The influence of the spatial resolution of the environmental predictors on model performances was also tested. As expected, not only ensemble models outperformed individual models, but ensemble models using high spatial resolution data proved to be consistently better than the ones using coarse resolution data.

Chapter 4 delves into the changing dynamics of 6 key European forest tree species between 2000 and 2020, building upon the high resolution distribution maps generated in Chapter 3. This chapter focuses on the response in the realized distribution of these species over the last twenty years by conducting a trend analysis on the distribution maps and then degrading the results to a coarser resolution to find the prevalent spatial patterns for each species. Findings indicate a predominant stability across species, with most areas exhibiting consistent distribution patterns. Nonetheless, the chapter also uncovers areas of concern, marked by declining species presence, especially at the latitudinal edges of their natural ranges. Additionally, the chapter examines the impact of forest disturbances on species distribution, noting a doubling in the range affected by disturbances such as wind and fire. The findings highlight the need for proactive conservation and long-term planning to ensure the resilience of European forests against climate-induced changes.

Chapter 5 investigates the impact of coordinate precision in National Forest Inventory (NFI) data on the accuracy of high-resolution tree species classification models. Focusing on 7 European tree species, the chapter compares machine learning models trained on true, precise plot coordinates with those trained on coordinates derived from publicly available data. The chapter describes a procedure to estimate the precise plot coordinates location with the aid of high resolution forest masks and forest types layers for the country of the Netherlands. The analysis demonstrates that models utilizing true coordinates consistently outperform those based on estimated coordinates. This is evident in the higher scores reached across performance metrics used for both hard classification and probabilistic classification problems across most species classes for the model using precise data. The chapter highlights the critical importance of coordinate precision in ecological modeling and underscores the need for high-quality, accurate NFI data for effective tree species classification using Earth observation data and machine learning methods.

Lastly, **Chapter 6** summarises the main findings and results of the previous chapters and provides an overview on how they contribute towards filling the research gaps identified by this thesis. Following that, it offers a reflection and outlook, exploring potential pathways for future applications and new research prospects in the integration of Earth observation data with machine learning and field data for vegetation modeling.

Chapter 2

Global potential biomes distribution under climate change scenarios

This chapter is based on:

C. Bonannella, T. Hengl, L. Parente, and S. de Bruin (2023). "Biomes of the world under climate change scenarios: increasing aridity and higher temperatures lead to significant shifts in natural vegetation". *PeerJ* 11, e15593. DOI: `https://doi.org/10.7717/peerj.15593`

Abstract

The global potential distribution of biomes (natural vegetation) was modelled using 8959 training points from the BIOME 6000 dataset and a stack of 72 environmental covariates representing terrain and the current climatic conditions based on historical long term averages (1979–2013). An ensemble machine learning model based on stacked regularization was used, with multinomial logistic regression as the meta-learner and spatial blocking (100 km) to deal with spatial autocorrelation of the training points. Results of spatial cross-validation for the BIOME 6000 classes show an overall accuracy of 0.67 and R^2_{logloss} of 0.61, with *"tropical evergreen broadleaf forest"* being the class with highest gain in predictive performances ($R^2_{\text{logloss}} = 0.74$) and *"prostrate dwarf shrub tundra"* the class with the lowest (R^2_{logloss} = -0.09) compared to the baseline. Temperature-related covariates were the most important predictors, with the mean diurnal range (BIO2) being shared by all the base-learners (i.e. random forest, gradient boosted trees and generalized linear models). The model was next used to predict the distribution of future biomes for the periods 2040–2060 and 2061–2080 under three climate change scenarios (RCP 2.6, 4.5 and 8.5). Comparisons of predictions for the three epochs (present, 2040–2060 and 2061–2080) show that increasing aridity and higher temperatures will likely result in significant shifts in natural vegetation in the tropical area (shifts from tropical forests to savannas up to 1.7×10^5 km^2 by 2080) and around the Arctic Circle (shifts from tundra to boreal forests up to 2.4×10^5 km^2 by 2080). Projected global maps at 1 km spatial resolution are provided as probability and hard classes maps for BIOME 6000 classes and as hard classes maps for the IUCN classes (6 aggregated classes). Uncertainty maps (prediction error) are also provided and should be used for careful interpretation of the future projections.

2.1 Introduction

Climate change is one of the biggest threats to human civilization, with slowly accumulating effects and unknown instabilities in front of us and future generations. To assess the potential impacts of climate change on the environment and to help us mitigate and prepare for negative effects, scientists offer predictions of possible futures including global maps of the Earth's environment in the future (Dow and Downing, 2016). Global datasets projecting the state of the Earth's environment include future climate predictions e.g. Representative Concentration Pathways (RCPs) (Hayhoe et al., 2017), future land use predictions e.g. Chen et al. (2020) and (Hurtt et al., 2020), human population scenarios (Jones and O'Neill, 2016), future terrestrial ecosystems maps (Nolan et al., 2018), future ecosystem productivity (Yin et al., 2023), and future gridded emissions (Fujimori et al., 2018). Even though the accuracy of these projections in the far future cannot currently be validated, such exercises are deemed useful as they help reveal patterns and assess the impact of scenarios. In essence, there are two main approaches to envision the future state of Earth's environment (Hayhoe et al., 2017; Reichstein et al., 2019):

1. Process-based mechanistic modeling: simulating evolution of the environment using biophysical process-based Earth System Models (ESM);
2. Data-based modeling: training predictive models using observations from the past and then extrapolating these models into the future;

Process-based modeling is often preferred by physicists as the relationships between model entities are explicitly defined. Examples of projected changes of land use based on the global Earth System Models are the LUH2 project (Hurtt et al., 2020) and Lund–Potsdam–Jena managed Land (LPJmL) model (Rolinski et al., 2018). In the case of data-based modeling, predictions and results of analyses are based on finding relationships between the target property and covariates and then fitting statistical models that are next used to predict values based on unseen combinations of states in feature space. Two common approaches here are: (1) use actual ground observations i.e. monitoring stations to fit spatiotemporal models (Hengl et al., 2018), and (2) use complete Earth observation data cubes and then basically all pixel combinations to visualize and model relationships (Mahecha et al., 2020). An advantage of the data-based modeling is that it is often computationally less demanding than process-based modeling and it can be extended by adding more covariates (Beigaitė et al., 2022). In addition, process-based modeling requires several assumptions and, in the case of chaotic behaviour or non-linear spatial scaling of features, it is often difficult to produce credible predictions. On the other hand, data-based modeling comes with the risks of producing poor predictions in the extrapolation space and the models are often difficult to interpret (Meyer and Pebesma, 2021). Yet, strict data-based modeling requires neither subjective parametrization nor model assumptions, and hence it can be considered less complex to start with. It is not to say that the approaches are mutually exclusive and

can't be combined: there is a full spectrum of models from process-based to data-based, which includes hybrid physics-based data-driven models. Different approaches exist in this sense: using data-driven models but constrain the results with boundary conditions derived from physics-based climate models (as suggested by Lindgren et al. (2021)), including the representation of natural processes in the data-driven model (Higgins et al., 2012) or using process-based models whose results have been parametrized and calibrated on real data (Higgins et al., 2023).

Predictions of future states of climate, land cover, terrestrial ecosystems, human population and similar have proven to be useful, with many of the datasets being frequently cited and used to communicate our possible futures (`https://probablefutures.org/`). Su et al. (2021) modeled yield gains under Conservation Agriculture (CA) and various practices for the future climate scenarios and found out that overall performance of CA will most likely decrease in the future in most temperate regions in South America, including Uruguay, southern Brazil and northern Argentina for barley, cotton, rice, sorghum and sunflower. Krause et al. (2022) has recently modelled the impacts of anthropogenic land cover changes on global Gross Primary Productivity (GPP) using maps of historical agricultural expansion and future land-use changes based on the 25 km resolution LUH2 dataset (Hurtt et al., 2020). Their results indicate that global GPP might get further reduced owing to agricultural expansion and to extents that depend on the prevailing scenario. Beigaitė et al. (2022) provides predictions of future distribution of MODIS vegetation types using machine learning and focusing on climate extremes (e.g. extreme cold days). Their results indicate that prediction accuracy can be improved by extending the averaged climatic conditions with maps of climate extremes e.g. bioclimatic variables and similar.

When it comes to mapping future vegetation, only few datasets are available and typically at coarse resolutions. Nolan et al. (2018) provides predictions of terrestrial ecosystems in the future as a function of annual temperature and simple logistic spline regression with ordered categories. Their results suggest that terrestrial ecosystems are at risk of major transformation. Despite these recent efforts, there is still no analysis of the main future trends in air temperature and precipitation and the magnitude of such change on potential vegetation on a global scale. Furthermore, most of these datasets are provided without per pixel uncertainty estimates. The existence of various biome classification schemes makes things even more confusing, since they can be overly subjective (Higgins et al., 2016) and in some cases they implicitly invoke climate (Moncrieff et al., 2015) in their definition: many of the early biome classification schemes included climate in their definition as a proxy for functional characteristics, traits and adaptations that were difficult to map properly at a global scale (Moncrieff et al., 2016) and only later on schemes based on Plant Functional Traits (PFTs) or ecosystem productivity have been developed; a paradygm shift has also taken place in the last decades, from considering biomes a deterministic entity to a more dynamic concept, a result of an ensemble of different processes and feedback loops (Mucina, 2019). However, the lack of datasets at high resolution that could be used to predict

biome envelopes that follow the functional-based classification scheme is a limitation for its application to a global scale. Scientific studies that use data-driven approaches to forecast the state of vegetation into the future are usually limited on the spatial scale, spanning one or more countries or one continent at most (Maksic et al., 2022; Zevallos and Lavado-Casimiro, 2022), while another limitation consists in the usage of mostly one algorithm only (Random Forest) to conduct the analysis. The purpose of this study is to use a data-driven approach to provide consistent projections of future potential natural vegetation under different climate scenarios, including uncertainty estimates: we provide projections of 20 biomes for three (3) climatic scenarios (RCP 2.6, 4.5 and 8.5) for the future 60 years. To do that, we extend the work of Hengl et al. (2018), which used a biome classification scheme based on PFTs and tried to spatialize it to the whole globe by using an ensemble of climatic, topographic and remotely-sensed predictor variables. Compared to Hengl et al. (2018), we apply the following three substantial improvements:

1. Instead of only using Random Forest, we use an ensemble of three learners of different types, which allowed quantifying the prediction uncertainty;
2. For each pixel we provide class probabilities and prediction errors computed by bootstrapping;
3. Modeling is done using a consistent set of covariates so that the effects of climate change are controlled purely by the climatic projections.

The paper is divided in four parts: (1) we first describe our predictive mapping framework based on using biome training points (Harrison, 2017); (2) we evaluate the accuracy of the fitted ensemble model using spatial cross-validation and generate predictions for the three future scenarios; (3) we next aggregate predictions according to the IUCN Global Ecosystem Typology classification system (Keith et al., 2020) to make our product comparable with an international standard and (4) we finally highlight the most pronounced changes per continent and biome type.

2.2 Material and methods

2.2.1 General workflow

We modeled the potential distribution of biomes on a global scale for current and future time periods using an ensemble machine learning approach. The model was trained on reference biome data compiled from pollen and fossil reconstructions (Harrison, 2017) along with regional environmental variables describing topography and long-term climatic averages. We used CHELSA climatological data (Karger et al., 2017) from the time period 1979–2013 to simulate the baseline potential natural distribution of biomes for the current (2022–2023) time period: since our goal was to model the potential natural vegetation, we tried to predict which PFT-based class of biome would be the dominant one in a specific

location based on environmental variables only. Future climatic conditions instead cover the epochs 2041–2060 and 2061–2080. For the future epochs we considered three different climate change scenarios using the concept of *"Representative Concentration Pathways"* (Van Vuuren et al., 2011), or, in short, RCPs. The ones used in this study are RCP 2.6, RCP 4.5 and RCP 8.5.

The output of the projections is provided as probability maps (0–100%) at 1 km spatial resolution, with the probabilities in each pixel summing to 100%. For each class we also provide model uncertainty maps. We excluded the continent of Antarctica, because of the presence of permanent ice areas and lack of training points. Also other areas covered by water bodies, barren land and permanent ice according to ESA's global land cover maps for the period 2000–2015 (ESA, 2017) were excluded from the analysis. We generalized the 20 biome classes analyzed in this study to 6 classes following the Global Ecosystem Typology classification system employed by the International Union for Conservation of Nature (IUCN) (Keith et al., 2020). We then compared the two epochs for each of the climatic scenarios with the current time period: using the latter as a baseline for the distribution of potential natural vegetation, the goal was to identify those areas where the change in climatological conditions could lead to a shift in the potential distribution.

2.2.2 Training points

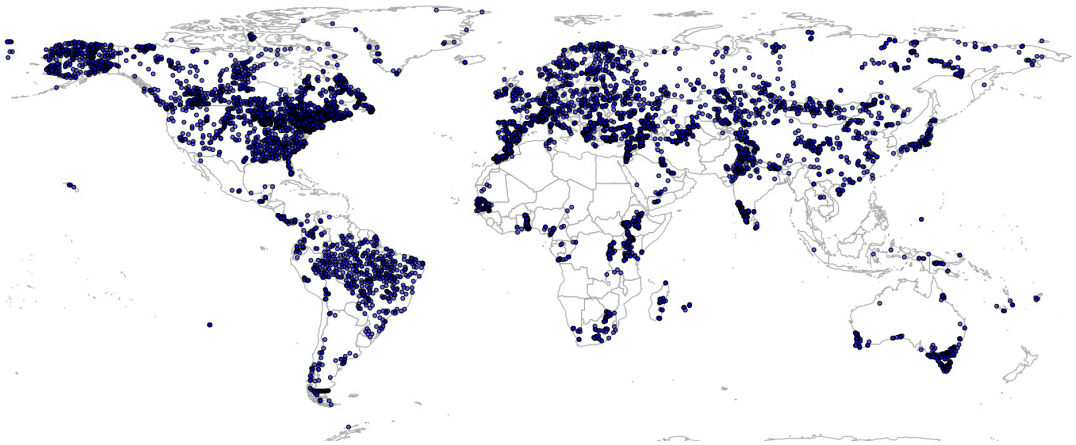

Figure 2.1: Global spatial distribution of the BIOME 6000 dataset enriched by Hengl et al. (2018). Despite the added pseudo points, there are still large areas (Patagonia, Sahara region, Central Africa and most parts of Australia) not covered by any observation.

We used the BIOME 6000 data set, compiled by Harrison (2017), with additional 350 pseudo observations to cover under-represented areas in South America (Fig 2.1) as described in Hengl et al. (2018), for a total of 8,959 points. The BIOME 6000 project aims to reconstruct past vegetation distributions from pollen and fossil records from different time periods, from the recent past (the last 50 years) to approximately 21 ka ago; in this study, following Hengl et al. (2018), we only used the points belonging to the most

recent time period. The method, described by Prentice and Webb III (1998), was used to assign the recovered *taxa* to PFTs, which were next ascribed to a specific biome following PFT-based biomes definitions. From the first version of the data set to its final publication by Harrison (2017), almost 20 years have passed: over this period, multiple surveys have been conducted on the same locations, resulting in more than one biome reconstruction per location. Furthermore, initially absent regions have been added to the original data set. To avoid issues with harmonization of nomenclature between biomes, Harrison (2017) provide a standardized classification legend that can be globally applied (32 biomes in total) and a *megabiome* classification legend (8 megabiomes in total). While the *megabiome* system implies a necessary loss of information due to generalization, the original standardized classification system devised by Harrison (2017) has been considered too detailed and location-specific to be used for global modeling (Hengl et al., 2018). We adopted the 20 classes (Fig. 2.2) system devised by Hengl et al. (2018) for the sake of data-model comparison.

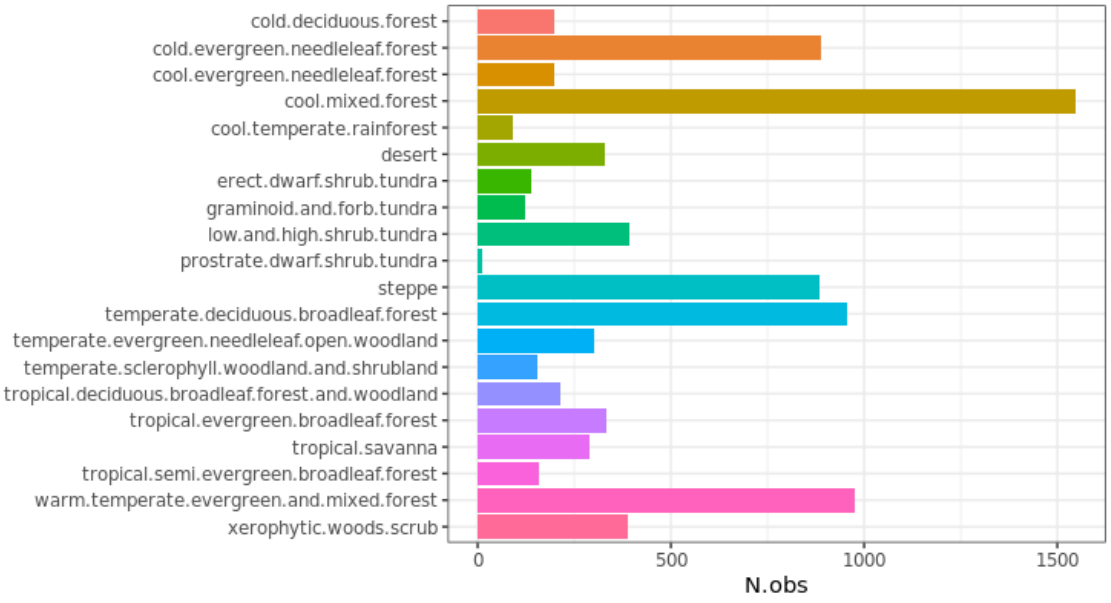

Figure 2.2: Number of observations per biome class. Note the strong imbalance between the different classes, with the most abundant class ("*cool mixed forest*") counting $> \sim$1500 observations while the least abundant ("*prostrate dwarf shrub tundra*") counts < 20 observations.

2.2.3 Predictor variables

A total of 72 spatially explicit and harmonized variables representing climatic, bioclimatic and topographic factors were used for modeling purposes. All the layers were resampled to a standard grid covering latitudes between 87.37°N and 62.0°S and reprojected to the coordinate reference system EPSG:4326 before the analysis. The original spatial resolution of the layers was used during the spatial overlay with the point dataset, while for the rest

of the calculations all the layers were resampled to a spatial resolution of 30 arcseconds (approximately 1 km at the equator).

We used long-term climate data and projections as provided by the CHELSA project (Karger et al., 2017). For future scenarios, we followed the work of the Intergovernmental Panel on Climate Change (IPCC) Assessment Reports (AR) based on narratives and outcomes of the Coupled Model Intercomparison Project (CMIP). IPCC AR5 (IPCC, 2014) featured CMIP5 model results using the concept of Representative Concentration Pathway (RCP), where each projected climatic scenario is labelled according to a possible increase in radiative forcing (from 2.6 to 8.5 W/m^2) values by 2100 due to increase in greenhouse gasses (GHG) emissions. The new IPCC AR6 (IPCC, 2021) featured instead CMIP6 model results while using a different concept, the "Shared Socioeconomic Pathways" (SSP): while RCPs did not include any socioeconomic factors in their modelization, SSPs included several assumptions on how population growth, technological development, climate policies and other similar factors would evolve by 2100. A subset of the new 50 CMIP6 models has been considered overly sensitive (i.e., "too hot") and with climate warming in response to carbon dioxide emissions that might be larger than supported by other evidence (Hausfather et al., 2022; Zelinka et al., 2020). For this reason, we decided to exclude CMIP6 models from our analysis and rely instead on CHELSA v.1.2 data with CMIP5 calculations, using an ensemble of 5 Global Circulation Models (GCMs): the Max-Planck-Institute Earth System Model (MPI-ESM-mr) (Giorgetta et al., 2013), the version 5 of the Model for Interdisciplinary Research on Climate (MIROC5) (Watanabe et al., 2010), the Community Earth System Model version 1 that includes the Community Atmospheric Model version 5 (CESM1-CAM5) (Neale et al., 2010), the version 5 of the Institut Pierre Simon Laplace Coupled Model (IPSL-CM5A-MR) (Dufresne et al., 2013) and the First Institute of Oceanography-Earth System Model (FIO-ESM) (Qiao et al., 2013). Since most of the GCMs are interdependent between each other, and not all of them include the three RCP scenarios we analyzed in this study, we followed the suggestions of Sanderson et al. (2015) for the selection process.

To train the model, we used average values for the period 1979–2013 for 17 bioclimatic variables, i.e., annual mean temperature, mean diurnal range, isothermality, temperature seasonality, maximum temperature of the warmest month, minimum temperature of the coldest month, temperature annual range, mean temperature of the wettest quarter, mean temperature of the driest quarter, mean temperature of the warmest quarter, mean temperature of the coldest quarter, annual precipitation, precipitation of the wettest month, precipitation of the driest month, precipitation of the wettest quarter, precipitation of the driest quarter, precipitation of the warmest quarter and precipitation of the coldest quarter. The Precipitation Seasonality (BIO15) was not included because of its excessive number of missing values in the layers of the future time periods. We also used monthly minimum, average and maximum temperature and monthly precipitation, for a total of 66 climatic and bioclimatic predictor variables. They can be downloaded from

https://chelsa-climate.org/downloads/. We used 6 additional predictor variables representing topographic conditions, i.e., sine and cosine of aspect, slope, upslope curvature and downslope curvature. These covariates have a 3 arcsecond resolution ($\sim$ 90 m at the equator) and they were derived from MERIT DEM (Yamazaki et al., 2017). MERIT DEM layers can be downloaded from http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/.

2.2.4 Model building and evaluation

We used an ensemble machine learning approach based on stacked generalization (Wolpert, 1992). Ensemble modeling techniques involve training several independent models with the same input data and then aggregating each of the model outputs into the final predictions. Stacked generalization uses the outputs of the individual models to train an additional model (*meta-learner* from here on) which then produces the final predictions. We used Random Forests (RF) (Breiman, 2001a), Generalized Linear Models (Nelder and Wedderburn, 1972) with Lasso regularization (Tibshirani, 1996) and Gradient-boosted trees (GBT) (Friedman, 2002) as component models for the ensemble model. To reduce overfitting in the training phase, we used a 5–fold spatial cross validation (Roberts et al., 2017): the out-of-fold predictions were used to train the *meta-learner*. Spatial cross validation was implemented by a 100 $\times$ 100 km grid and using the tile ID as the blocking variable during the training of the models. We used multinomial logistic regression (Wright, 1995) as the *meta-learner*.

Predictions are delivered as probability maps (0-100%) together with uncertainty maps: the standard deviation of the predicted values by the base learners serves as an indication of model uncertainty. The principle is that the higher the standard deviation, the more uncertain the model is regarding the proability to be assigned to the pixel (Brown et al., 2020). In contrast, for the hard class map we used the probability maps to calculate a per-pixel confidence metric. Contrary to Hengl et al. (2018), we chose not to use the per-pixel entropy (Shannon, 1948) but the margin of victory (Calderón-Loor et al., 2021) The margin of victory is defined as the difference between the first and the second highest class probability value in a given pixel. Potential values in this case would go from 0 (i.e. no difference between the first two classes, highest confusion possible) to 100 (the model is certain in the class probability value attribution, no confusion with other classes); in short, high values would be measures of low uncertainty, while low values would indicate a high uncertainty. All the analysis were performed using R (version 4.1.1) (R Core Team, 2021a) and, specifically, the *mlr* package (Bischl et al., 2016). For more details on the hyperparameter space used for the other component models and the overall architecture of the ensemble model, see **Bonannella** et al. (2022d).

We calculated the variable importance for each of the component models using Gini importance for RF, the gain metric for GBT (Shi et al., 2019) and the coefficients for

the minimum value of λ for GLM (Hastie et al., 2016): we took the 20 most important variables across the component models and retained the variables that these learners had in common. We then report these as the most important variables for the ensemble model. The predictive performance of the ensemble model was assessed through 5–fold spatial cross validation repeated 5 times with overall accuracy and the R^2_{logloss} (**Bonannella** et al., 2022d) as performance metrics. We then computed the R^2_{logloss} in addition to more classic metrics used for classification problems, like the True Positive Rate (TPR) and the F1 score (Van Rijsbergen, 1979) to assess model performances per class.

2.2.5 Shifts in potential biomes

While for data-model comparison we used the original 20 classes classification system from Hengl et al. (2018), to compare the model outputs we translated the classes in the IUCN Global Ecosystem Typology (Keith et al., 2020). This system classifies biomes based on functional characteristics and their structural role in the ecosystems rather than on climate, species distribution or vegetation patterns. Its principle is very similar to that of the BIOME 6000 classification system Prentice and Webb III (1998). The IUCN system comprises six hierarchical levels, with the three upper ones being *realms*, *biomes* and *functional groups*: the definitions of the functional groups are quite different from those of BIOME 6000, so we aggregated the 20 classes used in this study at the *biome* level according to the IUCN. We focused on the biomes present in the *terrestrial* realm, which include the following:

- **T1** - Tropical-subtropical forests biome;
- **T2** - Tempereate-boreal forests and woodlands biome;
- **T3** - Shrublands and shrubby woodlands biome;
- **T4** - Savannas and grasslands biome;
- **T5** - Deserts and semi-deserts biome;
- **T6** - Polar/alpine (cryogenic) biome;
- **T7** - Intensive land-use biome.

Since the focus of this paper is on potential biomes, the "*T7 - Intensive land-use biome*" class was not considered. The complete translation scheme is available in Table 2.1: we calculated the IUCN class by aggregating the per-class probability values of the BIOME 6000 classes according to the translation scheme. We computed the margin of victory for the IUCN classes as well and we used those maps to highlight areas with high confidence (i.e. low confusion) predictions. To assess change in potential biome class in fact, we calculated the difference in hard class between the potential biomes map of the current period and each of the future periods and RCP scenarios. We first reprojected all the IUCN classes and relative margin of victory maps to the Interrupted Goode Homolosine

projection, which is an equal-area composite projection. We chose it specifically to provide an unbiased (i.e. without geographical distortions) estimate of the areas subjected to change. In the results we discuss change dynamics only for the aggregated IUCN classes and only for pixels having a margin of victory $\geq 50\%$; pixels with a margin of victory $< 50\%$ are not considered.

Table 2.1: Overview of the translation scheme used to pass from BIOME 6000 to IUCN classes

BIOME 6000 class (from Hengl et al. (2018))	IUCN class
Tropical deciduous broadleaf forest and woodland Tropical evergreen broadleaf forest Tropical semi evergreen broadleaf forest	T1 - Tropical-subtropical forest biome
Cold deciduous forest Cold evergreen needleleaf forest Cool evergreen needleleaf forest Cool mixed forest Cool temperate rainforest Temperate deciduous broadleaf forest Temperate sclerophyll woodland and shrubland	T2 - Temperate-boreal forests and woodlands biome
Temperate evergreen needleleaf open woodland Warm temperate evergreen and mixed forest Xerophytic woods scrub	T3 - Shrublands and shrubby woodland biome
Tropical savanna	T4 - Savannas and grassland biome
Desert Steppe	T5 - Desert and semi-desert biomes
Erect dwarf shrub tundra Graminoid and forb tundra Low and high shrub tundra Prostrate dwarf shurb tundra	T6 - Polar/alpine (cryogenic) biome

2.3 Results

2.3.1 Model performances and variable importance

The hyperparameter tuning resulted in the following architecture for the ensemble model:

- Random forest: 452 trees, minimum node size 9, *mtry* 10, while the other hyperparameters were set to default;
- Gradient boosted trees: 20 boosting rounds, maximum depth per tree 5, learning rate 0.5, minimum loss reduction to split a leaf node 10, subsample ratio of the training instances 1, subsample ratio of columns when constructing each tree 0.5. The other hyperparameters were set to their defaults;
- Generalized Linear Models with Lasso: λ value 1.1×10^{-5};

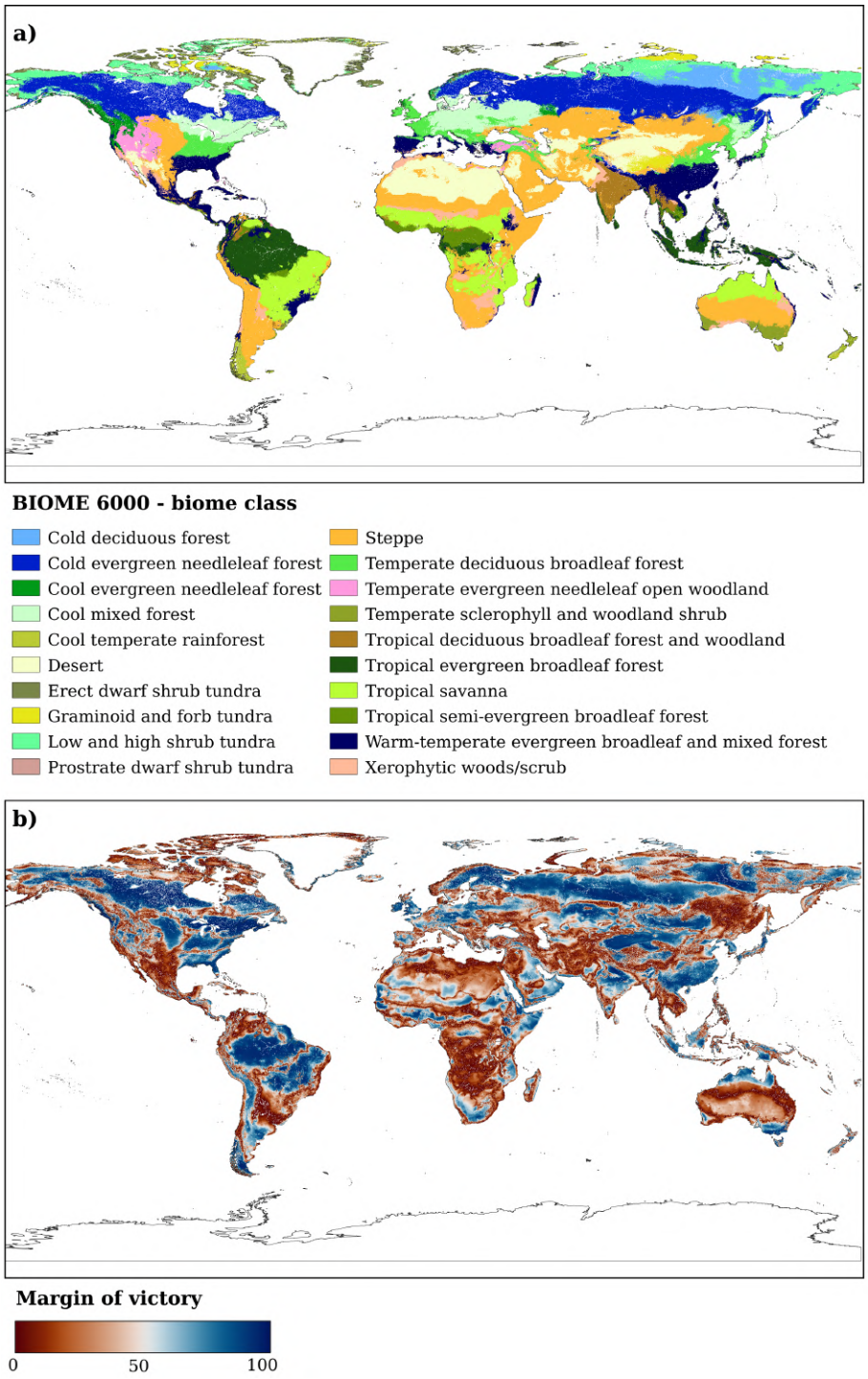

Figure 2.3: Distribution of **a)** the biome classes according to the BIOME 6000 classification scheme and **b)** the margin of victory for the current time period. The margin of victory is here used as an indication of uncertainty. High values (blue in figure) indicate high confidence in the attribution of dominant class by the model, while low values indicate high uncertainty.

- Multinomial logistic regression: multinomial function to minimize the loss.

The ensemble model had a moderate accuracy; according to the 5–fold spatial cross validation the overall accuracy is 0.67 and the R^2_{logloss} 0.61. Model performances per class are shown in Table 2.2. The "*tropical evergreen broadleaf forest*" is the class with the greatest gain in predictive performances ($R^2_{\text{logloss}} = 0.74$) compared to the baseline logloss, while the "*prostrate dwarf shrub tundra*" is the worst predicted class, with a negative gain in predictive performances compared to the baseline logloss (see Fig. 2.3).

The latter may be attributed to the very small (n.obs = 11) number of points in the training data for this specific class. It is also the only class with negative gain in predictive performances: all the other classes go from weak ("*cool evergreen needleleaf forest*", $R^2_{\text{logloss}} = 0.30$) to consistent ("*temperate sclerophyll woodland and shrubland*", $R^2_{\text{logloss}} = 0.71$) increase in predictive performances. The three models captured different parts of the feature space despite the relatively few (72) number of predictor variables. From the top-20 predictor variables, only one is shared across all component models, BIO2, the mean diurnal range. RF was the only component model which selected a topographic predictor (elevation) as one of the most important variables, while the other two models focused mostly on the climatic variables. RF and GBT shared 9 out of 20 predictor variables, with 7 out of these 9 being temperature-related (mean or maximum temperature and temperature-derived bioclimatic variables). GLM with Lasso differed mostly from the other two component models in the selected most important predictor variables. GLM was the only component model selecting variables from the group of the minimum temperatures.

Table 2.2: Results of the repeated 5–fold spatial cross validation per class

Class	**N.obs**	**TPR**	**F1**	$R^2_{\textbf{logloss}}$
Cold deciduous forest	199	0.51	0.57	0.53
Cold evergreen needleleaf forest	890	0.78	0.76	0.62
Cool evergreen needleleaf forest	198	0.23	0.31	0.30
Cool mixed forest	1548	0.81	0.79	0.62
Cool temperate rainforest	93	0.66	0.70	0.59
Desert	328	0.51	0.55	0.50
Erect dwarf shrub tundra	138	0.36	0.42	0.50
Graminoid and forb tundra	123	0.41	0.49	0.36
Low and high shrub tundra	391	0.68	0.66	0.63
Prostrate dwarf shrub tundra	11	0.00	0.00	-0.09
Steppe	884	0.67	0.66	0.46
Temperate deciduous broadleaf forest	958	0.62	0.62	0.47
Temperate evergreen needleleaf open woodland	302	0.58	0.59	0.52
Temperate sclerophyll woodland and shrubland	153	0.76	0.74	0.71
Tropical deciduous broadleaf forest and woodland	215	0.42	0.47	0.49
Tropical evergreen broadleaf forest	333	0.79	0.77	0.74
Tropical savanna	291	0.77	0.71	0.67
Tropical semi evergreen broadleaf forest	160	0.40	0.43	0.54
Warm temperate evergreen and mixed forest	976	0.73	0.67	0.52
Xerophytic woods scrub	387	0.45	0.48	0.42

2.3.2 Future predictions

Examining the biome state transitions from the current conditions to the future epochs, we found that most locations remained stable. Filtering the transitional areas with the margin of victory across all scenarios and epochs using 50% as a safety threshold value considerably reduced the predicted transitional area: less than 1% of Earth's surface showed signs of change, with the least changes found in the scenario RCP 2.6 (5.6×10^5 km^2) and the most in the scenario RCP 8.5 for the epoch 2061–2080 (5.0×10^6 km^2). For epoch 2040–2060, the main changes shown by all three scenarios are as follows: areas belonging to the polar/alpine biome will transition to the temperate-boreal forest biome and areas from the tropical forest biome will transition to more drier biomes, like the savannas and grasslands biome the shrublands biome and, in some cases, the deserts and steppes biome.

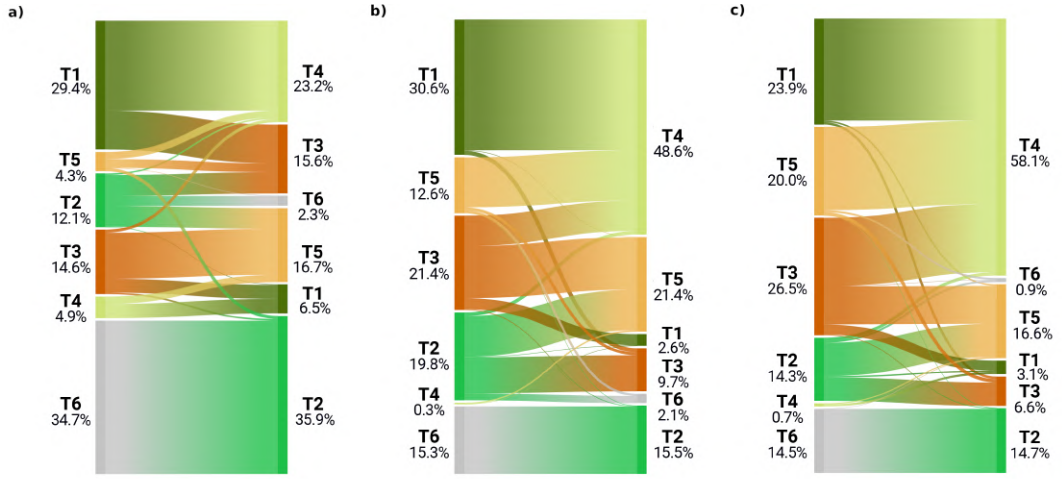

Figure 2.4: Biome transitions predicted for epoch 2040–2060 for the three climatic scenarios. Only the pixels that transitioned are represented in this diagram, so the percentages represent different amounts of surface area across the scenarios. For each plot, on the left axis the proportion of transitioned pixels in the current conditions and on the right axis the final state according to each climatic scenario: **a)** shows the transitional areas for scenario RCP 2.6, **b)** for scenario RCP 4.5 and **c)** for scenario RCP 8.5.

The same tendency can be observed for the temperate-boreal forest biome, with the difference that transitional areas are almost equally split between the shrublands biome and the deserts and steppes biome. It is interesting to notice that almost all the transitioned pixels from the tropical forest biome would change to savannas and grasslands biome in the RCP 4.5 and RCP 8.5 scenarios, while for scenario RCP 2.6 one third would change to the shrublands biome (see Fig. 2.4). According to scenario RCP 2.6, most of the changes would happen in the polar/alpine biome, so at higher latitudes, while for the other two scenarios the tropical areas seem to be the ones most affected. For scenario RCP 2.6, the transitional areas are also more equally split across the different classes, while for scenarios

RCP 4.5 and RCP 8.5 almost 50% and 60% of the transitional areas would shift to the savannas and grasslands biome.

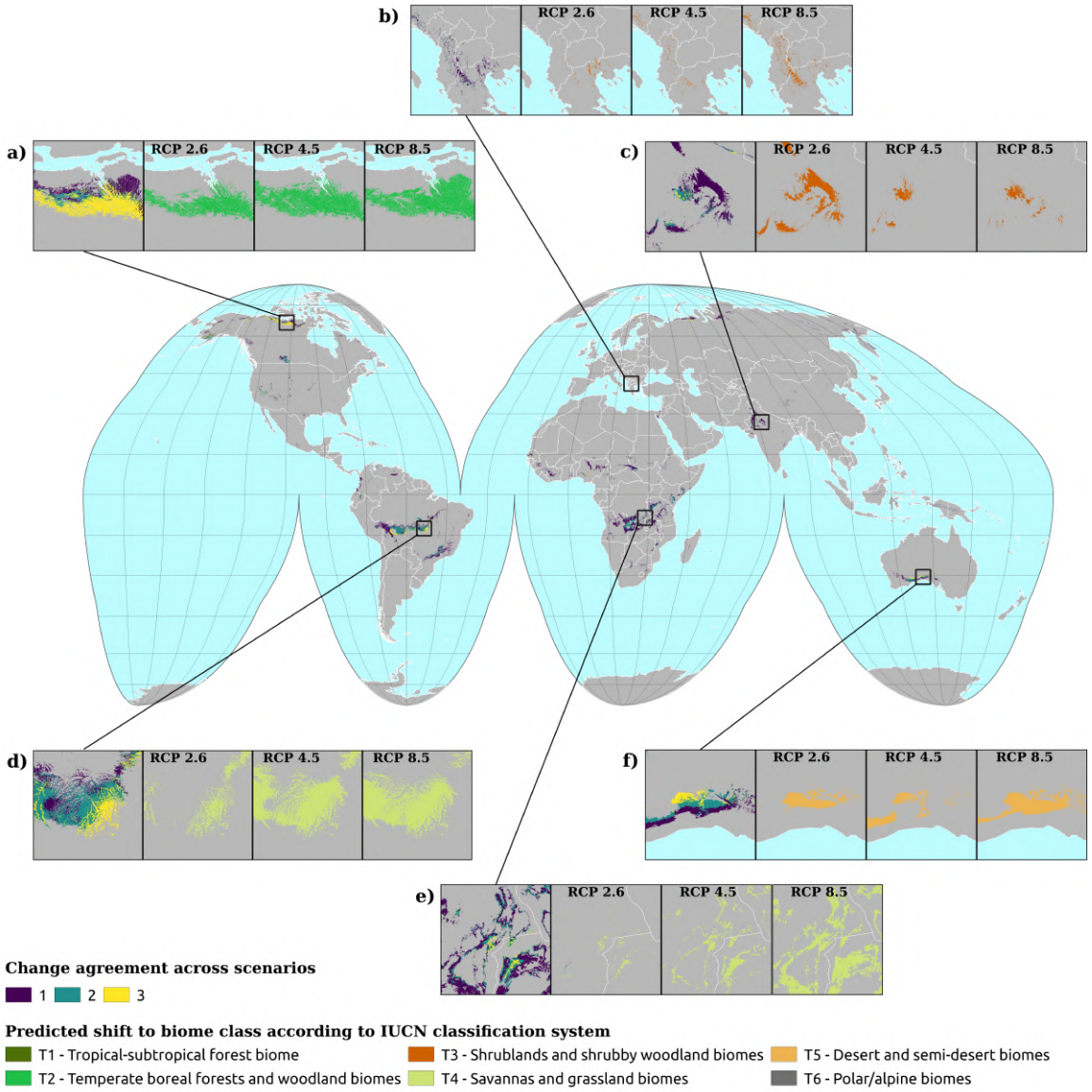

Figure 2.5: Spatial location of biome transitions as predicted by our ensemble model according to the three climatic scenarios for epoch 2040–2060. Colors on the main map show the degree of agreement between the three climatic scenarios: a value of 1 means that only one of the scenarios considers the pixel as transitioning, while a value of 3 shows complete agreement across the three scenarios. Inserts show towards which biome the current pixels are transitioning to according to the different scenarios. Inserts **a)**, **d)** and **e)** show the main trends, with transitions from T6 (polar) to T2 (boreal forest) in **a)** and from T1 (tropical forest) to T4 (savannas) in **d)** and **e)**

Fig. 2.5 shows the geographic locations of the biome shifts according to the three climatic scenarios. It is possible to discern different clusters where the changes are located: the

most noticeable is in the tropical area, between the Equator and 15°S; in South America, the region affected corresponds with the southern edges of the Amazon rainforest, which would shift from a tropical forest biome to savanna. In Central Africa, in the contiguous borders of Angola, Congo and Zambia, the shift goes instead from shrubland or steppic biomes to savanna. In the transitioning areas where all the scenarios agree in predicting change, there is one area that includes most of the shifts: between 60°and 75°N, just around the Arctic Circle. At this latitude, most of the areas currently in the polar/alpine class would shift to the boreal/temperate forest class. Big clusters can be observed in the northern parts of Canada and Alaska, while smaller clusters occur in Scandinavia, European Russia and some areas in Siberia. In most cases, scenario RCP 2.6 involves the smallest amount of transitions, while RCP 8.5 involves the greatest. However, there are also some areas where this does not hold, as can be seen in Fig. 2.5c and Fig. 2.5f. In the first case, scenario RCP 2.6 involves the greatest amount, while in the latter it is greater than scenario RCP 4.5 but smaller than scenario 8.5. In general, all three scenarios agree in predicting a change in some 3% of the cases, while if we consider the two most radical scenarios the percentage rises to 11%.

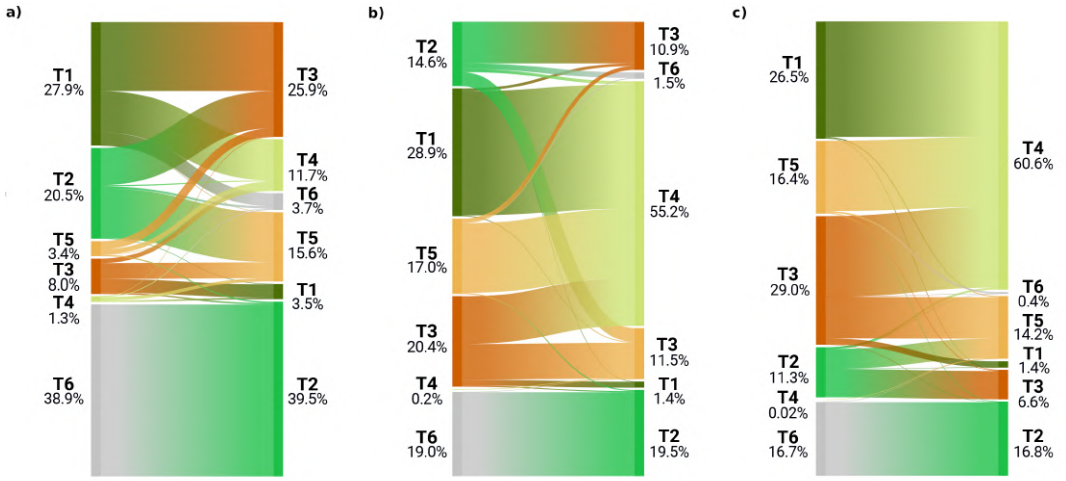

Figure 2.6: Biome transitions predicted for epoch 2061–2080 for the three climatic scenarios. Only the pixels that transitioned are represented in this diagram, so the percentages represent different amounts of surface area across the scenarios. For each plot, on the left axis the proportion of transitioned pixels in the current conditions and on the right axis the final state according to each climatic scenario: **a)** shows the transitional areas for scenario RCP 2.6, **b)** for scenario RCP 4.5 and **c)** for scenario RCP 8.5.

For epoch 2061–2080, we found similar trends to the ones observed in the previous epoch: all of the pixels from the polar/alpine biome tend to shift to the temperate-boreal forest biome and the pixels from the tropical forest biome would shift towards the savannas and grassland biome (see Fig. 2.6). The tendency shown in Fig 2.4a, with the transitioning pixels from the tropical forest biome split between the shrublands biome and the savannas and grassland biome, is in this epoch even more pronounced: the ratio is reversed, with one

third of the pixels transitioning to the savannas and grassland biome and the rest towards the shrublands biome. In the other two scenarios, once again, almost all transitioning pixels from the tropical forest biome would shift to the savannas and grassland biome.

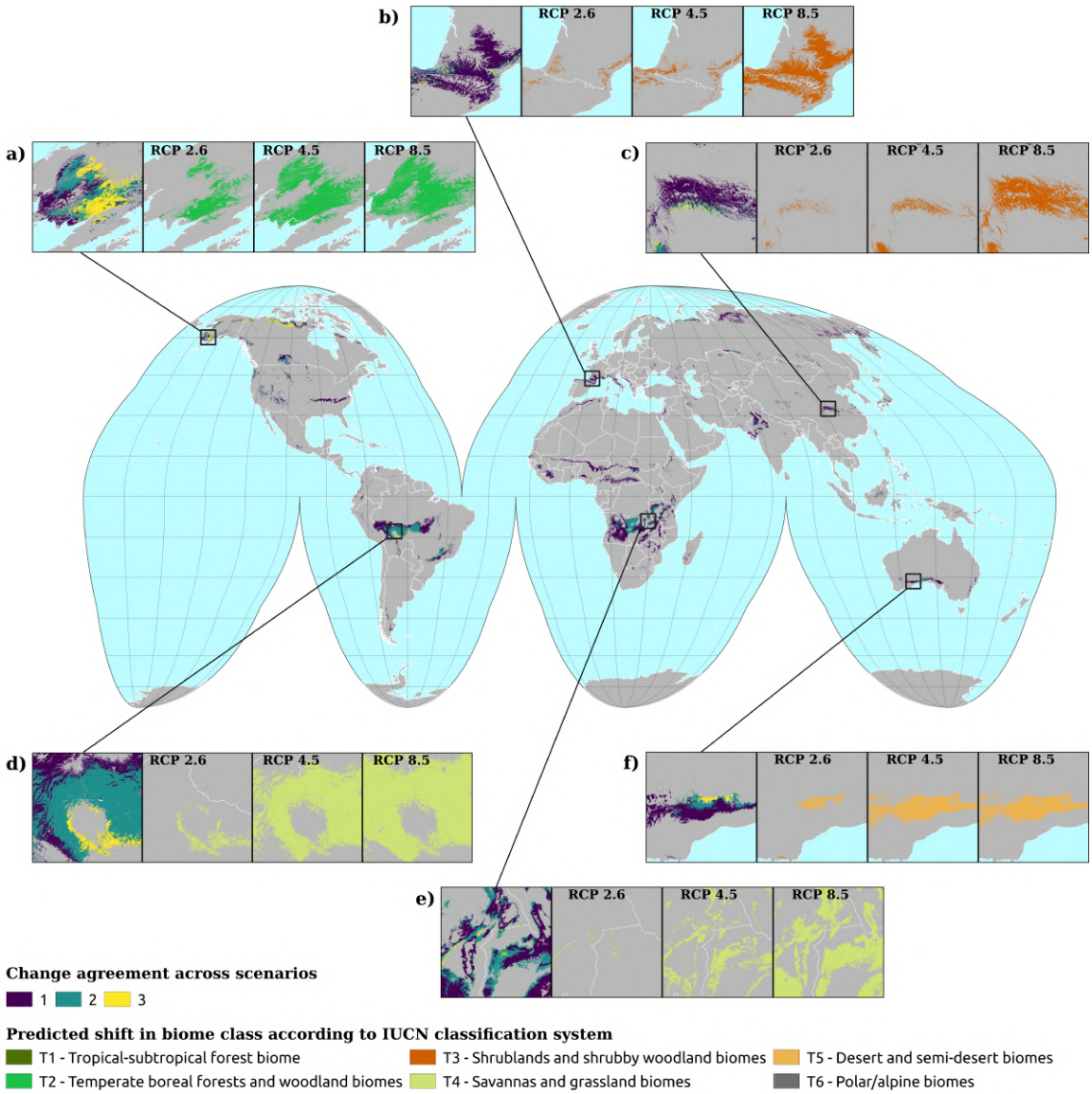

Figure 2.7: Spatial location of biome transitions as projected by our ensemble model according to the three climatic scenarios for epoch 2061–2080. Colors on the main map show the degree of agreement between the three climatic scenarios: a value of 1 means that only one of the scenarios considers the pixel as transitioning, while a value of 3 shows complete agreement across the three scenarios. Inserts show towards which biome the current pixels are transitioning according to the different scenarios. Inserts **a)**, **d)** and **e)** show the main trends, with transitions from T6 (polar) to T2 (boreal forest) in **a)** and from T1 (tropical forest) to T4 (savannas) in **d)** and **e)** Inserts **b)** and **c)** show instead the tendency to drier ecosystems in temperate areas.

Another recurring pattern is how most of the transitioning pixels from the temperate-boreal forest biome are equally split between the shrublands biome and the deserts and steppes biome, so either the canopy of those forests would become more open and the ratio between trees and shrub would increase in favor of the latter, or they become so dry that the trees are replaced by steppic vegetation. For scenario RCP 2.6, 80% of the transitioning pixels are part of the tropical forest biome, the temperate-boreal forest biome or the polar/alpine biome, with nearly 39% coming from the polar/alpine biome, while the classes which would gain most of this surface area are the temperate-boreal forest biome and the shrublands biome. In the other two scenarios, the polar/alpine biome covers a more marginal importance across the transitioning pixels: it is the third most important. Some differences can be observed in the transitioned classes as well: in scenario RCP 2.6, about 40% of the pixels go to the temperate-boreal forest biome, while in the other two scenarios the savannas and grassland biome takes from 55% to 60% of the transitioned classes.

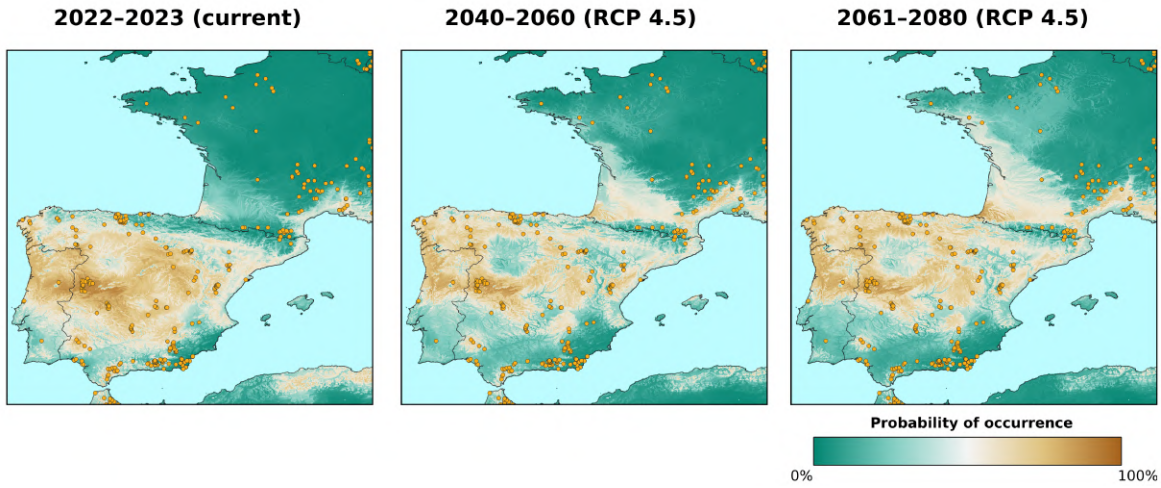

Figure 2.8: Predicted probability of occurrence of *"warm temperate evergreen and mixed forest"* class, zoom in on the area around the Pyrenees. The probability values over time show that the class is slowly shifting towards northern latitudes. Only the RCP 4.5 scenario is shown since it is considered as the "middle of the road" scenario. Points indicate training points from the BIOME 6000 dataset.

Fig. 2.7 shows the geographic locations of the shifts for the epoch 2061–2080. The two big clusters observed in the previous epoch remain, as well as the area around the Arctic Circle; the transitioning pixels where all three scenarios agree in predicting change are also located mostly in these two areas. Compared to the previous epoch, more small clusters of pixels, with no specific spatial pattern, are visible in the African continent from the Equator to 15°N (mostly around the Gulf of Guinea), Western India and Mediterranean Europe (mostly around the Pyrenees, see Fig. 2.7b); by checking the probability layers, it is possible to see the gradual shift (see Fig. 2.8) in vegetation conditions over time.

Contrary to the previous epoch, the mapped changes in the scenarios seem to be more consistent: all the inserts in Fig. 2.7 show that scenario RCP 2.6 is the one which projects the smallest number of transitioning pixels, while RCP 8.5 projects the most. The scenarios in this epoch also have a higher degree of agreement: all three scenarios agree in considering as changing 7% of all the transitioning pixels, while the value for agreement between two scenarios is just 8%. In general, if in the previous epoch most of the transitioning pixels were located either in the tropics or around the Arctic Circle, in this epoch we see them appearing in the temperate areas as well.

2.4 Discussion

2.4.1 Model evaluation and comparison with previous works

In this study we trained an ensemble machine learning model to classify the current and future potential distribution of biomes under different climate change scenarios. Our results show that it is possible to produce relatively accurate maps of natural vegetation using ensemble machine learning approaches and to reach consistent accuracy values even with a limited selection of predictor variables. Comparing our results with the previous work of Hengl et al. (2018), we achieved an increase in the overall accuracy by using an ensemble model and only 72 instead of 158 predictor variables: in the previous task the most accurate model (Random Forest) shows a spatial cross–validation overall accuracy of 0.33, less than half of what we estimate with the improved model (0.67). The performance values per class show some degree of agreement: when comparing TPR values both studies consider the *"prostrate dwarf shrub tundra"* class as the worst predicted, while they disagree in the best predicted class (*"temperate sclerophyll woodland and shrubland"* for Hengl et al. (2018), *"cool mixed forest"* in this study); while our TPR values are consistently lower across all classes, TPR values reported by Hengl et al. (2018) are for the model with no spatial partitioning. Model outputs are provided in probability values and not as hard classes in both studies; furthermore, the dataset is heavily imbalanced, as shown in Fig. 2.2. This makes logloss, and the R^2_{logloss}, a better metric to report model performances since it indicates how close the predicted value for an observation and its respective label is; logloss is also one of the most robust performance metrics when it comes to imbalanced data (Ferri et al., 2009). The fact that Hengl et al. (2018) used other performance metrics that do not fit the task at hand to report per class results, may have caused an overestimation of those values.

In machine learning, increasing the size of the feature space is expected to provide more discriminating power (Hall and Holmes, 2003), at the cost of higher computation time. However, it can also increase the complexity of the task at hand to the point that the added information is redundant or introduces noise in the model (Bellman and Kalaba, 1957). While feature selection procedures help considerably in tackling this problem, in

this case we used expert knowledge to select only climatic and topographic predictor variables. By doing that, we managed to achieve a twofold goal: reduce task complexity (i.e. less features) while maintaining consistent values of accuracy, and keeping the model simple enough to be able to transfer it to future epochs without introducing too many assumptions in the modeling framework; a similar approach to calculate future projections was used by Anjos et al. (2021), Maksic et al. (2022) and Zevallos and Lavado-Casimiro (2022), respectively, for the whole South America, Brazil and Peru. In the case of Zevallos and Lavado-Casimiro (2022), they used a Random Forest model and achieved higher levels of accuracy; however, they trained it on a smaller (6 bioclimatic) set of predictor variables, used a 80:20 train test split and did not use any spatial partitioning. Considering the huge differences in accuracy in the results obtained by Hengl et al. (2018) in their Random Forest model with and without spatial partitioning, there's the risk that predictions from Zevallos and Lavado-Casimiro (2022) may have been optimistic; on the other hand, their analysis is focused on just one country and not on a global scale, so it is still possible to reach high levels of accuracy on a limited study area.

Data-driven approaches in this topic mostly deploy Random Forest models, with predictions with high levels of agreement with process-based models: Lindgren et al. (2021) used Random Forest to reconstruct past global vegetation and compared their results with the LPJ-GUESS global dynamic vegetation model; the model was able to produce comparable results to the LPJ-GUESS when enough training data was available, with bad performances when predicting in the Last Glacial Maximum, the epoch with the least training data. Extrapolation and transferability are two common issues of machine learning models and data-driven approaches in general, which have limited the reliability of such methods in environmental modeling (i.e. invasive species modeling, past vegetation reconstruction, future vegetation forecasting etc) (Qiao et al., 2019). However, ensemble modeling is known to provide more advantages than using individual machine learning models, since ensemble models reduce model uncertainty (**Bonannella** et al., 2022d; Mehra et al., 2019). For future projections, "ensemble datasets" are more common than ensemble models: climate is assumed to be the major driving force for large-scale vegetation patterns (Whittaker and Marks, 1975); starting from this assumption, multiple studies create the training dataset by averaging together temperature and precipitation values as calculated by different GCM simulations (Anjos et al., 2021; Beigaitė et al., 2022), hence why we chose an ensemble of five independent GCMs. The model used in this study could benefit from using such datasets: while studies on performance comparisons between the different GCM simulations are available for the CMIP5 project (Sanderson et al., 2015), the same can't be said for the new CMIP6 simulations; future applications of the experimental design presented in this study that would use CMIP6 simulations, could benefit from using an ensemble dataset of all 50 of the GCM simulations instead of relying on the data of only five models.

2.4.2 Biome shifts: key emerging trends

We evaluated changes in potential biomes in two future epochs and across three different climatic scenarios: our results show that the distribution of the biomes on land in the future will mostly (≥99% land surface) remain the same. The limited geographic extent of the biome shifts under all three scenarios has probably to do with the chosen conservative threshold in margin of victory; despite that, the projections show specific emerging trends in biome shifts in precise locations of the globe that, while differing in size, are common across all the climatic scenarios.

One of these emerging trends is the transition from a polar to a boreal forest biome in the global north, around the Arctic Circle: in both of the epochs analyzed in the study, it is one of the most evident and consistent transitions in all three climatic scenarios, with its extension increasing in epoch 2061–2080. According to the future climatic projections, all three scenarios forecast either a modest or consistent rise in temperatures by 2100, from well below 2° C for scenario 2.6 and around 5° C for scenario RCP 8.5. Areas where this change was predicted currently present vegetation not belonging to the *"boreal forests and woodland biomes"* class mostly due to the fact that the low temperatures are a limiting factor for the presence of trees. The hypothesis that the thawing of permafrost would lead to the tree line advancing towards the North Pole finds more and more evidence, the last one provided by Berner and Goetz (2022) using Landsat time-series: their results on vegetation greenness for the period 1985–2019 showed a prevalence of greening in the pan-boreal vegetation. It is important to point out how this phenomenon is not uniform across the circumpolar arctic vegetation: while greening was more prevalent than browning phenomena, browning was still predicted in those areas where summers have become warmer and drier in the last 40 years. Our results, regardless of the scenario, show that while the change will happen, it won't be uniform across North America, Asia and Europe: North America seems to be the one that will be most affected, while only in few areas of Siberia all scenarios agree in a biome shift. Our maps can be used to locate *hot spots* of change, from which the shift can then expand: the model used in this study doesn't take into account many factors, such as the feedback effects on the carbon cycle caused by the permafrost thaw (Smith et al., 2022) or soil temperature, moisture and content, which, in turn, affect vegetation productivity and functional types (Berner et al., 2020). For this reason it is important to be cautious when assessing biome change implications. On top of that, we focus on the potential conditions that define a biome and not on what is currently on the ground: species that live in a biome may not be able to keep pace with the climate change advance (Rees et al., 2020), and while our maps may show that the conditions for a shift are present in a specific location, its reality may be different.

The second trend relates to the transition from tropical forest to savannas and grasslands, in particular in the southern edges of the Amazon and the Congo rainforests. Both the entity and the extension of this shift across the climatic scenarios follow the same pattern

that we found for the transition from polar to boreal, with the highest value of pixels shifting found in the epoch 2061–2080 for scenario RCP 8.5. The consequences of climate change on the Amazon rainforest are a critical area of research given its importance for global climate regulation and biodiversity (Foley et al., 2007; Lawton, 1998) and have been the subject of extensive research in the scientific community. While the full extent of these changes has not yet been completely understood, higher temperatures and variations in rainfall regime have been causing longer and more severe dry seasons (Agudelo et al., 2019; Arias et al., 2015; Xu et al., 2022), with an increase in frequency of droughts, floods and fires (Barlow et al., 2020; Lovejoy and Nobre, 2018; Marengo and Espinoza, 2016); a recent study by Gatti et al. (2021) has demonstrated how the southeastern edge of the Amazon rainforest has already reached the tipping point, acting as a net carbon source instead of a carbon sink. These findings agree with the projections showed in our results, which now are part of a long series of studies showing alarming signs of an incoming process of savannization in the area; the feedback loop created by a disruption in the carbon cycle such as the one showed by Gatti et al. (2021) could further exacerbate the savannization process. On the same note, Sampaio et al. (2007) were the first to show how when deforestation exceeds 40%, the savanna would become the new stable state of the ecosystem in south, east and partially central Amazonia due to altered precipitation patterns; climatological projections from Higgins et al. (2016) show how a rise by only 2° C in average temperature could lead to a loss of 50% of suitable areas for forest specialist species and an increase by 11%-30% for savanna species. While less studies on the Congo rainforest are available in literature, the projected savannization process can be attributed to the same causes: Giresse et al. (2023) have shown that the Congo rainforest has been resistant to change in the last 1000 years and it was not possible to identify any serious human impact over this period. However, the current increasing mix of climate change and human pressures (deforestation, agriculture expansion and other factors) may lead to unforeseen consequences: the current rainfall regime of much of the African rainforests is close to a threshold that favours savannas over rainforests (Malhi et al., 2013), so even a small alteration of this regime can cause large scale changes in the rainforest-savanna transition zone.

Overall, the shifts in biomes identified in this study portray a picture of minor or consistent changes across all the different biomes on the planet due to either an increase in temperature or decrease in precipitation/moisture conditions: this agrees with a similar existing dominating browning trend for global vegetation recently identified by Higgins et al. (2023) for the last four decades (1982–2015). Their approach is particularly relevant for the context of the present study, since they combined a process-based approach by using a dynamic plant growth model adjusted for climate with a data-driven approach by using Advanced Very High Resolution Radiometer (AVHRR) NDVI and EVI time series to describe vegetation activity. These shifts may have significant ecological implications for the distribution and diversity of plant and animal species, as well as societal implications

for human communities that depend on these ecosystems. It is likely that these shifts will also have economic impacts, as the distribution of resources such as timber, livestock grazing and agriculture or the potential displacement of human population; the alterations of the services that these ecosystems provide, such as climate regulation and flood control, have also to be considered. These shifts may contribute to the loss of biodiversity, as some species may not be able to adapt to the new conditions in their range: the loss of tropical and subtropical biomes can have a negative impact on the species that depend on these biomes for their survival, as many species have narrow habitat requirements and are not able to adapt to changes in their environment. The expansion of boreal biomes, on the other hand, can have a positive impact on the species that depend on these biomes, as they will have access to new areas with suitable habitat.

2.4.3 Technical limitations

There are some limitations to this study that should be considered. First, the dataset used to train the model was heavily imbalanced, with some classes having a very small number of observations: this has significantly affected the model performances for certain classes; on top of that, some locations are underrepresented, with limited or no observations. This is a serious limitation of the study, as the model may not be able to accurately predict the vegetation in those locations due to a lack of data. This highlights the importance of gathering more comprehensive ground truth data in the future to improve the model's accuracy and prediction abilities in those locations.

Secondly, while the use of expert knowledge to select the predictor variables let us to reduce the complexity of the task, it may also have introduced biases or limitations to the model's ability to accurately represent the full range of conditions present in different biomes; feedback loops (vegetation-climate interactions) or anthropogenic factors (human disturbances, deforestation) were also not considered in the study. One of these effects which are difficult to incorporate in data-driven models is the fertilization effect of increasing concentration of CO_2 in the atmosphere, or CO_2 fertilization effect (CFE): process-based models are known to be able to include and parametrize this factor, event though the mechanisms and limits of it are still not completely understood. Ballantyne et al. (2012) observed that since the 1960s the carbon uptake by both terrestrial and oceanic ecosystems has increased instead of declining, while a study by Chen et al. (2022) calculated how approximately 44% of the increase in global Gross Primary Productivity (GPP) since the 2000s can be attributed to the CFE. This proved to be especially important for the in the tropical region, with different satellite-based or in-situ studies showing patterns of greening in these areas (Anchang et al., 2019; Stevens et al., 2017). Thus, not including this effect in our model may lead to overestimate the amount and the type of shifts in the tropical region: for example, the chapter on the African continent of the last IPCC report (Trisos et al., 2022) mentions an overall continental trend in woody plant expansion, especially in the non-arid areas, with high confidence that the trend is attributable to the CFE. This

is in contrast with the desertification and contraction trend that was instead highlighted by the previous AR: since the AR6 had at its disposal longer time-series of observations, the trend captured by the previous AR, and hence the CMIP5 GCM projections, may have been overly pessimistic in some regions. Not including the CFE in our study and the usage of CMIP5 projections may have biased the results in favor of a desertification trend for some areas. All the simulations from the study from Friend et al. (2014) predicted a consistent increasing trend in CO_2 for every RCPs: so while our model may accurately capture a shift in the boreal region, due to the fact that the limiting factor for those biomes is temperature, it may capture a different type or extent of shifts in the tropical region, where neither temperature or precipitation are the limiting factors. On the other hand, however, Higgins et al. (2023) found the contribution to regreening from the CFE to be limited and Wang et al. (2020) observed a significant decline in the CFE on a global scale for the period 1982–2015. Overall, even if in this study we identified several *hot spots* of change in the tropical region, there is still high uncertainty for the future of biomes and their shifts there.

We also excluded from the study area those locations that are currently covered by permanent ice: glacial retreats due to climate change is a well known issue and the exclusion of these areas may also underestimate the potential changes in biomes in surrounding areas that may be influenced by the loss of glaciers. It is important to consider the inclusion of these areas in future studies to obtain a more comprehensive understanding of the potential impacts of climate change on the distribution of biomes. Another limitation is that the model can't be considered spatiotemporal: while spatial relationships are taken into account during the modeling through the use of spatial partitioning, the temporal relationship is not considered since the model is trained on only one point in time. The model doesn't know that the three epochs analyzed in the study have an order in the temporal dimension: the phenomenon we want to predict at location x for epoch 2061–2080 is not only a function of the predictor variables in the features space, but also of the realization of the phenomenon in previous states. The result, in the worst case scenario, is that the predicted values for one location may not be reliable over time: an example in our case would be a pixel labeled as *"desert"* for the current epoch, *"forest"* in the epoch 2040–2060 and *"desert"* once again in epoch 2061–2080, something that can be perfectly explained from a mathematical standpoint, but highly questionable from an ecological perspective.

Together with the heavily imbalanced dataset, this is another important reason to use the margin of victory to analyze the results. The inclusion of the margin of victory allows users to more accurately interpret the predicted maps and make more informed decisions based on the data. Without the margin of victory, the probability outputs could potentially be overinterpreted, leading to incorrect conclusions. The classification task examined in this study presented a total of 20 classes, with the model output per class constrained, for each pixel, to sum to 100%; the conservative threshold allowed us to focus our analysis

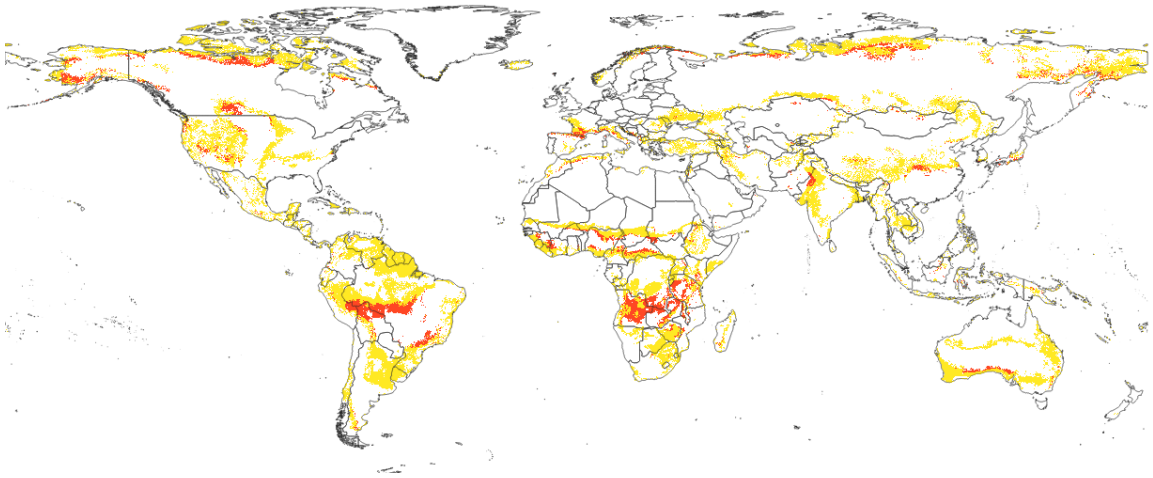

Figure 2.9: Spatial location of biome transitions for scenario RCP 8.5 epoch 2061–2080. In red: all the areas that will change according to our model with a *margin of victory* value ≥50%. In yellow: all the transitioning areas with *margin of victory* values <50%.

only on those areas where the model found considerable differences in probability output between the dominant class and the remaining classes (see Fig. 2.9). Even though the class probabilities in this study are model-based predictions and model fit is far from perfect, the margin of victory gives an impression of the uncertainty in our projections. We acknowledge our inability to predict the future and to make claims about biome shifts that would certainly happen. It is however possible to indicate the confidence of our projections through the uncertainty layers provided. That's why we recommend to use the uncertainty layers to filter the predicted areas using conservative (≥50%) thresholds in case of future use of these maps in other works: false positives may thus be avoided while still identifying the main patterns.

2.5 Conclusions

In this article we applied a methodological framework to predict current and future potential distribution of biomes under different climatic change scenarios using an ensemble machine learning approach. We focused our efforts on improving the caveats of previous work from Hengl et al. (2018), achieving greater accuracy in predicting current biomes and providing future distribution of biomes along with measures of prediction uncertainty to correctly interpret and use the maps. In general, our ensemble model achieved fairly accurate (overall accuracy = 0.67, $R^2_{\text{logloss}} = 0.61$) results. Using expert knowledge to select only a limited number of predictor variables, we were able to achieve reasonable accuracy values while keeping the model simple enough to be able to transfer it to future epochs without introducing too many assumptions. Temperature-related predictor variables were considered as the most important to produce accurate predictions. Overall, this study

demonstrates that an ensemble machine learning approach can be effective in modeling the potential distribution of biomes on a global scale and in identifying areas where climatological changes could lead to shifts in the distribution of these biomes.

Even though relatively small shifts in the distribution of biomes were projected under the RCP 2.6 and RCP 4.5 when compared to RCP 8.5, one of the significant findings of this study was the identification of areas where the change in climatological conditions could lead to a shift in the potential distribution of biomes regardless from which of the emission pathways analyzed will happen in the future. The biomes expected to shift the most are the tropical and subtropical biomes, particularly the tropical rainforests: these biomes are expected to experience a decrease in their potential distribution in the future time periods towards savanna and grassland biomes, a process called *"savannization"*. In contrast, biomes located at higher latitudes, such as boreal forests, are expected to experience an expansion in their potential distribution in the future time periods at the expense of the polar biomes.

Further research is needed to better understand the factors that drive these shifts and the potential consequences for the distribution and diversity of plant and animal species, as well as for human communities. We hope that this study will contribute to the broader field of study by providing a framework that can be used to better understand the potential impacts of climate change on the distribution of biomes and their associated ecosystems, and by identifying areas where these impacts could be particularly significant. We recommend that this information is used by policy makers and land managers to make informed decisions about the management and conservation of these ecosystems, and to take action to mitigate the negative consequences of climate change.

Chapter 3

High resolution distribution mapping of European forest tree species

This chapter is based on:

C. Bonannella, T. Hengl, J. Heisig, L. Parente, M. N. Wright, M. Herold, and S. de Bruin (2022d). "Forest tree species distribution for Europe 2000–2020: mapping potential and realized distributions using spatiotemporal machine learning". *PeerJ*. DOI: https://doi.org/10.7717/peerj.13728

Abstract

This paper describes a data-driven framework based on spatiotemporal machine learning to produce distribution maps for 16 tree species (*Abies alba* Mill., *Castanea sativa* Mill., *Corylus avellana* L., *Fagus sylvatica* L., *Olea europaea* L., *Picea abies* L. H. Karst., *Pinus halepensis Mill.*, *Pinus nigra J. F. Arnold*, *Pinus pinea* L., *Pinus sylvestris* L., *Prunus avium* L., *Quercus cerris* L., *Quercus ilex* L., *Quercus robur* L., *Quercus suber* L. and *Salix caprea* L.) at high spatial resolution (30 m). Tree occurrence data for a total of 3 million points was used to train different algorithms: random forest, gradient-boosted trees, generalized linear models, k-nearest neighbors, CART and an artificial neural network. A stack of 305 coarse and high resolution covariates representing spectral reflectance, different biophysical conditions and biotic competition was used as predictors for realized distributions, while potential distribution was modelled with environmental predictors only. Logloss and computing time were used to select the three best algorithms to tune and train an ensemble model based on stacking with a logistic regressor as a meta-learner. An ensemble model was trained for each species: probability and model uncertainty maps of realized distribution were produced for each species using a time window of 4 years for a total of 6 distribution maps per species, while for potential distributions only one map per species was produced. Results of spatial cross validation show that the ensemble model consistently outperformed or performed as good as the best individual model in both potential and realized distribution tasks, with potential distribution models achieving higher predictive performances (TSS = 0.898, R^2_{logloss} = 0.857) than realized distribution ones on average (TSS = 0.874, R^2_{logloss} = 0.839). Ensemble models for *Q. suber* achieved the best performances in both potential (TSS = 0.968, R^2_{logloss} = 0.952) and realized (TSS = 0.959, R^2_{logloss} = 0.949) distribution, while *P. sylvestris* (TSS = 0.731, 0.785, R^2_{logloss} = 0.585, 0.670, respectively, for potential and realized distribution) and *P. nigra* (TSS = 0.658, 0.686, R^2_{logloss} = 0.623, 0.664) achieved the worst. Importance of predictor variables differed across species and models, with the green band for summer and the Normalized Difference Vegetation Index (NDVI) for fall for realized distribution and the diffuse irradiation and precipitation of the driest quarter (BIO17) being the most frequent and important for potential distribution. On average, fine-resolution models outperformed coarse resolution models (250 m) for realized distribution (TSS = +6.5%, R^2_{logloss} = +7.5%). The framework shows how combining continuous and consistent Earth Observation time series data with state of the art machine learning can be used to derive dynamic distribution maps. The produced predictions can be used to quantify temporal trends of potential forest degradation and species composition change.

3.1 Introduction

Reforestation and forest restoration are considered key strategies for tackling climate change by enhancing CO_2 sequestration (Domke et al., 2020; Lefebvre et al., 2021; Nave et al., 2019). Under the European Green Deal and the European biodiversity strategy for 2030, the European Union has committed to plant at least 3 billion additional trees by 2030 (European Commission, 2021b). At the same time, tree deaths due to bark beetle infestations and increased drought fueled by a warming climate have reduced the total forest area of Germany by 2.5% since 2018 (Popkin, 2021). Obtaining reliable information on forest tree species distribution in both space and time is now urgently required for stakeholders and decision-makers in order to develop effective forest management and adaptation strategies (Keenan, 2015).

Understanding the range, constraints and drivers of species distribution has always been a primary goal of ecology (Andrewartha and Birch, 1954). However, only with the advent of Geographical Information Systems (GIS) and the usage of extensive digital maps of environmental variables were ecologists able to access powerful enough tools to study species distributions at landscape scales (Franklin, 1995). Progress in this direction has given rise to a new field called Species Distribution Modelling (SDM) (Franklin, 2010): maps of species ecological niches are made by associating values of different predictors to known locations of the target species and then used to predict distribution in geographic space where no field data for the target species is available. Commonly, SDMs rely on climatic or bioclimatic factors at a coarse spatial resolution ($\geq$1 km) while in the temporal dimension long time averages (30-–50 years) are often used (Iturbide et al., 2018a). Whilst forest distribution maps are often used to guide management decisions happening at local scales, the potential impact of differences in resolution of the predictor variables on the results is often overlooked (Porfirio et al., 2014). For conservation purposes, previous studies have shown how distribution maps with high spatial resolution ($<$ 100 m) and slightly lower prediction accuracy are actually more useful than coarser ($>$ 250 m) but more accurate maps (Gottschalk et al., 2011; Guisan et al., 2013; Manzoor et al., 2018; Prates-Clark et al., 2008). Therefore, even at the cost of overall map accuracy, finer spatial resolution maps are more valuable for practical use. For these reasons, Earth Observation (EO) data, and specifically the use of high spatial resolution data, have grown in use for SDM applications (Gelfand and Shirota, 2021; Hefley and Hooten, 2016; Pérez Chaves et al., 2018).

In addition to the clear need for finer spatial resolution mapping, there are similar needs to drive research towards producing finer temporal resolution mapping. This is due to the recent and relatively swift change in disturbance regimes and weather patterns, which are significantly altering the ecological niches of tree species on a temporal scale of less than a few decades instead of centuries. The impact of these changes on tree species has become more noticeable from year to year, with growth decline (Martinez del Castillo et al., 2022)

and increased mortality rates (Senf et al., 2018; Senf et al., 2021) demonstrated in literature as already occurring across large forested areas. Including the temporal domain in tree species distribution studies is therefore fundamental to capture the temporal evolution of these change processes. However no general consensus has yet been reached on the influence of these new high spatial and temporal resolution data sources on SDM performances. The inclusion of spatiotemporal data sources in SDM studies requires taking an additional effort when choosing the appropriate modeling technique, a task that has proved to be difficult even with traditional spatial-only data sources (Elith and Graham, 2009), let alone when also attempting to include the temporal dimension.

Aside from spatial and temporal considerations of predictor variables and species observations, in the last decade ecologists have conducted hundreds of studies purely to determine which modeling methods best suit the needs of SDM. Model choices have thus far proven to be highly impactful, with distribution maps derived with different models from the same dataset leading to quite opposite conclusions (Araújo and New, 2007; Pearson et al., 2006). Inter-model variability in projections has been tackled using ensemble modeling, where numerous independent models are fit using a range of methods applied to the same input data while the outputs of the individual models are aggregated into the final prediction. Ensemble modeling is a solution to high model variance and it has been demonstrated that reducing variance also reduces the effect of model overfitting and extrapolation (Zhou, 2019). This is achieved at the cost of increased model complexity, reduced model interpretability, and increased computational time (Zhou, 2019). As such, the few examples of ensemble modeling approaches that have been investigated for SDM applications are limited to *mean*, *median* and *weighted average* approaches (Hao et al., 2019). These approaches are intuitively simple to implement and interpret, and involve, in the first two cases, just taking the mean or median of the predictions of the individual models as the final prediction. The weighted average approach is similar but scales the predictions by weights assigned based on predictive performances of the models obtained from cross validation. A robust ensemble technique that, to our knowledge, has not been tested yet for SDM is *stacking* or *stacked generalization*. In this approach outputs made by the individual models are the inputs of a *meta-learner* (i.e. a model that learns from other models) which then produces the final prediction (Wolpert, 1992).

We tested this ensemble technique on European forest tree species distribution. There is no shortage of information on European tree species distribution: the European Atlas of Forest Tree species is among one of the largest data sources with information on forest tree species for Europe (San-Miguel-Ayanz et al., 2016). It describes in detail the autoecology of 76 different forest tree species and provides geographical information on each species in the form of chorological maps, probability of presence maps and maximum habitat suitability maps. Recently, the Atlas has been further expanded with future projections in different climatic scenarios (Mauri et al., 2022). While these predictions are certainly useful to determine potential species composition of European forests, new methods are now

needed to deal with the more and more attention to reproducibility of studies (Fidler et al., 2017), increasing spatial and temporal resolution of predictor variables (Zhu et al., 2019) and availability of ecological *"big data"* (i.e. gathered by multiple sources such as sensors, cameras etc.) (Hampton et al., 2013). Furthermore, SDM studies use high-dimensional data which is often non-linear and does not meet assumptions of conventional statistical procedures (Zhang and Li, 2017). For this reason, and thanks to the exponential increase in computing power of the last decade (Gorelick et al., 2017), solutions such as Machine Learning (ML) algorithms have recently become very popular for SDM studies. ML tries to learn the relationship between the response and the predictors through the observation of dominant patterns (Breiman, 2001b). Contrary to traditional statistical models, no kind of ecological assumptions are explicitly embedded in ML algorithms: ML can be especially useful when dealing with data gathered without a specific and rigorous sampling design (Ij, 2018). ML algorithms have great potential to analyze the large amount of data available nowadays, enabling the mapping and monitoring of changes on multiple geographical scales in a timely manner through reproducible research (Gobeyn et al., 2019).

In this sense, the objectives of this study were (**a**) to test different ML algorithms to develop a framework for modeling species distribution in space-time, (**b**) to assess the importance of various sources of EO data on model performances for mapping tree species distributions and (**c**) to explore and quantify the specific importance of high resolution data on model performances.

3.2 Material and methods

3.2.1 General workflow

We modeled potential and realized distribution for 16 forest tree species for continental Europe for the time period January 2000 – December 2020 using a spatio-temporal ML approach. The general workflow used to derive the distribution maps is shown in Fig. 3.1. We modeled the potential distribution as a baseline to assess the importance of EO data sources: we used only environmental predictors (i.e. temperature, precipitation, wind speed, water vapor and topographical variables) and environmental absences (i.e. location with known environmental conditions not suitable for the target species, following the definition used by Lobo et al. (2010)) to produce a neutral model for baseline species potential.

As an additional source of homogeneously distributed true absence data we used the Land Use/Cover Area Survey (LUCAS) (EUROSTAT, 2017) dataset: in-situ observations of land use and land cover distributed on a 2×2 km grid covering the whole European Union (see d'Andrimont et al. (2021) for more information and `https://ec.europa.eu/eurostat/web/lucas/data/lucas-grid` for the official grid). Final prediction maps show the probability of presence (0–100%) of at least one individual of the target species in the

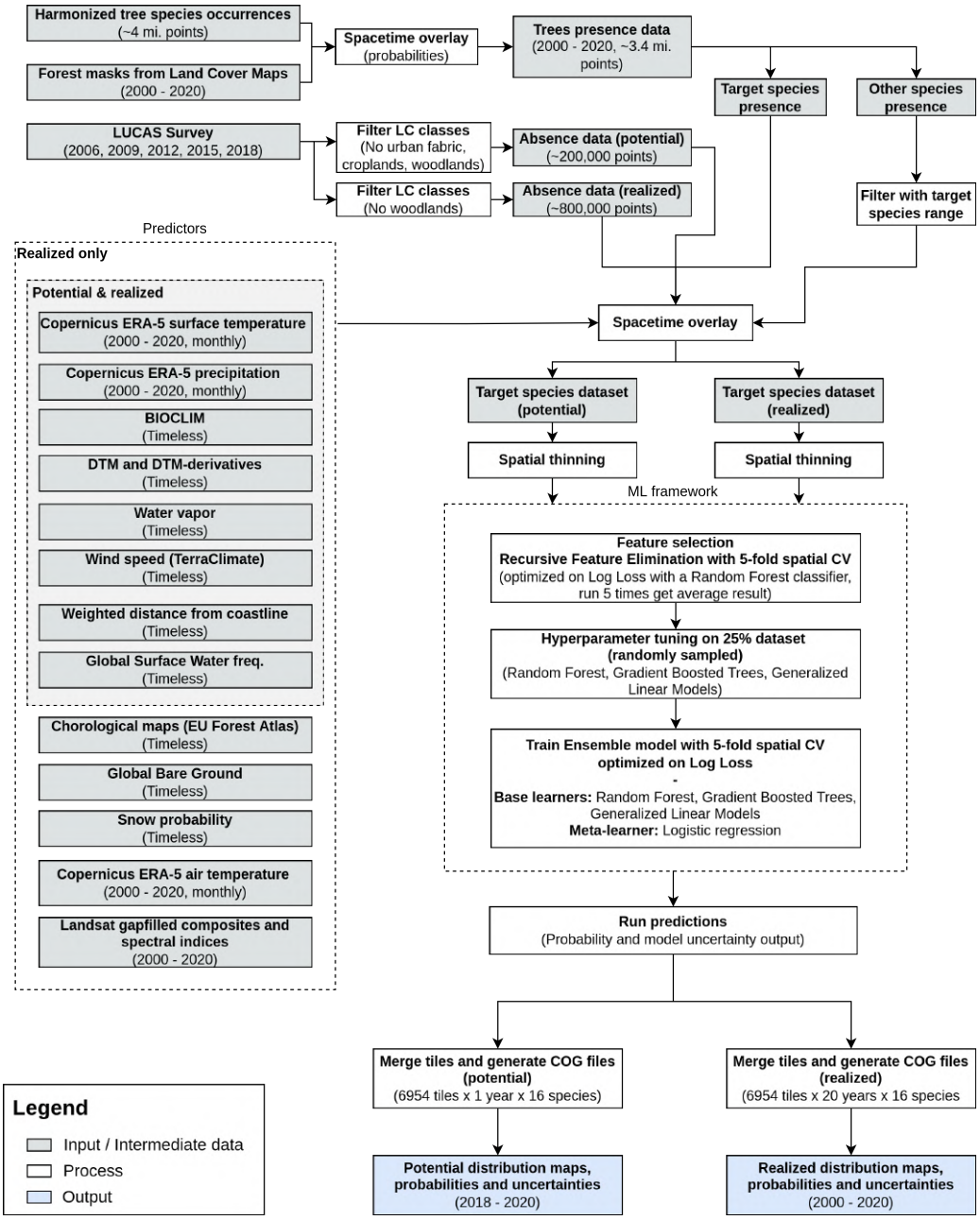

Figure 3.1: General workflow illustrating the preparation of the point data, the predictor variables used, model building (feature selection — hyperparameter optimization — training) and preparation of distribution maps for one species. The process was identically replicated for all the species.

area covered by a 30 m pixel. Probability of presence is relative to the mapped target species, irrespective of the potential co-occurrence of other species in the same 30 m pixel and should not be confused with the absolute abundance or proportion of each species in the pixel area. The sum of the presence probabilities of different species in the same pixel can thus exceed 100%. We produced one potential distribution map and six realized

distribution maps for each species: the assumption is that the conditions in the study area that determine the potential distribution of the species did not change over the time period analyzed; this does not hold for the realized distribution. We split the time period analyzed in six time windows according to the following scheme: (1) 2000–2002, (2) 2002–2006, (3) 2006–2010, (4) 2010–2014, (5) 2014–2018 and (6) 2018–2020.

To ensure transparent reporting and reproducibility, we described the dataset according to the ODMAP protocol suggested by Zurell et al. (2020). We implemented the workflow in the Python (Van Rossum and Drake, 2009) and R (R Core Team, 2021b) programming languages. More technical details on preprocessing steps and packages used according to ODMAP (Zurell et al., 2020) are presented in Table S1 (found in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.1`) (**Bonannella** et al., 2022c).

3.2.2 Study area

The study area covers the European continent, that is all countries included in the Corine Land Cover (CLC) database (Büttner et al., 1998) except Turkey (Fig. 3.2). European forests cover 33% of the continent's land area. Owing to the variety of climatic conditions across both latitudinal and longitudinal gradients, twelve out of the 20 FAO Forest Ecological Zones are represented in European forests (Rigo et al., 2016). The European Atlas of Forest Tree Species (San-Miguel-Ayanz et al., 2016) reports detailed information for a total of 76 forest tree species. From those, a selection of 16 were chosen and modelled in this study. The complete list of species is presented in Table S1.

3.2.3 Training points

Preparing and combining legacy occurrence points

A total of 2,454,997 tree species occurrence points from three different sources were gathered. The majority of points (71%) comes from the Global Biodiversity Information Facility (GBIF). National Forest Inventory (NFI) data from multiple EU member states published by Mauri et al., 2017 forms another 23% of the dataset. The remaining 6% comes from the LUCAS dataset.

Entries were filtered for species included in the European Atlas of Forest Tree Species (San-Miguel-Ayanz et al., 2016). Occurrences with a taxonomy rank other than species or genus were omitted. Same applies to points which had flags indicating serious location issues (i.e. missing coordinates). Geometries were re-projected to coordinate reference system ETRS89 / LAEA Europe (EPSG: 3035). A high resolution land mask for Europe (Hengl et al., 2020) was applied to further exclude misplaced occurrence points. GBIF taxon and genus keys were derived for the other two data sources. Quality flag variables for location accuracy and date were established from existing metadata to indicate potentially

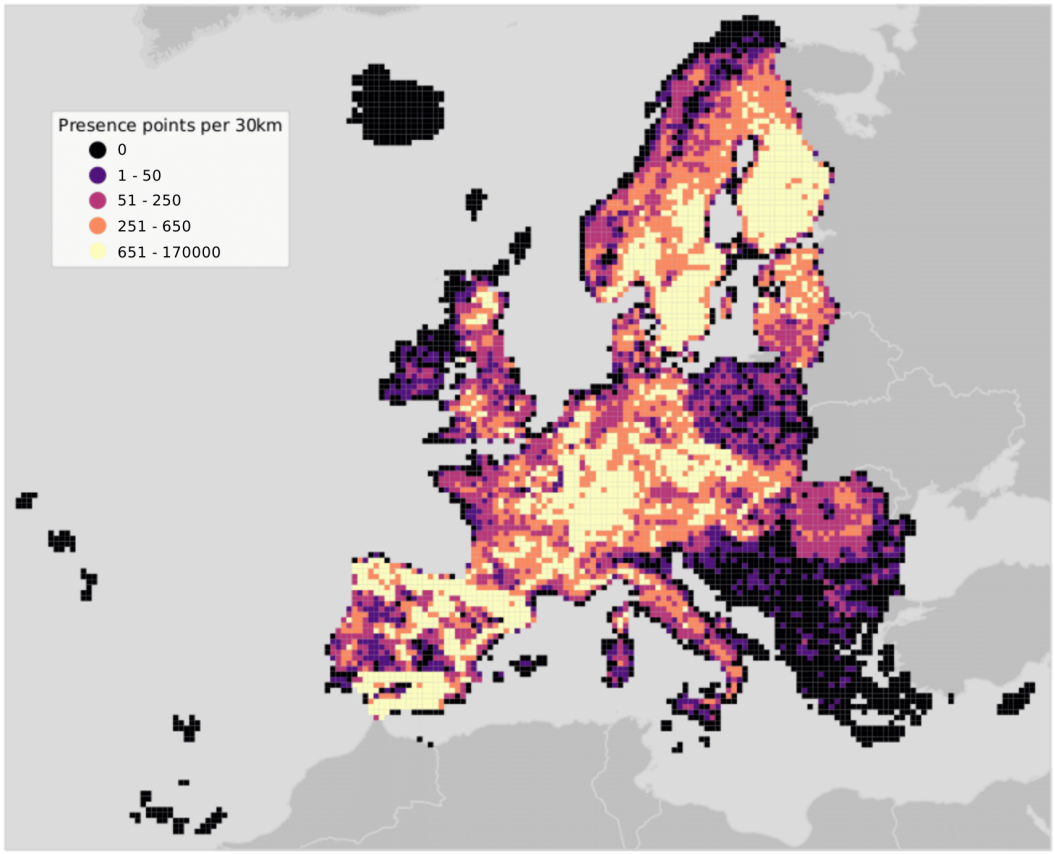

Figure 3.2: Map of the study area showing presence points only. Points are aggregated at a coarse resolution (30 km) scale and absence points are omitted for visualization purposes.

problematic entries. The harmonized point dataset has information on species and genus (including respective GBIF keys), year of observation, country, original data source, citation, and license among other auxiliary variables.The dataset was published separately (Heisig and Hengl, 2020).

We used yearly forest masks derived from Witjes et al. (2022) to decide upon including point data lacking the year of observation. Witjes et al. (2022) provides yearly probability maps at 30 m for the 2000–2020 for 43 land cover classes according to the CLC level 3 legend. We overlaid the points with the probability maps with prevalent forest (classes: 311, 312, 313 and 323) or woodland-shrub (324, 333) cover. Points were used only if the probability value extracted for at least one of the classes was $\geq$50% for all the years considered. Each unique combination of longitude, latitude and year was then considered as an independent sample. An additional quality flag was added to distinguish points coming from this operation and the points with original year of observation coming from source datasets.

Preparing non-occurrence points

A total of 883,630 land cover points were gathered from the LUCAS database as provided by Eurostat and used as absence data. All LUCAS survey data (2006, 2009, 2012, 2015 and 2019) was used: each survey was first downloaded individually and then aggregated. As for the occurrence points, spatial and temporal information were used to uniquely identify one observation. All main land cover classes were used for selecting observations for the absence dataset with the exception of class C (Woodland class), as points belonging to that class already served the selection of presence data. For potential distribution only, points coming from human influenced land cover classes (class A and B) were also excluded. This choice was taken assuming that cities (class A) and croplands (class B) could be suitable areas for the target species if only environmental criteria are met. Two presence-absence datasets were produced for each species, one to be used for potential distribution and one for realized distribution. Locations in space and time of the target species were considered to be true presences, while presence locations of other species and observations from the LUCAS dataset were assumed to be the true absence locations. Presence locations of other species used as true absence were additionally filtered by overlaying them with a rasterized chorological map downloaded from the European Atlas of Forest Tree Species portal for each of the target species. Only points falling outside the geographical extent of the target species chorological map were used as absence locations for modeling.

Spatial thinning

Combining different data sources to generate the tree occurrence points produced a dataset with unknown sampling design, while LUCAS points are regularly distributed across the whole study area. To overcome the problem of uneven sampling intensity and spatial clustering, we applied a spatial thinning procedure using the *spThin* R package (Aiello-Lammens et al., 2015). A distance of 2 km was considered as minimum distance between the points, to harmonize the sampling intensity between presence and absence data. The procedure was repeated 10 times: at each iteration, the algorithm randomly removes one observation from the dataset until no observation is left with a nearest neighbor closer than the thinning distance. Among the 10 datasets obtained, the one with the largest number of records was retained and used for modeling. However, the package was not developed for large datasets: the implementation of the thinning algorithm cannot be processed in parallel and computation time can take even days with a number of observations ≥ 3000. Due to these computational constraints, we first overlaid the points with a 10×10 km grid and ran the thinning procedure per tile. Results of this operation are shown in Table S2 and Fig. S1 (found in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.2`) (**Bonannella** et al., 2022c).

3.2.4 Predictor variables

A total of 305 harmonized variables covering continental Europe at different spatial resolution were used as predictors to model the realized distribution of the species. In this study we included both dynamic (i.e. time-series of data of different temporal resolution) variables covering the time period January 2000 – December 2020 and static (i.e. variables not expected to change during the modelled time period) variables. A subset of only 103 variables were used instead to model the potential distribution (see Fig. 3.1). All data was reprojected in the coordinate reference system ETRS89 / LAEA Europe (EPSG: 3035) before the analysis.

Dynamic data

We used a reprocessed version of Landsat ARD data provided by Global Land Analysis and Discovery (GLAD) (Potapov et al., 2020): time series used in this study covers the period 1999–2020. Cloud and cloud shadow pixels were removed from the images, maintaining only the quality assessment-QA values labeled as clear-sky. Afterwards, individual images were averaged by season according to three different quantiles (25th, 50th and 75th) and the following calendar dates for all periods:

- Winter: December 2 of previous year until March 20 of current year,
- Spring: March 21 until June 24 of current year,
- Summer: June 25 until September 12 of current year,
- Fall: September 13 until December 1 of current year.

84 images (3 quantiles $\times$ 4 seasons $\times$ 7 Landsat bands) were produced for each year. Missing values were imputed using the *Temporal Moving Window Median* algorithm. For more details on the preprocessing of Landsat data for this study see Witjes et al. (2022). 7 different spectral indices (see Table 3.1) were computed for each year and season using the 50th quantile only, for a total of 7 x 4 = 28 spectral indices variables per year.

A reprocessing of the ERA5 Land hourly dataset has been used to have monthly aggregates of air temperature (2 meters above ground), surface temperature and precipitation. Original ERA5 data was aggregated to daily data, and subsequently to monthly data, with increased resolution (1 km) using CHELSA data (Karger et al., 2020): in this way the general spatial and temporal pattern of ERA5 Land dataset was kept while using the fine spatial detail coming from the CHELSA dataset. For air and surface temperature we obtained the monthly minimum, mean and maximum, while for precipitation the monthly sum for a total of 84 climatic time series layers.

Table 3.1: Table with Landsat-derived spectral indices used in this study.

Spectral Index	Abbreviation	Formula	Reference
Enhanced Vegetation Index	EVI	$2.5 \times \frac{NIR - RED}{NIR + 6 \times RED - 7.5 \times BLUE + 1}$	(Huete et al., 2002)
Enhanced Vegetation Index 2	EVI2	$2.5 \times \frac{NIR - RED}{NIR + 2.4 \times RED + 1}$	(Jiang et al., 2008)
Modified Soild Adjusted Vegetation Index	MSAVI	$\frac{(2 \times NIR + 1) - \sqrt{(2 \times NIR + 1)^2 - 8 \times (NIR - RED)}}{2}$	(Qi et al., 1994)
Normalized Burned Ratio	NBR	$\frac{NIR - SWIR2}{NIR + SWIR2}$	(Key and Benson, 1999)
Normalized Difference Vegetation Index	NDVI	$\frac{NIR - RED}{NIR + RED}$	(Tucker, 1979)
Normalized Difference Wetness Index	NDWI	$\frac{NIR - SWIR1}{NIR + SWIR1}$	(Gao, 1996)
Soil Adjusted Vegetation Index	SAVI	$(1 + 0.5) \times \frac{NIR - RED}{(NIR + RED + 0.5)}$	(Huete, 1988)

Static covariate datasets

As additional static covariates, we used the following datasets:

- 19 bioclimatic variables (Hijmans et al., 2005) for the period 1979 - 2013 to provide a baseline of the actual state of the climate; we used bioclimatic variables from the CHELSA dataset since it has been claimed to better match data from meteorological stations than WorldClim (Karger et al., 2017),
- 50 chorological maps downloaded from the European Atlas of Forest Tree Species web portal. Chorological maps provide a qualitative overview of the spatial distribution of a species over an area, differentiating between native and introduced. We considered both the native and introduced areas as the potential distribution of a species for the time period covered by the study,
- Global bare ground cover from Hansen et al. (2013). The layer provides information on bare ground cover on a percent (1–100) scale,
- Solar direct and diffuse irradiation,
- 13 cloud fraction layers (monthly averages and annual average) derived from MODIS (Wilson and Jetz, 2016),
- Digital terrain model (DTM) for Europe (Hengl et al., 2020) and DTM-derived (slope, hillshade) variables,
- Easterness, northness (Olaya, 2009), and positive and negative openness,
- Probability of surface water occurrence at 30 m resolution derived from Landsat time series (Pekel et al., 2016),
- Height above nearest drainage (HAND) and flow accumulation area at 90 m resolution from the MERIT Hydro global hydrography datasets,

- Long-term flood hazard map calculated on a 500 years time period (Dottori et al., 2016),
- Water vapor pressure (kPa) based on the WorldClim2.1 dataset (Fick and Hijmans, 2017),
- Long-term snow probability based on MODIS (MOD10A2) and available at `https://doi.org/10.5281/zenodo.5774953`,
- Monthly wind speed (1998–2018) from TerraClimate.

For more details on spatial and temporal resolution, preprocessing and data sources see the supplementary material.

3.2.5 Feature selection

Features for potential and realized distribution for each species were selected using the Recursive Feature Eliminitation (RFE) strategy, implemented in the scikit-learn library (Pedregosa et al., 2011). For each combination of species and modelled distribution we trained a random forest classifier (num.trees = 50, default values were used for the other parameters): RFE fits the model and removes the weakest feature until a specified number of features is reached, then ranks the importance of the features based on the model's coefficients (for regression-based models) or feature importance (for random forest).

The minimum number of features was not known before hand: to select this number, we ran the Recursive Feature Elimination with a spatial 5–fold Cross Validation (RFECV), using the logarithmic loss, or logloss, as a scoring estimator. Logloss is one of the most robust performance metrics when it comes to imbalanced datasets (Ferri et al., 2009). Logloss is indicative of how close the predicted probability for an observation i is to the corresponding label y. For binary classification with label $y \in 0, 1$ the overall logloss was calculated as:

$$Logloss = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \ln\left[p(y_i)\right] + (1 - y_i) \cdot \ln\left[1 - p(y_i)\right] \tag{3.1}$$

where N is the total number of observations and $p(y_i)$is the predicted probability for an observation with $y_i = 1$. It follows that values close to 0 indicate high prediction performances, with logloss = 0 being a perfect match, and values that are positive to infinite are progressively worse scores. For comparison, the value of logloss for random assignment depends on the number of classes (a) and the prevalence of the classes (b): for binary classification and a balanced (50:50) dataset with $N = 10$ observations, the equation (3.1) gives a value of 0.69.

We ran the RFECV on a 25% random subsample for each species and modelled distribution; this operation was replicated 5 times. For each iteration we selected the minimum of the logloss function (see Fig. S2 in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.3`) and the averaged result was then used as the minimum number of features for the RFE.

3.2.6 Model building and evaluation

Modeling methods

To build an ensemble model, we decided to compare predictive performances and computing time (hyperparameter tuning — cross validation — prediction time) of different machine learning algorithms on a random 25% subset of observations for both potential and realized distribution datasets. A detailed workflow of this process is shown in Fig. S3 (in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.4`). We decided to conduct this test on seven different species: choice of the species was based on the spatial distribution of the training points and the ratio between presence and absence points. In this way, algorithms performances could be tested on different ecological conditions (latitudinal and longitudinal gradient) and imbalance of classes. The species selected were: *A. alba*, *C. sativa*, *F. sylvatica*, *P. abies*, *P. halepensis* and *P. sylvestris*.

We compared seven different algorithms: Random Forests (RF) (Breiman, 2001a), Gradient-boosted trees (GBT) (Friedman, 2002), Classification trees (CART) (Therneau and Atkinson, 2011), Generalized Linear Models (Nelder and Wedderburn, 1972) with Lasso regularization (Tibshirani, 1996) (just GLM from now on), C5.0 (Quinlan, 1986), K-nearest neighbor (KNN) (Fix and Hodges, 1989) and Artificial Neural Network (ANN) (Ripley and Venables, 2017). Analyses were conducted using the *mlr* package (Bischl et al., 2016). For each algorithm, a hyperparameter space was defined: combinations of hyperparameters were generated per model based on a grid search of 5 steps per hyperparameter. More details on the hyperparameter space are available in Table S3 (found in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.5`).

Selecting component models

We evaluated each combination of hyperparameters by comparing logarithmic loss values during a 5–fold spatial cross validation replicated 5 times: we used spatial instead of normal cross validation for hyperparameter tuning because it reduces overoptimistic performance results in the presence of strong data clustering (Schratz et al., 2019). We used the tile ID produced in the tiling system for Europe as the blocking parameter in the training function in *mlr*. All the compared algorithms were used in *"probability"* mode, that is, predicting for each observation in the dataset a probability value for presence (class 1) and absence (class 0). Besides the performance achieved in the logloss metric, computing time

for the hyperparameter tuning, a 5–fold spatial cross validation and prediction time for a 30 km tile were also considered as additional criteria: we calculated the computing time only for the species that had the highest computational costs (*P. sylvestris*). This gave us an estimate of how long the process of training each component model could take during the building of the ensemble model. We used logloss performance as the first criteria to choose the component models: only in the case of two or more methods performing within one standard deviation from the average performance, we chose the computationally fastest.

Training ensemble model using stacking

Stacked generalization involves combining predictions made by level 0 models and using them as training data for a level 1 model (or meta-learner from now on) (Wolpert, 1992). To limit overfitting in the training data, we used a 5–fold spatial cross validation: the out-of-fold predictions were used to build a level 1 training dataset for the meta-learner. We used logistic regression as a meta-learner, which is usually the most used model for classification problems (Gomes et al., 2012). Final predictions are delivered as probability maps (0–100%) for presence together with model uncertainty maps: we consider as model uncertainty the standard deviation of the predicted values of the base learners. The principle is that the higher the standard deviation the more uncertain the model is regarding the right value to assign to the pixel (Brown et al., 2020).

Variable importance assessment

To assess to what extent the three level 0 models used different parts of the available feature space and the agreement between these models, we compared the variable importance when possible. For RF and CART we used Gini importance, for C5.0 the *"percentage of training set samples that fall into all the terminal nodes after the split"*(Quinlan, 1986), for GBT the gain metric (Shi et al., 2019) and for GLM the coefficients for the minimum fitted value of λ (Hastie et al., 2016). To better analyze the results, we aggregated the whole set of variables in 7 macro-classes:

- Climate (i.e. precipitation, wind speed, water vapor, snow probability etc.),
- Temperature (i.e. time series of recorded temperatures for the observed time period),
- Bioclim (i.e. bioclimatic variables from CHELSA),
- Topography (i.e. DTM and DTM-derivative variables),
- Landsat band (i.e. all percentiles, all seasons),
- Distribution (i.e. species distribution maps from European Atlas of Forest Tree Species),
- Spectral index (i.e. spectral indices derived from Landsat bands).

Model evaluation

Predictive performance of the ensemble model was assessed through spatial 5–fold cross-validation repeated 5 times (Roberts et al., 2017) with logloss as performance metric. To investigate if the ensemble model outperformed the component models, we compared results of the spatial cross validation of the ensemble with the results of the component models. To be able to compare performances between different species, we converted logloss performances used the following formula:

$$R^2_{\text{logloss}} = 1 - \frac{Logloss_{\text{m}}}{Logloss_{\text{r}}} \tag{3.2}$$

where $Logloss_m$ is the performance achieved by the model and $Logloss_r$ is the value for random logloss, used as a baseline for predictive performances. Values close to 1 indicate high predictive performances, while values close to 0 indicate lower performances, with 0 meaning that the model is no better than a random guess. We also reported a threshold-dependent metric, the True Skill Statistic (TSS) and a threshold independent metric, the area under ROC curve (AUC), as they are commonly used metric to evaluate SDMs predictive performances (Chakraborty et al., 2021; Shabani et al., 2018). TSS was computed using the default threshold value (0.5) when assigning predicted probabilities values to the presence or absence class. Logloss is one of the least sensitive metric to prevalence (Ferri et al., 2009), hence our choice of logloss as a primary performance metric to compare different models coming from different training datasets.

To assess the effect of high resolution products on predictive performances, we excluded Landsat bands and Landsat-derived spectral indices from the list of predictors used for realized distribution. We then applied our spatio-temporal machine learning framework (feature selection — hyperparameter tuning — ensemble model training) on each species and ran a 5–fold spatial cross validation repeated 5 times to evaluate model performances. For the ensemble model we used the same component models (RF, GBT and penalized GLM) and meta-learner (logistic regression). Results of this analysis were then compared with the performances achieved by the ensemble models using Landsat data.

3.3 Results

3.3.1 Spatio-temporal machine learning framework

Table 3.2 shows that RF on average had the highest predictive performances for all species, with GLM coming closer. RF scored the lowest logloss among all the other algorithms in 9 cases out of 14 and scored the same as GLM in 1 case out of 14. In the remaining cases, GLM scored the lowest logloss value, with RF scoring the second lowest. On average, GLM performed better in in potential distribution tasks, with RF clearly outperforming

every other algorithm in realized distribution tasks. Overall, GLM and RF always scored the lowest logloss values, from two to three times lower than all the other algorithms in some cases.

Table 3.2: Average logloss for the compared algorithms and for the subset of seven target species. In bold are highlighted the best performing learners for each task.

Species	Distribution	ANN	C5.0	GBT	GLM	KNN	RF	CART
A. alba	Potential	0.242±0.024	0.053±0.009	0.097±0.003	**0.027±0.003**	0.120±0.021	0.033±0.006	0.063±0.008
C. sativa	Potential	0.210±0.019	0.118±0.020	0.128±0.003	0.058±0.006	0.197±0.033	**0.057±0.008**	0.132±0.011
F. sylvatica	Potential	0.516±0.027	0.114±0.011	0.138±0.003	**0.055±0.004**	0.108±0.012	0.060±0.004	0.184±0.014
P. abies	Potential	0.390±0.017	0.176±0.010	0.199±0.005	0.144±0.006	0.314±0.022	**0.114±0.005**	0.292±0.009
P. halepensis	Potential	0.220±0.023	0.053±0.012	0.092±0.003	**0.019±0.002**	0.075±0.015	0.023±0.003	0.070±0.013
P. sylvestris	Potential	0.655±0.041	0.370±0.017	0.358±0.005	0.318±0.008	0.569±0.030	**0.232±0.006**	0.430±0.013
Q. robur	Potential	0.422±0.019	0.117±0.012	0.144±0.004	**0.065±0.004**	0.152±0.026	0.068±0.007	0.154±0.010
A. alba	Realized	0.383±0.030	0.145±0.017	0.140±0.007	0.106±0.009	0.245±0.059	**0.059±0.009**	0.151±0.021
C. sativa	Realized	0.316±0.025	0.175±0.031	0.161±0.010	0.118±0.017	0.351±0.065	**0.077±0.010**	0.173±0.017
F. sylvatica	Realized	0.654±0.055	0.118±0.010	0.147±0.005	0.129±0.007	0.209±0.039	**0.057±0.011**	0.200±0.025
P. abies	Realized	0.666±0.053	0.180±0.011	0.208±0.009	0.177±0.009	0.467±0.042	**0.125±0.008**	0.280±0.035
P. halepensis	Realized	0.292±0.034	0.065±0.014	0.104±0.004	**0.025±0.005**	0.074±0.022	**0.025±0.005**	0.084±0.013
P. sylvestris	Realized	0.656±0.043	0.451±0.013	0.473±0.009	0.478±0.009	0.776±0.052	**0.304±0.006**	0.549±0.012
Q. robur	Realized	0.642±0.049	0.105±0.016	0.133±0.005	0.091±0.011	0.200±0.057	**0.053±0.008**	0.200±0.030

The absolute difference between values scored by GLM and RF is lower than when RF had the advantage over GLM. This indicates a high reliability of RF performances even when other models outperform it. The ANN scored the highest logloss values in all tasks, so it was immediately excluded from the pool of level 0 models to choose from. It was time consuming to find a common hyperparameter range well suited for different tasks, since neural networks are often extremely situation-dependent. After a preliminary selection, we used the range shown in Table S3: despite that, our results remained inferior to those obtained with the other learners. On top of that, the *mlr* implementation of neural networks, based on the *deepnet* R package (Rong and Rong, 2014), doesn't allow the use of ReLU (rectified linear activation function) as an activation function, which would have been beneficial for our purposes. Based on logloss performances, we selected RF and GLM as the first two components of the ensemble. Based on similar values of logloss (within one standard deviation of the average performance) scored by C5.0, GBT, KNN and CART, we used computational costs to choose the third component model (Table 3.3).

KNN was excluded due to computing time values being from one to two order of magnitude higher than the ones scored by the other models. Even though CART scores very low values in cross validation and prediction time in both potential and realized tasks, tuning time is the second highest, just behind KNN. C5.0 is faster than GBT in the whole potential workflow (446.8 seconds against 741.6) but slower in the realized workflow (1327.7 seconds against 1006.4). Considering both workflows, GBT proved to be faster and more consistent in cross validation and prediction time, showing an increase in tuning time of just 30% with double the amount of training data (see Table S2).

Table 3.3: Hyperparameter tuning, cross validation and prediction time for each model and distribution task. Time values are reported in seconds. Tests were conducted in a parallel computing setup on a CPU server running 2 x Intel(R) Xeon(R) Gold 6248R - 3.00GHz (96 threads) with 504 GB RAM.

Distribution	Process	ANN	C5.0	GBT	GLM	KNN	RF	CART
Potential	Tuning	561.2	310.7	527.2	448.9	2433.6	104.4	576.5
Potential	Cross validation	57.2	44.7	192.5	620.4	356.9	240.2	26.5
Potential	Prediction	24.1	91.4	21.9	14.8	19272.9	35.5	15.4
Realized	Tuning	1031.6	964.2	688.4	859.1	12321.9	396.6	1650.2
Realized	Cross validation	119.2	165.4	290.8	1372.9	1445.1	805.5	114.9
Realized	Prediction	26.1	198.1	27.2	17.3	$>$ 1 day	52.4	17.4
	Total	**1819.2**	**1774.5**	**1748.0**	**1455.5**	**$>$ 1 day**	**1634.6**	**2400.9**

3.3.2 Variable importance

Of all the features used in both potential and realized distribution, 60 are considered important for both tasks. For potential distribution, diffuse irradiation, precipitation of the driest quarter (BIO17) and precipitation of the driest month (BIO14) were the most important and most frequent predictors across all component models and species (see Fig. 3.3).

The density distributions per macro-class help understanding how the Bioclim macro-class was the one with on average both most important and most frequent variables. Other variables are more species-specific: the minimum surface temperature of April records the highest absolute value in relative importance but it was important for only one species (*Q. robur*, see Fig. S7 in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.7`). The Temperature macro-class accounts the highest numbers of predictors, but the values recorded in both variable importance and frequency are the lowest among all the macro-classes. The Climate macro-class had the largest variety in predictors and variables in this class are homogeneously spread out across all the species in both variable importance and frequency.

For realized distribution, the summer aggregates of Landsat green (25th, 50th and 75th quantiles) were the three most important and most frequent variables across all models and species, closely followed by the summer aggregates of Landsat red and fall aggregates of NDVI (Fig. 3.3). While the Spectral Index macroclass clearly outperformed the other ones in relative importance, the Bioclim class scored as the most frequent across all the species. The distribution maps scored the highest values for variable importance (distribution of the *F. excelsior* and the *Tilia spp.*) but they were species-specific.

Overall, the component models show more differences in variable importance in the potential distribution models than in the realized ones. On average, RF and GBT selected the same variables in the top–10 but not always in the same order, while GLM tended to

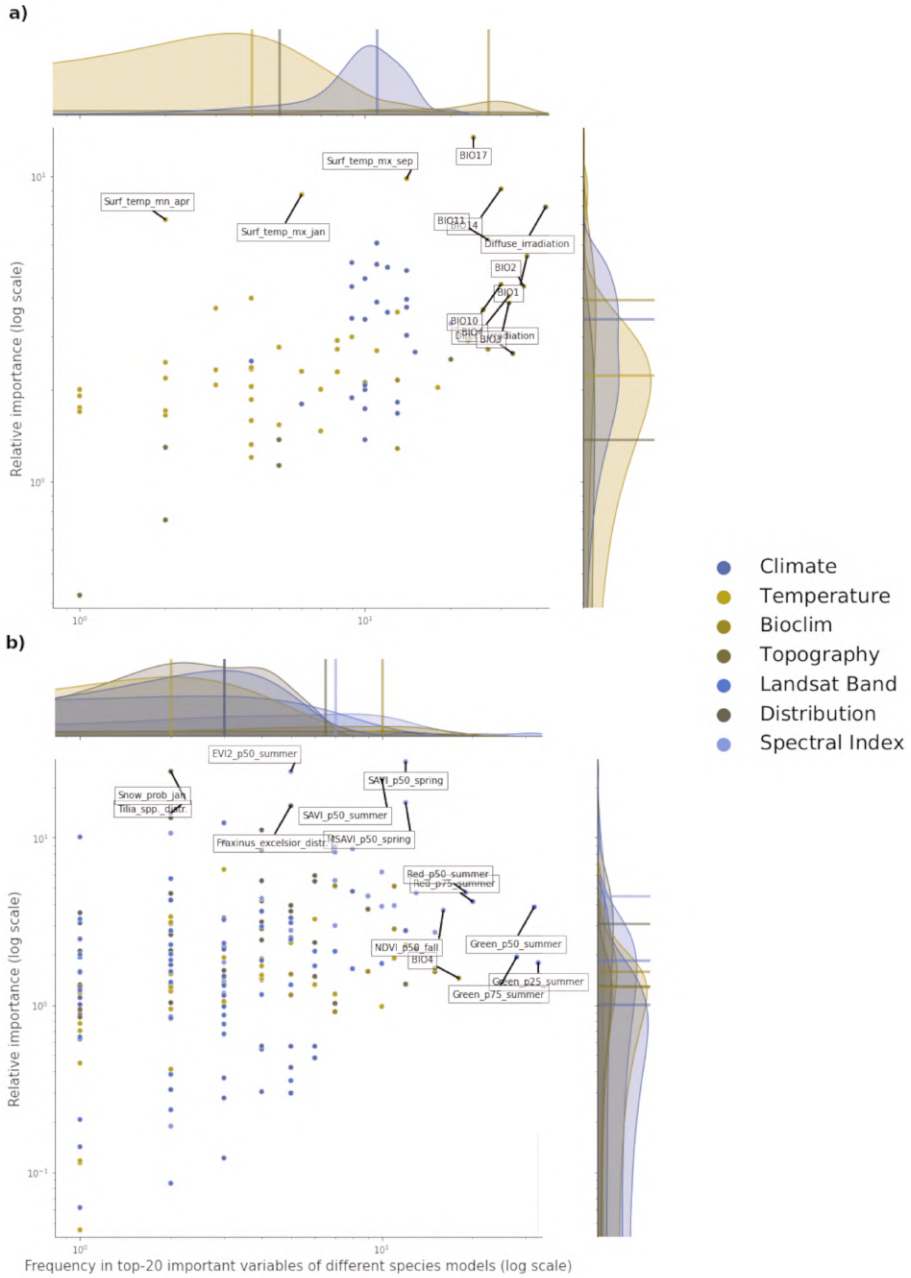

Figure 3.3: Relative variable importance vs frequency of the variables of the top–20 most important across the component models and all species for potential **(a)** and realized **(b)** distribution. Each plot can be divided in four quadrants, from the top left clockwise: variables with high relative importance but low frequency (i.e. important for one or few species), variables with high importance and high frequency (i.e. important for all species), variables with low importance and high frequency (i.e. they occured often but were not important) and variables with low importance and low frequency. Labeled dots are variables that recorded high values of relative variable importance or frequency.

choose completely different variables (i.e. spectral indices for realized distributions and wind speed for potential distribution). This suggests how the ensemble models tend to use a wider proportion of the feature space than single models. This tendency is most apparent in the potential distribution models. In the realized distribution models, the component models agree in selecting the top-10 most important variables predictors from Landsat bands or Spectral indices. RF and GBT considered on average the Landsat bands as the most important, while GLM selected the spectral indices more often.

3.3.3 Accuracy assessment

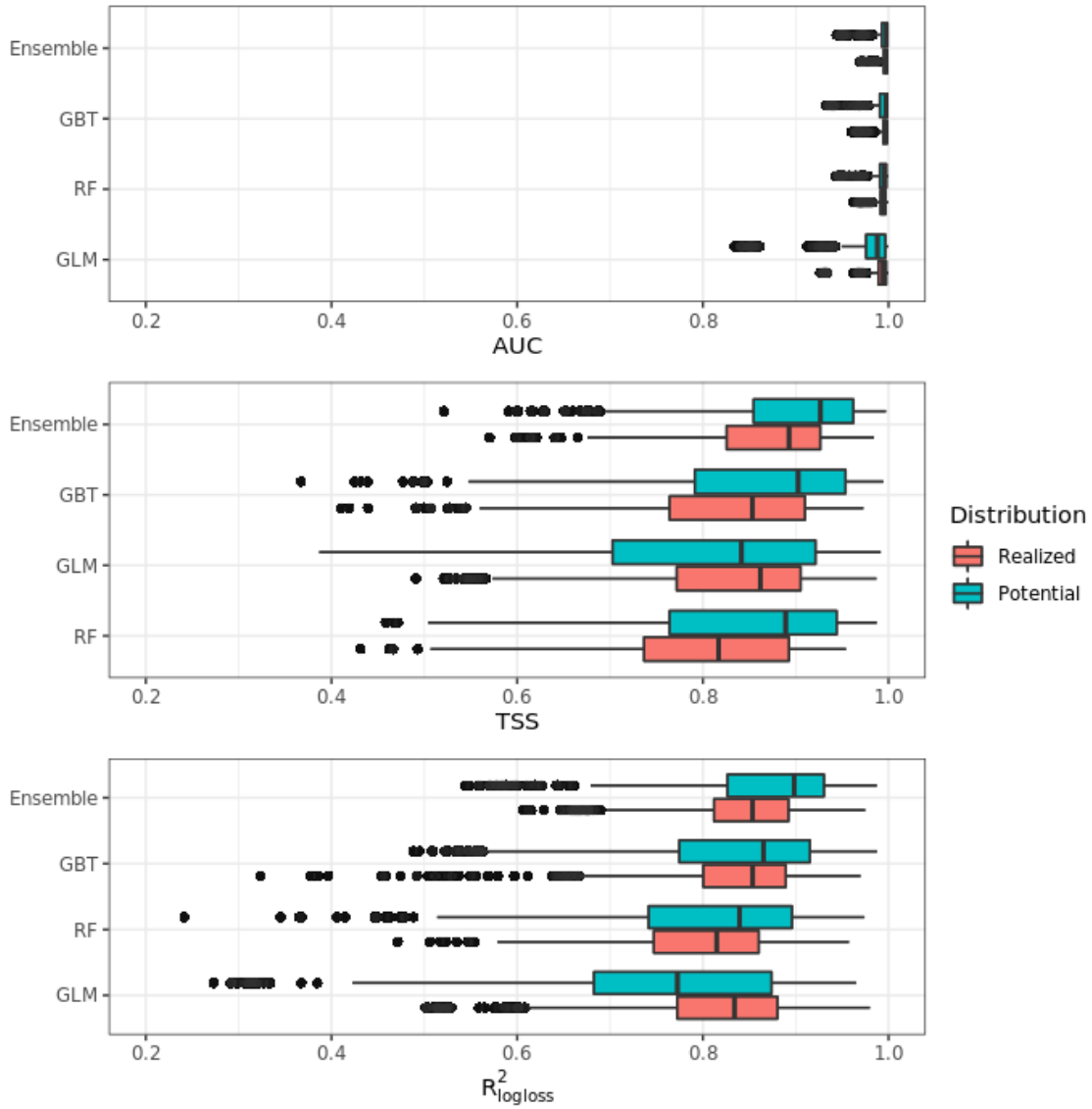

Figure 3.4: Aggregated results of the accuracy assessment per model and distribution expressed using AUC, TSS and $R^2_{logloss}$

Fig. 3.4 shows that on average the ensemble model outperformed all component models in both potential and realized distributions. AUC values seem to be overoptimistic and with low variability for all algorithms and distributions, with the largest interquartile range (IQR) being GLM - potential, going from 0.97 to 0.99. Values for TSS and $R^2_{logloss}$ seem to be more conservative, with the ensemble still having the highest average (TSS = 0.898, 0.874 and $R^2_{logloss}$ = 0.857, 0.839, respectively, for potential and realized distribution) values, and lowest IQR (TSS = 0.85 - 0.96, 0.82 - 0.92 and $R^2_{logloss}$ = 0.82 - 0.93, 0.81 - 0.89, respectively, for potential and realized distribution). While results from our modeling framework proved GLM and RF being the best models in both potential and realized tasks (see Table 3.2), GBT achieved overall better performances than both algorithms. In general, the models for potential distribution achieved better predictive performances than those for realized distribution; however potential distribution has greater IQRs as well as larger outliers.

Fig. 3.5 shows the performances of the ensemble model per species and distribution. Results for the component models are available in Table S4 and S5 (found in `https://zenodo.org/record/6516728/preview/Supplementary_material.pdf#subsection.0.8`). Following the trend shown in Fig. 3.4, differences in performances are minimal if we look at the AUC results, while they grow significantly if we look at the TSS and $R^2_{logloss}$ ones. We see that for both potential and realized distribution, models for *Q. suber* achieved the best performances (TSS = 0.968, 0.959 and $R^2_{logloss}$ = 0.952, 0.949, respectively, for potential and realized distribution), while *P. sylvestris* (TSS = 0.731, 0.785 and $R^2_{logloss}$ = 0.585, 0.670) and *P. nigra* (TSS = 0.658 0.686 and $R^2_{logloss}$ = 0.623, 0.664) achieved the worst. Furthermore, while *Q. suber* model has overall best performances in all metrics, in both potential and realized distribution AUC grades *P. sylvestris* as the worst and *P. nigra* as the second worst; the opposite is true for TSS scores. For $R^2_{logloss}$, *P. sylvestris* scored as the worst in potential distribution and second worst in realized distribution, with the opposite happening for realized distribution.

3.3.4 Influence of high resolution on predictive performances

Fig. 3.6 shows that ensemble models for realized distribution including Landsat data consistently outperformed models without Landsat data. In all metrics, values scored by the Landsat models show higher median and average values than the ones without Landsat and lower IQR range. The same trend is shown by all metrics: like in the previous cases shown in section 3.3.3, AUC values are high, reaching 1 for some species, and with low IQR. TSS and $R^2_{logloss}$ show a larger IQR and lower median and average values than AUC: given the large differences in values scored by the models in TSS and $R^2_{logloss}$, these two metrics proved to be more helpful to discriminate model performances across multiple species. On average, including Landsat data increases TSS performances by 6.5% and $R^2_{logloss}$ by 7.5%, while when comparing median values the increase in performances is higher in TSS (+7.5%) and lower in $R^2_{logloss}$ (+7%).

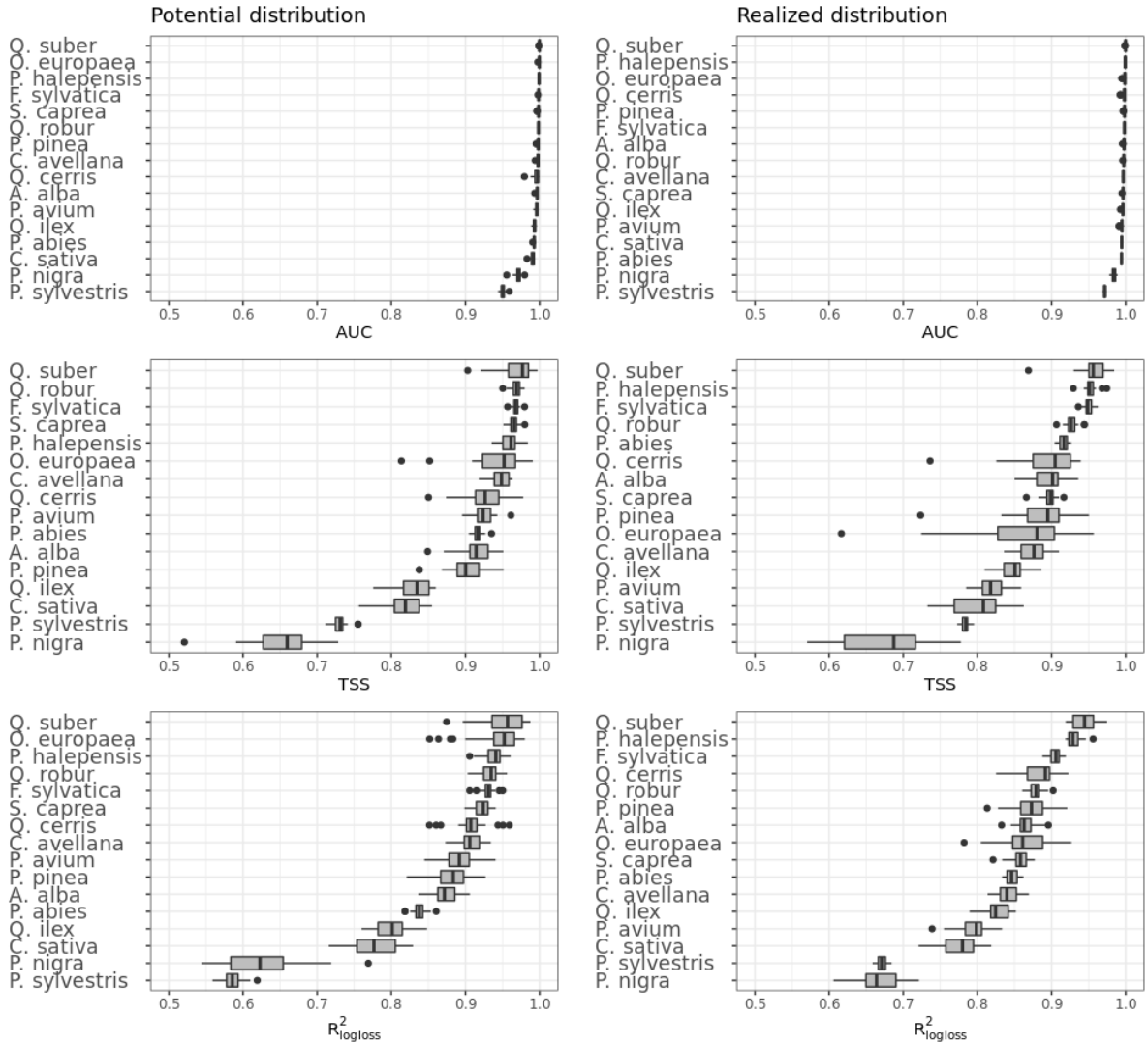

Figure 3.5: Results of the accuracy assessment per model and distribution for the ensemble model only expressed using AUC, TSS and $R^2_{logloss}$

3.4 Discussion

3.4.1 Modeling framework

Combining models using the ensemble approach is thought to reduce model uncertainty and increase its robustness in modeling species distributions (Araújo and New, 2007). We used ensemble with stacked generalization as ensemble approach, which has not been tested yet for species distribution modeling. We also trained the models in a spatio-temporal framework, expecting the models to generalize better when predicting in a temporal window not included in the training data.

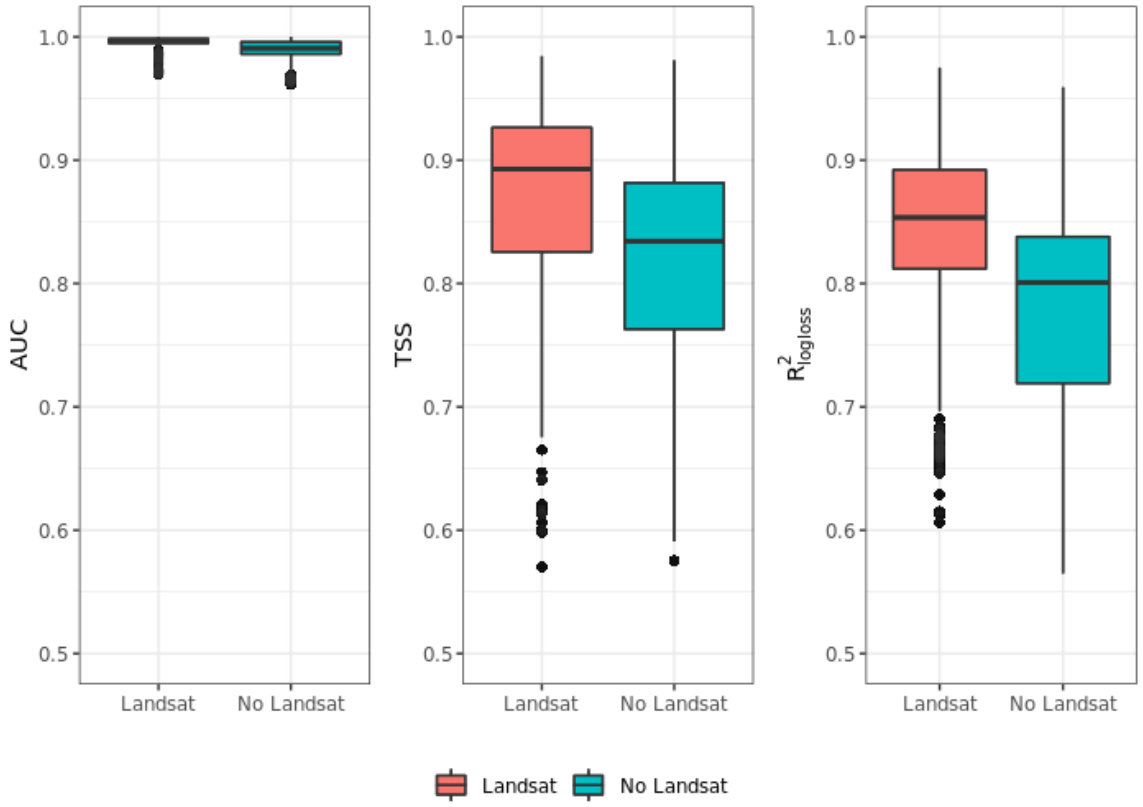

Figure 3.6: Aggregated results of the accuracy assessment for modeling realized distribution with and without the Landsat bands and spectral indices expressing using AUC, TSS and $R^2_{logloss}$

Part of the intent of the paper was to provide a robust reproducible framework to model species distributions based on ensemble ML. Hao et al. (2020) used a similar methodological framework to the one used in this study. They modelled the distribution of 13 species of the genus *Eucalyptus* in South Australia and tested performances of ensemble model against individual models; they used mean and weighted average as ensemble strategies. They also tested cross validation versus spatial cross validation for model performances. The study doesn't specify which type of distribution was modelled: according to the definition provided in our study, we can compare their results with our potential distribution results. Their results show how spatial cross validation performances were more conservative than non spatial cross validation ones when compared with performances on independent validation sets. This supports and reinforces our use of spatial cross validation as a validation strategy for the modeling framework. Ensemble models performed well but were outperformed by untuned individual models and by a tuned GBT. There was also no clear advantage in predictive performances when using different ensemble strategies. This is in contrast with our results, where the ensemble based on stacking outperformed even tuned component models (13 cases out of 16), performed as good as the best component model (2 out 16) and in just one case performed worse (tuned GBT was better than the ensemble).

However, this is true only when comparing results from the Logloss and $R^2_{logloss}$: AUC and TSS both show that the ensemble outperformed or performed as good as tuned GBT in all cases and never performed worse. Given the very few occurrences in which the ensemble performed worse, this may be an indication of stacking being a better ensemble strategy when modeling species distribution. Valavi et al. (2021) reported an ensemble of tuned individual models as outperforming all other ML and regression based algorithms when benchmarking model performances on potential distribution of 225 different species; their results also show nonparametric technques outperforming traditional regression methods. Among SDM studies focused on testing and comparing different SDM methodologies, the study from (Valavi et al., 2021) is also one of the few reporting computation time for all the models: this is a metric seldomly reported, but relevant when considering the optimal trade-off between accuracy and time, a well-known issue in the ML field (Hosseinzadeh et al., 2021).

In our cross validation estimates, AUC proved the least useful of the performance metrics used in this study, with low variability in AUC scores among different species and distributions despite the difference in predictor variables used and amount and source (i.e. LUCAS, other tree species) of absence data. A general trend shown by the other metrics is mostly picked up (i.e. ensemble superior to base models, models for *Q. suber* being the most accurate), but AUC scores are too similar to ascertain critical problems or possible artifacts in our models. Both TSS and the $R^2_{logloss}$ provided more useful metrics, showing models for *P. sylvestris* and *P. nigra* performing poorly compared to other models in both potential and realized distribution. Our results seem to agree with the ones from Chakraborty et al. (2021), who predicted current and future potential distribution of tree species over Europe using an ensemble framework based on averaging. We compared our results only with species present in both studies (*A. alba, F. sylvatica, P. abies, P. sylvestris, Q. robur*): final model AUC values are all $\geq$0.94 on both test set and external validation set, while TSS values start from 0.80.

RF and GLM are the best component models to map both potential and realized distributions when trained on a data sample, but GBT often outperforms RF or even the ensemble when tuned and trained on the whole dataset. In general, differences in predictive performances between the ensemble and the component models are also higher in potential distribution than in realized distribution. The list of variable importance per component model, species and task may give an insight to this: in the potential tasks, the component models use different parts of the feature space before the predictions are combined by the meta-learner. All the models select as most important variables for the task different predictors. For realized distribution tasks, the models all agree in selecting either Landsat bands or spectral indices as most important variables, resulting in predictions that are highly correlated and with less variance between the models.

Ensemble modeling is known to perform best when there is a high diversity between the base models and no or negative correlation between their outputs (Zhou, 2019). The introduction of Landsat bands and spectral indices in general greatly increased the predictive performances of the models for realized distribution compared to potential distribution models. However, this also homogenized predictions, which makes the second condition reported above not always respected. We separately compared the repeated spatial cross validation performances of ensemble and component models excluding the Landsat bands and spectral indices. In this case, the ensemble never performed worse than the best component model. In general, if the ensemble provides predictive performances as good as or worse than the best component model, the best component model must be preferred (Zhang and Ma, 2012). However, ensemble models can still provide more advantages than individual models since they reduce model uncertainty and are more robust towards extrapolation (Mehra et al., 2019).

High resolution or hyperspectral data have not been used so far for SDMs but are extremely popular in tree species classification: the usage of such predictors has consistently increased over the years the predictive performances of ML tree species classifiers (Deur et al., 2020). Such data is not always available on a large spatial and temporal scale, so studies including these predictors usually cover a limited area compared to the one covered by our study. Bridging this gap may help having operational continental scale species distribution maps. Similarly, ML methods have mostly been used for tree species classification, where predictor variables such as temperature or precipitation are seldomly included and environmental variables not throughly considered, rather than for SDM. Despite that, we found in literature several studies which agree with our results: when classifying five (three broadleaves and two conifers) forest tree species in Portugal, Łoś et al. (2021) found out that GBT outperformed RF and KNN, reaching accuracy values $\geq$90% using Sentinel-2 reflectance bands only. Wessel et al. (2018) similarly reached high level of accuracy using only Sentinel-2 bands in German forests, but their best results were achieved using an object-based multitemporal Support Vector Machine (SVM) classifier: despite SVM being a very powerful ML method, it was was purposefully excluded from this study due to its computation intensity and lack of parallelization.

In our study, the ANN tested performed poorly, mostly due to the limited implementation options in the R environment of this method. Contrary to our results, Raczko and Zagajewski (2017) found ANN outperforming RF and SVM when including hyperspectral data; however, they also showed that ANN seemed to be the algorithm which predictions were strictly dependent from the dataset used, while RF and SVM showed more stable performances: the extreme sensibility of ANN to perturbations is a well known issue in the ML field (Colbrook et al., 2022). This issue is partially solved by Convolutional Neural Networks (CNNs), which have achieved considerable results when applied for SDM purposes, even when compared with ML methods such as GBT and RF. CNNs also showed to be particularly promising when commonly used remotely sensed predictor variables

such as LiDAR (Light Detection and Ranging) and hyperspectral high resolution data are available (Fricker et al., 2019; Zhang et al., 2020). In some cases, CNNs have even outperformed tuned ML methods given their ability to grasp how local landscape structure affects prediction of species occurrence, in contrast with more conventional ML methods which cannot acknowledge the influence of environmental structure in local landscapes (Deneu et al., 2021; Sothe et al., 2020). The ML framework presented in this study could greatly benefit from the inclusion of CNNs.

3.4.2 Species distributions

Our cross-validation accuracy assessment results indicate high predictive performances for all species, in both potential and realized distributions. In the case of mapping potential distribution, diffuse irradiation and precipitation of the driest quarter (BIO17) are the most important predictors overall. These results are partially in contrast with Dyderski et al. (2018), who modelled current and future potential distribution of 12 tree species over Europe. We compared our results only with species present in both studies (*A. alba, F. sylvatica, P. abies, P. sylvestris, Q. robur*): in their case, temperature-related bioclimatic variables (BIO1, BIO5, BIO7 and BIO10) were more important than precipitation-related bioclimatic variables. Few peer-reviewed studies have reported on the importance of predictors other than bioclimatic ones in shaping species' potential distributions. We found that, on average, each component model considers two or more predictors from the Bioclim macro-class among the top-10 most important variables to predict the potential distribution. Previous findings in literature have shown the importance of bioclimatic variables when modeling species distributions (Fourcade et al., 2018), but this may also be a consequence of bioclimatic variables and elevation being the most employed, if not the only, predictors in numerous SDM studies (Fois et al., 2018). Bucklin et al. (2015) compared the influence of different sets of environmental predictors on model performances, but the list of predictors used in the study included human influenced factors, so their results cannot be used to assess the driving factors for potential distributions. Even if our results show the bioclimatic variables as the most important predictors for potential distributions, further studies in this direction may be needed. The scale of the study may affect the importance of predictor variables: on a large scale, distribution may be influenced by macro environmental factors, while at a local scale, other environmental factors may limit distribution more significantly. Walthert and Meier (2017) and Weigel et al. (2019) proved that soil properties are more important than either bioclimatic or only climatic variables when modeling potential tree species distribution at, respectively, country and regional scale.

Variable importance confirms that Earth Observation layers such as the 25th, 50th and 75th quantile summer aggregates for the Landsat green and red band and the 50th quantile fall aggregates of NDVI are overall the most important layers for mapping realized distribution of species. The inclusion of Landsat data and derived spectral indices increases

predictive performances and contains more detailed information on species distribution ranges. Importance of NDVI is well known since it is one of the most used proxies in vegetation studies such as biodiversity estimation (He et al., 2009; Madonsela et al., 2017), net primary productivity (Schloss et al., 1999) and land degradation (Easdale et al., 2018), phenology (Fawcett et al., 2021) and species composition changes (Wang et al., 2021). NDVI incorporates information from the red and the near-infrared (NIR) portion of the electromagnetic spectrum. Vegetation's behavior in this portion of the spectrum has long been used in vegetation mapping to distinguish between coniferous and deciduous tree species (Hoffer, 1984). The green band, although usually less important than the red and NIR band, has already proved useful in vegetation mapping to classify forest types (Gao et al., 2015), predict forest variables (stem volume, diameter and tree height) at species level (Astola et al., 2019) and forest biomass at community level (Nandy et al., 2017).

Comparing our results with chorological maps from the European Atlas of Forest Tree Species (San-Miguel-Ayanz et al., 2016), we can see that in general both potential and realized distribution correctly capture the species ranges. Overall, potential distribution maps show homogeneous patterns of high probability values for all target species, while realized distribution maps show very fragmented patterns. The realized distribution model helps discriminating the presence or absence of the species due to biotic or other external factors. A high geographical overlap between probability maps of realized distribution of different species may reflect co-existence within the same forest stands and could help in clearly define forest communities.

However, our results have to be interpreted and analyzed carefully: contrary to process based models, correlative models describe the patterns, not the mechanisms, in the association between species occurrences and predictor variables; for this reason, correlative SDMs risk overlooking potentially important driving factors determining species distributions since they cannot distinguish between direct and indirect effects (Sirén et al., 2022). A known issue in this sense is the masking effect of abiotic factors on competition and predation: SDMs could estimate abiotic predictors as the most important for species abundance, even in those cases when distribution is strongly affected by competition (Godsoe et al., 2017) or when biotic interactions are strictly correlated with abiotic factors (Filazzola et al., 2020).

ML methods strongly depend from high quality datasets, so a considerable effort was spent in creating two different presence-absence datasets for each target species, one for potential and one for realized distribution: while the same geographical extent was used for both datasets, different rules were used to select true absence (LUCAS dataset) or pseudo-absence (other tree species occurrences) points. This modeling choice may be one of the causes of cross validation estimates for potential distribution being higher than for realized distribution. Restricting the study area from which true and pseudo

absence points are collected reduces the applicability of the models for predictive purposes (Pearson and Dawson, 2003), with unpredictable effects on future projections (over- or under-prediction, see Thuiller et al. (2004)). On the other end, no spatial constraint leads to unwanted situations where larger scale differences rather than local ones are picked, leading to the infamous SDM case of "there-are-no-polar-bears-in-the-Sahara" (Lobo et al., 2010). Chefaoui and Lobo (2008) proved how best hypotheses for potential distribution are obtained using absence points that are placed farther apart than the ones needed for the best hypotheses on realized distribution, hence using the same study area for both distributions may have lead to an overstimation of presence in potential distribution. Furthermore, Lobo et al. (2010) proved that for realized distribution best practice would be to avoid the absences from nearest localities due to possible contamination with methodological absences; while this was not an issue when considering absences coming from the LUCAS dataset, using information of other tree species occurrences as pseudo-absence may have affected our models. The criteria used to select pseudo-absence occurrences in SDM represent a big challenge in SDM, such that the topic has been the focus of multiple studies in the last decade (Iturbide et al., 2018a; Iturbide et al., 2018b; Senay et al., 2013).

3.4.3 High resolution contributions: is finer always better?

Bioclimatic variables available only at coarse spatial resolution were used as predictor variables in both potential and realized distribution. The Landsat bands and the spectral indices were not the only high resolution layers used in this study: terrain and terrain-derived predictors were also included at 30 m resolution. However, despite the terrain data high resolution, the tree species potential distribution patterns mostly reflect the original spatial resolution of the bioclimatic variables. Thus, climate influences species distribution at the European scale. Even though this might indicate that mapping potential distributions at high resolution may not be necessary, it can still be useful for different case studies. For example, comparing the difference, and hence mapping the gap, between potential and realized distribution at the same fine scale may prove to be an invaluable tool for both forest managers and conservation planners that work on the local level.

Potential distribution maps can be used to identify suitable areas for species in reforestation and restoration programs; realized distribution maps can inform the forest managers on the presence or absence of said species in those areas at a particular point in space and time (Fig. 3.7). By removing the biotic factors that limit the presence of the species in a potential reforestation site, using multiple distribution maps and including expert knowledge on species synecology, structurally complex forest stands could be planned and developed in a much more informed and data-driven way. A similar approach could be used by conservation planners. Potential distribution is modelled by studying the relationship between a species and the environmental conditions found in its native range, where the species is at equilibrium (Jiménez-Valverde et al., 2011). Invasive species are usually more abundant and productive in the introduced range than in their native ranges (Hierro et al.,

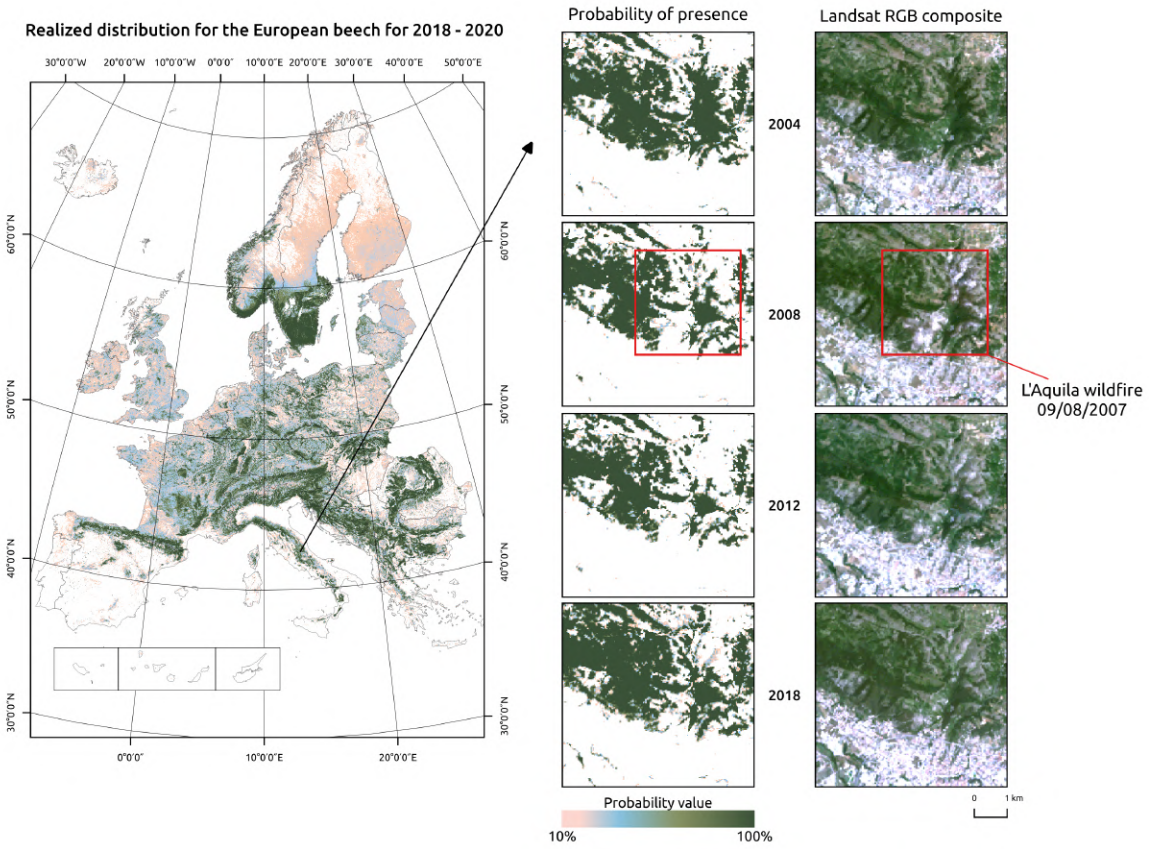

Figure 3.7: Realized distribution of *Fagus sylvatica* for the period 2018-2020. Detailed insets show a region around L'Aquila city, in Central Italy. The *Fagus sylvatica* forest on the northern outskirts of the city was affected by a serious wildfire in 2007. The realized distribution maps can be used to track compositional changes through time.

2005). This is due to the absence of biotic factors that normally limit species distribution in their native range in the introduced range. Thus, a species that occupies only 10% of its potential distribution in its native range may end up occupying a bigger percentage of it in the introduced range. Estimation of potential distribution in the introduced range that depends only on environmental factors are conservative by definition, potential distribution maps may provide a good indication to conservation planners of how much the invasive species could spread in the introduced range.

For realized distribution, including high resolution predictor variables in the model not only increases predictive performances but also lowers overall and local values of uncertainty. For forest management purposes, a large, consistent, standardized, long-term and high resolution image collection such as the one provided by the Landsat program can help extending in space and time information on tree species presence, composition and abundance. A spatial resolution of 30 m is particularly well suited for NFI applications: Strickland et al. (2020) derived probability maps of forest tree species for a 25 years time period (1985–2010) using yearly Landsat composites to extend missing information from

the Canadian NFI and estimating changes in forest cover, species composition and forest disturbances. The increasing availability of even higher-spatial resolution satellite data from the European Copernicus program (i.e. Sentinel 1 and 2) and commercial providers (i.e. Planet) can potentially further enhance predictions by including more data and a better spatial matching of in-situ and satellite-derived information.

3.5 Conclusions

In this paper we have developed, tested and reported a methodological framework for predicting potential and realized distributions of 16 forest tree species using ensemble ML and analysis-ready EO data. In general, our ensemble models achieved better predictive performances than individual models when modeling both potential and realized distribution, while performing as good as the best individual model in the worst cases. Bioclimatic variables proved in general to be the most important and frequent predictors for potential distribution across Europe, mainly through precipitation-related predictors (BIO17 and BIO14) even at high resolution (i.e. on a local scale), while reflectance-based covariates were the most important predictors of the realized distributions. Overall, realized distribution proved to be more complex to map accurately than potential distribution and, among the species analyzed, distributions of specialist species proved easier to classify than pioneer species. In general, the ensemble and component models achieved better predictive performances for the potential distributions than for the realized distributions as judged from the cross-validation estimates. Our results indicate a consistent increase in predictive performances for realized distribution when adding high resolution data, especially Landsat data at 30 m resolution and spectral indices to the list of predictors. Significant findings of our work include: (**a**) distribution mapping for forest tree species can be efficiently automated to the level of full automation, but this assumes high quality / artifact free training points with a homogenous distribution of occurrence and absence points whenever possible; (**b**) complexity of ML methods can be significantly reduced by implementing efficient hyperparameter tuning and feature selection; (**c**) analysis-ready, high resolution reflectance time-series layers are maybe cumbersome to prepare and gap-fill for clouds and artifacts, but overall come as the most important inputs for maximizing predictive performances of realized tree species distribution.

We have released the maps and the code under open data / open source licenses to enable other similar research and to help speed up land restoration and reforestation projects in Europe. The code is publicly available in our GitLab repository at `https://gitlab.com/geoharmonizer_inea/spatial-layers/-/tree/master/veg_mapping`, while the datasets and predictions of tree species are available as Cloud-Optimized GeoTIFFs on Zenodo (see `https://doi.org/10.5281/zenodo.5818021`) and can be dis-

played in 2D and 3D using the compare tool on the EcoDataCube.eu viewer (see Fig. 3.8).

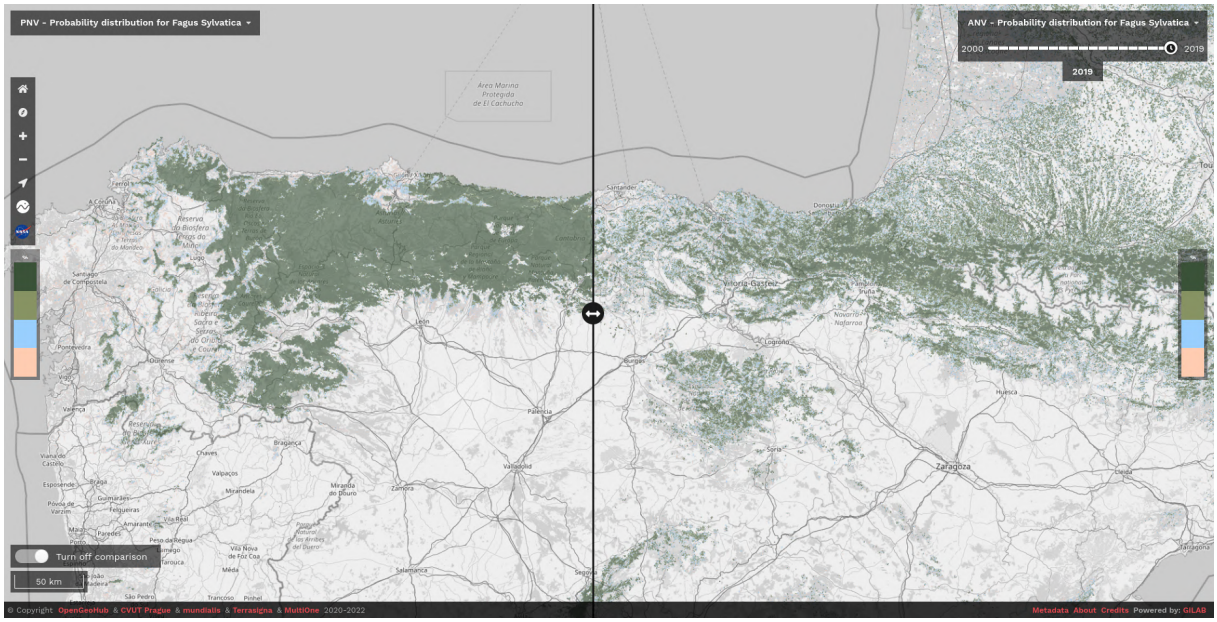

Figure 3.8: Difference between potential and realized distribution for *Fagus sylvatica* in Northern Spain for the period 2018–2020 visualized using slider in the EcoDataCube.eu viewer (`https://ecodatacube.eu`). ©Copyright OpenGeoHub & CVUT Prague & mundialis & Terrasigna & MultiOne 2020–2022.

Even though we achieved high values of predictive performances, we still recognize many future areas of improvements. Given the importance of Landsat data for the results of this study, using a larger and higher resolution stack of reflectance-based predictors could help to improve precision of the predictions. A good example in this direction would be fusing all EO data currently available such as Harmonized Landsat Sentinel-2 (HLS) (Claverie et al., 2018a) and eventually Sentinel 1 datasets. Hyper-spectral images (i.e. from future hyper-spectral missions such as ENMAP, see `https://www.enmap.org/`) are also proving to be useful to discriminate between different tree species and especially those that grow under dominant species (Fricker et al., 2019; Shen and Cao, 2017). As any ML-derived product, our predictions would benefit from having more and better quality data on tree species, in particular those that come from NFI plots: it is now crucial to have such data freely available to monitor processes such as species compositional changes, niche shifts, forest regrowth and degradation, as recently stated by Nabuurs et al. (2022). Exploring more sophisticated and different ML algorithms such as Deep Learning (DL) techniques (Lakshminarayanan et al., 2016) to our ensemble framework is also another area of improvement given the wide variety of applications these methods possess and the results obtained in comparison with other conventional ML algorithms (Anand et al., 2021; Choe et al., 2021; Deneu et al., 2021).

European forest dynamics, even though some recent results indicate increased mortality in European forests (Popkin, 2021; Senf et al., 2018; Senf et al., 2021), are probably among the least troubling in comparison to other continents. Our methodological framework could potentially be implemented at a global scale, and possibly through Google Earth Engine (GEE) (Hoogen et al., 2021) or through the European Space Agency's OpenEO platform (`https://openeo.cloud/`) to produce high resolution (10–30 m) predictions of forest dynamics. Globally, there are many more tree species which are more important for forest management and monitoring. For example, South America as a whole has 4 times the amount of tree species present in Europe and 50% of all tree species on Earth (Cazzolla Gatti et al., 2022); in Brazil, it has been estimated that about 220 tree species cover most of the land and represent over 95% of the biomass (i.e. so called *"hyper-dominant species"* (Draper et al., 2021)). Scaling up the approach described in this paper to help producing objective predictions, to assist with monitoring forest dynamics and to support reforestation efforts globally is part of our next objectives.

Chapter 4

Current trends in European forest tree species distribution shifts

This chapter is based on:

C. Bonannella, L. Parente, S. de Bruin, and M. Herold (2024a). "Multi-decadal trend analysis and forest disturbance assessment of European tree species: concerning signs of a subtle shift". *Forest Ecology and Management* 554, 121652. DOI: https://doi.org/10.1016/j.foreco.2023.121652

Abstract

Climate change poses a significant threat to the distribution and composition of forest tree species worldwide. European forest tree species' range is expected to shift to cope with the increasing frequency and intensity of extreme weather events, pests and diseases caused by climate change. Despite numerous regional studies, a continental scale assessment of current changes in species distributions in Europe is missing due to the difficult task of modeling a species realized distribution and to quantify the influence of forest disturbances on each species. In this study we conducted a trend analysis on the realized distribution of 6 main European forest tree species (*Abies alba* Mill., *Fagus sylvatica* L., *Picea abies* L. H. Karst., *Pinus nigra* J. F. Arnold, *Pinus sylvestris* L. and *Quercus robur* L.) to capture and map the prevalent trends in probability of occurrence for the period 2000–2020. We also analyzed the impact of forest disturbances on each species' range and identified the dominant disturbance drivers. Our results revealed an overall trend of stability in species' distributions (85% of the pixels are considered stable by 2020 for all species) but we also identified some hot spots characterized by negative trends in probability of occurrence, mostly at the edges of each species' latitudinal range. Additionally, we identified a steady increase in disturbance events in each species' range by disturbance (affected range doubled by 2020, from 3.5% to 7% on average) and highlighted species-specific responses to forest disturbance drivers such as wind and fire. Overall, our study provides insights into distribution trends and disturbance patterns for the main European forest tree species. The identification of range shifts and the intensifying impacts of disturbances call for proactive conservation efforts and long-term planning to ensure the resilience and sustainability of European forests.

4.1 Introduction

Tree species have repeatedly demonstrated remarked adaptability over time, changing their geographical distributions in response to large scale fluctuations in the climate (Cheddadi et al., 2016; San-Miguel-Ayanz et al., 2016; Svenning et al., 2010; Svenning and Skov, 2007). Over centuries up to millennia, these adjustments have highlighted the delicate balance between species survival and the environmental conditions that either support or negate their existence. Owing to their long generation time, trees are however particularly sensitive to climate change. The current, unprecedented rapid climate change has the potential to drive major shifts in tree species distributions (Hanberry and Hansen, 2015; IPCC, 2022) to the point that it could alter the global location of entire biomes (Berner and Goetz, 2022; **Bonannella** et al., 2023; Gonzalez et al., 2010).

Two decades of studies have significantly improved the detection and attribution of biome shifts to climate change (Higgins et al., 2023; Lindner et al., 2010). Numerous studies address the expected shifts in distributions of tree species using Species Distribution Models (SDMs) (Franklin, 2010), which map the ecological niche of the species in space under current conditions, and then include different climatic scenarios projections to predict the geographical boundaries of the niche in the future (Dyderski et al., 2017; Mauri et al., 2022; Thuiller et al., 2008; Zhang et al., 2017). There is general consensus regarding the response of species' ranges to climate change, pointing towards poleward (Berner and Goetz, 2022; Zhu et al., 2012; Zhu et al., 2014) and upward (Feeley et al., 2011) shifts in distribution ranges. Studies using tools like SDMs have mainly focused on predicting potential species distributions. In contrast, modeling the realized distribution is an inherently complex task and extrapolating this model into the future is even more challenging. Despite its complexity, modeling the realized distribution in current conditions may still be feasible, for example using Earth Observation (EO) data, but there are only few examples of such exercises (**Bonannella** et al., 2022d; Gelfand and Shirota, 2021; Hefley and Hooten, 2016). Compared to previous studies focusing on long term future predictions using climate models, using EO data based models allows for timely and consistent assessments of ongoing changes (Strickland et al., 2020): EO data does not only provide higher temporal resolution compared to, for example, periodic surveys such as National Forest Inventories (NFIs) programs, but also offers a broader spatial coverage, including remote or inaccessible regions which are difficult to monitor from the ground. EO data also offers a compelling advantage to extensive field surveys in terms of cost efficiency: accessing and processing EO data is nowadays remarkably streamlined due to the free availability of large collections of data, like the Landsat mission (Zhu et al., 2019), or tools able to analyze such large collections like the cloud computing platform Google Earth Engine (Gorelick et al., 2017). While field campaigns like NFI programs have been used to detail current ongoing changes in species distributions at regional or national scale (Ewald, 2012; Rigling et al., 2013; Scherrer et al., 2022), the decadal temporal resolution,

the limited geographic extent (national at most) and access (only few users can actually harvest all the useful information these collections of plot data offer) of these surveys makes their use for tree species shifts assessment challenging (Nabuurs et al., 2022). Due to these reasons, a comprehensive understanding of changes in tree species distributions trends at large (i.e. continental) scale in the face of climate change is still lacking.

Particularly European forests have recently faced a barrage of severe threats as an effect of previous forest management decisions and climate change (Senf and Seidl, 2021a). These disturbances have increased consistently over the last century (Patacca et al., 2023; Schelhaas et al., 2003; Seidl et al., 2011) with the most pronounced intensification occurring in the last two decades (Senf et al., 2018). This intensification has caused extensive tree loss and drastic changes in the structure and dynamics of European forests (Maes et al., 2023; Popkin, 2021; Senf and Seidl, 2021a). Several studies identified positive correlations between tree species diversity and abundance and the provision of ecosystem services (Brockerhoff et al., 2017; Gamfeldt et al., 2013; Himes et al., 2020). Hence, their decline not only negatively affects the environment but human societies as well (FAO, 2022; FOREST EUROPE, 2020). Given the importance of the topic, numerous studies have analyzed European forest disturbances qualitatively (Forzieri et al., 2021; Seidl et al., 2017; Sommerfeld et al., 2018), quantitatively (Patacca et al., 2023) and spatially (Senf and Seidl, 2021a; Senf and Seidl, 2021b). While some researchers focused on dissecting specific disturbance events, seeking to unravel the causes or the impacts of individual storms (Chirici et al., 2019; Kronauer, 2007), fires (Ganteaume et al., 2013; San-Miguel-Ayanz et al., 2012) or pests (Hlásny et al., 2021b), others analyzed the species-specific effects of these disturbances. Schmidt et al. (2010) tried to estimate the risk of storm damage for different groups of species, spruce and Scots pine in particular in Southwestern Germany; Hlásny et al. (2021a) and Kautz et al. (2023) analyzed drivers and symptoms of European spruce bark beetle outbreaks on the spruce in Central Europe and Moris et al. (2023) tested the resilience of beech stands to fire in Northern Italy. These localized or context-specific studies offer valuable insights into the species' behavior in response to disturbances, especially if the study area is at the edges of the species range: in this case, they can provide valuable knowledge in understanding what the species range response could be in the future, if it will advance or recede. While insightful, these studies primarily concerned localized (regional or national) contexts; in contrast, the pivotal aspect of our study is to explore disturbances at a continental level, matching the scale of European forest tree species' ranges. Our understanding of disturbances on species' ranges at this scale remains currently uncharted territory.

Considering this knowledge gap and the relevance of European forests, this paper characterizes changes in the distribution of key European forest tree species and the contribution of forest disturbance events to species' ranges for the period 2000–2020. Our research addressed two specific research questions: (a) are there observable range shifts in European forest tree species? (b) To what extent do forest disturbance events impact different

tree species? To answer these we conducted a trend analysis on high spatial resolution (30 m) time series maps of six main European forest tree species realized distribution. We next aggregated the results of the trend analysis to coarser resolution (1 km) maps to capture the prevalent trends. Furthermore, we used available high resolution forest disturbance maps detailing both the year and type of disturbance to quantify how much of each species' range was affected by disturbances and identify the dominant disturbance types. By following this methodology, our analysis provides spatial evidence of tree species distributions dynamics, uncovers potential shifting patterns and provides species-specific insights on disturbance regimes.

4.2 Material and methods

4.2.1 Tree species probability of occurrence datasets

We used the realized distribution maps produced by **Bonannella** et al. (2022d) to quantify the probability of occurrence over the twenty years period of analysis. In their study, they used a combination of EO data and Machine Learning (ML) to produce a time series for the period 2000–2020 covering the whole European continent for 16 forest tree species at 30 m resolution. To account for the multiple conditions that affect the realized distribution, the predictor variables used in the model include well known static (i.e. not changing over time) variables such as long term climatic and topographical data, but also dynamic (i.e. time series) variables such as multispectral data coming from Landsat. A novelty of their approach was to use chorological maps of different tree species as well to account for mutualism/competition effects. The presence–absence point dataset used for model training was published separately and is available on Zenodo (`https://doi.org/10.5281/zenodo.6516590`): presence data comes from a harmonized and preprocessed version of existing tree species data from the Global Biodiversity Information Facility (GBIF), the EU-Forest project (Mauri et al., 2017) and the LUCAS survey (EUROSTAT, 2017); absence data comes instead exclusively from the LUCAS survey. The authors then tested and combined the predictions of several ML algorithms using an ensemble approach known as stacked generalization (Wolpert, 1992) to deliver the final predictions. The probability of occurrence (0–100%) is provided for each species individually, irrespective of co-occurrence, meaning that the sum of the probabilities of different species in the same pixel can exceed 100%. The 2000–2020 time period is split in time windows of different extent: 2000–2002, 2002–2006, 2006–2010, 2010–2014, 2014–2018 and 2018–2020, for a total of six observations in time for each species for the period of analysis. The reason of these uneven time windows is due to design decision (i.e. a time window large enough to see changes in realized distribution over time and short enough to have multiple points in the period 2000–2020) and data constraints and irregularities (i.e. the SLC failure from Landsat 7 (Zhang et al., 2007)). The data is available on Zenodo, with a different entry for each species (example for *Abies alba*

Mill. at: https://doi.org/10.5281/zenodo.5818021), and in European Environmental Data Cube (Witjes et al., 2023) at https://ecodatacube.eu. For the purposes of this paper we focused our analysis on only six of the tree species from **Bonannella** et al. (2022d): *Abies alba* Mill. (silver fir), *Fagus sylvatica* L. (European beech), *Picea abies* L. H. Karst. (Norway spruce), *Pinus nigra* J. F. Arnold (black pine), *Pinus sylvestris* L. (Scots pine) and *Quercus robur* L. (Common oak). The species were chosen because of their common occurrence (Barbati et al., 2014) in European forest, their economic importance (Hanewinkel et al., 2013) and the fact that they represent more than 70% of the growing stock volume in the region (FOREST EUROPE, 2020).

4.2.2 Forest disturbances datasets

We used a combination of two datasets to collect data on the the year of disturbance (Senf and Seidl, 2021a) and type of disturbance (Senf and Seidl, 2021b). Both datasets have a pan-European extent, 30 m spatial resolution and cover the time period 1986–2020. However, they do not cover all European countries (i.e., Iceland is missing); our analysis was therefore limited to the countries covered by the disturbance datasets. Firstly, the forest disturbances map only lists the year of the greatest disturbance over the whole 1986–2020 time series, regardless of the disturbance agent. We used the latest version of this product, available on Zenodo at https://zenodo.org/record/7080016 (v. 1.1.4). A combination of reference data from Landsat satellite images of visually interpreted disturbed and non disturbed patches (available on Zenodo at: https://zenodo.org/doi/10.5281/zenodo.3561924) and outputs from the well known LandTrendr algorithm (Kennedy et al., 2010) is used to train a random forest classifier; the model classifies each pixel in non-forest, undisturbed forest or disturbed forest. Information on the year of disturbance is provided from the trend analysis of spectral bands and indices.

Secondly, the forest disturbances type map attributes to each disturbed pixel of the forest disturbances map a code based on type of event detected, as described in Senf and Seidl (2021b). We used the latest version of this product, available on Zenodo at https://zenodo.org/record/8202241 (v. 1.2). Information on forest disturbance agents was gathered from the FORWIND database for storms (Forzieri et al., 2020) and the European forest fire information system (EFFIS) for fire related disturbances (San-Miguel-Ayanz et al., 2012); areas classified as disturbed but with low probability for wind or fire disturbances were assigned to the "other" class. A set of ten predictors (four spectral, three spatial and three landscape-based ones) was used to model the probability of a disturbed pixel to be affected by storm, fire or neither through a random forest model. Probabilities were then converted into hard classes in the final version of the disturbance maps, with the following legend: "0" for no disturbances, "1" for other disturbances (i.e. logging, drought, biotic disturbances such as bark beetles etc), "2" for wind-related disturbances and "3" for fire-related disturbances. In both cases, maps are provided on a country scale: for our analysis we first downloaded all the countries available and then merged them in

a single file with European-wide coverage matching the same spatial properties (extent, pixel size, coordinate reference system) used by **Bonannella** et al. (2022d) to produce the tree species probability of occurrence maps. We disregarded disturbances outside the time period of the analysis (prior year 2000).

4.2.3 Trend analysis and forest disturbances attribution

To analyze the change in probability of occurrence over time, we fitted simple OLS regression models with the probability of occurrence as the dependent variable and time as the independent variable. We applied this procedure at pixel level and for all six species individually. We then calculated the t-test statistic to determine the presence of an increasing or decreasing trend according to the OLS model. By combining the regression slope (β) of the linear model and the p-value from the t-statistics, we assigned each pixel to one of three categories:

- *positive*: $\beta > 0.25$ AND p-value < 0.05
- *negative*: $\beta < -0.25$ AND p-value < 0.05
- *no trend / stable*: $-0.25 \leq \beta \leq 0.25$ OR p-value > 0.05

The chosen threshold of ± 0.25 for the regression coefficient denotes a total change of 5% in the probability of occurrence over the 20 years period of the analysis, which is considered a substantial change. Pixels in the *stable* class have relatively constant values of probability of occurrence over the time period analyzed. Pixels in the *positive* class have an increase in probability of occurrence over the time of the analysis while the opposite is true for pixels in the *negative* class; *stable* pixels could also experience fluctuations in the probability of occurrence, alternating favorable and unfavorable conditions for the species, but ultimately the net effect of these events would balance out by the end of the time series. The focus of the analysis was on identifying the direction of change in the probability of occurrence for each pixel and species, thereby revealing overall patterns in species distribution changes rather than quantifying the magnitude of the trend.

To better identify and analyze these patterns, we prepared species-specific 1 km resolution layers displaying the proportion (0-100%) of *positive*, *stable* and *negative* pixels within each cell: this aggregation of the 30 m results at 1 km allowed us to filter out local mature stand changes or other local variations, distinguishing genuine signals from isolated ones or false positives. Each species was modelled independently in **Bonannella** et al. (2022d), so model accuracy varies between different species; the aggregation at 1 km helps in removing false signals caused by this variability. In addition, in areas undergoing intense forest management, due to clearcuts or relatively aggressive thinning operations, 30 m pixels may display very diverse spatial patterns, with a pixel having a strong positive slope ($\beta > 1$) next to a pixel with strong negative slope ($\beta < -1$) (Fig. 4.1). While from a trend analysis perspective those isolated pixels may provide information on, respectively, the

species advancing or retreating, the spatial context is instrumental in discerning weak signals from true niche shifts. The aggregation at 1 km filters those weak signals out and ensures us that the identified patterns are robust, providing a clearer understanding of the ecological dynamics at play.

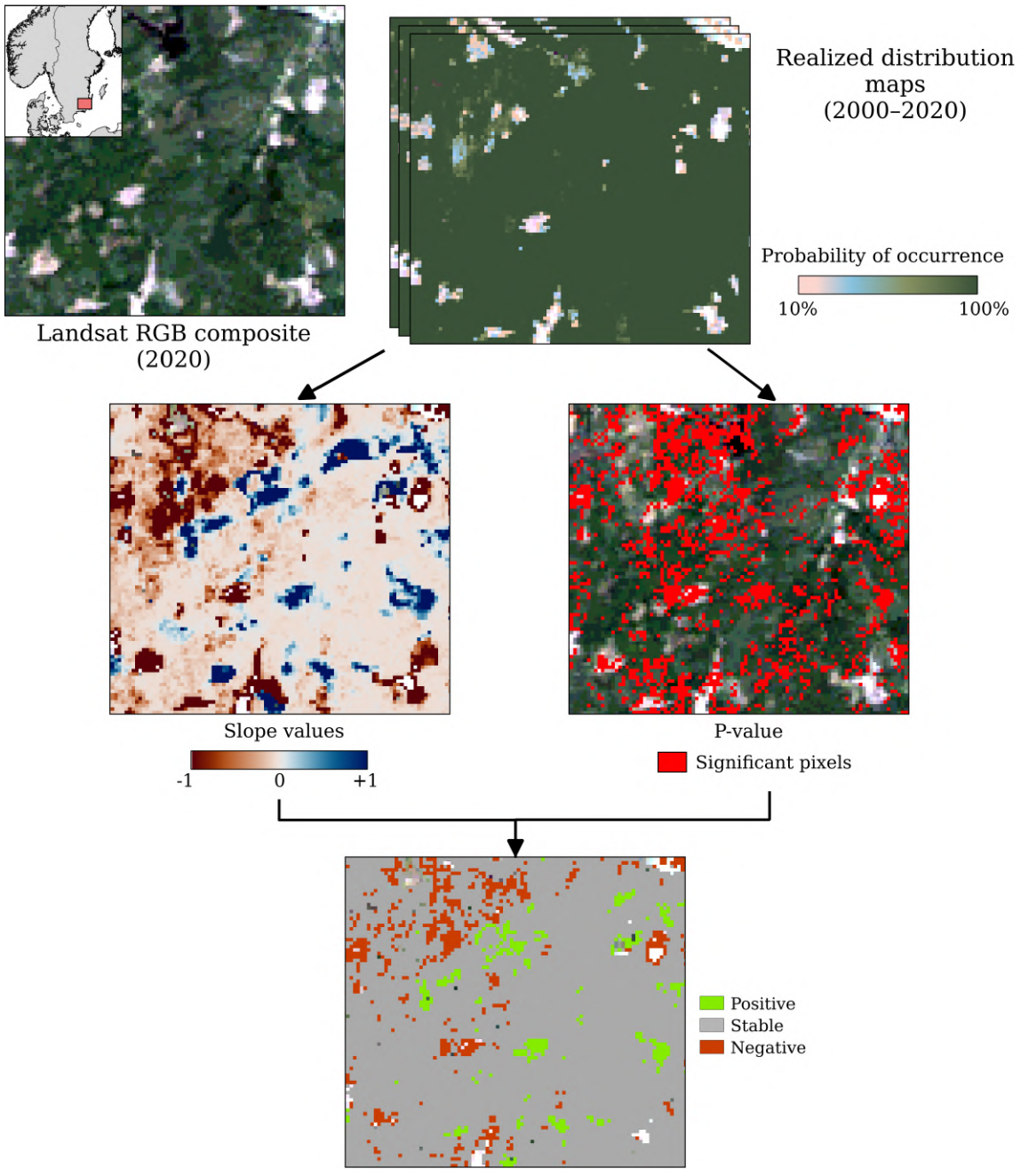

Figure 4.1: Example of the 30 m workflow for a production forest in Southern Sweden. Production forests are areas intensively managed: this creates local variations in probability of occurrence over time. Results of the trend analysis would then reflect this through chessboard-like spatial patterns in the slope values. The aggregation of the results at 1 km resolution is then necessary to correctly interpret the results of the analysis.

As a final step, following the trend analysis and the spatial quantification, we overlayed slope values of each species with the forest disturbances maps. This allowed us to differentiate between disturbed and non-disturbed pixels and, for the disturbed pixels only, assign the year of disturbance and disturbance type.

4.3 Results

4.3.1 Changes in probability of occurrence

For each species we visually inspected a histogram of the slope values. The modes of all histograms corresponds to zero slope (i.e., stable), as shown in Fig. 4.2. However, for four of the six species investigated, the modal histogram bin is very different from its direct neighbors. For these four species, the zero slope bin holds 40% to 50% of the total number of pixels, with beech having the greatest proportion, closely followed by spruce. Between 50% to 70% of the pixels are distributed in the "-0.1", "0", and "0.1" bins, indicating that the majority of the pixels is classified as *stable* and not having a clear increasing or decreasing trend. For the two pine species instead, the "0" bin holds less than 20% of the pixels and the frequency distributions are remarkably flatter, with the majority distributed over the bins left of the mode.

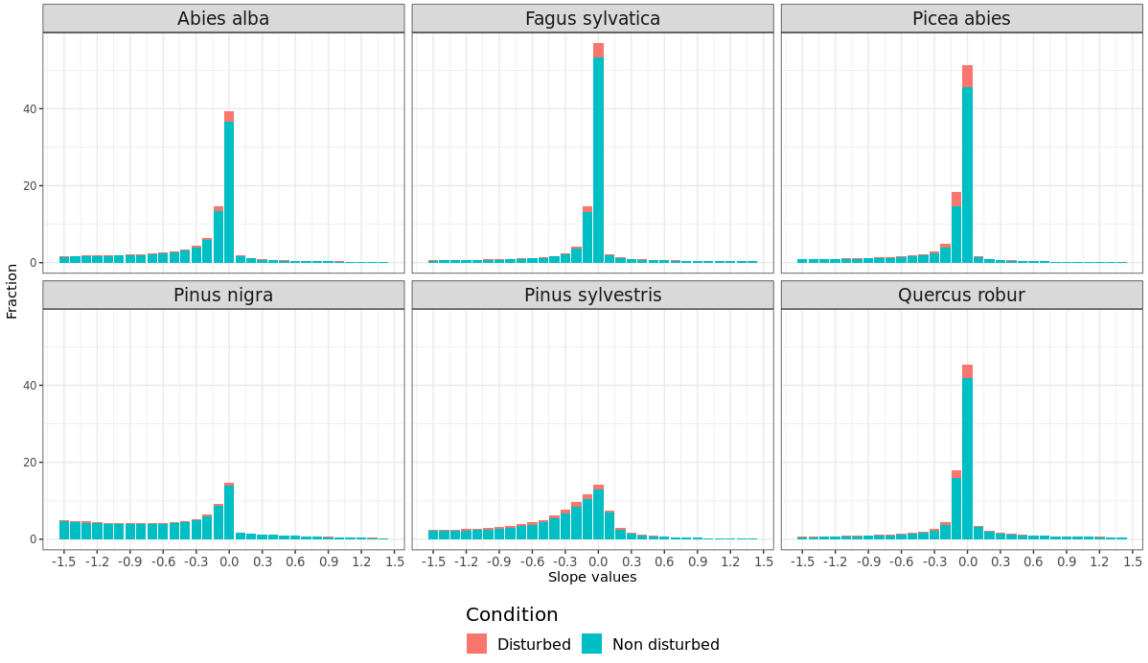

Figure 4.2: Distributions of slope pixels per species; given the significant difference in absolute number of pixels across the species analyzed, the values are normalized. The color of the stacked histogram bins represents the frequency of pixels affected or not by a disturbance event.

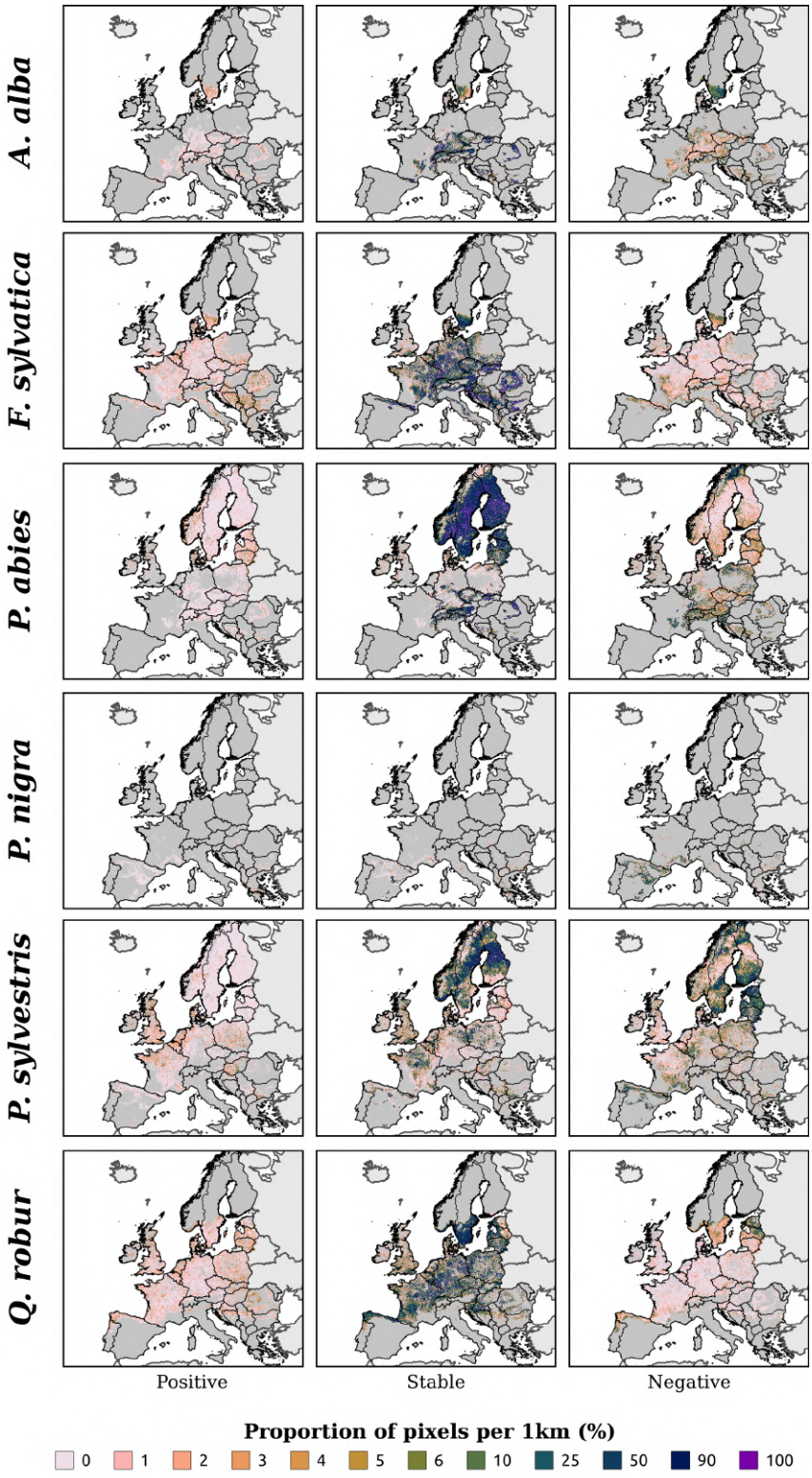

Figure 4.3: Proportions of pixels with, from left to right, increasing (*positive* class) trend in probability of occurrence, no trend (*stable* class), and decreasing (*negative* class) trend in probability of occurrence over 1×1 km blocks for all species. The dark grey background marks the countries included in the study area. Note the legend employs unequal interval widths to facilitate distinguishing varying proportions and comparing spatial patterns within or across species.

Fig. 4.3 depicts the spatial patterns of positive, negative or no trends (*stable* in the figure): the column representing pixels with stable probability of occurrence over time is the one that has a more evenly distributed presence throughout the study area. The mapped distributions (Fig. 4.3) confirm the general trends in Fig. 4.2 that despite a large fraction of stable situations, there are much more negative than positive trends. Furthermore, the pixels in the *stable* column that have high values form large clusters that cover the extent of one or more countries rather than being concentrated in specific localized regions. This suggests that the areas considered stable cover a large portion of the total realized distribution of each species compared to areas with majority of positive or negative trends, which exhibit more distinct and localized hot spots. While this is true for almost all the species, the spatial patterns for black pine match what was highlighted previously in the distribution of the slope values: when compared with the other species, black pine is the sole species that lacks hot spots for positive trends and has a much more dispersed distribution of pixels with prevalent negative trends than pixels with a prevalent stable probability of occurrence over time. Except for black pine, the hot spots for negative trends consistently exhibit a notable pattern characterized by their localization at the latitudinal edges of each species distribution.

Even though five of the six species share this pattern, there are some differences in the negative hot spots' spatial locations; with most of them being located in Central and Northern Europe. For the silver fir, beech and oak, the hot spots seem to coincide with the northeastern edge of their distribution in the study area. The proportion values vary between 25 and 60%. Spruce presents a consistent negative hot spot towards the northernmost edge of its distribution like the other species, but also exhibits additional hot spots in the southern edge and in other areas. Notably, as the range extends further north, these hot spots are mostly located in regions with relatively low (500–1000 m.a.s.l.) elevation for the species (i.e. Carpathians, Ardennes) or even at sea level (i.e. Poland coastline). Similar considerations can be made for Scots pine, where the negative hot spots are located in the north (Finland, Baltic states) or at the lowest elevations of places located in the southern edge of the range (Provence in France, hilly areas surrounding the Pyrenees, the Cantabrian and Leon mountains in Spain). Areas with prevalent negative trends exhibit in general moderate values (40-50%) but certain portions of the negative hot spots for spruce, Scots pine, and black pine stand out with very high values (95-100%). Notably, these three species show elevated values primarily in the southern regions such as Spain, France, Slovenia, and parts of Romania, while the northern regions' hot spots remain in the moderate range.

Regarding positive trends, no clear hot spots can be discerned: pixel values vary between very low (5-10%) to low (20-30%) proportions and there are only few exceptional cases where they reach moderate (40-50%) proportions. Most of the pixels with low to moderate values correspond to specific mountain ranges (Alps, Norwegian Alps and Carpathians). Other areas with pixels in that value range are more species-specific: the Dinaric Alps for

beech and oak, Scotland and Ireland for spruce and Scots pine, with Scots pine having some additional concentrations of positive trends in Galicia and Normandy.

4.3.2 Effects of forest disturbances

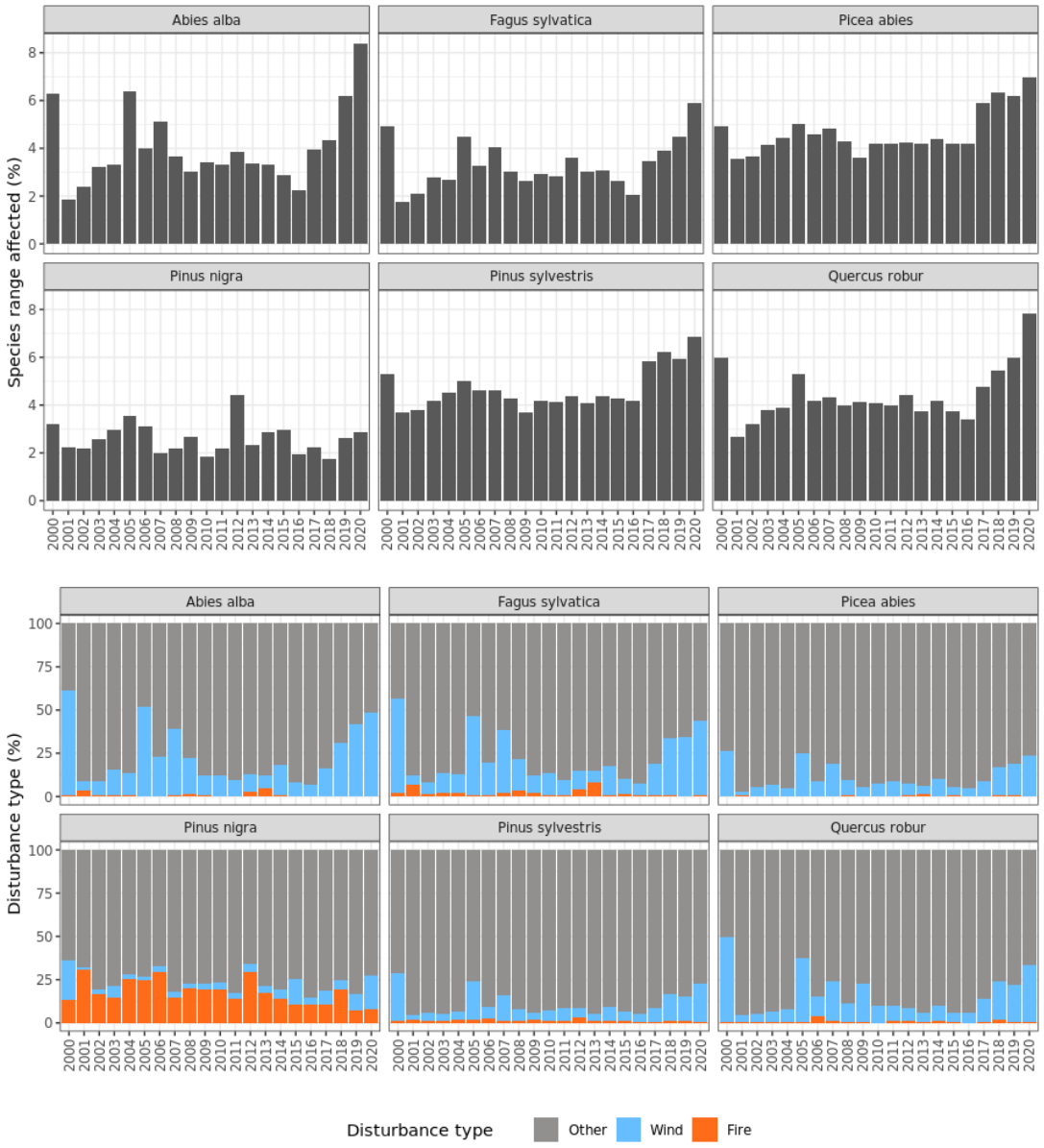

Figure 4.4: Species specific distribution of disturbances over time. Top: percentage of 30 m pixels affected by disturbances relative to the species total realized range. Bottom: disturbance types normalized to the total number of disturbed pixels.

In addition to analyzing the slope values distributions, we flagged pixels affected by disturbances and used color-coded bars in the histograms to have a visual representation of the disturbance status of the pixels (see Fig 4.2). Pixels affected by disturbances fall

into the "0", "-0.1" and "-0.2" histogram bins, with a prevalence of disturbed pixels in the "0" bin. The similarities stop there, since each slope distribution seems to behave differently and no consistent trend can be picked up. Beech and spruce have the greatest proportion of pixels in the "0" histogram bin. Both black pine and Scots pine have an equal proportion of disturbed pixels across all histogram bins, mostly left of the mode (until bin "-0.9"), while on the right it stops much earlier (bin "0.3").

The distribution of disturbed pixels over the years shows some common patterns among all the species (see Fig. 4.4 top); at least 2% of the total range of all the species is affected by disturbances every year, with most species having peaks in the year 2000, 2005 and 2020, which corresponds to a sudden increase of wind-related disturbance events. It is especially evident for the silver fir, where the percentage of pixels disturbed by wind goes from a value of 10-15% every year, up to more than 50% in the year 2020. There is an increasing trend of disturbed pixels over time, with all the species having on average 2% more of their range affected by disturbance events by the end of the time series. The *other* class is the most prevalent across all the years, reaching peaks of 90% in some years, with the exception of years affected by intense wind events where the percentage drops to 40-50% for some species. Among the coniferous species, spruce and Scots pine are the ones most affected by disturbances, with 5% of the range on average every year and peaks around 6.5%; while the silver fir has a lower average score. Silver fir is the coniferous species with the highest percentage of its range affected over the time period of analysis (8%). Black pine is the only species which is consistently affected by fire events (10% of the disturbed pixels on average every year, with peaks around 30%), while the other coniferous species are affected by fire much more sporadically. Both broadleaved species follow almost the same pattern in both proportion of affected range and distribution of disturbance types, with beech slightly more affected by fire events than oak.

4.4 Discussion

4.4.1 Species' ranges assessment

In this study we conducted a trend analysis of the time series of probability of occurrence for different forest tree species on a European scale to identify potential shifts in tree species distributions in the last two decades. Our results show that, on average, the probability of occurrence among the pixels for all the species is rather stable for significant areas in Europe (no significant positive or negative trend), suggesting a predominant status quo in the species distributions over the study area. However, despite this high stability, there are some ongoing changes in species distribution patterns shared across the species and for different regions. While among these subtle variations we have identified both positive and negative trends, the latter are more prevalent than the former and that Northern and Central Europe shows more negative hot spots than other parts. Furthermore, the

various patterns among the species analyzed suggest a different response to the wide range of environmental stressors each species is subjected to. These results still need to be approached with caution as for some of the analyzed species from **Bonannella** et al. (2022d), the model performance is not very high. This could lead to underestimation or overestimation of the species' realized distribution and incorrect conclusions about increasing or receding species ranges. While the authors provide model uncertainty maps together with the probability maps, their analysis did not involve plot data such as NFI plots with precise locations. They used a filtering procedure to achieve a high spatial confidence (i.e. enough to be used for 30 m resolution maps) for datasets whose locational precision is either difficult to assess (i.e. GBIF) or is degraded to coarse resolution by definition (EU-Forest). For this reason, the precision of the distribution maps is expected to be somewhere between the species level and forest type level. In the current study we aggregated the results from 30 m to 1 km to filter out false signals (i.e. areas with intense forest management, local mature stands changes etc.); this process may also have canceled out the effect of potentially misclassified areas in the distribution maps.

In our analysis, it is interesting to note that areas with the highest values (90-100%) of pixels with no increasing or decreasing trend in probability of occurrence across all the species roughly align with what is reported in literature as their native range (**Bonannella** et al., 2022d; San-Miguel-Ayanz et al., 2016): the places where species have historically persisted and adapted over time despite external perturbations are associated with a high degree of stability. While it is true that the current state of European forests is far from being natural (Strona et al., 2016) and that they have been influenced by humans for centuries (Kaplan et al., 2009), the identification of these stable regions within the native range of the species highlights the importance of preserving and protecting these areas to safeguard the diversity of European forests. For some species such areas can be considered as ecological *refugia* (Ashcroft, 2010), potentially providing valuable insights for conservation strategies. For instance, areas with stable probability of occurrence for the silver fir are localized in the Alpine regions of Central Europe (Tinner et al., 2013), well known as the species native range; for beech and oak vast parts of temperate Europe are classified as stable, which are areas where the two species have thriven for centuries (Fang and Lechowicz, 2006; Jones, 1959); for spruce in the central and northeastern Alps, the Carpathians and Scandinavia (Aarrestad et al., 2014; Latałowa and Knaap, 2006) and for Scots pine the stable areas are mostly located in Scandinavia, with some clusters on the border between Germany and Poland (Eckenwalder, 2009; Krakau et al., 2013).

While this overall trend is in agreement with the results of Maes et al. (2023), who showed that forest conditions across Europe have been even improving from 2000 to 2018, the areas with a prevalence of negative trends can't be ignored. Multiple studies have shown that climate change, by the end of the century, will considerably alter the distribution of European tree species (Mauri et al., 2022) and will be especially threatening for coniferous species (Dyderski et al., 2017), so it becomes crucial to scrutinize any signs of possible range

contractions to differentiate between localized phenomena and large-scale shifts. The areas with a prevalence of negative trends can represent hot spots where this phenomenon could spread or, alternatively, they could be isolated episodes caused by extremely particular events occurring in those locations.

Our results reveal that the hot spots of negative trends generally lie at the edges of each species latitudinal range, suggesting these areas might be more exposed to unfavorable factors (climate change impacts, natural disturbances or lack of thereof, habitat fragmentation and species-specific disturbances). These factors could push the species beyond their optimal conditions, leading to range contractions or reduced fitness. Examples of the effects of these stressors on a European scale are numerous: Neumann et al. (2017) localized hot spots of tree mortality in both Northern (Scandinavia, mainly Finland) and Southern (Cantabrian and Leon mountain ranges in Spain) Europe for the period 2000–2012 due to extreme climatic variability, hot spots that align with our results for some tree species, mainly spruce, black and Scots pine; Senf et al. (2018) later proved how tree mortality rate, mainly due to changes in climate and land-use, has doubled in the last decade compared to the previous two.

Quantifying climate change impacts on tree species and their range may be difficult, as each species may respond differently to these stressors (Fei et al., 2017). Previous studies on the subject have shown a variety of responses, the most common being upward (Feeley et al., 2011; Lenoir et al., 2008; Morin et al., 2008) and poleward (Berner and Goetz, 2022; Hanberry and Hansen, 2015) shifts due to temperature increases, but downward shifts are also not uncommon (Crimmins et al., 2011). While in some cases a slight increase in temperature outside the species optimal range has proven to be beneficial for tree growth due to the CO_2 fertilization effect (Martínez-Vilalta et al., 2008), a steady increase has proven to be one of the most important drivers of tree mortality, especially in Mediterranean Europe (Archambeau et al., 2020). Because of their economical importance, most of the species analyzed in this study have been historically planted outside of their native range, as is the case for spruce and Scots pine.

It is no coincidence that several of the negative trends hot spots that also have the highest proportion (90-100%) values for these two species are found at low elevations in Central (i.e. Slovakia, Czech Republic, Poland) and Southern (all Spain except the Pyrenees) Europe that are outside of their native range, where the species have been naturalized through continuous forest management practices (Caudullo et al., 2017; San-Miguel-Ayanz et al., 2016). These results align with a long list of studies assessing either stress or decline for the two species in those areas (Aguadé et al., 2015; Kolář et al., 2017; Rigling et al., 2013; Sedmáková et al., 2019). In contrast, the negative trends hot spots in the moderate (40-50%) range found in the northernmost part of Scandinavia for the same two species align with the results of Maes et al. (2023), which indicate a general reduction in ecosystem productivity (NDVI) and soil organic carbon (SOC) in the area rather than

a species-specific issue. Hot spots for Scots pine extend further in the northern part of Finland, while spruce stays stable in the area: both trends have been confirmed by a tree-ring analysis conducted by Mäkinen et al. (2022) on the data from the last Finnish NFI, which found a declining trend for Scots pine especially in sites with mineral soils and in peatlands. Hot spots for very stable species like beech in south-eastern Sweden are also well known: beech does well especially in the south-west, where water availability is high, and gets progressively worse moving north, as shown by Martinez del Castillo et al. (2022).

Our results indicate that the proportion of positive trends is notably smaller compared to stable and negative trends for all the tree species analyzed, reaching moderate (40-50%) values at best and without consistent hot spots. This was also expected: colonizing new suitable areas through upward or poleward shifts to compensate for range erosion is a difficult task for most of the European tree species due to their low dispersal ability (Mauri et al., 2022; Svenning and Skov, 2004). Evidence of latitudinal (i.e. poleward or southward) shifts in communities of adult trees are still rare (Hanberry and Hansen, 2015) and trees natural dispersion capabilities are deemed to be insufficient to face the future risks posed by climate change (Jump and Peñuelas, 2005; Zhu et al., 2012), with assisted migration being a possible solution to the issue (Mauri et al., 2023). Our results for the species analyzed in this study suggest a higher degree of stability in the species' ranges, with signs of localized negative trends in the last two decades. A combination of *lean* and *crash* phenomena, as described by Lenoir and Svenning (2015), characterizes the situation observed in the species analyzed. The presence of overall stable range edges and shifting optimum, alongside localized declines, within the existing range, suggests significant dispersal limitations and the early stages of a shifting process that lags behind climate change. This phenomenon is often referred to as *climatic debt* in the scientific literature (Dullinger et al., 2012; Tilman et al., 1994). As we see in our results, the species can persist within their current range and even their native range, as indicated by the areas with prevalent stable probability of occurrence, despite environmental conditions being sub-optimal, for instance, for sexual reproduction and expansion. However, the species may eventually confront this *climatic debt* through abrupt range shifts, potentially leading to widespread forest dieback, particularly when critical thresholds are crossed due to repeated extreme climatic events or climate-induced pest outbreaks (Bałazy, 2020; Menezes-Silva et al., 2019; Pietrzykowski and Woś, 2021).

4.4.2 Species disturbances patterns

Our results show a consistent increase over time in disturbances over the species realized distribution, which is in line with the reported intensification of disturbance regimes in the last two decades for European forests (Hlásny et al., 2021b; Patacca et al., 2023; Senf and Seidl, 2021b). While our study did not focus on the magnitude of these disturbance events on each species, the slope distributions show that most of the disturbed pixels have

values of β coefficient close to 0 for the time period analyzed. This could be interpreted as disturbances having a negligible effect on the probability of occurrence during this time period; however, it might also be a result of the majority of the disturbance events being concentrated at the beginning or the end of the time series. The OLS model is fitted over only six points in time, with each point capturing the average situation of 3-4 years, and the information on the year of disturbance is limited to the greatest disturbance only. Information on recurring events over the same area is therefore not available, even though the probability of occurrence could experience fluctuations due to other disturbance events; this limitation in the data could lead to a severe underestimation of the effect of disturbances on species probability of occurrence. Furthermore, as for the species' ranges assessment, some areas of Europe the distribution maps may depict forest types rather than forest species. The disturbance analysis was conducted at 30 m and is not affected by the 1 km aggregation we applied to the species range analysis; any potential misclassification in the distribution maps directly influences species-specific conclusions of the disturbance analysis, with once again possible under- or overestimation of disturbance events per species. This is an additional reason on why we decided to focus on qualitative trends only regarding disturbances and we advice caution when reporting these results for future research.

Results for both the affected range and disturbance types for spruce and Scots pine align with what in general is reported in literature: well known for their high susceptibility to wind-induced disturbances (Mitchell, 2013; Peltola et al., 1997), information on the effects on these species of the major wind storms (Kronauer, 2007; Kronauer, 2000; Seidl and Blennow, 2012; Valinger et al., 2014) or bark beetles outbreaks (Jaime et al., 2019; Mezei et al., 2017; Wermelinger et al., 2021) are also well documented. Despite this, our findings indicate that for other species, like silver fir or even the broadleaved species, the percentage of pixels affected by wind disturbances in the years of severe storms is higher ($\geq$ 50%) than for spruce and Scots pine. While it has been proven that silver fir is more sensitive to wind than Scots pine and less than the spruce (Schmidt et al., 2010), broadleaved species, and in particular the beech – oak group, are in general less sensitive to wind than coniferous species (Klaus et al., 2011; Seidl et al., 2017; Wohlgemuth et al., 2022). This discrepancy can be attributed to the significant impact of the *other* disturbance class, which includes logging, drought and bark beetles, for spruce and Scots pine: not only these two species have a relatively larger distribution range compared to other species, but also higher harvesting rates (FOREST EUROPE, 2020). Coniferous species have also been more affected by drought (Pardos et al., 2021) due to forest management practices (i.e. monospecific stands) than broadleaved species and by bark beetles (Hlásny et al., 2021a), with spruce in particular being affected by multiple severe *Ips typographus* (L.) outbreaks in the last decade (Patacca et al., 2023; Romeiro et al., 2022). A combination of all these factors may overshadow the prominence of wind-induced disturbances for spruce

and Scots pine. This highlights once again the importance of having accurate and up to date spatiotemporal information on forest disturbance drivers.

Black pine stands out among the other species not only for keeping an overall stable trend in affected range but also for being affected by fire more frequently and intensely: compared to the other species analyzed in this study, black pine native range includes regions where wildfires have historically been a regular occurrence (Isajev et al., 2004). When compared to other Mediterranean pines (i.e. *Pinus halepensis* Mill. or *Pinus pinaster* Ait.) however, black pine lacks fire-adaptation mechanisms (Rodrigo et al., 2004); it can tolerate low intensity surface fires, but crown fires are usually either stand replacing or lead to the development of non-forest ecosystems (i.e. grasslands or shrublands) (Ganatsas, 2010; Lucas-Borja et al., 2021). This behavior is consistent with the negative trends observed for the species in Catalonia, which, according to the disturbance type map we used from Senf and Seidl (2021b), is also one of the two hot spots mostly affected by wildfires in Mediterranean Europe. There is a chance that the majority of the pixels in the area classified as undisturbed but with negative trends were previously affected by wildfires in the years not covered by our analysis. Large fires in the area are common and the last one before 2000 was in 1998, with more than 24,000 ha of forest cover completely burned (Rodrigo et al., 2004). While fire can explain a significant portion of the disturbances and its higher proportion compared to the other species is significant, it is still limited in comparison to the prevalence of the *other* class. As we know that black pine is also fairly harvested (FOREST EUROPE, 2020), there is a significant component of logging in the *other* class. However, due to the lack of clear distinction within this class, further interpretation on the effects of the *other* class on black pine is challenging, making it a limitation in our ability to clearly infer the individual impact of the different disturbance types within the *other* class.

Broadleaved species show distinct patterns in their response to disturbances compared to coniferous species. While they exhibit similarities with other conifers in terms of the percentage of affected range and the proportion of wind and *other* disturbances, the distribution of disturbed pixels for broadleaved species is concentrated either in the mode of the slope distribution or evenly spread to the right of the mode (reaching "1.2" – "1.3" bins for the oak), indicating their ability to cope with disturbances. Other studies have highlighted how disturbances seem to be of minor concern for broadleaved species, to the point of being advantageous to them: Moris et al. (2023) showed that beech has high resilience to fires due to its fast resprouting ability and Scherrer et al. (2022) observed how beech outlived and even replaced spruce and Scots pine in those areas affected by wind or other disturbances. Salvage logging, a common practice for windthrown trees (Lindenmayer et al., 2012), has also contributed in the past to conversion from coniferous to broadleaved stands through natural regeneration only (Götmark and Kiffer, 2014; Jonášová et al., 2010). However, recent studies have pointed out that despite their resilience, broadleaved species may face challenges under increased frequency and intensity

of disturbances, especially during prolonged drought spells. Mathes et al. (2023) analyzed the effect of three consecutive years of drought (2018–2020) on beech dominated forests in Germany and their study revealed a significant reduction in radial growth, indicating that severe droughts can override the effects of competition on tree growth. Overall, while disturbances might pose a danger by themselves for coniferous species, an increase in their frequency and intensity could potentially act as a tipping point for broadleaved species.

4.5 Conclusion

In this study, we conducted a comprehensive trend analysis of the probability of occurrence for the realized distribution of different forest tree species across Europe to investigate potential shifts in their distributions over the last two decades. Our results demonstrate an overall trend of stability in the distributions for many regions in Europe, with stable regions largely aligning with the native ranges of each species. These stable areas hold ecological significance and should be prioritized for conservation efforts to safeguard the diversity of European forests. However, we also identified areas with a high proportion of negative trends mostly situated at the edges of each species latitudinal range, specifically located in Central and Northern Europe for the species investigated. These areas may serve as potential hot spots of range contractions or reduced fitness for the species due to exposure to unfavorable factors, so they should be monitored closely in the coming years. Lastly, we identified areas with localized negative trends that were more species-specific, indicating that the response patterns to environmental stressors vary among different species. While our results align with several related studies, the potential limitations of the distribution maps used in the current assessment calls for caution and also highlights the need for more precise distribution maps that include useful information provided by detailed field surveys such as National Forest Inventories. Our analysis on forest disturbances focused on quantifying affected range and most prominent disturbance drivers for each species. Disturbance patterns have shown an increase over time in species affected range for all species, with some disturbance agents being very species-specific. The *other* disturbance class, remains the most prevalent disturbance type for all the species: this highlights the difficulty in precisely discerning the individual impact of different disturbances due to the absence of information on each disturbance agent other than wind and fire.

In conclusion, our study provides insights into distribution trends and disturbance patterns for some of the most important tree species in European forests in the recent past. Notably, this research introduces a novel approach by making use of Species Distribution Models mostly based on Earth Observation data, which enables consistent assessments of the ongoing changes in tree species distributions. This methodology contributes to understanding the observed trends, particularly the emerging changes in certain areas, which appear to be more pronounced in the recent years. Furthermore, the observed trends

emphasize the significance of species-specific responses to environmental stressors and the need for more accurate spatiotemporal information on forest disturbances to monitor and understand these responses in the long term. It is imperative to proactively plan for the future species composition of European forests, alongside with preserving what we already have or adopting assisted migration practices. Careful evaluation and consideration of which species to plant where are essential to strike a balance between fostering adaptability and safeguarding the integrity of natural ecosystems. Our results for the last two decades, however, underscore the urgency of taking prompt actions, especially considering the long time period forest will require to adapt to climate change.

Chapter 5

Impact of NFI coordinate precision on high resolution tree species classification

This chapter is based on:

C. Bonannella, L. Parente, B. Lerink, S. de Bruin, and M. Herold (2024b). "Impact of NFI coordinate precision on high resolution tree species classification". *In preparation*

Abstract

This study assesses the impact of coordinate precision on high-resolution tree species classification in the Netherlands by comparing two models: a True Coordinates Model (TCM), using precise National Forest Inventory (NFI) coordinates, and an Estimated Coordinates Model (ECM), using publicly available, approximate coordinates. We developed an estimation procedure to identify a set of locations for each NFI plot potentially capturing the true location. However, this procedure included the actual true location pixel in only 10% of the cases; the ECM was then trained with all the subsampled pixels. This highlights the limitations of using estimated data and underscores the need for access to precise NFI data. In our analysis, the TCM demonstrated superior performance over the ECM. Specifically, the TCM achieved an Overall Accuracy of 0.71, compared to just 0.45 for the ECM. Additionally, the TCM exhibited a R^2_{logloss} of 0.40, indicating a moderate increase in performance compared to a naive classifier, whereas the ECM's R^2_{logloss} barely reached 0.03. The Prediction Interval Coverage Probability (PICP) for the 95% PI for the TCM was 0.88, while for the ECM it was only slightly lower (0.86), so both models were overconfindent in their predictions. In terms of species classification, "Fagus spp" and "Picea spp" showed the highest balanced accuracy in the TCM, whereas "Picea spp" presented the lowest balanced accuracy in the ECM. These results demonstrate the significant impact of coordinate precision on the accuracy of tree species classification models. Our study highlights the challenges in integrating remote sensing data with NFIs, especially in terms of data precision and availability, advocating for a combined approach of precise NFI coordinates with advanced remote sensing techniques for high resolution forest management products. Overall, this research underscores the need for precise spatial data and improved data-sharing practices in forestry and remote sensing, contributing to more effective environmental management and conservation efforts.

5.1 Introduction

Accurate knowledge of tree species distribution is fundamental for a wide range of applications in forest inventory and both forest and environmental management. This information is crucial for underpinning sustainable forest management practices and informing necessary reporting requirements (Bontemps et al., 2022; Bussotti and Pollastrini, 2017; Gillis et al., 2005), aids in species-specific growth and yield estimation (Pretzsch et al., 2015; Tompalski et al., 2014), plays a vital role in carbon modeling (Temesgen et al., 2015) and is critical for understanding responses and adaptations to climate change (**Bonannella** et al., 2024a; Hof et al., 2017; Mauri et al., 2023).

The traditional approach of collecting tree species data mainly involves National Forest Inventories (NFIs) conducted at periodic intervals, as information on tree species is one of the basic forest inventory attribute (Tomppo et al., 2010). However, this approach frequently results in gaps in both spatial and temporal coverage. While these inventories typically detail species assemblages based on criteria such as basal area or canopy cover (Barbati et al., 2014; Barbati et al., 2007; Tomppo et al., 2010), they often fall short in providing necessary information for in-depth ecological assessments and informed decision-making. The periodic nature of NFIs often results in gaps in spatial and temporal coverage, limiting their utility: the spatial sampling of NFIs, based on permanent plots, poses challenges in a priori stratified sampling or assessing commission errors, among other limitations. For example, the permanent plots might not be distributed in a way that adequately covers and represents these varied strata, potentially leading to skewed or biased data. Additionally, the extended temporal gaps between successive plot measurements, often ranging from 5 to 10 years (Gschwantner et al., 2022), hinder the ability to deliver timely information.

In response to these limitations, remote sensing technologies, particularly Earth observation, have emerged as crucial tools offer a promising complementary approach. The European Union's Forest Strategy for 2030, a key component of the European Green Deal, recognizes the potential of such technologies. It envisages actions for enhancing forest protection and sustainable management, with a specific focus on improving forest monitoring. This includes proposals for an EU-wide forest observation framework, providing open access to detailed, regular, and timely information on forest conditions (European Commission, 2023). This strategy aligns with the increasing use of high-resolution satellite data, which have significantly advanced the capabilities for forest monitoring at various scales.

Thus, remote sensing provides an opportunity for spatially comprehensive assessments at an increased temporal frequency to fill spatial and temporal gaps coming from using traditional approaches; it also enables comprehensive assessments with systematic and transparent methodologies (Coops et al., 2023; Wulder et al., 2024). In this sense, tree species classification using remote sensing data is an evolving research field marked by

diverse methodologies and varying scales of application, to the point that comprehensive reviews recently published, such as the one by Fassnacht et al. (2016), are already partially outdated. Studies have employed different remote sensing data types, including moderate to high-resolution satellite imagery from Landsat and Sentinel-2 missions, more specialized data like hyperspectral and LiDAR or a combination of all of them. The choice of data often influences the scale and precision of the analysis. For instance, very high-resolution data (pixel size <1 m), allows for detailed tree species analysis at the individual tree level but are often constrained by operational costs and limited areal extents (Waser et al., 2014). Conversely, medium-resolution data from Landsat and Sentinel-2 satellites have been utilized for broader regional or national studies, offering cost-effective solutions for extensive area coverage despite challenges like cloud cover and limited image acquisition dates (Breidenbach et al., 2021; Hermosilla et al., 2022; Welle et al., 2022). Furthermore, continent-scale analyses have utilized high-resolution hyperspectral data, coupled with field data, to develop models for tree species classification, assessing both general and site-specific models (Marconi et al., 2022). The methods employed in these studies are as varied as the data sources. Machine learning techniques, notably Random Forest classifiers, have been widely used due to their effectiveness in handling complex datasets and producing reliable classifications (Axelsson et al., 2021; Dalponte et al., 2012). The inclusion of additional environmental variables, such as topographic predictors, climatic variables, spectral indices or phenological metrics, has been found to enhance classification accuracy significantly (Grabska et al., 2020; Waser et al., 2021; Ye et al., 2021).

Despite significant advancements in remote sensing technologies and their integration with NFIs, a critical challenge remains in the effective utilization of NFI data for species distribution modeling. Coops et al. (2023) demonstrate the potential of multi-source remote sensing data for near real-time forest inventory, underscoring the importance of timely and accurate data for informed decision-making. However, the limited availability and accessibility of precise, plot-level NFI data, as noted by Fassnacht et al. (2024), poses a significant barrier. This limitation is particularly pronounced given the common practice of degrading the precision of geographical coordinates for public release, primarily to protect landowner privacy (Nabuurs et al., 2022). Ground surveys remain fundamental for assessing forest resources: Wulder et al. (2024) showcase the feasibility and benefits of implementing a stand-level satellite-based forest inventory, yet the successful application of such methodologies hinges on the availability of accurate ground-truth data. The current scenario, where remote sensing practitioners often rely on publicly available, spatially imprecise NFI data, raises concerns about the accuracy of tree species distribution derived from these datasets. This gap in precise, publicly accessible NFI data not only hampers the potential synergies between NFIs and remote sensing but also limits the scope of ecological insights that can be derived from these products (**Bonannella** et al., 2022d; **Bonannella** et al., 2024a). The ultimate goal would be to achieve a synergistic integration of NFIs with

satellite data, harnessing the potential of an integrated approach for advanced monitoring of forest resources (Ceccherini et al., 2022; Fassnacht et al., 2024).

Addressing this research gap, our study focuses on evaluating the effectiveness of tree species classification models by comparing outcomes derived from precise coordinates against those generated using publicly available coordinates. To ensure a balanced comparison, we developed and tested a novel procedure aimed at identifying a set of locations for each NFI plot potentially capturing the true location from the publicly available data. These approximated points are then employed as training data in our tree species classification models. Rather than focusing on model accuracy and producing reliable tree species maps, in this study our objective is to rigorously assess the influence of coordinate precision on the accuracy of tree species models through a comparative analysis. We tested this approach on the Netherlands. The results of this analysis can provide insights in understanding the constraints imposed by current data-sharing protocols and underscore the necessity for enhanced data accessibility and sharing practices. Additionally, our research aims to offer a critical perspective into the integration of NFIs with Earth observation data, thereby contributing to the advancement of more precise and reliable forest inventory methodologies.

5.2 Material and methods

5.2.1 Dutch National Forest Inventory

The Dutch National Forest Inventory (NFI) is a valuable source of information regarding the state of forests in the Netherlands. Its historical roots date back to 1938, but, in this study, we used NFI measurements conducted between 2000 and 2020. Consequently, our dataset comprises data from the 5th Dutch NFI (NBI-5), carried out during 2001-2005, the 6th Dutch NFI (NBI-6), conducted in 2012-2013, and the 7th Dutch NFI (NBI-7), executed in 2017-2021 (Schelhaas et al., 2022).

The 7th Dutch NFI employed an unaligned systematic sampling approach with a density of one plot per square kilometer. Initially, a digital forest map was generated based on LULUCF (Land Use, Land-Use Change, and Forestry) maps and aerial photographs (Dirkse et al., 2001; Nabuurs et al., 2003). The forest definition used for this map adheres to FAO guidelines (FAO, 2004). Subsequently, a 1×1 km INSPIRE-compliant grid (European Parliament, 2007) was superimposed over the forest map. Within each grid cell, a plot was randomly selected for data collection, with the condition that only plots within forested areas were considered for field visits. While each iteration of the NFI ideally aims to measure approximately 3600 plots, the actual number of plots varies over the years. This can be attributed to several factors, primarily related to accessibility issues (Schelhaas et al., 2022). In earlier iterations of the NFI (NBI-5 and NBI-6), the dataset comprised 50% permanent plots and 50% temporary plots (Tomppo et al., 2010). However,

in the 7th NFI, all sampled points were treated as permanent (Schelhaas et al., 2022). In this study we used plots having at least two observations in time; this left us with a total of 8586 NFI plots almost equally spread over the different forest inventory campaigns (Fig. 5.1).

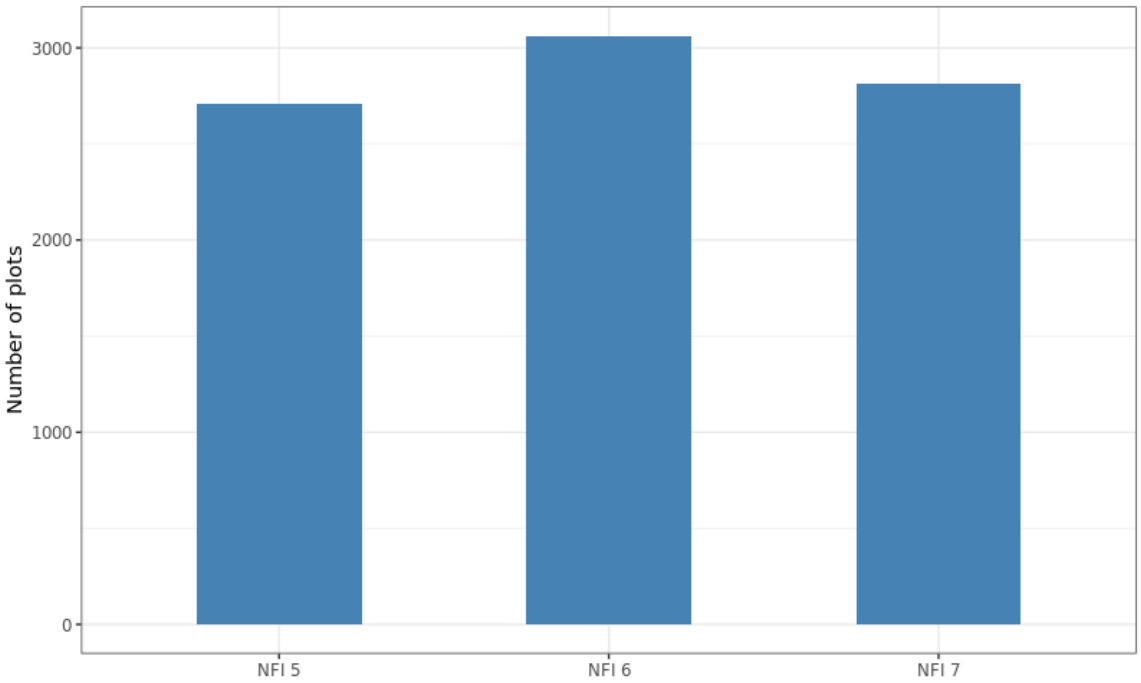

Figure 5.1: Distribution of NFI plots across the forest inventory campaigns used in this study.

The original coordinates of these plots were reported in the Dutch coordinate reference system (EPSG:28992). To align these coordinates with the INSPIRE-compliant grid, we re-projected them to the coordinate reference system ETRS89 / LAEA Europe (EPSG:3035). The exact plot locations are not publicly disclosed. Instead, the publicly available coordinates correspond to the centroids of the INSPIRE-compliant grid cells encompassing the plots. The plot locations are typically situated within a radius of approximately 0.5 kilometers from the publicly disclosed grid cell centroids. This arrangement is designed to protect sensitive forest inventory data while still allowing for certain levels of analysis.

Given the vast individual tree species list recorded in the NFI, species were aggregated into main dominant tree species groups where possible; for example *Quercus petraea*, *Quercus robur*, *Quercus rubra* and other oaks were all grouped in the class "Quercus spp". This was necessary because not all individual species are equally represented (e.g., only 1–2 plots across the three NFI surveys). Consequently, certain plots had to be grouped within miscellaneous categories, such as "other broadleaved", to accommodate these variations in species representation. To classify the NFI plots accordingly we used the information provided by the tree-level data of the NFI. Specifically, we considered attributes such as the number of trees, tree species, and basal area. Within each plot, we identified the dominant

species group, which served as the reference data for our modeling. The dominant species group was determined as the one with the highest fraction of basal area relative to the total basal area of the trees in the plot. This information allowed us to assign a stand purity coefficient to each plot. The stand purity coefficient informs the algorithm about the relevance of a data point for dominant species modeling. Pure stands are deemed most informative because they exhibit a more homogeneous spectral signature compared to mixed stands. For this reason, our modeling only considered plots with a stand purity coefficient >= 0.8, which further reduced the number of NFI plots to 4256.

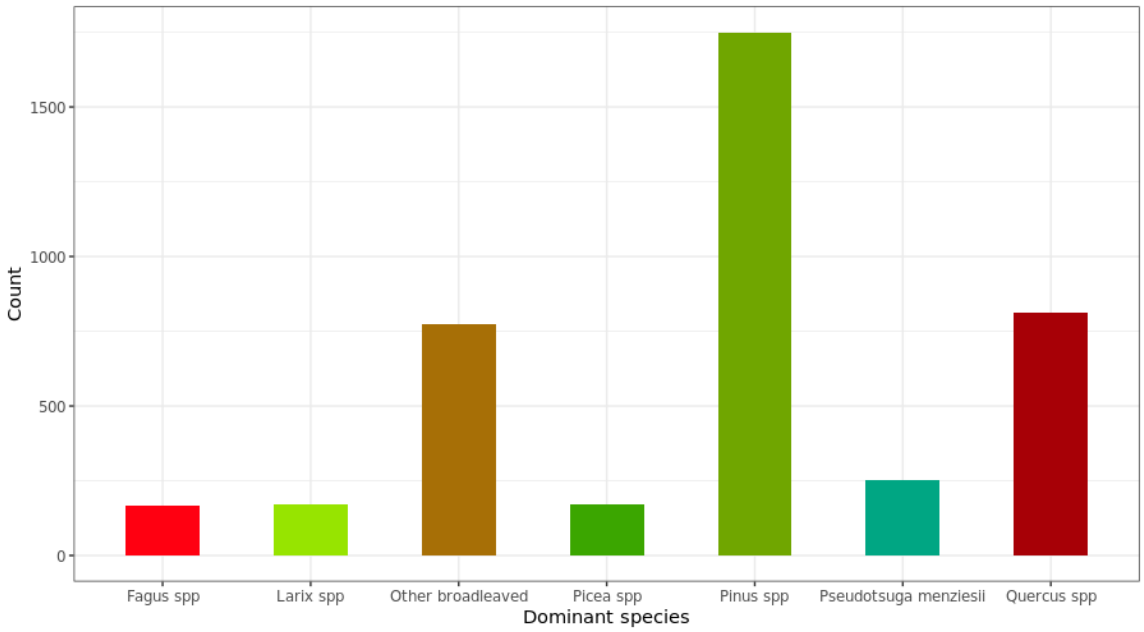

Figure 5.2: Distribution of dominant species. Note the strong imbalance between the different classes, with Pinus spp counting >1500 observations while most of the other classes count < 200 observations.

5.2.2 Tree canopy extent

In our procedure using approximate NFI plot coordinates, we utilized the European tree canopy extent dataset from Turubanova et al. (2023). The layers have 30 m spatial resolution and cover the time period 2001–2021 at a yearly temporal resolution. This dataset was used as a proxy for a forest mask. The product defines tree canopy extent by a land cover class having at least 5 m tree canopy height. This comprehensive time series was generated by integrating annual tree canopy height and removal maps. The tree canopy height data were derived using an empirical modeling approach based on the bagged regression tree method, utilizing multidecadal spectral data from the Landsat archive. This was coupled with calibration data from Airborne Laser Scanning (ALS) collected from mainly Scandinavian countries and spaceborne Global Ecosystem Dynamics Investigation (GEDI) lidars (Dubayah et al., 2020).

5.2.3 Forest types layers

Witjes et al. (2022) developed a spatiotemporal ensemble machine learning framework to generate land use/land cover (LULC) yearly time-series maps for Europe for the period 2000–2019 at 30 m resolution. They used a harmonized legend based on the CORINE land cover product (Bossard et al., 2000) and the LUCAS survey (d'Andrimont et al., 2020) for a total of 43 predicted land use/land cover classes. A distinct feature of their framework is the generation of not just yearly dominant land cover classes, but also probability layers for each of the 43 classes. This approach marks a significant departure from conventional LULC products, offering a nuanced and multifaceted understanding of land cover dynamics. The framework predicts the probability of each land cover class per pixel, providing also per-pixel uncertainty values. In our research, we focused specifically on two classes only: "Broad-leaved forest" and "Coniferous forest".

5.2.4 Predictor variables

Landsat GLAD ARD collection 2

The Landsat Analysis Ready Data (ARD) from the Global Land Analysis and Discovery (GLAD) team at the University of Maryland is one of the few globally consistent archives for historical time series of normalized surface reflectance derived from the Landsat satellites collections Potapov et al., 2020. The original GLAD Landsat ARD, however, often contains cloud contamination and image artifacts, necessitating extensive preprocessing for effective use in modeling and analysis. To address these limitations, we developed an enhanced version of the GLAD Landsat ARD-2.

The GLAD Landsat ARD-2 is a comprehensive dataset encompassing 16-day interval tiled composites of images. Spanning from 1997 to 2022, it provides 23 images per year, culminating in a total of 598 global images. This dataset includes six reflectance (blue, green, red, near-infrared, short-wave infrared 1, and short-wave infrared 2) bands, one thermal band and a detailed quality flag for each pixel. This flag categorizes pixels as land, water, cloud, cloud shadow, topographic shadow, hill shade, snow, haze, cloud proximity, shadow proximity, other shadows, or as buffered proximity of these categories. Based on its quality flag, we derived a space-time overcast-sky mask, identifying points classified as cloud, cloud shadow, haze, cloud proximity, shadow proximity, or other shadows as '*gaps*' for further processing.

First, from the 16-days time interval time series, a bimonthly product was derived performing weighted temporal aggregation. The procedure combines four images from the original product into one aggregated image. The last image of a year (covering November-December), is produced using the last three images of the current year and the first images of the next year of the GLAD Landsat ARD (for the last two months of 2022 only three images were used). The temporal aggregation involves a weighted averaging of the four

images, where each weight corresponds to the clear sky fraction of the tile in the temporal frame. The so obtained images are scaled and stored as byte in different compressed GeoTIFF files for each band, tile and time-frame.

In addition to the six reflectance bands from the enhanched GLAD Landsat ARD-2 dataset, we also incorporated several key vegetation and water indices as predictor variables for our modeling purposes. These indices include the the Enhanced Vegetation Index (EVI), the Normalized Burn Ratio (NBR), the Modified Normalized Burn Ratio (NBR2) also called Normalized Tillage Index (NDTI), the Normalized Difference Snow Index (NDSI), the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI). Each of these indices is derived from different combinations of the reflectance bands and provides unique insights into vegetation health, moisture content, burn severity, and overall ecological conditions. We further used the following temporally aggregated indices:

- Bare Soil Fraction (BSF), determined by the proportion of time the NDVI is < 0.35 (Castaldi et al., 2019)
- Number of Seasons (NOS), a measure indicating the annual count of crop cycles (Li et al., 2014)
- Crop Season Ratio (CSR), representing the percentage of the year during which a pixel shows active crop growth (Estel et al., 2016)

These indices are essential for capturing natural or anthropogenic processes that require a longer temporal frame for sensible quantification. This approach is particularly pertinent in a country like The Netherlands, characterized by its intricate agricultural spatial patterns.

Climate data

The ERA5 Land hourly dataset was temporally aggregated to obtain monthly data of air temperature (2 meters above ground), surface temperature and precipitation. The original ERA5 data was first aggregated to daily data, and subsequently to monthly data, with increased resolution (1 km) using CHELSA climatic data (Karger et al., 2020): in this way the general spatial and temporal patterns of ERA5 Land dataset were retained whilst using the finer spatial detail from the CHELSA dataset. For air and surface temperature we obtained the monthly minimum, mean and maximum, while for precipitation the monthly sum for a total of 84 climatic time series layers. Additionally, we used 17 bioclimatic variables (Hijmans et al., 2005) for the period 1979–2013 to provide a baseline of the actual state of the climate. We used variables from the CHELSA dataset since it has been claimed to better agree with data from meteorological stations than WorldClim (Karger et al., 2017).

Topographic data

The elevation data utilized in this study was derived from the Digital Terrain Model (DTM) for Europe, as developed by Hengl et al. (2020). This model represents a comprehensive integration of various high-resolution digital surface models (DSMs) and elevation data points. Specifically, it combines GEDI Level 2B (Dubayah et al., 2020) and ICESat-2 (ATL08) (Markus et al., 2017) elevation points with well-known DSMs, including ALOS AW3D30 (Tadono et al., 2016), GLO-30 (Strobl, 2020), and MERIT DEM (Yamazaki et al., 2017). An ensemble machine learning approach was employed to harmonize these diverse data sources, resulting in a refined 30 m resolution elevation product covering the entirety of Europe. We also used other DTM-derived topographical variables such as slope, aspect, easterness, northness, and Topographic Wetness Index (TWI).

5.2.5 Sampling-based estimation of precise NFI plot locations from public coordinates

We developed a multi-step procedure utilizing publicly available NFI plot coordinates, dominant species information, and auxiliary raster layers to approximate the precise coordinates of the NFI plot. This procedure operates within the spatial context of the 1×1 km INSPIRE-grid cell, delimitating a subsample of 30 m pixels from the broader population of pixels within the cell. All the pixels within this subsample are treated as training points for our model. The dominant species information used as reference data for the model is derived from the original NFI plot. It's worth noting that this process can result in the inclusion of hundreds of pixels for each NFI plot. To account for potential inaccuracies in the spatial location of these training points, we assign each newly generated training point a weight of $1/n$, where n represents the total number of pixels in the final subsample. This weighting approach helps inform the model about the reliability of these training points, considering the possibility of high spatial location inaccuracies. To match the characteristics of the NFI plots with precise coordinates in the feature space for modeling purposes, for this procedure we selected only the NFI plots with stand purity coefficient >= 0.8.

In our analysis, we utilized the tree canopy extent maps developed by Turubanova et al. (2023) to generate dynamic forest/non-forest masks for each 1×1 km km grid. This approach is particularly crucial given the non-permanent nature of some of the NFI plots in our study. For instance, plots measured in the 5th Dutch NFI may not necessarily be present in the 6th Dutch NFI or viceversa. On top of that, the dominant species of the plot may change over different surveys. To accommodate this variability, our forest masks are designed to be dynamic, necessitating yearly tree extent layers. For a given plot, the forest mask is derived based on the tree extent layer corresponding to the year of the plot's survey, as well as layers from up to five years prior. For example, if a plot was surveyed in 2006, the tree extent layers from 2001 to 2006 are utilized. In determining forest status

within these masks, we apply a stringent threshold: a pixel is classified as *forest* only if it is consistently identified as forested in all the selected yearly layers.

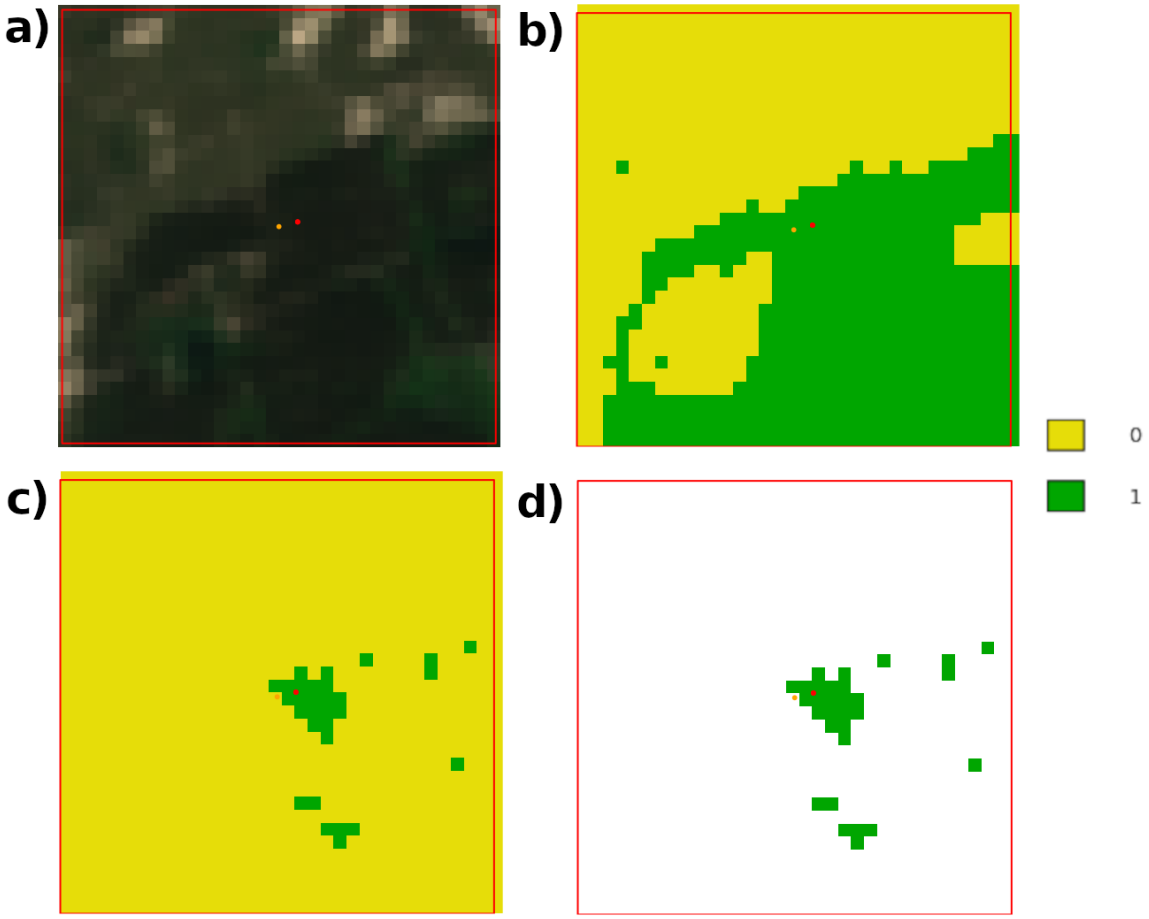

Figure 5.3: Workflow to determine set of potential locations containing exact NFI plot coordinates: initial public plot coordinates (orange) and precise plot coordinates (red) within a 1×1 km grid cell (red border), over a Sentinel-2 RGB composite **(a)**; refinement using dynamic forest mask highlights potential precise plot locations in those pixels classified as "1" **(b)**; refinement using combined forest and forest type masks (coniferous) narrows down potential plot locations **(c)**; the ultimate subset of pixels selected as candidates for the precise plot locations **(d)**. The precise plot location (red dot) is captured within this final pixel subset.

We further refined this selection of pixels using the forest type layers from Witjes et al. (2022); in particular, the choice between "Broad-leaved forest" and "Coniferous forest" layers was determined by the dominant species present. After that, we analyzed the distribution of probability values within the grid cell for the chosen forest type layer. To identify the most likely pixels representing the actual location of an NFI plot, we set a threshold based on the 90% of the maximum pixel value found in the distribution. Only pixels with values equal to or exceeding this threshold were considered as potential candidates. This approach ensures that we sample from the most representative pixels

within the grid cell, significantly enhancing the accuracy of our plot location approximation. In the next stage of our procedure, the pixels selected based on the 90% threshold are reclassified as "1" within our forest type mask. This reclassification is a critical step towards pinpointing the precise location of the NFI plots. The final phase involves integrating this forest type mask with the forest mask derived from Turubanova et al. (2023). By multiplying these two masks, we effectively isolate the pixels that are most representative of the actual forest conditions and the specific forest type of each NFI plot. The pixels that emerge from this multiplication process are then regarded as the prime candidates for accurately representing the precise locations of the NFI plots (Fig. 5.3).

5.2.6 Model building and evaluation

In assessing the efficacy of classification models to compare the effects of using true coordinates or estimated coordinates, we prioritized interpretability and robust evaluation. Random Forests emerged as the optimal choice, offering an accessible yet sophisticated machine learning solution: Random Forests is a non-parametric, ensemble-learning algorithm based on classification trees where each tree is grown with a different subsample of training data and predictor variables (Breiman, 2001a). This method's intrinsic randomness effectively counters overfitting. Compared to other powerful machine learning algorithms such as Gradient Boosting, Support Vector Machine or Artificial Neural networks, Random Forests is also less sensitive to hyperparameter tuning: this means that default hyperparameter values often result in near-optimal performance (Probst et al., 2019). Nevertheless, we kept all hyperparameter values as default but we decided to tune the *num.trees* parameter, setting an hyperparameter search space between 100 and 1000 trees, using random search as a search strategy and logarithmic loss as our scoring estimator to evaluate the model's performance across different settings. The weights generated during the estimation procedure for the sampled pixels were used for the estimated coordinates model. Since we were interested in probabilities per class, the model was trained with the "*prob*" option, so that the final model would output probabilities instead of hard classes.

For the purpose of validation, we utilized the remaining records from the 7th Dutch National Forest Inventory as an independent validation set from the 4236 NFI precise plot locations previously selected. The validation set consisted of 1305 observations in total. Given that the tree species classification problem consists not only of probabilistic assessments but also traditionally addresses the need for hard class mapping, we evaluated our models using a set of both classical hard labeling metrics like Overall Accuracy (OA), as well as probabilistic/scoring rules such as Brier the score (Brier, 1950) and a pseudo R^2 based on logarithmic loss (**Bonannella** et al., 2022d). To compare class performances we instead used balanced accuracy (Soleymani et al., 2020), F1 score, Brier score and R^2_{logloss}. After that we generated 100 bootstrapped iterations for both training sets (using true and estimated coordinates) and trained an equivalent number of Random Forest

models. This process enabled us to compute a 95% prediction interval (PI) for each class in the dataset. We then utilized these PIs to calculate the Prediction Interval Coverage Probability (PICP), a measure that evaluates the accuracy of the predictions within the PI in the probability space.

5.3 Results

5.3.1 Sampling-based estimation procedure

The estimation procedure used a total of 2930 plots coming from the 5th and 6th Dutch NFI, using the public coordinates as input generated from the centroid of the 1×1 km grid. This approach generated a total of 83714 potential pixel coordinates across the considered time period (2001–2013). Given the dynamic nature of several plots during this timeframe, each set of spatial and temporal coordinates was treated as a distinct entry in the estimation process. Consequently, given the same spatial coordinates but different time reference, the potential locations associated with a specific plot number varied over time, as illustrated in Figure 5.4.

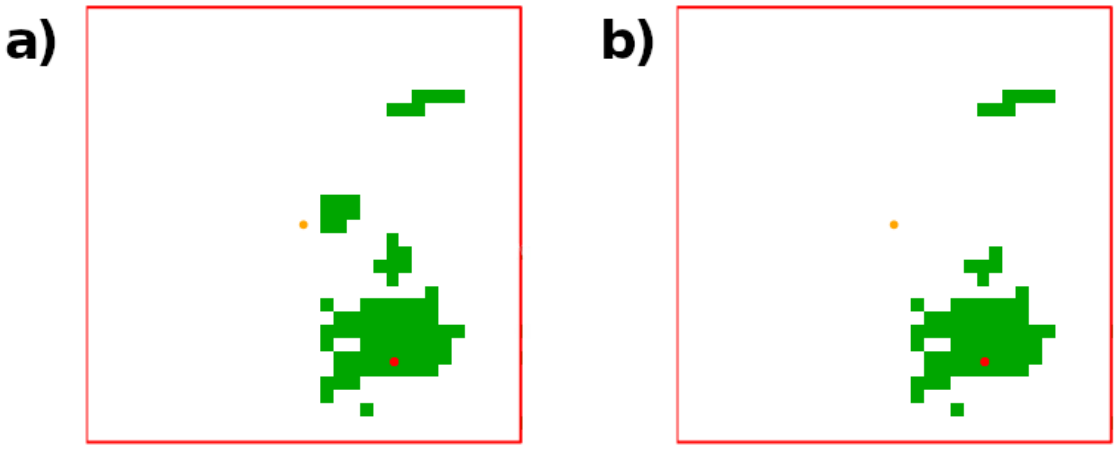

Figure 5.4: Selected pixels as potential candidates for NFI plot locations in year 2003 **(a)** and in year 2013 **(b)**. Notice the forest patch next the centroid coordinates being absent in 2013 and that in both cases the precise NFI coordinates (red dot) are captured by the selection.

In the estimation process, for 87% of the plots with public coordinates, a sample of pixels meeting the procedure's filtering criteria was selected as potential candidates for the true location. However, for the remaining 13% of the plots, no pixel satisfied these criteria. Upon comparing these estimated locations with the actual, true coordinates, it was found that the procedure accurately identified the true location in only 10% of the cases. The average number of pixels sampled for each public coordinate plot was 41, with the median being 17 and the mode being 1. In very few cases the procedure sampled a number of

pixels in the order of hundreds, reaching almost 700 pixels per public coordinate plot (Fig. 5.5

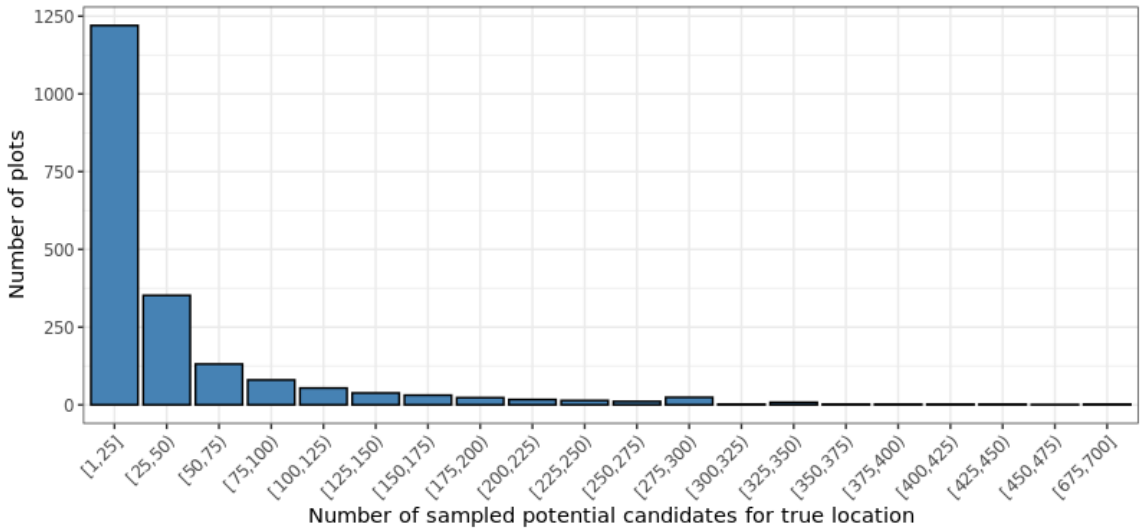

Figure 5.5: Distribution of potential sampled pixels per number of plots. Sampled pixels are grouped in bins of 25 elements. Notice how for more than 50% of the plots used for the estimation procedure the pixels sampled are below 50.

5.3.2 Model performances and uncertainty assessment

The hyperparameter tuning resulted in a total of 557 trees for true coordinates model (TCM from now on), while this number increased up to 772 for the estimated coordinates model (ECM from now on). On average, the performances of the TCM on the validation set achieved moderate OA values (0.71), while, as expected, the ECM had much lower performances, reaching OA values of only 0.45. The Brier score indicates good overall performances in the probability space by both models, as the score for the TCM is 0.07 while is only 0.1 for the ECM. The R^2_{logloss} provides additional information on the model performances, as the metric reaches values of 0.40 for the TCM, meaning a moderate increase in performances of up to 40% compared to a naive classifier, while the same metric for the ECM barely reaches 0.03. The two models performances across the different evaluated performance metrics per class are shown in Table 5.1.

In evaluating the classification performance by class, the TCM uniformly outperformed the ECM in terms of balanced accuracy. The TCM achieved notably high balanced accuracy scores, especially for the classes "Fagus spp" and "Picea spp," both reaching 0.98, which reflects a substantial equilibrium between sensitivity and specificity, as well as a pronounced ability to differentiate these species accurately. In contrast, the ECM's performance peaked at a balanced accuracy of 0.70 for the "Other broadleaved" class,

demonstrating its relative struggle to maintain the same level of discriminative power, with "Picea spp" presenting the lowest balanced accuracy across all classes.

The F1 scores were markedly higher in the TCM, especially for "Pinus spp" and "Other broadleaved" classes; in contrast, the ECM's F1 scores were generally lower, with particularly low scores for "Pseudotsuga menziesii", while it was even impossible to calculate it for the "Picea spp" class. The R^2_{logloss} for the TCM, although not extremely high as no class reaches values > 0.5, were positive for all classes, suggesting good improvements of the TCM model compared to a naive classifier; conversely, the ECM reached even negative R^2_{logloss} values for several classes, implying a performance worse than random guessing for those classes, particularly for "Larix spp". The Brier scores for both models were relatively low across species, suggesting reasonable calibration of probabilistic predictions. However, the TCM generally reported slightly lower Brier scores, indicating more accurate probability estimates for the presence of species compared to the ECM.

Table 5.1: Comparison of model performace metrics per class.

	Bal. Accuracy		F1 Score		R^2_{logloss}		Brier Score	
Class	TCM	ECM	TCM	ECM	TCM	ECM	TCM	ECM
Fagus spp	0.98	0.55	0.29	0.18	0.48	-0.22	0.04	0.05
Larix spp	0.87	0.54	0.23	0.12	0.37	-0.36	0.03	0.04
Other broadleaved	0.78	0.70	0.68	0.42	0.32	-0.13	0.10	0.11
Picea spp	0.98	0.48	0.08	NaN	0.48	0.12	0.03	0.03
Pinus spp	0.86	0.64	0.84	0.62	0.49	0.19	0.11	0.19
Pseudotsuga menziesii	0.95	0.51	0.34	0.05	0.40	0.02	0.05	0.05
Quercus spp	0.77	0.56	0.67	0.25	0.33	-0.18	0.12	0.17

Passing on to the PICP scores calculated over the 100 bootstrapped iterations, it's evident that the overall PICP scores for both models (TCM: 0.88, ECM: 0.86) fall below the target set for this analysis (95% PI). This discrepancy indicates that the models' prediction intervals are narrower than intended, reflecting an overestimation of their predictive accuracy (Table 5.2). The TCM PICP score is closer to the target than the ECM one, suggesting a marginally better estimation of the prediction interval. The TCM consistently demonstrates PICP scores close to the target for the "Fagus spp", "Larix spp" and "Picea spp" classes, with the last two having model based PIs slightly too wide (PICP score: 0.96). The remaining classes however have PICP scores close to 0.8 or even 0.7 ("Quercus spp"), suggesting that the model based PIs for these classes are too narrow and not well calibrated. The ECM follows the same trend but performs worse than the TCM, with "Picea spp" and "Larix spp" having PICP scores close to target while exhibiting much narrower PIs for the other classes. The TCM PICP scores are all closer to target than the ECM PICP scores except for the "Other broadleaved" and "Quercus spp" classes, which are also the classes with the narrowest PIs for both models. Overall, the TCM shows

better calibrated PIs than the ECM, although they are still too narrow compared to the target PI.

Table 5.2: PICP scores computed over 100 bootstrapped iterations

Class	TCM	ECM
Fagus spp	0.95	0.92
Larix spp	0.96	0.94
Other broadleaved	0.81	0.83
Picea spp	0.96	0.95
Pinus spp	0.82	0.70
Pseudotsuga menziesii	0.93	0.93
Quercus spp	0.71	0.77
Overall	0.88	0.86

5.4 Discussion

5.4.1 Estimation procedure evaluation

The results from the sampling-based estimation procedure, which managed to include the true location pixel in the selected subsample in only 10% of the cases, highlight significant challenges in ecological data accuracy when using distorted plot coordinates for modeling purposes. Furthermore, the methodology employed in building the forest masks, which involved the combined use of a general forest mask derived from Turubanova et al. (2023) and a more specific forest type mask derived from Witjes et al. (2022), had significant implications on the results. On one hand, this approach had the advantage of selecting only those pixels that were identified with high confidence by both products, potentially enhancing the reliability of the pixels chosen for analysis. This dual-filtering process could lead to a higher degree of precision in terms of the spatial locations being truly representative of forested areas, thus strengthening the validity of the extracted data.

The method inherits the biases and limitations of both products used to derive the forest mask and the forest type mask. The biases of each product, when combined, may have led to a systematic exclusion or inclusion of certain areas inside the 1×1 km grid cell, thereby skewing the results of the sampling procedure when selecting the potential true location pixels. This dual reliance not only amplifies the inherent inaccuracies of each individual product but also introduces new complexities into the analysis. Therefore, while the use of two distinct products for forest mask construction may have seemed advantageous in theory, in practice, it appears to have contributed to the paradox of achieving high precision (i.e. a small number of sampled pixels) but overall low accuracy (i.e. including

in only 1 case out of 10 the pixel with the true location in the sampled pixels). The outcomes of our sampling-based estimation procedure, may be further contextualized by considering the unique spatio-temporal forest context of the Netherlands. As per the 7th Dutch National Forest Inventory, forests cover only 11% of the country's land area (Schelhaas et al., 2022). This limited forest coverage, coupled with the fact that 54% of these forests are even-aged and only about 30% are relatively single-species stands (Schelhaas et al., 2022), presents a distinct challenge. In contrast to the more homogeneous forest landscapes typical of Scandinavian regions, where productive forests dominate, the Dutch forests' heterogeneous nature likely exacerbates the difficulty of accurate location estimation. Even with precise raster layers employed in the estimation process and filtering the NFI plots by stand purity coefficient, these conditions inherently complicate the task, potentially contributing to the less than optimal results observed in our procedure. This highlights the importance of considering regional forest characteristics and their impact on the efficacy of sampling methodologies in ecological studies.

Comparing our results with the ones from **Bonannella** et al. (2022d), which also utilized a forest type layers from Witjes et al. (2022) to filter species presence plot data, reveals some interesting parallels and divergences. Their approach, like ours, grappled with the challenges of data accuracy and spatial resolution. However, not only they benefitted from a relatively more comprehensive dataset, including harmonized data from various sources like GBIF or the LUCAS survey, but their response variable differed significantly in its fundamental objectives and methodologies. While our study aimed at tree species classification, a task that typically involves mapping all species collectively and then assigning each pixel to the species with the highest probability, **Bonannella** et al. (2022d) focused on species distribution modeling (SDM) and, in particular, mapping both presence and absence of different tree species. This distinction is crucial as tree species classification presents inherently greater challenges than SDM. In classification, the complexity lies not only in accurately identifying forested areas but also in differentiating between multiple species within these areas. This process requires a more nuanced understanding of species-specific characteristics and their spatial distribution, often leading to higher levels of uncertainty and complexity in the model.

In addition, the approach of **Bonannella** et al. (2022d) was characterized by mapping each tree species individually. This method, while potentially offering more detailed insights into the distribution of each species, contrasts with the common practice in tree species classification where a collective mapping of all species is followed by the probabilistic assignment of the pixels. This difference in approach can have significant implications for the accuracy and applicability of the results. Mapping species individually allows for a more focused analysis, potentially yielding more precise distribution patterns for each species. However, it might also overlook the interactions and competitive dynamics among different species, which are crucial elements in understanding forest ecosystems. Furthermore, it's important to note that **Bonannella** et al. (2022d) did not employ a coordinate estimation

procedure akin to ours. Instead, they utilized a filtering process to include or exclude plots with degraded coordinates. By directly using the distorted coordinates without attempting to reverse-engineer the degradation, their study possibly incorporated the inherent inaccuracies of the public datasets. While this approach simplifies the data processing workflow, it may also result in the perpetuation of systematic errors present in the original data sources. In contrast, our attempt to estimate the true coordinates, despite its limited accuracy, represents an effort to mitigate the impact of such inaccuracies on our analysis. These methodological differences underscore the varied challenges and trade-offs inherent in ecological modeling. The choice between using degraded coordinates directly or attempting to correct them, as well as the decision to map species individually versus collectively, fundamentally shapes the nature and utility of the resulting models. As such, these approaches reflect differing priorities and constraints, each with its unique set of advantages and limitations.

5.4.2 TCM vs ECM

Using distorted NFI plot coordinates and their effects on modeling outputs has been investigated in previous studies in particular for SDM: McRoberts et al. (2005) for example demonstrated that the degradation of coordinates significantly affects model estimates; this negative effect diminishes when the distortion exceeds a certain threshold, such as 32 km. This finding suggests that while high resolution modeling is desirable for accuracy, it becomes counterproductive if the underlying spatial data is significantly distorted. Or rather, it raises a fundamental question: why would one engage in high resolution modeling if accurate estimates are only achievable at such a broad scale? In essence, the benefit of high resolution data is lost if the coordinate accuracy is not up to par. It's a clear indication that the scale of modeling must be carefully considered in relation to the quality of spatial data available. In their study, Gibson et al. (2014) examined the implications of using accurate versus degraded coordinates in SDMs, building on the same NFI dataset previously analyzed by McRoberts et al. (2005) from the Forest Inventory and Analysis (FIA) database. This investigation was focused on how the scale of input data affects model outputs in the context of climate change scenarios. They discovered that enlarging the grid size in SDMs can significantly skew the predicted distribution ranges of species, highlighting the necessity of aligning the spatial resolution of the model with the precision of the input data. Their findings indicated minimal differences between the models using true and altered coordinates: this can be attributed to the alignment of the coordinate alteration radius with the grid size of both the input data (climatic and bioclimatic layers at 1 km resolution) and the output (distribution maps at the same resolution), suggesting that the scale of analysis can mitigate the effects of coordinate accuracy.

It becomes evident that while high resolution modeling offers a granular view of ecological dynamics, its effectiveness is heavily contingent on the precision of the underlying spatial coordinates. Degraded coordinates can significantly undermine the utility of high resolution

data. On the other hand, models built on true coordinates, even if at a somewhat coarser resolution, may provide more dependable insights, particularly for broad-scale ecological studies like those examining the impacts of climate change (i.e. future projections, forecasting and similar approaches). Our study's comparative analysis of the TCM and the ECM exemplifies this, demonstrating the superior accuracy and reliability of the TCM due to its use of true coordinates. The overall results of the TCM model merit discussion, particularly considering their moderate performance despite certain limitations in the modeling approach. Notably, the model achieved a commendable level of accuracy even without the implementation of advanced techniques such as hyperparameter tuning, ensemble modeling, feature selection, or the integration of more complex covariates like hyperspectral and radar data, or other multispectral-derived covariates such as tasseled cap transformations or phenology proxies. This outcome underscores the inherent strength of the modeling framework used, suggesting that even basic implementations can yield significant insights in tree species classification.

However, it's important to acknowledge the potential impact of our methodological choices on the model's performance. The decision to filter points by survey alongside the application of a stand purity coefficient, inevitably led to a reduction in the number of training points available to the model, totaling less than 3000. This reduction in training data size may have influenced the model's effectiveness, as machine learning algorithms like random forests typically benefit from larger datasets. More data often equates to improved model robustness and predictive accuracy, as it allows the model to capture a wider range of variability and nuances present in the target species' distribution and characteristics. In light of these considerations, while the models demonstrated good accuracy with the available data and techniques, future enhancements in data quantity and quality, as well as the incorporation of more sophisticated modeling techniques, could further elevate their performance. This suggests a promising avenue for advancing tree species classification models, leveraging more comprehensive datasets and refined analytical methods to achieve even greater accuracy and ecological insight.

The challenges we identified in our study regarding the limitations of using public or estimated coordinates in tree species classification are echoed and expanded upon in the broader context of forestry research. Both Wulder et al. (2024) and Coops et al. (2023) illustrate the significant benefits of combining precise NFI coordinates with advanced remote sensing techniques, proposing innovative, dynamic, and comprehensive approaches to forest inventory. This approach aligns with our findings, indicating a notable accuracy gap when using estimated coordinates. Hermosilla et al. (2022) further supports this, identifying geographic coordinates as one of the most important variables in their tree species classification model, emphasizing the complications in species classification that arise from less accurate spatial data. Fassnacht et al. (2024) explores the evolving role of Earth observation data in forestry, stressing the importance of high-resolution and precise spatial data in enhancing remote sensing applications. This perspective resonates with

our study's observation of the limitations of degraded coordinates, bolstering our call for the necessity of using precise NFI coordinates to improve the effectiveness of classification models. In summary, our study's challenges reflect a broader requirement in forestry and geo-information science for the integration of precise NFI coordinates into not only tree species classification models, but forest monitoring as a whole. We advocate for wider access and public sharing of NFI true coordinate data, recognizing its crucial role in enabling more accurate, efficient, and impactful forestry research. This collaborative effort in data sharing is vital for fully leveraging remote sensing technologies, fostering innovation, and effectively responding to environmental challenges and transformations.

5.5 Conclusions

In this study, we compared two different models for classifying forest tree species at high resolution in the Netherlands and focused on the outcomes derived from using the National Forest Inventory plot precise coordinates against those generated through an estimation procedure from the publicly available coordinates. Our findings highlighted several important points.

Firstly, a significant part of our research involved developing a method to identify a set of locations for each NFI plot potentially capturing the true location from public data. The estimation procedure demonstrated that it successfully incorporated the actual location pixel within the selected samples for only 10% of the cases, reflecting the inherent difficulty of the task. This result underlines the challenge and limitations inherent in using estimated or degraded data for ecological studies. This finding is critical, as it reflects the limitations of current data-sharing protocols and emphasizes the need for enhanced access to precise NFI data. Secondly, the comparison between the TCM and ECM revealed a stark difference in performance. While it was expected that the TCM would perform better than in classifying the tree species, the difference across the whole spectrum of performance metrics used highlights a considerable disparity between the two models and is a testament to the significant impact of coordinate precision on the accuracy of high resolution tree species models. The precise location data allowed for a more detailed, reliable prediction of tree species distribution, which is crucial for informed decision-making in forestry.

Moreover, the research shed light on the challenges of integrating remote sensing data with NFIs, particularly concerning the availability and precision of NFI data. The limited accessibility of precise, plot-level NFI data poses a significant barrier to maximizing the synergies between NFIs and remote sensing. This gap not only hampers the potential of these integrated approaches but also limits the ecological insights that can be derived from such models. The study's findings also have broader implications for forest inventory and management. The integration of precise NFI coordinates with advanced remote sensing techniques, as proposed by Wulder et al. (2024) and Coops et al. (2023), presents a promising direction for enhancing forest monitoring and management. This approach

aligns with our observation of the limitations of degraded coordinates and reinforces the necessity for true NFI coordinates to improve the effectiveness of classification models and forest inventories.

In conclusion, this research highlights the critical need for accurate spatial data for high resolution tree species classification. It also emphasizes the importance of improved data-sharing practices and collaboration in the field of forestry and remote sensing. This study demonstrates that while models using estimated coordinates can provide baseline insights, access to precise NFI data is essential for more accurate and impactful forestry research and environmental management. The study advocates for a synergistic integration of NFIs with remotely sensed data, recognizing the potential of such an integrated approach for advanced monitoring of forest resources. The findings call for a collective effort in data sharing and transparency, which is vital for leveraging remote sensing technologies effectively, fostering innovation in forest inventory, and responding to environmental challenges and transformations.

Chapter 6

Synthesis

This thesis has contributed towards advancing the exploration of integrating Earth observation data, machine learning techniques and field plot data to deepen our understanding of ongoing vegetation dynamics and improve monitoring and management of forest ecosystems. More specifically, the thesis contributed to this topic with assessing the impact of climate change on biomes distributions under different climate change scenarios and identifying potential hot spots of change (R.O.1) regardless of the upcoming climatic scenario. Furthermore, it helped with developing a data-driven framework combining EO data and ML techniques for high resolution mapping of both potential and realized tree species distributions (R.O.2) at a continental scale. Such methodology made possible the investigation of ongoing changes in forest ecosystems and the impact of forest disturbances on tree species distributions (R.O.3) at a continental scale, identifying areas with signs of ecological niche shifts. Lastly, it examined the effect of coordinate precision in field plot data for high resolution tree species classification modeling (R.O.4), showing the necessity of using precise field plot coordinates to fully exploit the opportunities offered by EO data and ML technologies in forest and environmental monitoring.

In this general discussion, the key findings and insights from the previous chapters will be synthesized to offer actionable guidance for forest modelers and remote sensing practitioners on effectively integrating EO data, ML techniques and field data. For this reason, the outcomes of this thesis are not limited to this dissertation only but also include open-source code, computational notebooks with tutorials, training datasets, blog articles and time series maps of biomes and tree species distributions. These resources, distributed under the CC BY 4.0 license, are freely accessible online, facilitating further research and application in the field.

The upcoming chapter is divided into two sections: the first summarizes contributions to the research objectives formulated in Section 1.3, comparing methodologies and findings; the second reflects on these contributions, offering perspectives on future research opportunities and collaborative endeavors.

6.1 Main findings

6.1.1 What is the impact of climate change on potential biomes distribution based on Earth observation and machine learning methods? And what are the projected shifts in vegetation under various climate change scenarios?

Research assessing the impact of climate change on vegetation typically employs two main approaches: global mechanistic vegetation models, which simulate the dynamics of vegetation including carbon fluxes, sinks and sources, and empirical models that establish correlations between climatic patterns and observed vegetation characteristics, such as biomass, species distributions or other physical traits. Empirical studies on this subject

typically use a mix of EO data, ML methods, and field data in their experimental setups, yet it is rare to find a study that integrates all three elements together. Studies that do not incorporate ML methods often produce results with coarse resolution and limited depth due to the absence of advanced computational analysis, resulting in moderate model performance and broad/general findings. Those excluding EO data typically face constraints in extending their findings across various temporal and spatial scales, hindering the ability to conduct comprehensive analyses. Meanwhile, research omitting field data and focusing solely on EO data and ML methods tends to offer conclusions with a degree of abstraction, as they rely on indirect vegetation indicators like NDVI, lacking direct observational support.

In Chapter 2, the developed ensemble ML model based on stacked regularization used for biomes distribution classification took advantage of climate projections derived from different Global Circulation Models (GCMs), EO data in the form of high resolution topographic layers and observational support on plant communities from field surveys. This comprehensive approach enhances model performances by leveraging the strengths of each data source in conjunction with the flexibility, scalability, and predictive power of ML models. This strategy not only surpassed previous studies in terms of accuracy despite the limited amount of training data, but also kept the experimental design relatively simple and free from assumptions, distinguishing it from other empirical studies. Given the limitations in model extrapolation and transferability common to ML models, addressing these through the spatial blocking during the cross validation, alongside the use of the margin of victory for model uncertainty quantification and the production of uncertainty maps, has been crucial. This approach allowed the spatial visualization of areas predicted by the model with high confidence, enhancing the practical utility of the outputs produced in the analysis despite its inherent limitations and offering a better understanding of where these outputs could be most trusted.

Applying a 50% margin of victory threshold to model outputs across the different climate scenarios and epochs, the analysis in Chapter 2 indicated that by the end of the century most of the spatial configuration of vegetation communities on the planet will stay the same; despite that, the tundra and the tropical-subtropical forest biomes were identified as some of the vegetation communities most at risk and susceptible to climate-induced changes. In fact, part of the land surface area occupied by such communities is projected to be occupied by either boreal forests, in the case of the tundra, or by savannas and grassland biome in the case of the tropical-subtropical forest biome. While the amount of land expected to change varies between climate change scenario and epoch, the methodology implemented in Chapter 2 allowed not only to identify which vegetation communities and region of the planet would be most vulnerable to changing climatic conditions but also pinpoints areas at the highest risk of significant shifts. It also allowed to identify regions within the same biome with higher risk, indicating that shifts in vegetation due to climate variations will not be uniformly distributed and some areas will be more affected than

others. Lastly, it allowed to identify specific locations that can be considered *hot spots* of change, meaning areas that are projected to shift from one biome to another regardless of the climatic scenario forecast; these *hot spots* are mostly located in North America, close to the Arctic Circle, and at the southern edges of the Amazon and Congo rainforests.

6.1.2 What combination of Earth observation and machine learning methods allows to map and analyze the distribution of forest tree species at high resolution?

To address this research question in Chapter 3, a more detailed data-driven framework than the one presented in Chapter 2 was developed to map the potential and realized distribution of 16 different tree species at European scale for the period 2000–2020. Instead of treating all 16 species as classes of a multiclass classification model like in Chapter 2, in this case each species was mapped independently and two different models were trained for each species, one for the potential distribution and one for the realized distribution. While the core of the framework was still the ensemble ML model based on stacked regularization, like in Chapter 2, given the significant computational resources needed to train two models and map the whole of Europe at 30 m resolution for multiple species, several improvements were incorporated in the modeling framework to achieve a good compromise between model performances and computational demands.

First, an automatic feature selection process was implemented to reduce the feature space; this step was necessary as the list of covariates used for the analysis included up to 300 different predictor variables, 4 times more than in the previous chapter. Secondly, before training the ensemble ML model we compared the model performances of 7 distinct common ML algorithms configurations, each undergoing independent hyperparameter tuning to ensure a fair comparison. This preliminary analysis allowed to keep a good coverage of the feature space while employing ML algorithms with different configurations or approaches (e.g. tree based versus regression, bagging versus boosting). This is also shown in model performances, where the ensemble model outperformed or performed as good as the best of the component models. After training the final ensemble model, the standard deviation of the predicted probabilities by the component models was used to quantify the model uncertainty at pixel level, to provide a measure of uncertainty in the probability space for the model predictions. Like for the analysis in the previous chapter, the provision of uncertainty maps is uncommon in this type of studies and represents an added value, offering map users insight into where and when the model outputs are most reliable. Not only this ML-based workflow was able to provide detailed and accurate predictions ($R^2_{\text{logloss}} = 0.857$, 0.839, respectively, for potential and realized distribution averaged across all 16 species) at different geographical and temporal scales, but is also fully reproducible and robust enough to be applied to multiple tree species.

The inclusion of EO data at different spatial resolutions was of paramount importance throughout different steps of the workflow. Firstly in the preprocessing phase: the field data recording tree species presence and absence used in the analysis was sourced from datasets with different degrees of spatial offset. Notably, public NFI plot coordinates coming from the EU-Forest dataset, potentially offset by up to 1 km, required careful handling. To address this, EO land cover layers were employed to discern and exclude plots located outside forested areas throughout the study period, ensuring the analysis focused only on relevant areas. This allowed to use even plots with imprecise coordinates for high resolution SDM. Secondly, during the modeling phase, and in particular in the variable importance analysis: the EO layers at 30 m (i.e. Landsat spectral bands and indices) proved to be the most important predictors across all 16 species for the ensemble model for realized distribution. The importance of EO high resolution data for such modeling exercises was furtherly assessed by conducting a comparative analysis. Models that integrated Landsat data were pitted against those reliant on up to 250 m resolution data. The findings were clear: models leveraging the Landsat data consistently outperformed their counterparts across all species and evaluated performance metrics. This comparison not only highlighted the superiority of high resolution EO data in modeling efforts but also reinforced the necessity of such detailed data for accurate species distribution modeling. The integration of high resolution EO data, ML techniques, and field data, even with partially degraded spatial coordinates, within a robust modeling framework, enabled accurate spatiotemporal mapping of forest tree species distributions across Europe.

6.1.3 How can these methods be applied to capture trends and disturbance impacts on forest tree species distributions and how do these reflect the ongoing changes in forest ecosystems?

When dealing with ecological niche shifts, SDM research traditionally focus on long term ecological forecasting based on climate change scenarios and equilibrium assumptions. This approach typically results in a bitemporal analysis over long time frames, which often overlooks ongoing dynamics in species distributions within shorter time periods and neglects potential insights gained from analyzing observed data using a multitemporal approach. While the bitemporal approach was preferred due to the lack of information on time series maps of species realized distribution, the integration of EO data and ML methods in SDM studies offers the chance for a more granular and data-drive approach. Forest disturbances significantly impact tree species realized distribution, yet while EO and ML technologies have advanced the detection, analysis, and spatial mapping of these disturbances, a comprehensive understanding of their effects on specific tree species distributions remains lacking.

Of central interest for this thesis was not only to provide clear methods, like in the previous chapters, but also to derive insights from the integration of EO data, ML methods and field data. For this reason, in Chapter 4, the realized distribution time series maps for 6 forest

tree species from Chapter 3 were analyzed to detect the presence of ongoing distribution shifts processes over Europe. These maps were combined with existing continental high resolution products mapping forest disturbances drivers and year of disturbance to assess how such disturbances have influenced the range of each species analyzed and to identify the dominant disturbance drivers. Having per-pixel information on the probability of occurrence every 4 years for the time period 2000–2020 allowed to discern whether there was an increase, decrease, or no significant change in trends for each species examined. Given the similar spatial resolution between the distribution maps and forest disturbance maps (i.e. 30 m resolution), this made it possible to categorize each pixel as either disturbed or undisturbed during the time period analyzed, and, if disturbed, to clearly assign disturbance type and year of disturbance. While findings on forest disturbances were analyzed at the native resolution of the products, to derive information on species ranges advancing or retreating results were aggregated at 1 km resolution. This step was taken to filter out weak signals from the analysis, thereby highlighting clear and robust patterns in species distribution changes. To make the qualitative spatial analysis even clearer, a 1 km map was produced for each identified phenomenon - specifically, increases or decreases in species probability of occurrence, as well as areas of stability. This approach facilitated a more detailed and understandable presentation of the spatial dynamics affecting species distributions.

The results showed that, across Europe and the time period analyzed, the probability of occurrence for all species studied predominantly displayed stability, indicating no significant changes in species distributions. These stable areas roughly align with the native range of each species, namely those places where the species have historically persisted and adapted over time; the identification of these areas highlight their importance for conservation purposes of the species analyzed. However, some general or species-specific hot spots of ongoing changes have been identified in the continent, with a prevalence of areas with decreasing probability of occurrence over those with an increasing ongoing phenomenon. Those areas were mainly located in Northern and Central Europe, particularly at each species latitudinal range edges or, in some cases, outside of their original native range due to human intervention. In such instances, species have been introduced or cultivated outside of their native ranges for various purposes such as wood production. Identifying these areas is crucial for distinguishing between localized incidents and broader range contractions, as they could indicate the onset of widespread changes or merely isolated events. Areas with an increasing ongoing phenomenon were found much rarer and scattered, showing no clear spatial pattern. Additionally, a steady increase in disturbance events was identified in each species range by disturbance (affected range doubled by the end of the time period analyzed, from 3.5% to 7% on average) and highlighted species-specific responses to forest disturbance drivers such as wind and fire. This thorough analysis of species ranges shifts and species-specific forest disturbances was made possible only by the extensive usage of EO data and ML methods. The insights garnered from this

application hold significant potential for implementing broad measures aimed at preserving the resilience of European forests, showcasing the pivotal role of advanced technologies in ecological conservation efforts.

6.1.4 What is the effect of coordinate precision in NFI data on the accuracy of high resolution tree species classification models?

To fully exploit the opportunities EO data and ML for high resolution ecological modeling, precise ground information is crucial. Imprecise coordinates can significantly compromise the ability to accurately match labeled or quantitative ground information in the feature space during the modeling phase, potentially skewing model outcomes. NFI data is one of the most comprehensive ground source for forest analysis but the needed access to the precise coordinates is often challenging due to privacy and legal constraints. Chapter 5 investigates the impact of using spatially degraded coordinates versus precise coordinates for the classification of dominant tree species in the Netherlands. To limit the confusion in the feature space, an estimation procedure was developed to identify a set of location potentially containing the precise coordinates for each NFI plot; while one model was trained with the precise coordinates, the other model was trained on the subsample of pixels coming out of this estimation procedure.

Results showed that the estimation procedure did not have much success: not only the procedure managed to include the precise coordinates in the training subsample only up to 10% of cases, but the model trained on this subsample of pixels achieved very low performance scores when compared with the one trained on the precise coordinates. Although the model using precise coordinates performed better, analysis of the PICP scores indicated an overconfindence in predictions for both models. This highlights that, despite employing advanced ML techniques, high quality field data and predictor variables, the task remains inherently complex; given these challenges even with the precise coordinates data, relying on publicly available coordinates only exacerbates the difficulty, highlighting the critical need for making the precise location accessible to researchers and practitioners. With greater accessibility, research could shift focus towards developing more accurate models to address similar ecological challenges, rather than navigating the added complexity of privacy-related data restrictions. This shift would significantly enhance the integration of EO data, ML technologies, and field data, streamlining the model development phase.

6.2 Reflection and outlook

This work was motivated by the growing need for methods capable to prepare and deliver timely, temporally continuous and spatially explicit data products of vegetation attributes at high spatial resolution. Leveraging the opportunities offered by EO data and ML algorithms, these methods and products are designed to complement traditional vegetation

modeling and monitoring approaches. In this context, this thesis has demonstrated the benefits of integrating EO data and ML methods with field data to improve our understanding of vegetation dynamics, be it either ongoing changes as explored in Chapter 3 to 5 or potential future shifts as in Chapter 2.

6.2.1 The role of integrated approaches for vegetation modeling

Understanding vegetation dynamics is crucial for economic and environmental sustainability, especially under the threat posed by climate change. Precise, timely knowledge of vegetation structure and composition, as well as the ability to make forward-looking projections, is essential for sustainable management of environmental resources. EO data provides wall-to-wall, spatially explicit characterizations of areas covered by vegetation regardless of management, ownership, or protection status, with up to several decades of spectral information (Wulder et al., 2019), while ground-based measurements, often in the form of extensive networks, provide detailed baseline information on the state of vegetation ecosystems and form the foundational level of most traditional vegetation monitoring frameworks (Fassnacht et al., 2024; Zweifel et al., 2023). Co-registration of ground-based measurements with EO data, coupled with the pattern-recognition ability in a multidimensional data space offered by advanced machine learning techniques, could be the backbone of new standardized methods for vegetation modeling and monitoring (Gessler et al., 2024; Zweifel et al., 2023), to the point that ML for satellite data has become a branch of ML in and of itself (SatML - (Rolf et al., 2024)). While ground-based measurements have historically been crucial for developing, calibrating, and validating EO-based approaches (e.g. FLUXNET (Baldocchi et al., 2001), U.S. FIA program (Lister et al., 2020), Swedish NFI (Olsson, 2015)), ML models developed from EO and field data are now transcending traditional monitoring methods. These models have evolved into comprehensive systems capable of monitoring, simulating, and predicting environmental and human activity interactions, showcasing their potential to redefine ecosystem analysis and management strategies.

In their review, Radeloff et al. (2024) discuss the importance of producing systematic satellite data products for monitoring and managing Earth's surface changes due to global climate and land use changes and categorize these products into five thematic areas, based on their relevance to sustainable management, societal benefits, and global change challenges. These products are further prioritized into three levels: *essential*, which are of the highest importance and currently feasible with mature algorithms; *desirable*, which are technically achievable but not as urgently needed; and *aspirational*, which require further research and development. Information on forest types and tree species is included in the *aspirational* tier: while some examples of national level of tree species information are reported by the review, an integrated approach based on EO data, ML algorithms and ground measurements for large scale tree species mapping has not been developed yet. Thus, to address this methodological gap, this thesis has contributed in providing a

framework for high resolution spatiotemporal mapping of presence-absence of individual tree species at continental scale (Chapter 3) alongside an analysis on ongoing tree species shifts (Chapter 4). Notably, the proof-of-concept examples highlighted in the review related to tree species mapping all rely on precise NFI plot coordinates for their machine learning frameworks; in this context, Chapter 5 has contributed with a practical demonstration on stressing the critical need for researchers to have access to such precise coordinates. For any progress to be made in this direction, such access is essential.

Contributions from Chapters 3 to 5, using free and open EO data and being full reproducible, stand to significantly improve national forest monitoring strategies specifically within the European Union. In fact, this alignment with the European Commission's initiative for a new forest monitoring system (European Commission, 2023) underscores the potential for these scientific advancements to be adopted and implemented across member states. This proposal from the European Commission seeks to integrate advanced EO technologies with ground observations to develop a comprehensive, high quality monitoring system. By providing a blueprint for such integration, this research facilitates the creation of a unified monitoring framework. Such contributions are pivotal for countries looking to enhance their forest resilience strategies and meet reporting requirements under EU policies, ensuring that decision-making is supported by the most accurate and comprehensive data available. This approach not only aligns with the EU's goals for forest monitoring but also exemplifies the thesis's practical impact on shaping future policies and strategies for sustainable forest management.

Forecasting the vegetation dynamics over the long term and at an appropriate spatial scale is essential for decision making to reduce the impacts of environmental disasters such as land degradation and drought. However, the response of vegetation to climate change remains poorly understood. Despite a globally consistent greening trend has been observed since 1980 (De Jong et al., 2011; Zhu et al., 2016), it is not possible for traditional linear approaches or mechanistic models, which rely on the assumption of stationarity, to detect nonlinear changes in vegetation and then forecast them into the future. Forward-looking projections using mechanistic models show large discrepancies between each other (Anav et al., 2013), but even when using approaches that allow to smooth these discrepancies like the multi model ensemble mean (MME) (Zhu et al., 2017), these models tend to overestimate the impact of the CO_2 fertilization effect on the greening trends (Zhu et al., 2016), effect which has been shown to plateau over the last decades (Higgins et al., 2023; Wang et al., 2020; Zhao et al., 2020). Integrated approaches combining EO data, ML methods and ground observations do not suffer from the same limitation (Ferchichi et al., 2022). A practical example of such integrated approaches for forward-looking projections was presented with the simulations conducted in Chapter 2: the chapter described a fully reproducible data-driven framework able to identify which vegetation communities and regions would be most vulnerable to changing climatic conditions that can be updated at any time given more field data observations or new GCM climatic projections. While

the model is relatively simple compared to mechanistic models providing similar type of information (Anderegg et al., 2022; Nolan et al., 2018) and does not include feedback loops, the predicted shifts are in line with ongoing greening and browning phenomena observed by EO data for the last four decades in those same areas (Berner and Goetz, 2022; Gatti et al., 2021; Higgins et al., 2023).

Contributions from Chapter 2 and 4 could be useful at European scale to reach the targets set by the new Nature Restoration Law (European Commission, 2022): this law sets binding targets to restore 20% of EU land and sea areas by 2030, extending to all ecosystems by 2050, to combat biodiversity loss and climate change. It requires member states to develop National Restoration Plans, focusing on habitats crucial for carbon storage and mitigating natural disasters, directly supporting the EU's biodiversity and climate goals. By offering a predictive analysis of vegetation community vulnerabilities and detailing current and future species shifts, contributions from this thesis could help ensuring that National Restoration Plans are both proactive and grounded in comprehensive scientific understanding.

6.2.2 Towards seamless integration: overcoming data challenges in vegetation modeling

Modeling and monitoring spatiotemporal dynamics of vegetation at both high spatial and temporal resolution is fundamental for informing decision-makers and shaping policy-making processes, providing insights for effective climate change mitigation and adaptation strategies. Recent advancements in in Earth Observation Programmes (EOP), like the Landsat and Sentinel series, have significantly improved vegetation analysis capabilities across a vast spectrum of scales, facilitating both local and global studies. The integration of ML with EO data, supported by advances in computing power and algorithm development, has simplified the analysis of large datasets, revealing patterns and insights in the data that were previously difficult if not impossible to detect. The implementation of a free data policy for major continuous EOPs has democratized data access, propelling the use of EO data for vegetation modeling applications and emphasizing the importance of open-access data in forest management and conservation efforts.

The development of both commercial (i.e. Google Earth Engine, Microsoft Planetary Computer, Amazon Web Services) and non-commercial cloud computing platforms (SEPAL, openEO, Joint Research Centre Big Data Analytics Platform), has revolutionized access to EO data and the application of ML algorithms for global scale analysis. These platforms offer robust infrastructures that eliminate the need for individual researchers to worry about physical infrastructure or computing power, making it significantly easier to perform comprehensive analyses. On these platforms, EO data is frequently provided in an Analysis Ready Data (ARD) format, streamlining the workflow by eliminating the need for extensive preprocessing and allowing researchers to dive straight into analysis (Dwyer et al., 2018;

Potapov et al., 2020). The SpatioTemporal Asset Catalog (STAC) further simplifies accessing EO data by standardizing how geospatial data is catalogued and queried, thereby enhancing the efficiency of data discovery and utilization across these cloud platforms (STAC Community, 2024).

Google Earth Engine (GEE), for instance, has emerged as a leading platform in this field, offering access to a vast archive of satellite imagery from multiple EOP and a high performance computing (HPC) environment to process and analyze data at scale. This has opened up new possibilities for monitoring environmental changes, land cover classification, and water resource management, among other applications: Venter and Sydenham (2021) for example reported producing a 10 m resolution land cover map for the whole of Europe in just 4 days of computing on a single research user account. The Horizon 2020 project openEO has proposed a harmonized API (Application Programming Interface) facilitating access to various cloud infrastructures using essentially the same source code (Schramm et al., 2021). By reducing complexity and fostering interoperability across platforms, these developments have democratized EO data analysis, allowing for in depth, cost effective environmental assessments accessible to a broader range of researchers and practitioners.

However, the effectiveness of these advanced tools is fundamentally constrained by the availability and quality of reference data. Despite the wealth of EO data and the analytical power of ML, the scarcity of accurate, high resolution ground information emerges as a critical bottleneck. Reference data are essential for calibrating and validating remote sensing models, yet they are often limited in spatial coverage, outdated, or inaccessible due to proprietary restrictions.

A notable example in literature that underscores the importance of the proper usage of reference data involved the analysis by Ceccherini et al. (2020), who used the Global Forest Change (GFC) maps from Hansen et al. (2013) to prepare remote sensing based estimates of harvested forest area over EU26, comparing these estimates with the national statistics on harvest removals officially provided by individual countries. Their findings revealed large inconsistencies between the remote sensing estimates and national statistics, particularly after 2015 and for Nordic and Baltic countries, with an increase in harvested area over the continent of about 49%. This claim started a long discussion in the media and in scientific journals, which ended only recently with the analysis provided by Breidenbach et al. (2022). By comparing the GFC maps with a series of 120,000 NFI plots from the Finnish and Swedish NFI, they demonstrated that the claimed increase in harvested area was actually due to improvements in the map's detection capabilities after 2015 and not an actual increase in harvest activities in either Finland or Sweden. This incident underscores a critical gap between the potential of EO and ML technologies and the limitations imposed by inadequate reference data. It accentuates the urgent need for comprehensive,

open access datasets that can support the next generation of forest monitoring solutions, ensuring that such analyses are grounded in reliable and up to date information.

In response, efforts like the Forest Observation System (FOS) (Schepaschenko et al., 2019) and the OpenForest catalogue (Ouaknine et al., 2023) have emerged, targeting the critical data challenges that hinder effective forest data collection and utilization. The FOS, a collaborative international effort, focuses on creating and sustaining a global database of in-situ forest biomass and canopy height measurements. In contrast, OpenForest aggregates 86 open access forest datasets from various spatial scales, enhancing machine learning applications in forest biology; for instance, contributions from Chapter 3 of this thesis are included within OpenForest. Additionally, the Global Biodiversity Information Facility (GBIF) compiles biological collections from a wide array of species, beyond just vegetation, supporting SDM applications, while the Global Forest Biodiversity Initiative (GFBI) amasses over 1.2 million plot records globally, with its data enabling the mapping of tree species diversity on Earth as demonstrated by Cazzolla Gatti et al. (2022). On a European scale, the EU-Forest dataset (Mauri et al., 2017) significantly expands available information on tree species distribution across Europe by harmonizing forest plot surveys from European NFIs for a total of more than 700,000 plots. The International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) (ICP Forests, 2010) was instead devised to provide information on air pollution on forests and collects another 18,000 plots across 42 countries within the European continent. Shifting from multinational towards national datasets, and specifically NFIs, a prime example of this category is the U.S. Forest Inventory and Analysis (FIA) program; this dataset is not only open access but also includes geospatially referenced information.

Given this wealth of open access datasets containing forest information, it might seem puzzling that ground reference data still represents a significant bottleneck for integrated approaches. The core of this issue lies in the accessibility of precise geographical coordinates for plot data from various networks and frameworks, especially NFIs. These datasets often withhold exact locations or provide them with diminished precision, severely limiting their utility for high resolution modeling and analysis as demonstrated in Chapter 5. All the open access datasets mentioned in the previous paragraph indeed provide spatially explicit information, but the highest level of detail for the such coordinates is an area of 1×1 km surrounding the precise plot coordinates; for some datasets the spatial resolution can be significantly coarser, with areas as large as 16×16 km (ICP Forests) or more. Having the precise coordinates for the co-registration with EO data is a fundamental step in the workflow of integrated approaches, as co-registration of EO data with spatially degraded coordinates leads to significant misalignments. These inaccuracies in spatial matching directly impact the feature space identified by ML models, resulting in models that, when trained on such compromised feature spaces, are often entirely inaccurate at high resolutions. The discrepancy between the actual and received by the model locations

distorts the model's understanding of vegetation patterns, leading to unreliable outcomes when applied to high resolution analyses.

A closer examination of the methodologies employed in Chapters 3 and 5 reveals differences that further illustrate these challenges. Despite seeming similarities, these chapters diverge not only in their input data points but also in their response variable. For instance, Chapter 3 used a harmonized dataset characterized by plots where the spatial coordinate precision varies (very precise for LUCAS, up to 1 km for EU-Forest and GBIF datasets), in contrast to Chapter 5, which employed precise coordinates and compared them with estimated/spatially degraded coordinates. Chapter 3 focused on mapping potential and realized distribution, while Chapter 5 was focused on dominant tree species classification: even though the response variable for "realized distribution" and "dominant tree species" might seem similar at first, they represent fundamentally different aspects of forest ecology and modeling purposes; by definition, a model used to solve a tree species classification problem faces always a more complex and overall needs more data and assumptions than a model used for a realized distribution problem. In Chapter 3, the analysis on potential distribution greatly benefitted from climatic data inputs which aligned closely with the spatial bias of the plot data, approximately 1 km; the resulting maps and spatial patterns, together with model performances and uncertainty quantification, could be considered reliable and of easy interpretation. However, the analysis on realized distribution may contain some challenges, the main one arising from mismatches between the resolution of the most important predictor variables, Landsat data at 30 m, and the plot data coordinates. While the misalignment in this case is not as impactful as for tree species classification problems, it may have caused optimistic model performances and uncertainty estimates. Furthermore, Chapter 3's methodology includes a direct use of distorted coordinates without attempting to reverse-engineer the degradation, a strategy that simplifies data processing but may perpetuate systematic errors from the original data sources. Using the precise location of NFIs plot data would have greatly improved the quality of outputs produced in this chapter.

After these considerations, a straightforward conclusion would be to publicly release the precise NFI plot locations; however, the situation is complex, as there is an ongoing debate in the scientific community regarding the openness of such data (Gessler et al., 2024; Nabuurs et al., 2022; Päivinen et al., 2023; Schadauer et al., 2024), with arguments on both sides. The concerns about releasing NFI plot locations revolve around the potential compromise of the randomness and representativeness of these plots. Critics argue that making geocoded plot data public could encourage changes in plot management, thereby introducing a bias that is difficult to detect. Permanent plot data derive their validity from being random samples, and revealing their exact locations could lead to targeted management practices or alterations by landowners aware of plot locations, thus skewing the data intended for unbiased national forest assessments (Päivinen et al., 2023). To address the need for using NFI data without publicly disclosing plot locations, several

solutions have been proposed. One approach is conditional access, where researchers can access precise plot locations under strict non-disclosure agreements that ensure the data is used responsibly and ethically. Since the procedure to access the data, even with data-sharing agreements, is quite bureaucratic and not straightforward, varying from country to country, another suggested solution would be to include the plot locations into secure data analysis platforms, enabling researchers to work with sensitive data within a controlled environment without having direct access to the raw data (Gessler et al., 2024); solutions such as having the organizations managing NFIs conducting the co-registration procedure with EO data, as identified by Schadauer et al. (2024), are impractical from a logistical perspective. Often, these organizations may lack the technical capacity required for co-registration across potentially hundreds of datasets and large scale, a task that becomes exponentially complex and resource intensive; for the same privacy reasons that would not allow them to share the plot locations, the co-registration cannot be executed on cloud computing platforms such as GEE. Additionally, the timeframe within which such tasks are processed frequently extends beyond practical limits for research, particularly in applications demanding timely or near real-time data analysis. For example, in monitoring projects where rapid data processing is essential, the co-registration request could take months to fulfill. By the time researchers receive the data, the situation on the ground may have changed, necessitating a restart of the data collection and analysis cycle. This lag not only hampers the agility and responsiveness of research but also risks rendering the findings obsolete before they can be applied, highlighting the need for more efficient data-sharing mechanisms that accommodate the dynamic nature of environmental research. For such purposes, the solution proposed by Schadauer et al. (2024) is not applicable and their claim that the quality of models would not be compromised has to be refused.

Going back to the controversy sparked by Ceccherini et al. (2020), the authors prepared a detailed response to the NFI analysis by Breidenbach et al. (2022), addressing each of the main claims. In their response, Ceccherini et al. (2022) acknowledged the vital role of NFI data in assessing forest resources but highlighted its significant limitations for validating and calibrating EO products. They argued that, despite adhering to the state-of-the-art methodologies for sample-based validation of EO products as outlined by Olofsson et al. (2014), the fundamental issue lies in the design of the NFI sampling schemes. These field protocols were not conceived with the integration of EO data in mind, posing challenges for their effective use in combination with remote sensing technologies, especially in the case of validation of land cover change products. This implies that even with access to precise NFI plot locations provided by the respective organizations, researchers could indeed conduct co-registration, training ML models and generate new EO based products of forest attributes; however, the independent validation of these products would be challenging, if not entirely unfeasible, depending on the forest attribute mapped and due to the inherent limitations of the NFI sampling schemes not aligning with EO data integration requirements.

To address these needs within the EU, several Horizon Europe projects have been initiated, each with distinct goals and methodologies aimed at enhancing forest monitoring and policy-making. PathFinder, which started in September 2022, focuses on enhancing the efficiency of forest monitoring by creating high resolution maps and precisely estimating forest attributes; in this context, one of the main goals of the project is to establish a new European sampling scheme that thoughtfully integrates NFI and ICP Forests plots for seamless EO data incorporation. MoniFun, started in January 2024, aspires to develop a harmonized European Forest Multifunctionality Monitoring System (EFMMS); in this context, one of its most relevant contributions would be the development of protocols and routines that facilitate the usage of NFI plot data in remote sensing applications while safeguarding the confidentiality of plot coordinates. Given that the projects are in their early stages, their impact and contributions to forest monitoring and ground data integration practices remain to be seen. While concrete outcomes are yet to emerge, these projects embody a forward-looking approach to address the challenges of plot confidentiality and ground data integration. Their progress and eventual findings will undoubtedly contribute valuable insights into effectively leveraging NFI data for integrated approaches using EO data and ML methodologies.

6.2.3 Enhancing the reliability of integrated approaches

Producing timely, continuous, and spatially explicit data products of vegetation attributes at high spatial resolution is essential for informing and guiding decisions and policies effectively. The advent of higher quality EO data, coupled with increased computational capabilities and usage of ML methods, has significantly simplified the task of generating predictive maps. However, the challenge lies in assessing the reliability of such maps, namely assessing the uncertainty estimates of the target vegetation attribute. Without the ability to produce trustworthy uncertainty estimates, the practical utility of these maps in decision-making processes is severely limited. The recent progress in remote sensing and computational technology has refined the presentation of vegetation attributes, making them more accessible to end-users. For species distribution, for instance, contemporary methods treat the suitability in the geographical space as a continuous spectrum instead than as a binary variable (i.e. suitable / unsuitable), assigning suitability percentages ranging from 0 to 100%, as illustrated in Chapter 3. Similarly, for identifying dominant tree species or other categorical variables, it is now common to deliver both the hard class map and accompanying probability maps that depict the likelihood of each species' presence across pixels, as demonstrated in Chapter 2. Additionally, there's a growing trend of accompanying these estimates with pixel-based uncertainty maps, enhancing the transparency regarding potential variances and bolstering confidence in the models, as evidenced in Chapters 2 and 3. While such products are now more frequently requested by the users, the practice of producing uncertainty maps is still uncommon.

Despite the utility of such new mapping methods, model-based approaches still serve only as supplementary tools in decision-making processes, lacking the robust statistical uncertainty estimates characteristic of traditional design-based approaches (McRoberts, 2006). In model-based approaches, statistical inference refers to the process of using the ML models to make predictions or estimates about a population, based on sample data: model-based inference relies on the assumption that the chosen model accurately represents the underlying data-generating stochastic process. This allows for the estimation of population parameters even in the absence of comprehensive sampling designs, by leveraging the relationships encoded within the model. The strength of model-based approaches lies in their flexibility to incorporate various data sources and their ability to provide the response variable in a spatially continuous form, making them especially valuable in contexts where traditional sampling data might be limited or unavailable. In those areas where there is reference data available from probability samples, design-based or model-assisted approaches can be used to derive such uncertainty estimates; however, in regions where probability samples of reference data is lacking or where data from multiple non harmonized sampling designs need to be integrated, the reliance on model-based approaches becomes not just a matter of convenience but of necessity.

For the former case, the utilization of globally produced mapping products, such as the GFC maps by Hansen et al. (2013), exemplifies the necessity of model-based approaches in regions lacking reference data. These maps, despite their well known limitations, have been instrumental for governments, researchers, and organizations to have a qualitative overview of deforestation trends, serving as critical tools for environmental monitoring and policy development in the absence of localized, officially sanctioned data. The latter case fits the previously examined European situation quite well, especially regarding the quantification and reporting of spatially explicit products of forest attributes. The European National Forest Inventory Network (ENFIN) was established in 2004 to promote the harmonization of NFIs across Europe, aiming to make data and estimates comparable across borders and over time (Vidal et al., 2016); while ENFIN has achieved good results in harmonizing forest biomass at a European level (Avitabile et al., 2024), the full potential of these advancements has yet to be realized in creating universally accepted mapping products, relegating many of the existing vegetation attribute maps to stratification tools (Ceccherini et al., 2022; Goodbody et al., 2023; Latifi et al., 2015; Latifi et al., 2012; McRoberts et al., 2012) rather than direct reporting resources, as also evidenced in the experimental designs of Chapter 3 and 5 of this thesis.

While model-based estimates offer valuable insights, they inherently require assumptions about the data and the model that can introduce biases or uncertainties, which are more challenging to quantify in a statistically rigorous manner (McRoberts et al., 2022). This potential for increased uncertainty makes model-based approaches less desirable for critical decision-making processes where the stakes are high and the demand for precision and reliability is paramount (McRoberts et al., 2010). Additionally, these approaches are

significantly limited by the quality and quantity of training data, where insufficient or low-quality data leads to inaccuracies: in such extrapolation cases, the model may not generalize well. For instance, in Chapter 2, the uneven geographical and class distribution in the BIOME 6000 dataset highlights the challenge of ensuring reliable predictions across different ecological zones. This uneven representation increases the risk of inaccurate predictions in unrepresented areas and could have caused the notably low performance metrics reported for the "*prostrate dwarf shrub tundra*" class.

For these reasons, stakeholders tend to prefer design-based approaches to model-based ones: design-based approaches derive their validity from the random selection of samples, ensuring that the results are representative of the population. The strength of this method lies in its ability to provide unbiased estimates of population parameters, as the random selection process minimizes systematic errors and biases that can distort findings. One of the main advantages of design-based statistical inference is its robustness against assumptions about the underlying population distribution. Since the inference is based on the design of the study rather than assumptions about the population, it can provide reliable results even when little is known about the population characteristics. This makes it particularly valuable in fields where the population distribution is unknown or difficult to model. The recent emphasis of the FAO on establishing good practices for design-based area estimation (Jonckheere et al., 2024), of the European Union on establishing a harmonized forest monitoring system (European Commission, 2023) and the allocation of substantial funding for projects like MoniFun reflect this preference for design-based confidence intervals and uncertainty estimates over model-based alternatives. This preference is rooted in the design-based approach's ability to provide robust and statistically reliable measures of uncertainty that are directly linked to the probability sampling design.

However, the requirement for a random sampling design can be difficult or costly to implement in practice and conducting such field sampling campaigns can be a limitation, particularly in remote or inaccessible areas. The inclination to invest in enhancing sampling designs, despite their associated high costs, originates from the imperative need for reliable data capable of supporting legally binding international reporting obligations and informing policy decisions (Gregoire and Valentine, 2007). In this context, integrating ML-produced maps with design-based and model-assisted estimators emerges as a compelling solution (Fassnacht et al., 2024). While model-assisted approaches effectively utilize these maps to refine predictions of environmental attributes, combining the detailed spatial predictions possible with ML with the statistical reliability of design-based sampling, they primarily enhance the precision of estimates derived directly from sampled data.

Hybrid inference (Corona et al., 2014) represents a further evolution, explicitly combining design-based, model-based, and model-assisted methodologies in a comprehensive framework. Unlike model-assisted approaches, which focus on reducing estimation variance

by leveraging predictions from external models, hybrid inference integrates these models with the probabilistic foundation of design-based sampling and the predictive power of model-based approaches. This integrated framework allows for a more holistic treatment of uncertainties by considering the variance from the sampling design, the model's predictive error, and any additional errors that may arise from mismatched spatial scales or misalignments between field data and remotely sensed inputs.

The studies conducted in the miombo woodlands of Tanzania (Næsset et al., 2020) and the Peruvian Amazon (Málaga et al., 2022) exemplify the superior capability of hybrid inference over traditional model-assisted approaches in enhancing the accuracy of aboveground biomass (AGB) estimates. In Tanzania, the calibration of global biomass maps using local NFI data under a hybrid inference framework not only substantially reduced systematic errors inherent in original remote sensing maps but also markedly improved the precision of AGB estimates. This improvement demonstrates hybrid inference's unique ability to simultaneously tackle both random and systematic errors, providing a more holistic error correction method than that offered by purely model-assisted approaches. Similarly, the application of hybrid inference in the Peruvian Amazon further underscored its potential to refine ML-derived AGB maps by comprehensively addressing a broad spectrum of uncertainties, including measurement errors, model assumption violations, and spatial heterogeneity. This rigorous uncertainty accounting led to more accurate, transparent, and reliable AGB stock assessments, essential for fulfilling stringent international reporting requirements and informing critical policy decisions. Although the precision gains appeared negligible due to the comprehensive incorporation of uncertainties, the resultant estimates were significantly more robust and trustworthy. Therefore, hybrid inference not only addresses the limitations of field sampling and the challenges of integrating EO data and ML with NFI data but also enhances the statistical reliability of the resulting estimates.

While the comparative review of design-based, model-based, model-assisted and hybrid inference methodologies reveals that each has its strengths and is best suited to particular scenarios in large-area forest surveys, there is compelling evidence to suggest that under the right conditions, hybrid inference emerges as a particularly powerful strategy. Specifically, when comprehensive datasets are available, including EO and ML-produced maps alongside robust NFI data, hybrid inference stands out for its ability to rigorously account for various sources of uncertainty, thereby enhancing the reliability of the resulting estimates. However, this endorsement does not diminish the value of the other approaches; rather, it highlights the importance of selecting the most appropriate methodology based on the survey objectives, data availability, and specific challenges at hand (Fig. 6.1). In sum, while recognizing the inherent strengths of all four methodologies, the advantages of hybrid inference under optimal conditions suggest that it could indeed offer the best strategy for increasing the reliability of EO data, ML tools, and NFI data-derived maps. This

conclusion is contingent upon a clear understanding of each method's applicability and the careful consideration of the survey's specific context and goals.

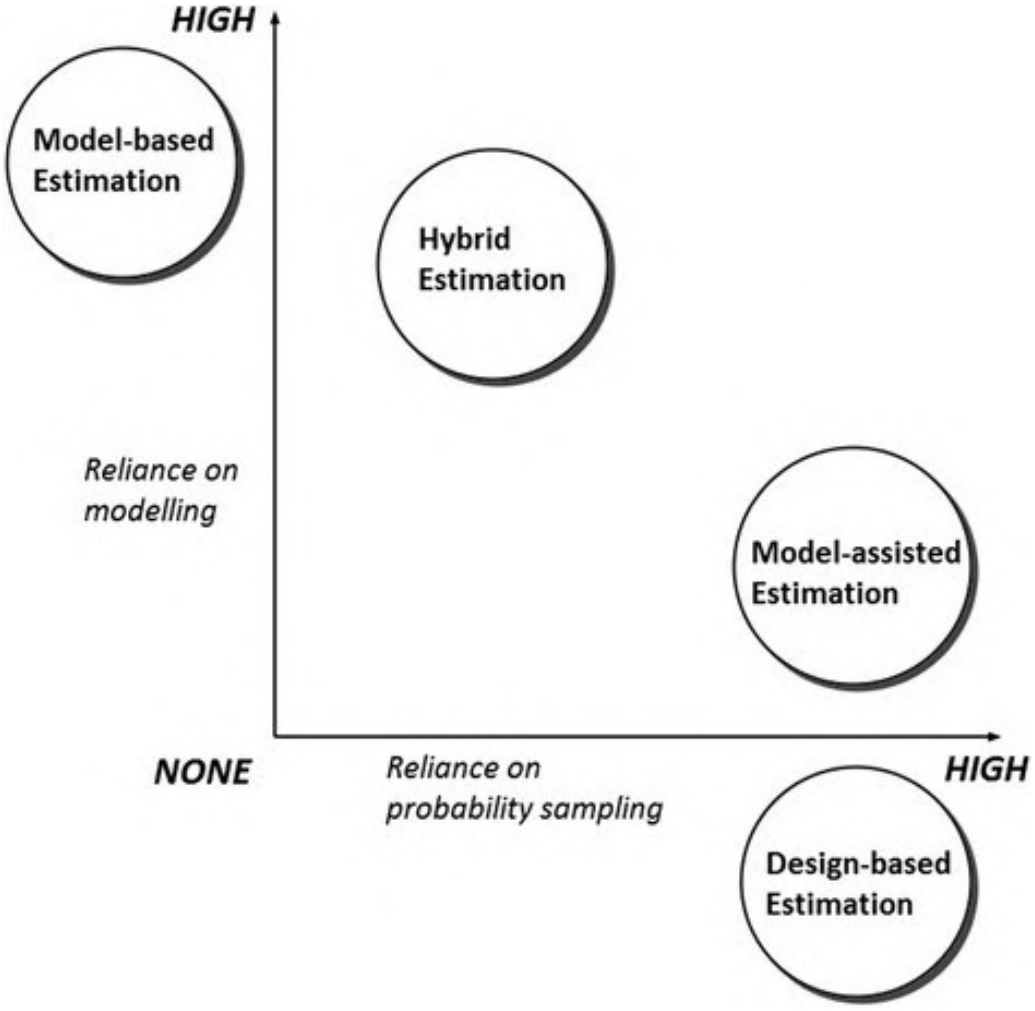

Figure 6.1: Diagram depicting the spectrum of statistical inference approaches for forest monitoring, illustrating the continuum from model-based estimation, which does not rely on probability sampling, to design-based estimation, which does not rely on modeling. Hybrid and model-assisted estimations are positioned between these extremes. From Ståhl et al. (2016).

6.2.4 Prospects for future research

In the constant evolving domain of vegetation modeling and monitoring, the integration of EO data, ML methods and field data has emerged as a pivotal approach for improving the production, accuracy and applicability of spatially explicit data products of vegetation attributes. This thesis highlights the ongoing need for advancements across these three key fields, not only to maintain but also to improve the quality and relevance of such maps. Most of the focus of this thesis regarding EO data has been on the Landsat program due to its longevity (more than 50 years of continuous observations at the moment of writing

this thesis (Wulder et al., 2022)); however, for near real-time monitoring applications, EO programs with a higher temporal resolution like Sentinel have been preferred. To combine the longevity of the Landsat program with the high temporal resolution of the Sentinel program, multi-sensor EO data collections such as the NASA Harmonized Landsat and Sentinel-2 (HLS) have been developed (Claverie et al., 2018b). NASA HLS harmonizes Landsat 8/9, Sentinel-2A and -2B data to a 30 m product from 2015 onwards in the Sentinel-2 tile grid. Landsat Next exemplifies a strategic push towards enhancing the synergy between NASA and ESA's EO programs, further streamlining the integration of their datasets for comprehensive global monitoring efforts. Scheduled for launch around 2030, Landsat Next represents the forthcoming phase in the Landsat program, following Landsat 9's 2021 deployment. This marks a significant advancement in the program: Landsat Next introduces a trio of satellites, each equipped with 26 spectral bands, more than its predecessors, which featured only 11 spectral bands (Fig. 6.2).

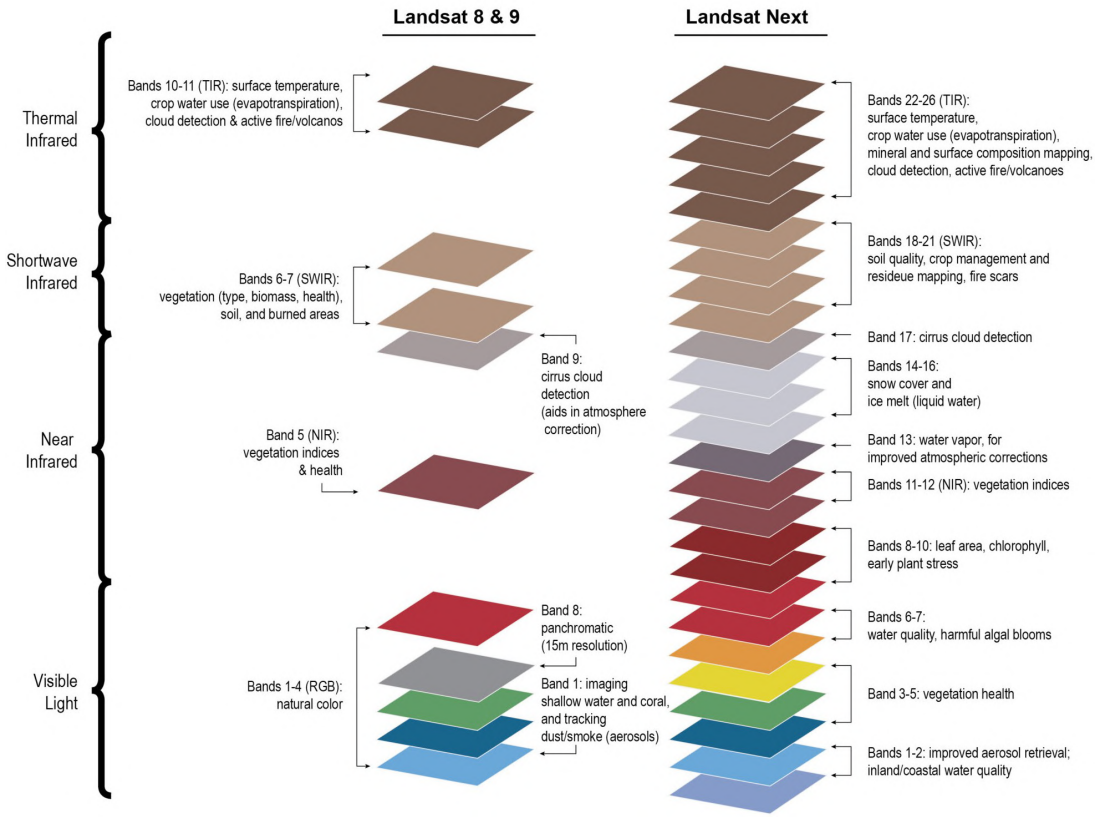

Figure 6.2: Spectral comparison between Landsat 8/9 and Landsat Next. Image credit: NASA Landsat Communications and Public Engagement Team.

This expansion allows for refined versions of "heritage" bands to maintain continuity with past data, alongside new bands aimed at enhancing Sentinel-2 data compatibility and supporting emerging applications. Spatial resolutions will be improved, with most bands having resolutions between 10 to 60 meters, facilitating more detailed observations. A key innovation of Landsat Next is its enhanced temporal resolution. The constellation,

with satellites spaced 120 degrees apart, ensures an aggregate 6-day revisit time for the globe. This is a substantial improvement over the 16-day interval of either Landsat 8 or 9, increasing the likelihood of acquiring cloud-free images and enabling more frequent monitoring of dynamic environmental changes.

Another interesting EO program that could improve the quality of vegetation mapping products is the German Environmental Mapping and Analysis Program (EnMAP); launched in 2022, this mission provides global scale hyperspectral data in the range of visible (418.2 nm) to shortwave infra-red region (2445.5 nm) of the spectrum across a total of 224 bands at a spatial resolution of 30 m. Its design allows for a rapid revisit time of just 4 days, promising detailed and frequent monitoring of environmental changes (Storch et al., 2023). The importance of hyperspectral data in mapping vegetation attributes, such as tree species classification, is significant (Fricker et al., 2019; Marconi et al., 2022; Nezami et al., 2020; Shen and Cao, 2017; Zhang et al., 2020). However, the challenge of achieving large-scale data coverage has been a major limitation. EnMAP, with its comprehensive hyperspectral capabilities, stands out as a particularly valuable tool in overcoming this challenge.

ML applications of Chapter 2 and 3 were focused on ensemble modeling, specifically in combining the predictions of multiple shallow ML models; ensemble modeling enhances prediction accuracy, reduces risk of overfitting and overall model variance, so it can be a way of tackling the issues of model generalization. In this sense, deep learning (DL) approaches could provide an alternative, if not more powerful, solution for several reasons. Unlike shallow ML models like decision trees, DL models employ multiple layers of nonlinear processing units, or *neurons*, to autonomously learn complex, hierarchical features from data. This methodology enables DL models to grasp and process data in a human-like manner, recognizing both intricate patterns and overarching concepts. Such an approach facilitates *transfer learning*, where a model, trained on one feature space, adapts its learned insights to a distinct, albeit similar, feature space by tweaking its task-specific layers while preserving the general knowledge of its initial layers (Hamrouni et al., 2021; Tuia et al., 2016). However, DL models often raise concerns among scientists for their "black box" nature, primarily due to their limited interpretability. There is a preference for models that, while possibly less accurate, offer greater transparency and thus, reliability, which severely limits further research in DL applications; to mitigate this, hybrid or physics-informed modeling presents a promising solution. These approaches integrate domain-specific knowledge or physical laws into the ML framework, enhancing model transparency and interpretability. By doing so, they bridge the gap between high accuracy and interpretability, offering models that are both effective and trustworthy (Ferchichi et al., 2022).

Forest ground data is essential for training and validation of remotely sensed based ML model: in this sense, Zweifel et al. (2023) proposed an innovative approach to enhancing

global forest monitoring by integrating existing forest plot infrastructures into a real-time monitoring network. They proposed using automated, standardized linking methods to connect monitoring sites, thereby overcoming the current fragmentation of forest monitoring efforts, enhancing the detection of patterns and facilitating near real-time assessments of forest conditions. Meanwhile, Nesha et al. (2021) reported substantial improvements in national forest monitoring capacities, showcasing a notable rise in the use of EO and NFI for accurate and comprehensive forest area monitoring. From 2005 to 2020, the adoption of EO methodologies expanded across 99 countries, while NFI utilization saw enhancements in 102 countries, collectively covering approximately 85% of the global forest area with high-quality monitoring data. This progress highlights a significant move towards mitigating the fragmentation of forest monitoring efforts. These developments collectively underscore the necessity of persisting in global efforts to enhance the accuracy and reliability of forest ground data, fundamental for the application of ML models reliant on EO data. The framework proposed by Wulder et al. (2024) for a satellite-based forest inventory (SBFI) for Canada exemplifies how the integration of EO data, ML methods, and field data can revolutionize forest resource assessment. Leveraging satellite observations to deliver a comprehensive, stand-level inventory across Canada's forested ecosystems, this approach not only represents a significant leap towards establishing a national standard for forest reporting but is also designed to be adaptable to other regions. It supports both national and international reporting requirements by offering detailed forest attribute data, paving the way for its adoption as a universal model for forest resource assessment.

As these technologies for vegetation modeling and monitoring evolve, they not only have the potential to enhance mapping capabilities, providing more accurate and timely data, but also offer an opportunity to drive policy innovation. Their progress provides a robust foundation for data-driven decision-making. The next step involves a concerted effort to bridge these technological capabilities with policy frameworks, ensuring that the insights gained from cutting-edge research directly contribute to the development of informed, effective environmental strategies.

This thesis brings to light a critical gap in the policy landscape, especially the imperative for seamlessly integration of NFIs with EO-based monitoring. While the EU forest monitoring proposal law (European Commission, 2023) signals a move towards incorporating remote sensing, the contributions from Chapter 5 highlight how enhanced integration of NFI data with EO technologies and ML models could revolutionize forest monitoring and underscore the inadequacy of current practices that rely on spatially degraded coordinate data. The provision of degraded coordinates starkly contrasts with the ambition for high resolution mapping products of forest attributes, highlighting a pressing need for policies that leverage precise, spatially explicit data for environmental monitoring. This necessity points toward an imperative shift in policy design, advocating for a more informed and technology-driven approach to meet the contemporary demands of forest resource assessment and conservation.

Supplementary material

Supplementary material for Chapter 3

Table S1: ODMAP protocol

ODMAP element	Contents
OVERVIEW	
Authorship	**Authors:** Carmelo Bonannella, Tomislav Hengl, Johannes Heisig, Leandro Parente, Marvin N Wright, Martin Herold, Sytze de Bruin **Contact email:** carmelo.bonannella@opengeohub.org **Title:** Forest tree species distribution for Europe 2000–2020: mapping potential and realized distributions using spatiotemporal Machine Learning **DOI:** `https://doi.org/10.7717/peerj.13728`
Model objective	**Objective:** Mapping/interpolation **Target outputs** continuous occurrence probabilities of potential and actual presence
Taxon	16 tree species native/commonly found in Mediterranean, temperate or boreal forests in Europe
Location	Europe
Scale of analysis	**Spatial extent (Lon/Lat):** Longitude -48.52° E, 60.65° E, Latitude 26.77° N, 72.02° N **Spatial resolution:** 30 m x 30 m **Temporal resolution and extent:** we modelled a single time slice for all potential distributions (2018–2020). We split the a 20 years time period (2000–2020) in 6 time slices (2000–2002, 2002–2006, 2006–2010, 2010–2014, 2014–2018 and 2018–2020) for realized distributions. **Type of extent boundary:** rectangular
Biodiversity data overview	**Observation type:** standardised monitoring and volunteer-based surveys **Response/Data type:** presence/absence data
Type of predictors	Climatic, topographic, lithological, hydrological, vegetation (binary species distribution maps of other tree species), remotely sensed (spectral reflectance)

Conceptual model / Hypotheses	We modelled realized distribution using environmental and remotely sensed variables, while we used only environmental variables for potential distribution. We wanted to test the influence of (i) remotely sensed and (ii) high resolution variables on predictive performances of realized distribution.
Assumptions	We assumed that species are at pseudo-equilibrium with the environment. We assumed that distribution maps of other tree species could be used as a proxy for interspecific interactions. Probability of presence is relative to the mapped target species, irrespective of the potential co-occurrence of other species in the same 30 m pixel and should not be confused with the absolute abundance or proportion of each species in the pixel area. The sum of the presence probabilities of different species in the same pixel can thus exceed 100 %
SDM algorithms	**Algorithms:** The final SDMs were fitted using three different algorithms: generalised linear models (GLM) with Lasso regularization, gradient boosted trees (GBT), and random forests (RF). **Model complexity:** We chose different modelling parameters to optimise each algorithm. Model settings were chosen to yield intermediately complex response surfaces but prevent excessive overfitting. **Ensembles:** The results from the three learners were combined in an ensemble model based on stacked generalization. We used logistic regression with Lasso regularization as ensemble model.
Model workflow	Hyperparameter tuning was conducted for all models on a a random 25% subset of observations. We evaluated each combination of hyperparameters by comparing logarithmic loss values during a 5–fold spatial cross validation replicated 5 times. Models were then fitted on the entire training dataset using the best (lowest logloss) combination of hyperparameters. Probability values obtained from the three models were then used to train the meta-learner. 5-fold spatial cross validation was used for the level 0 models. The out-of-fold predictions were used to build a level 1 training dataset for the meta-learner.

Software, codes and data	**Software:** All analyses were conducted using R version 4.1.1 (R Core Team, 2021b) with packages caret (Kuhn, 2021), data.table (Dowle and Srinivasan, 2021), dplyr (Wickham et al., 2021), minio.s3 (Leeper, 2017), mlr (Bischl et al., 2016), rgdal (Bivand et al., 2021), sf (Pebesma, 2018), sp (Pebesma and Bivand, 2005), spThin (Aiello-Lammens et al., 2015), stringr (Wickham, 2019), terra (Hijmans, 2021) and custom functions and Python version 3.8.6 (Van Rossum and Drake, 2009) with packages matplotlib (Hunter, 2007), numpy (Harris et al., 2020), pandas (McKinney et al., 2010) and scikit-learn (Pedregosa et al., 2011). **Code availability:** All codes are available on GitLab (`https://gitlab.com/geoharmonizer_inea/spatial-layers/-/tree/master/veg_mapping`) **Data availability:** All data is available through Zenodo (**Bonannella** et al., 2022b)
DATA	
Biodiversity data	**Taxon names:** *Abies alba* Mill., *Castanea sativa* Mill., *Corylus avellana* L., *Fagus sylvatica* L., *Olea europaea* L., *Picea abies* L. H. Karst., *Pinus halepensis* Mill., *Pinus nigra* J. F. Arnold, *Pinus pinea* L., *Pinus sylvestris* L., *Prunus avium* L., *Quercus cerris* L., *Quercus ilex* L., *Quercus robur* L., *Quercus suber* L., *Salix caprea* L. **Details on taxonomic reference system:** standard biologic taxonomy **Ecological level:** individual **Biodiversity data source:** Global Biodiversity Information Facility (GBIF), EU-Forest (Mauri et al., 2017), LUCAS (EUROSTAT, 2017) **Sampling design:** uniform (EU-Forest, LUCAS) or unknown (GBIF); latest instance in case of multi-temporal observations **Sample size per taxon:** Abies alba Mill. (), Castanea sativa Mill. (), Corylus avellana L. (), Fagus sylvatica L. (), Olea europaea L. (), Picea abies L. H. Karst. (), Pinus halepensis Mill. (), Pinus nigra J. F. Arnold (), Pinus pinea L. (), Pinus sylvestris L. (), Prunus avium L. (), Quercus cerris L. (), Quercus ilex L. (), Quercus robur L. (), Quercus suber L. (), Salix caprea L. () **Country mask:** European Union, Iceland, Norway and Switzerland

	Data cleaning/filtering: Point locations were filtered using a high-resolution land mask for Europe and by leveraging existing quality flags indicating serious location issues. A second filtering step involved the usage of yearly high-resolution forest masks. To overcome the uneven sampling intensity and potential point clustering, we applied an additional spatial filter to the presence points using the spThin R package. A distance of 2 km was considered as minimum distance between the points, to harmonize the sampling intensity between presence and absence data. The procedure was repeated 10 times: at each iteration, the algorithm randomly removes one observation from the dataset until no observation is left with a nearest neighbor closer than the thinning distance. Among the 10 datasets obtained, the one with the largest number of records was retained and used for modeling. **Absence data:** We used the land cover points from LUCAS database as absence data. All land cover classes except class C (woodland) were used as absence data for realized distribution, while for potential distribution points belonging to class A (artificial land) and class B (cropland) were also excluded. Other tree species presence locations were also used as absence data: these points were first overlaid with a chorological map of the target species and only the ones falling outside the species range were then used as absence. **Potential biases:** Spatial observation density for presence data varies greatly throughout the study area and across different species resulting in over- or underrepresented areas.
Data partitioning	A 5-fold spatial cross validation was used to partition the data for model fitting. Predictive performances were assessed using a 5-fold spatial cross validation repeated 5 times. Spatial allocation for the points was conducted using spatial blocking with a blocking factor of 30 km.
Predictor variables	**Predictor variables:** • *Climate:* 19 bioclimatic predictors, solar direct and diffuse irradiation, 13 cloud fraction layers, 12 snow probability layers, 12 water vapor pressure layers, 12 wind speed layers, time series of precipitation, monthly time series of surface temperature (minimum, average, maximum), monthly time series of air temperature (minimum, average, maximum) • *Topography:* elevation, slope, aspect, positive and negative openness, eastness and northness, hillshade

•*Hydrology:* surface water occurrence, height above nearest drainage (HAND), flow accumulation area and long-term flood hazard map
•*Vegetation:* 50 chorological tree species maps
•*Remotely sensed:* 7 seasonal time series of Landsat bands (blue, green, red, NIR, SWIR1, SWIR2 and thermal), 7 seasonal time series of spectral indices (EVI, EVI2, MSAVI, NBR, NDVI, NDWI, SAVI), long term bare ground cover.

Data sources:

•*Climate:* bioclimatic predictors were downloaded from `https://chelsa-climate.org/bioclim`, solar direct and diffuse irradiation from `https://globalsolaratlas.info/download`, 13 cloud fraction layers from `https://www.earthenv.org/cloud`, 12 snow probability layers from `https://doi.org/10.5281/zenodo.5774953`, 12 water vapor pressure layers from `http://www.worldclim.com/version2`, 12 wind speed layers from `https://www.climatologylab.org/terraclimate.html`, time series of precipitation, monthly time series of surface temperature (minimum, average, maximum), monthly time series of air temperature (minimum, average, maximum) were downloaded and reprocessed from reprocessed Copernicus ERA5 data (`https://doi.org/10.24381/cds.bd0915c6`)
•*Topography:* elevation, slope, aspect and hillshade were downloaded from `https://doi.org/10.5281/zenodo.4724549`, positive and negative openness, eastness and northness from `https://doi.org/10.5281/zenodo.4486135`
•*Hydrology:* surface water occurrence was obtained from `https://global-surface-water.appspot.com/`, height above nearest drainage (HAND) and flow accumulation area from `http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/` and long-term flood hazard map from Dottori et al. (2016)
•*Vegetation:* 50 chorological tree species maps were downloaded from `https://forest.jrc.ec.europa.eu/en/european-atlas/atlas-data-and-metadata/`
•*Remotely sensed:* time series of Landsat bands were downloaded from `https://glad.umd.edu/ard/landsat-ard-download` and the spectral indices were derived from them; the long term bare ground cover was downloaded from `https://glad.umd.edu/dataset/global-2010-bare-ground-30-m`

Spatial resolution and spatial extent of raw data: Original resolution of raw data goes from 5 km to 30 m. Data were collected and overlaid with the points at their original resolution and then resampled at 30 m resolution for map production. Data with global coverage were clipped at the spatial extent of the scale of analysis.
Map projection: All layers were reprojected in EPSG:3035 coordinate reference system
Temporal resolution and temporal extent of raw data: Data with long-term aggregates or considered as not changing for the temporal scale of the analysis (i.e. elevation) were treated as static variables. The long-term aggregates cover the period 1979–2013 (bioclimatic layers) or 2000–2016 (cloud fraction, snow probability, water vapor). Landsat data and derived spectral indices cover the period 1999–2020, with a temporal resolution of 16 days. Time series of precipitation, surface temperature and air temperature cover the period 2000–2020 with a hourly temporal resolution.

	Data preprocessing: Cloud and cloud shadow pixels were removed from Landsat images, maintaining only the quality assessment-QA values labeled as clear-sky. Individual images were averaged by season according to three different quantiles (25th, 50th and 75th) and the following calendar dates for all period: winter (December 2 of previous year until March 20 of current year), spring (March 21 until June 24 of current year), summer (June 25 until September 12 of current year) fall (September 13 until December 1 of current year). Missing values were imputed using the *Temporal Moving Window Median* algorithm. Spectral indices were computed for each year and season using the 50th quantile only. Time series of precipitation, surface temperature and air temperature were aggregated on a monthly scale. The following steps were used for temperature data: (i) aggregate CHELSA to ERA5 spatial resolution, (ii) calculate difference between ERA5 Land and aggregated CHELSA, (iii) interpolate differences with a Gaussian filter to 30 arc seconds and (iv) add the interpolated differences to CHELSA. A different approach was used for precipitation, with proportions instead of differences: using proportions ensures that areas without recorded precipitation remain areas without precipitation; only in the case of actual precipitation in a given area, precipitation was redistributed according to the spatial detail of CHELSA: aggregate CHELSA to ERA5 spatial resolution, calculate proportion between ERA5 Land and aggregated CHELSA, interpolate proportion with a Gaussian filter to 30 arc seconds, multiply the interpolated proportion with CHELSA.
MODEL	
Multicollinearity	Collinearity analysis was not conducted.

Model settings	**Model settings:** Grid search with a 5 step was used to conduct hyperparameter tuning. Hyperparameters used in the final models change based on the species and the distribution type. Common settings across all models was the *"probability"* mode, to have a probabilistic (0–100) output for presence and absence. Spatial 5-fold cross validation was used in the inner loop (training of the three component models for the ensemble) and in the outer loop (training of the meta-learner). All random forests model were trained with the same number of trees ($ntree = 85$) due to computational constraints. All GLM models used in the final model were fitted with the $\lambda_{\min}$. Lasso regularization was used for the GLM component model and for the meta-learner (logistic regression).
Model estimates	We did not analyzed the coefficients of the meta-learner in depth.
Model selection / Model averaging / Ensembles	We used an ensemble model based on stacked generalization (Wolpert, 1992). Outputs made by the component models are the input of a meta-learner which then produces the final prediction.
Non-independence correction/analyses	None
Threshold selection	None
ASSESSMENT	
Performance statistics	Predictive model performance on validation data (based on the 5-fold spatial block cross-validation repeated 5 times) was assessed using two different performance measures: area under the receiver operating characteristic curve (AUC) and logarithmic loss (log loss)
Plausibility check	Produced maps were compared with distribution maps from European Atlas of Forest Tree Species.
PREDICTION	
Prediction output	**Prediction unit:** We used continuous predictions of occurrence probability per species.
Uncertainty quantification	Uncertainty for the final ensemble model was accounted as the standard deviation of the predicted probabilities of the base learners. The principle is that the higher the standard deviation the more uncertain the model is towards the right value to assign to the pixel.

Table S2: Number of presence points before and after the thinning operation.

Species	Distribution	Presence	Absence	Presence thinned	Absence thinned	Prevalence	Prevalence thinned
A. alba	Potential	45422	793484	38622	264833	0.05	0.15
C. sativa	Potential	77363	796588	35068	327986	0.09	0.11
C. avellana	Potential	32134	533362	28436	120445	0.06	0.24
F. sylvatica	Potential	197107	750752	147655	240030	0.26	0.62
O. europaea	Potential	50656	667416	5794	418689	0.07	0.01
P. abies	Potential	360112	820472	284594	270365	0.44	1.10
P. halepensis	Potential	233961	916627	41758	446998	0.26	0.09
P. nigra	Potential	139497	1091032	29711	467060	0.13	0.064
P. pinea	Potential	239254	853945	19314	464513	0.28	0.042
P. sylvestris	Potential	507664	763571	283882	279289	0.66	1.00
P. avium	Potential	22852	575044	21616	184178	0.04	0.12
Q. cerris	Potential	13774	1106501	13072	450317	0.01	0.03
Q. ilex	Potential	57213	1045184	36294	422799	0.05	0.09
Q. robur	Potential	113044	655526	97113	181057	0.17	0.54
Q. suber	Potential	419975	840378	12877	454942	0.5	0.03
S. caprea	Potential	45753	488610	42544	75030	0.094	0.57
A. alba	Realized	45422	1240411	32144	455595	0.037	0.07
C. sativa	Realized	77363	1215288	29985	528953	0.064	0.06
C. avellana	Realized	32134	1044035	22919	383698	0.031	0.06
F. sylvatica	Realized	197107	1196053	128325	457652	0.16	0.28
O. europaea	Realized	50656	1019575	5249	597376	0.05	0.01
P. abies	Realized	360112	1207099	260011	448564	0.3	0.58
P. halepensis	Realized	233961	1261372	36844	598154	0.19	0.06
P. nigra	Realized	139497	1425066	28380	608094	0.098	0.05
P. pinea	Realized	239254	1192136	17181	615999	0.2	0.03
P. sylvestris	Realized	507664	1178176	239508	427508	0.43	0.56
P. avium	Realized	22852	1058215	17093	430548	0.022	0.04
Q. cerris	Realized	13774	1455524	11536	602902	0.0095	0.02
Q. ilex	Realized	57213	1401501	32826	580366	0.041	0.06
Q. robur	Realized	113044	1156574	76487	433485	0.098	0.18
Q. suber	Realized	419975	1188462	12538	610216	0.35	0.02
S. caprea	Realized	45753	1007707	36704	342866	0.045	0.11

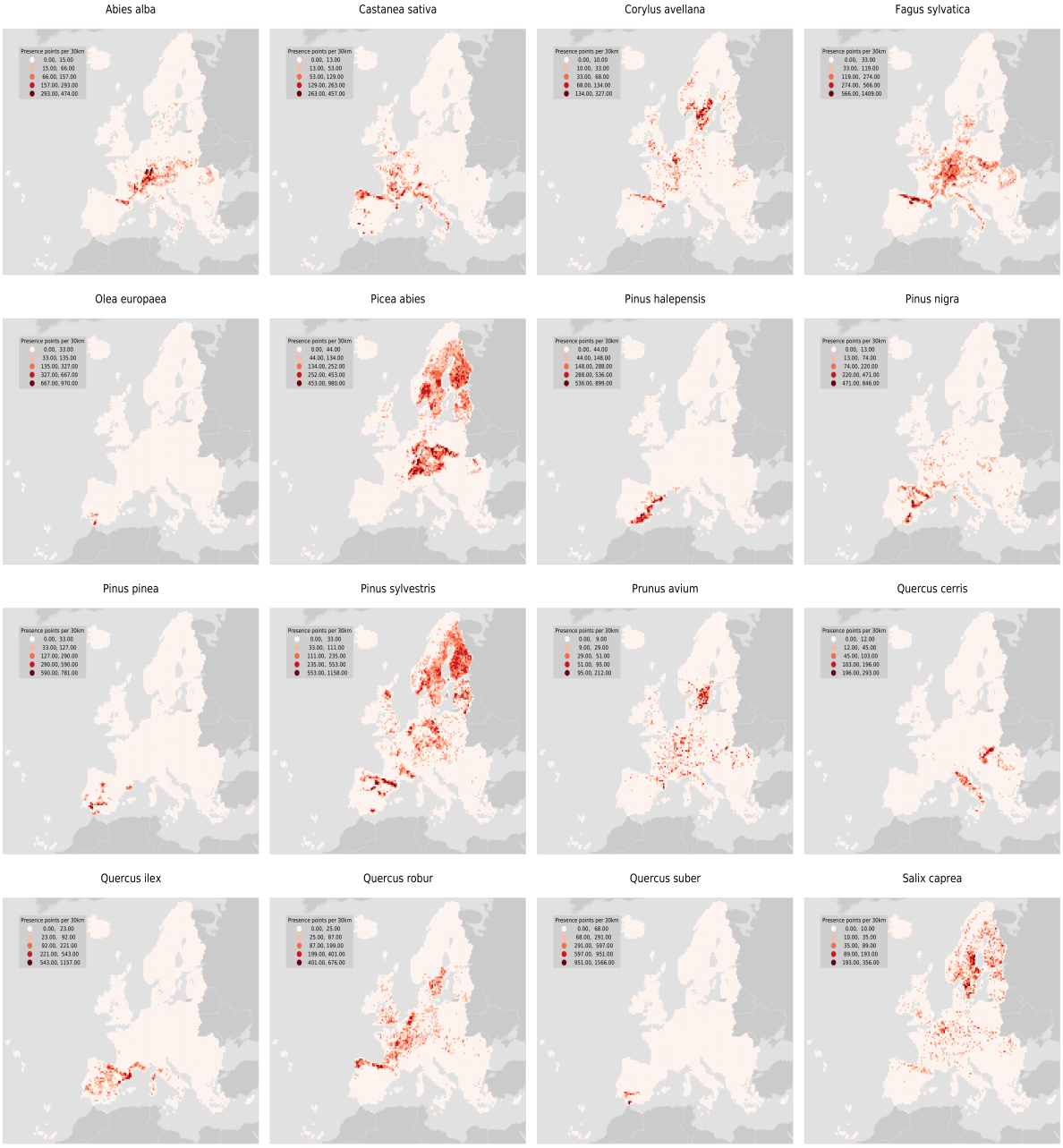

Figure S3: Distribution of presence points per species after thinning. Absence points are omitted for visualization purposes.

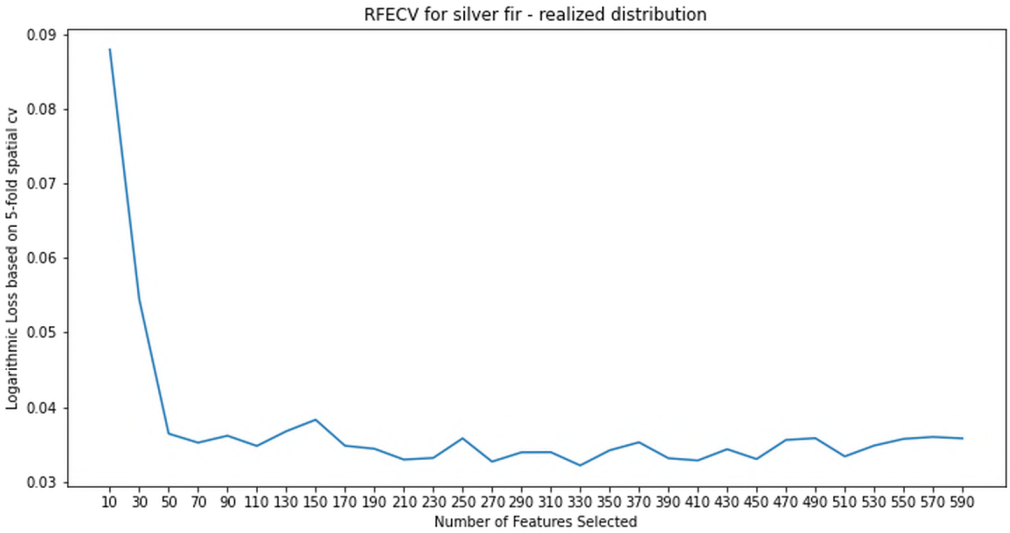

Figure S4: Log Loss performances by number of selected features.

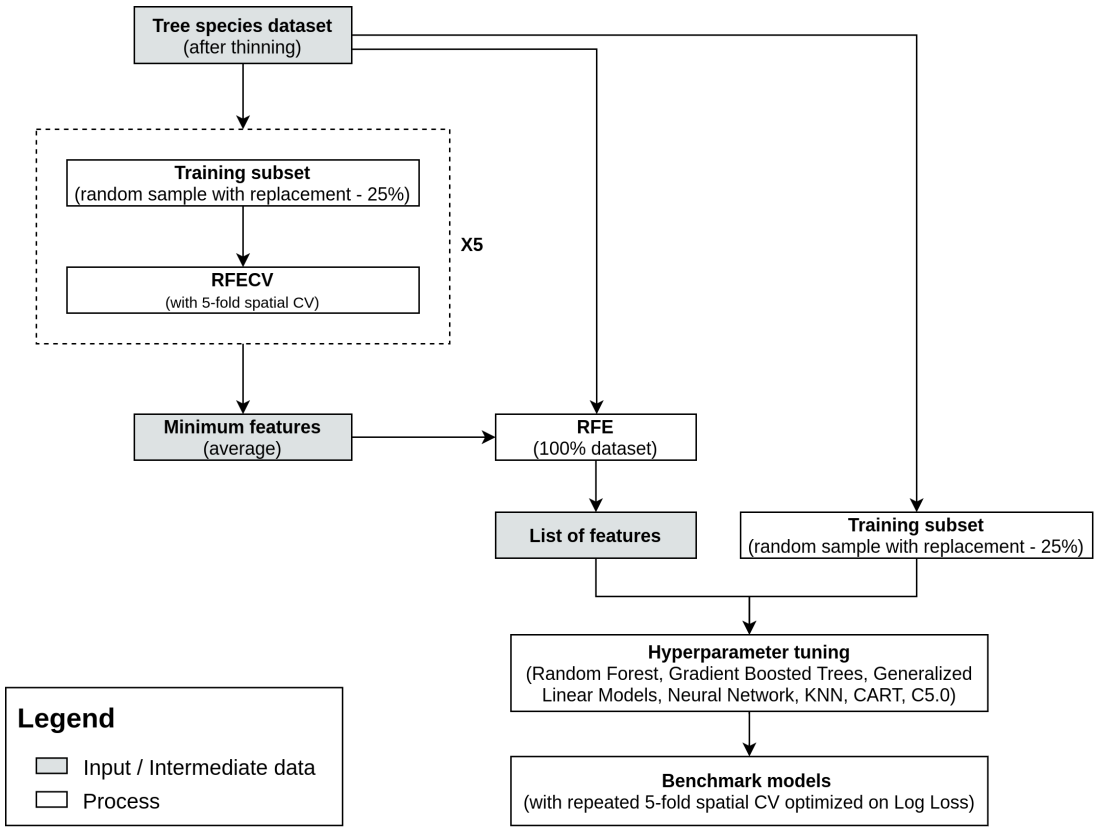

Figure S5: Example workflow illustrating the feature selection and benchmarking process for one species

Table S3: Hyperparameter space for the analyzed algorithms. In light gray the name of the R package used to implement the algorithm is reported, in brackets the name of the algorithm. p refers to the number of predictor variables, while columns *Lower* and *Upper* indicate the bounds of the regions in the hyperparameter space. Due to computational constraints, we set the *num.trees* parameter for Random forest to 85. The tested activation functions for the neural network were *sigmoid* and *tanh*, while for the output we tested both *sigmoid* and *softmax*. For GLM we used the automatically generated λ sequence and selected the $\lambda_{\min}$.

Algorithm	Hyperparameter	Type	Lower	Upper
C5.0				
(Classification trees)	minCases	integer	0	10
	CF	numeric	0	0.5
kknn				
(k-nearest neighbor)	k	integer	1	50
deepnet				
(Artificial neural network)	learning rate	numeric	0.0001	0.00001
	numepochs	integer	10	20
	batchsize	integer	50	150
	hidden_dropout	numeric	0.1	0.3
	activationfunction	discrete	-	-
	output	discrete	-	-
	momentum	numeric	0	0.05
	number.of.layers	integer	2	4
	units	integer	32	64
ranger				
(Random forest)	mtry	integer	$\sqrt{p}/3$	p
rpart				
(CART)	minsplit	integer	20	25
	minbucket	integer	5	10
	cp	numeric	0.01	0.1
	maxcompete	integer	3	4
	maxsurrogate	integer	4	5
	usesurrogate	discrete	-	-
	surrogatestyle	discrete	-	-
	maxdepth	integer	5	15
xgboost				
(gradient-boosted trees)	nrounds	integer	10	20
	max_depth	integer	3	5
	eta	numeric	0.01	0.1
	subsample	numeric	0.5	0.9
	min_child_weight	integer	10	20
	colsample_bytree	numeric	0.5	0.9

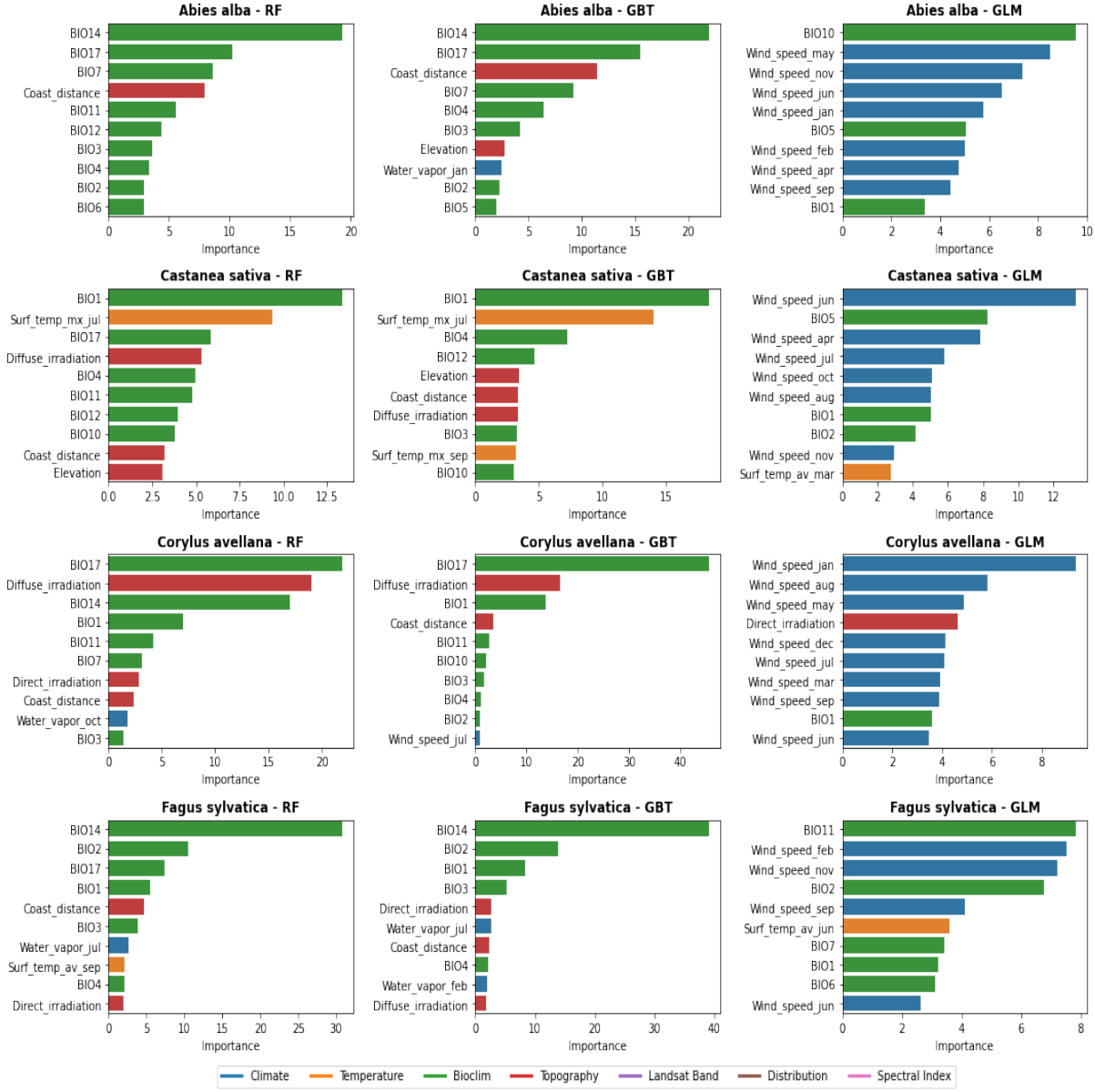

Figure S6: Relative importance of the 10 most important variables across the three component models for **potential** distribution of *A. alba, C. sativa, C. avellana, F. sylvatica*

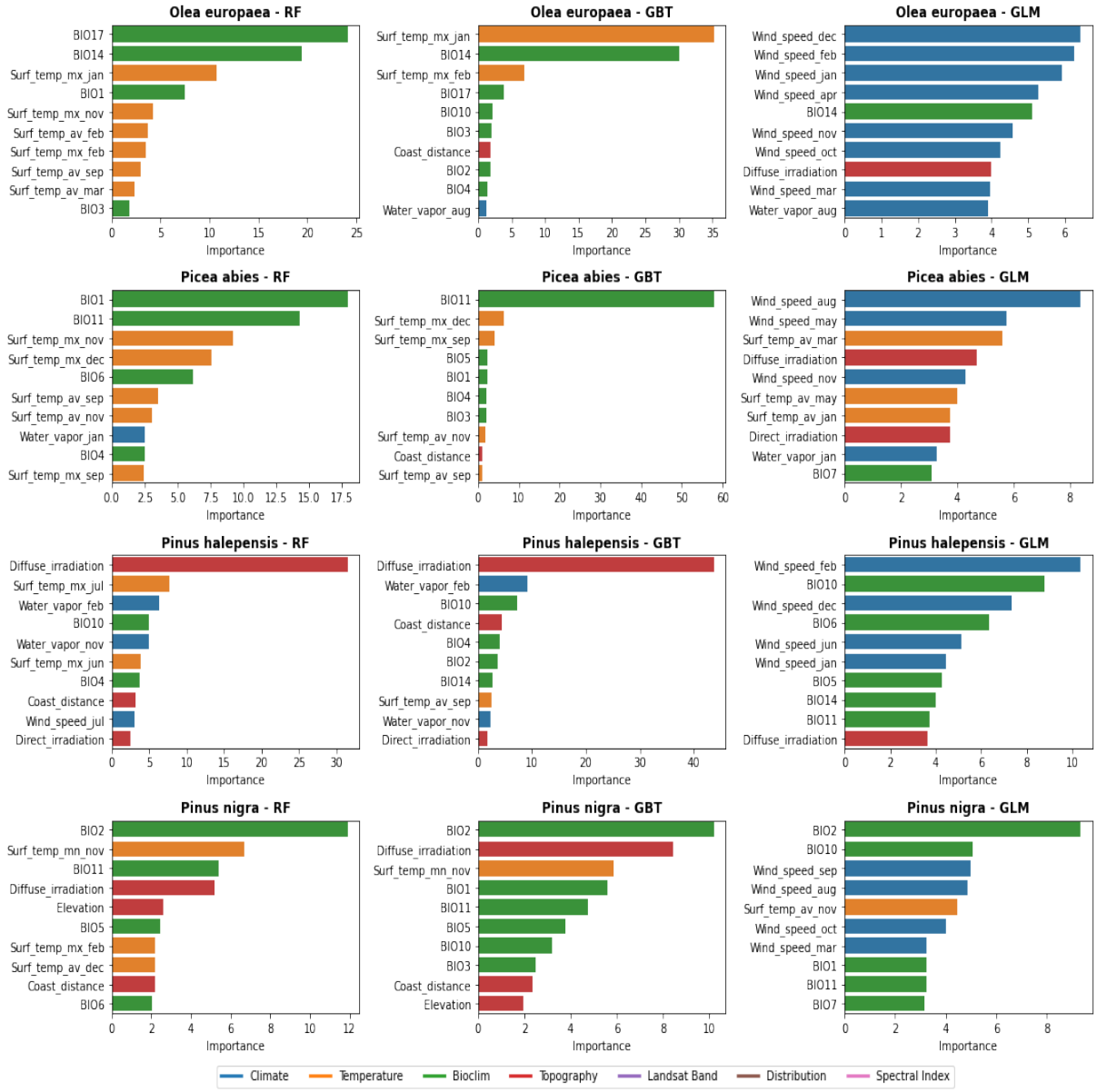

Figure S7: Relative importance of the 10 most important variables across the three component models for **potential** distribution of *O. europaea, P. abies, P. halepensis, P. nigra*

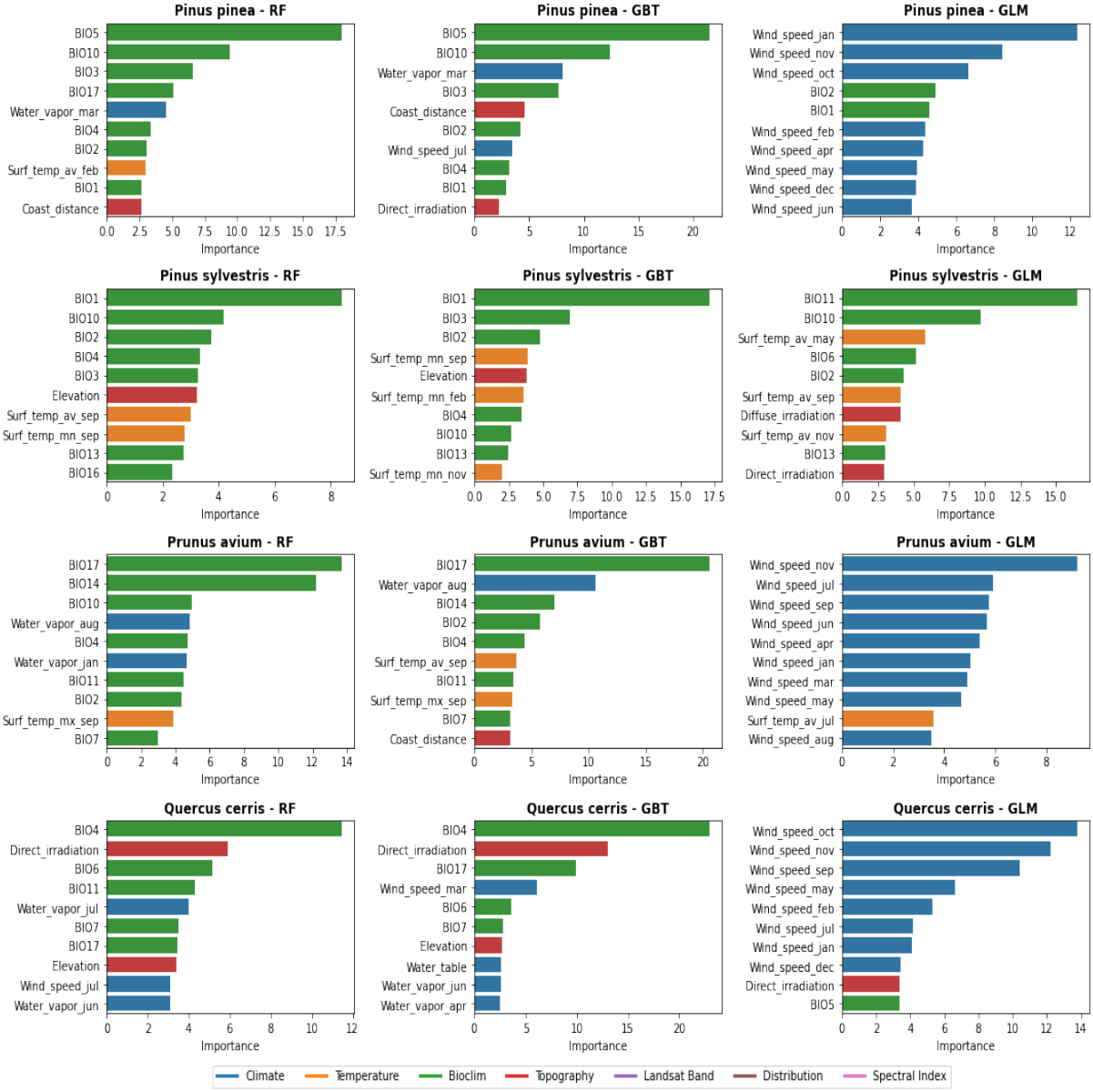

Figure S8: Relative importance of the 10 most important variables across the three component models for **potential** distribution of *P. pinea, P. sylvestris, P. avium, Q. cerris*

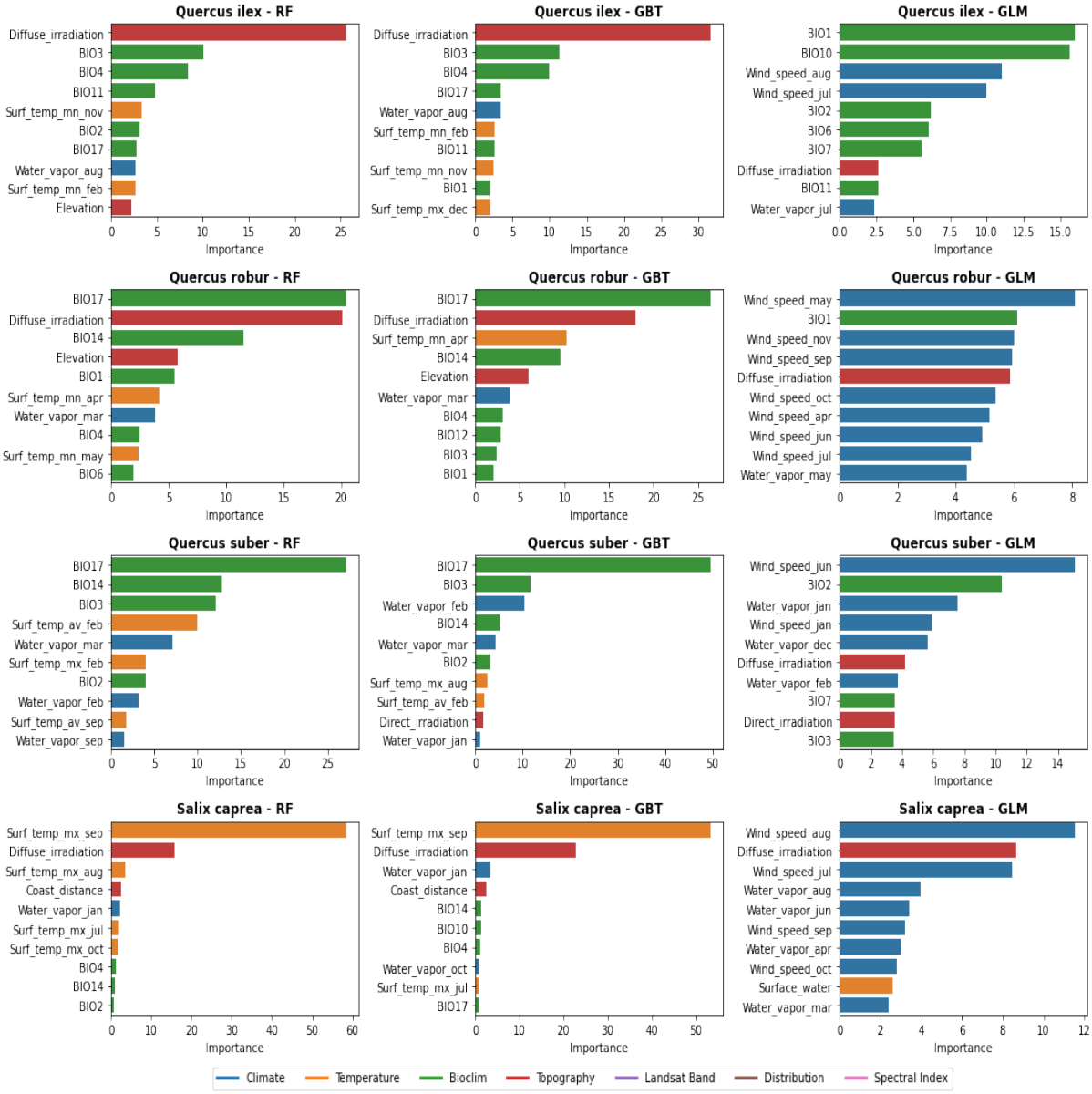

Figure S9: Relative importance of the 10 most important variables across the three component models for **potential** distribution of *Q. ilex, Q. robur, Q. suber, S. caprea*

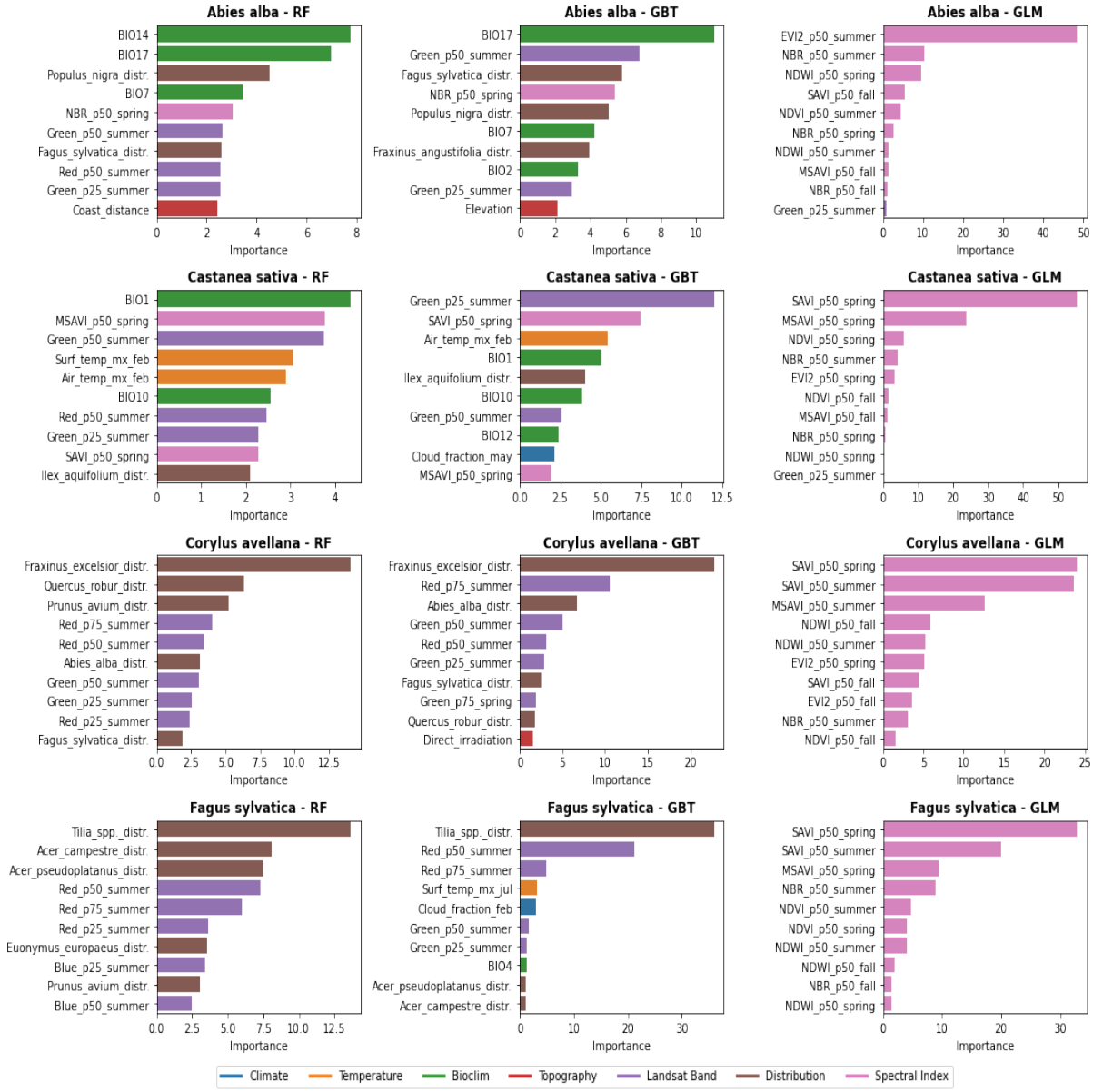

Figure S10: Relative importance of the 10 most important variables across the three component models for **realized** distribution of *A. alba, C. sativa, C. avellana, F. sylvatica*

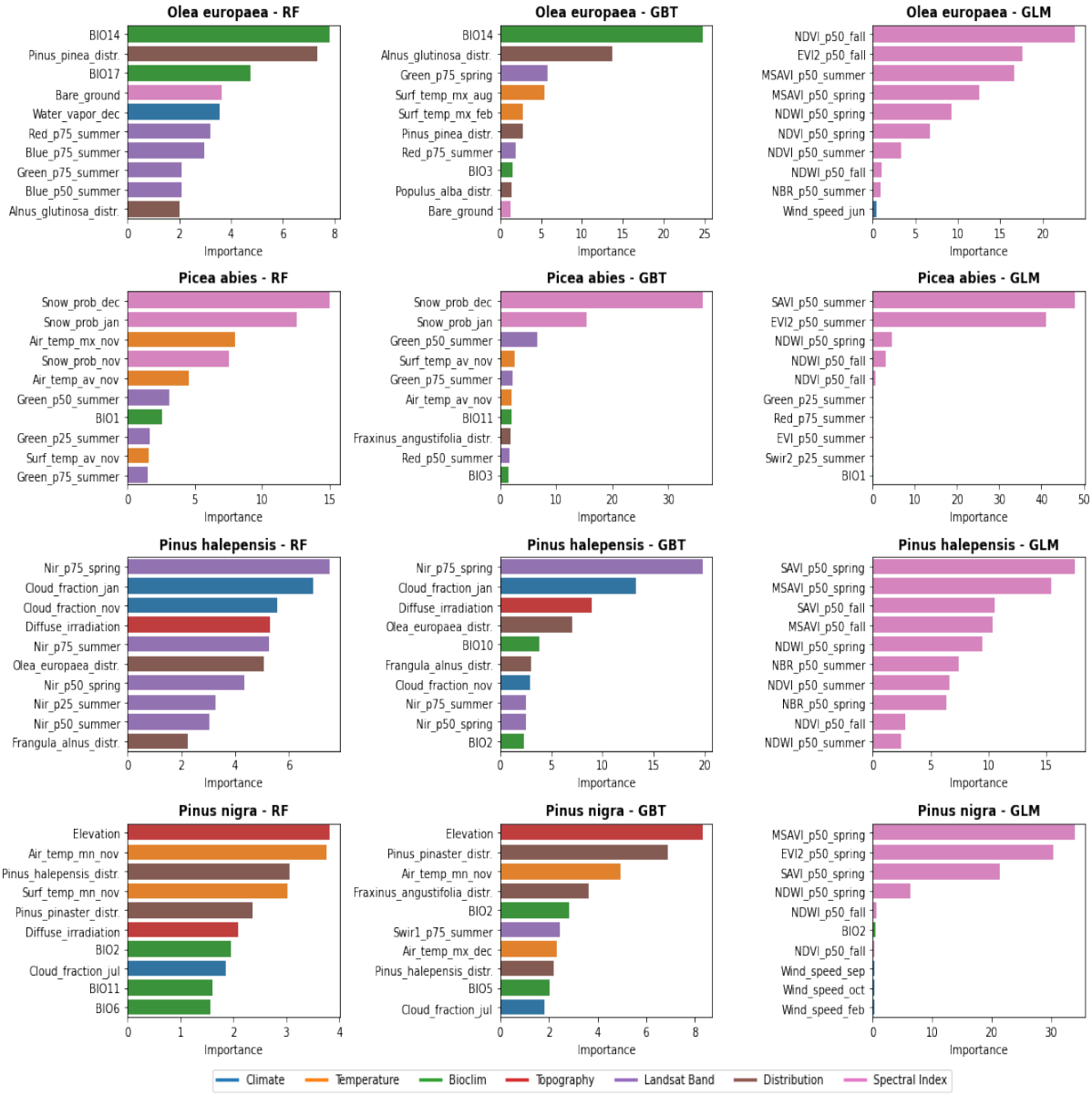

Figure S11: Relative importance of the 10 most important variables across the three component models for **realized** distribution of *O. europaea, P. abies, P. halepensis, P. nigra*

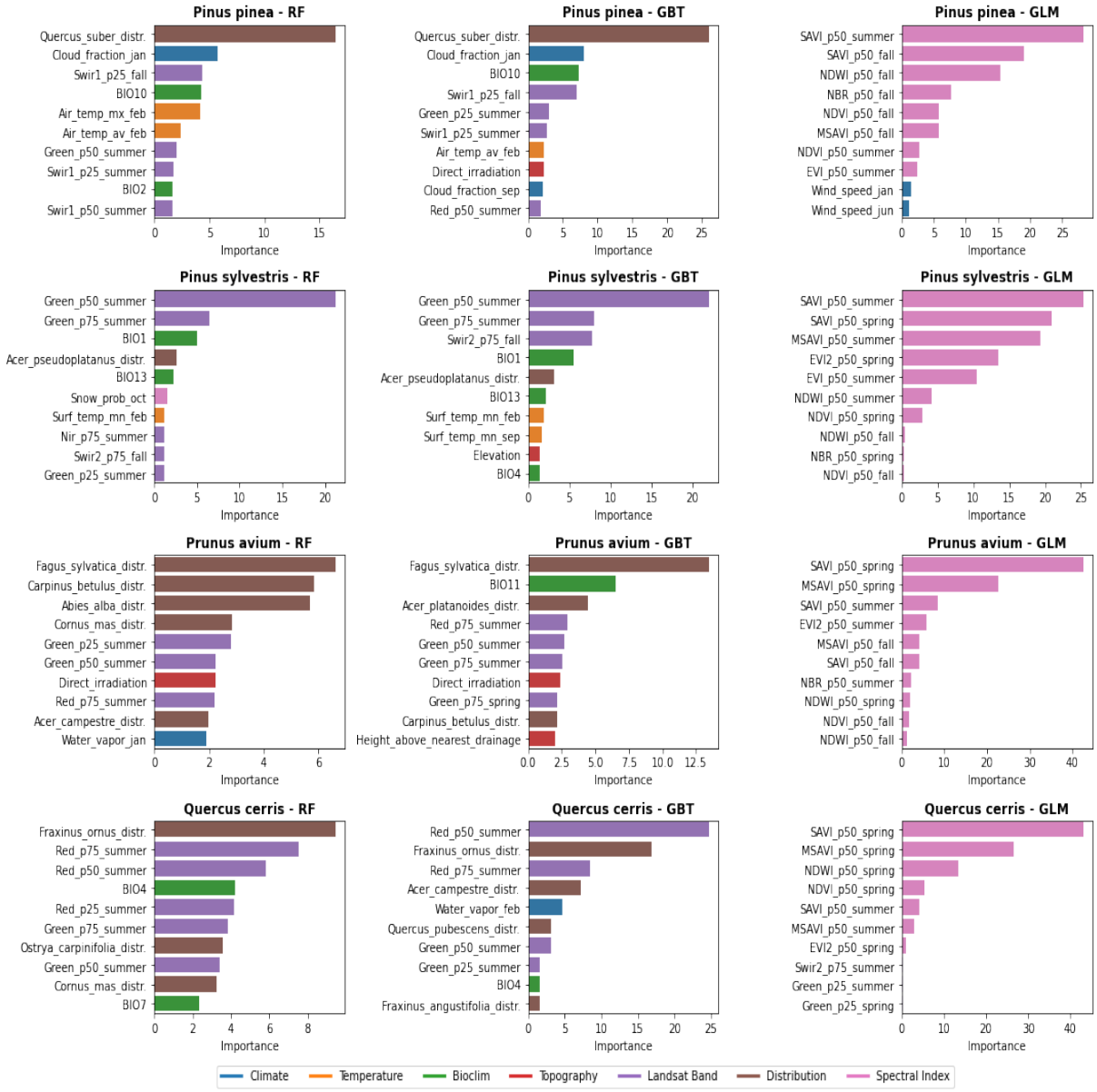

Figure S12: Relative importance of the 10 most important variables across the three component models for **realized** distribution of *P. pinea, P. sylvestris, P. avium, Q. cerris*

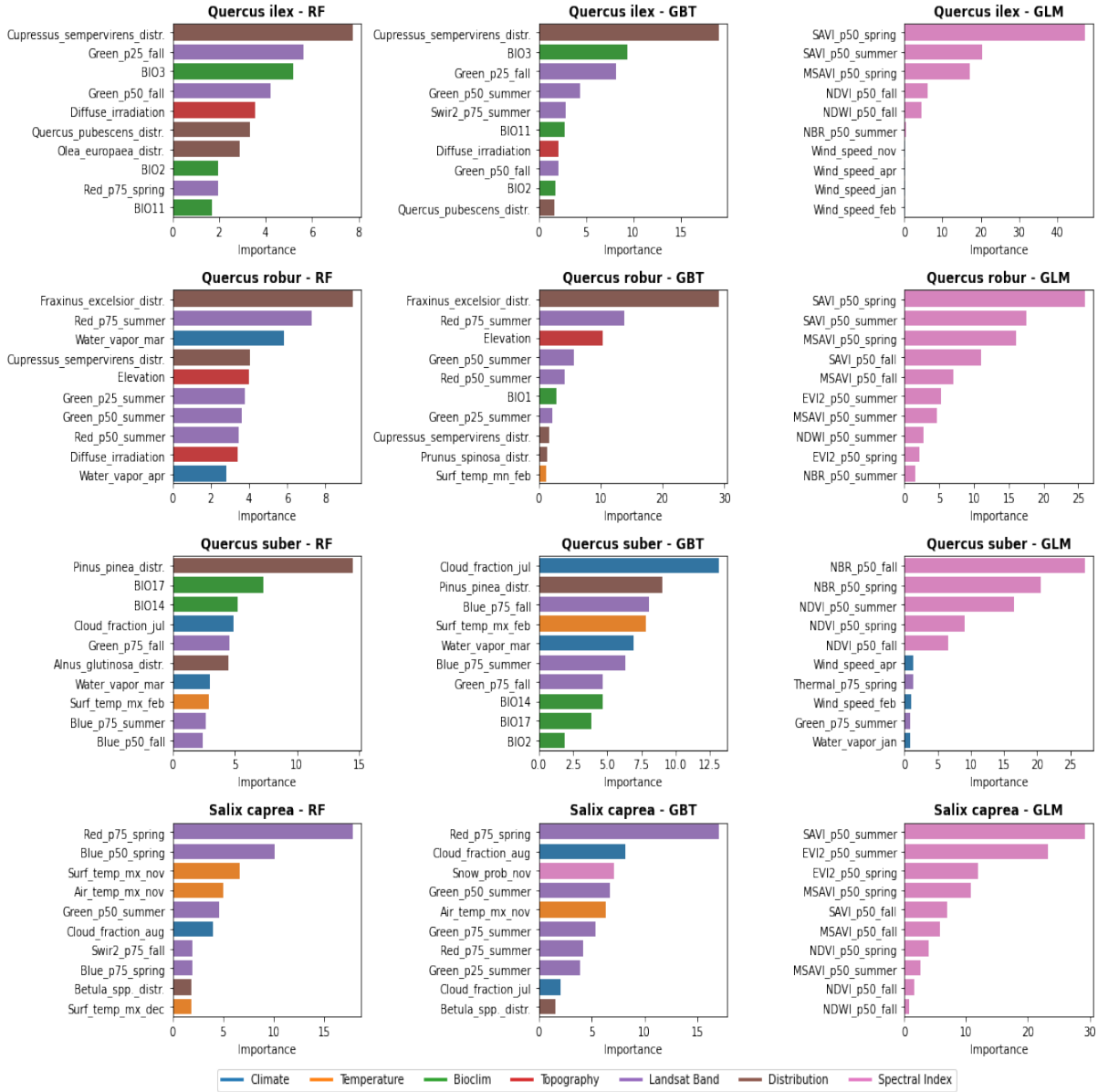

Figure S13: Relative importance of the 10 most important variables across the three component models for **realized** distribution of *Q. ilex, Q. robur, Q. suber, S. caprea*

Table S4: Average logloss and $R^2_{logloss}$ for the component learners and the ensemble model. Base logloss values are dependent on presence-absence ratio in the species dataset and are here used as a baseline for predictive performances comparison.

		Logloss					$R^2_{logloss}$			
Species	Distribution	GBT	GLM	RF	EML	Base	GBT	GLM	RF	EML
A. alba	Potential	0.058	0.107	0.063	0.048	0.381	0.85	0.72	0.83	0.87
C. sativa	Potential	0.102	0.129	0.099	0.071	0.318	0.68	0.59	0.69	0.78
C. avellana	Potential	0.052	0.080	0.072	0.045	0.488	0.89	0.84	0.85	0.91
F. sylvatica	Potential	0.060	0.141	0.065	0.046	0.664	0.91	0.79	0.90	0.93
O. europaea	Potential	0.004	0.005	0.006	0.004	0.072	0.94	0.93	0.92	0.94
P. abies	Potential	0.127	0.176	0.118	0.112	0.693	0.82	0.75	0.83	0.84
P. halepensis	Potential	0.025	0.027	0.029	0.018	0.292	0.91	0.91	0.90	0.94
P. nigra	Potential	0.097	0.120	0.119	0.085	0.226	0.57	0.47	0.47	0.62
P. pinea	Potential	0.028	0.029	0.035	0.020	0.168	0.83	0.83	0.79	0.88
P. sylvestris	Potential	0.310	0.478	0.295	0.287	0.693	0.55	0.31	0.57	0.59
P. avium	Potential	0.043	0.101	0.067	0.037	0.336	0.87	0.70	0.80	0.89
Q. cerris	Potential	0.025	0.031	0.025	0.012	0.128	0.80	0.76	0.80	0.91
Q. ilex	Potential	0.068	0.094	0.071	0.055	0.276	0.75	0.66	0.74	0.80
Q. robur	Potential	0.058	0.096	0.071	0.044	0.647	0.91	0.85	0.89	0.93
Q. suber	Potential	0.005	0.010	0.009	0.006	0.126	0.96	0.92	0.93	0.95
S. caprea	Potential	0.053	0.062	0.083	0.050	0.654	0.92	0.91	0.87	0.92
A. alba	Realized	0.030	0.036	0.039	0.033	0.243	0.88	0.85	0.84	0.86
C. sativa	Realized	0.055	0.056	0.055	0.047	0.209	0.74	0.73	0.74	0.78
C. avellana	Realized	0.032	0.038	0.045	0.034	0.217	0.85	0.82	0.79	0.84
F. sylvatica	Realized	0.047	0.063	0.057	0.050	0.526	0.91	0.88	0.89	0.90
O. europaea	Realized	0.007	0.005	0.011	0.007	0.050	0.86	0.90	0.78	0.86
P. abies	Realized	0.097	0.144	0.105	0.101	0.657	0.85	0.78	0.84	0.85
P. halepensis	Realized	0.016	0.018	0.023	0.015	0.221	0.93	0.92	0.90	0.93
P. nigra	Realized	0.091	0.071	0.072	0.061	0.182	0.50	0.61	0.60	0.66
P. pinea	Realized	0.021	0.016	0.027	0.016	0.125	0.83	0.87	0.78	0.87
P. sylvestris	Realized	0.224	0.316	0.224	0.215	0.653	0.66	0.52	0.66	0.67
P. avium	Realized	0.032	0.038	0.043	0.033	0.162	0.80	0.77	0.73	0.80
Q. cerris	Realized	0.014	0.012	0.018	0.011	0.093	0.85	0.87	0.81	0.88
Q. ilex	Realized	0.041	0.045	0.047	0.036	0.209	0.80	0.78	0.78	0.83
Q. robur	Realized	0.050	0.064	0.064	0.051	0.423	0.88	0.85	0.85	0.88
Q. suber	Realized	0.005	0.005	0.009	0.005	0.099	0.95	0.95	0.91	0.95
S. caprea	Realized	0.043	0.058	0.059	0.046	0.318	0.86	0.82	0.81	0.86

Table S5: Average AUC and TSS for the component learners and the ensemble model.

		AUC				TSS			
Species	Distribution	GBT	GLM	RF	EML	GBT	GLM	RF	EML
A. alba	Potential	1.00	0.98	1.00	1.00	0.89	0.78	0.87	0.92
C. sativa	Potential	0.99	0.96	0.99	0.99	0.75	0.64	0.71	0.82
C. avellana	Potential	1.00	0.99	1.00	1.00	0.94	0.90	0.93	0.95
F. sylvatica	Potential	1.00	0.99	1.00	1.00	0.96	0.88	0.96	0.97
O. europaea	Potential	1.00	1.00	0.99	1.00	0.95	0.94	0.93	0.95
P. abies	Potential	0.99	0.98	0.99	0.99	0.90	0.86	0.92	0.92
P. halepensis	Potential	1.00	1.00	1.00	1.00	0.94	0.94	0.92	0.96
P. nigra	Potential	0.97	0.93	0.97	0.97	0.54	0.44	0.56	0.66
P. pinea	Potential	1.00	0.99	0.99	1.00	0.85	0.82	0.81	0.90
P. sylvestris	Potential	0.94	0.85	0.95	0.95	0.70	0.53	0.73	0.73
P. avium	Potential	1.00	0.98	0.99	1.00	0.90	0.76	0.86	0.93
Q. cerris	Potential	0.99	0.99	0.99	0.99	0.82	0.75	0.76	0.93
Q. ilex	Potential	0.99	0.98	0.99	0.99	0.78	0.66	0.78	0.83
Q. robur	Potential	1.00	0.99	1.00	1.00	0.96	0.93	0.95	0.97
Q. suber	Potential	1.00	1.00	1.00	1.00	0.97	0.94	0.95	0.97
S. caprea	Potential	1.00	1.00	1.00	1.00	0.96	0.95	0.95	0.97
A. alba	Realized	1.00	1.00	1.00	1.00	0.88	0.88	0.84	0.90
C. sativa	Realized	0.99	0.99	0.99	0.99	0.71	0.73	0.67	0.80
C. avellana	Realized	1.00	0.99	0.99	1.00	0.85	0.85	0.81	0.87
F. sylvatica	Realized	1.00	1.00	1.00	1.00	0.94	0.93	0.94	0.95
O. europaea	Realized	1.00	1.00	0.99	1.00	0.82	0.87	0.76	0.88
P. abies	Realized	0.99	0.99	0.99	0.99	0.91	0.88	0.91	0.92
P. halepensis	Realized	1.00	1.00	1.00	1.00	0.94	0.94	0.91	0.95
P. nigra	Realized	0.97	0.97	0.98	0.98	0.54	0.56	0.55	0.69
P. pinea	Realized	1.00	1.00	0.99	1.00	0.82	0.87	0.79	0.89
P. sylvestris	Realized	0.97	0.93	0.97	0.97	0.76	0.68	0.77	0.79
P. avium	Realized	0.99	0.99	0.99	0.99	0.76	0.77	0.73	0.82
Q. cerris	Realized	1.00	1.00	0.99	1.00	0.84	0.88	0.79	0.90
Q. ilex	Realized	0.99	0.99	0.99	1.00	0.79	0.78	0.75	0.85
Q. robur	Realized	1.00	0.99	1.00	1.00	0.92	0.91	0.90	0.93
Q. suber	Realized	1.00	1.00	1.00	1.00	0.94	0.96	0.89	0.96
S. caprea	Realized	1.00	0.99	0.99	1.00	0.88	0.86	0.87	0.90

References

Aarrestad, P. A., T. Myking, O. E. Stabbetorp, and M. M. Tollefsrud (2014). *Foreign Norway spruce (*Picea abies*) provenances in Norway and effects on biodiversity*. Tech. rep. 1075. Norwegian Institute for Nature Research - NINA, 1–39.

Afuye, G. A., A. M. Kalumba, and I. R. Orimoloye (2021). "Characterisation of vegetation response to climate change: A review". *Sustainability* 13.13, 7265.

Aguadé, D., R. Poyatos, T. Rosas, and J. Martínez-Vilalta (2015). "Comparative drought responses of *Quercus ilex* L. and *Pinus sylvestris* L. in a montane forest undergoing a vegetation shift". *Forests* 6.8, 2505–2529. DOI: https://doi.org/10.3390/f6082505.

Agudelo, J., P. A. Arias, S. C. Vieira, and J. A. Martínez (2019). "Influence of longer dry seasons in the Southern Amazon on patterns of water vapor transport over northern South America and the Caribbean". *Climate Dynamics* 52.5, 2647–2665.

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson (2015). "spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models". *Ecography* 38.5, 541–545.

Allen, C. D., A. K. Macalady, H. Chenchouni, D. Bachelet, N. McDowell, M. Vennetier, T. Kitzberger, A. Rigling, D. D. Breshears, E. T. Hogg, et al. (2010). "A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests". *Forest ecology and management* 259.4, 660–684.

Anand, A., M. K. Pandey, P. K. Srivastava, A. Gupta, and M. L. Khan (2021). "Integrating Multi-Sensors Data for Species Distribution Mapping Using Deep Learning and Envelope Models". *Remote Sensing* 13.16, 3284.

Anav, A., G. Murray-Tortarolo, P. Friedlingstein, S. Sitch, S. Piao, and Z. Zhu (2013). "Evaluation of land surface models in reproducing satellite Derived leaf area index over the high-latitude northern hemisphere. Part II: Earth system models". *Remote Sensing* 5.8, 3637–3661.

Anchang, J. Y., L. Prihodko, A. T. Kaptué, C. W. Ross, W. Ji, S. S. Kumar, B. Lind, M. A. Sarr, A. A. Diouf, and N. P. Hanan (2019). "Trends in woody and herbaceous vegetation in the Savannas of West Africa". *Remote Sensing* 11.5, 576.

Anderegg, W. R., C. Wu, N. Acil, N. Carvalhais, T. A. Pugh, J. P. Sadler, and R. Seidl (2022). "A climate risk analysis of Earth's forests in the 21st century". *Science* 377.6610, 1099–1103.

Andrewartha, H. G. and L. C. Birch (1954). *The distribution and abundance of animals.* The University of Chicago Press.

Anjos, L. J., E. B. de Souza, C. T. Amaral, T. K. Igawa, and P. M. de Toledo (2021). "Future projections for terrestrial biomes indicate widespread warming and moisture reduction in forests up to 2100 in South America". *Global Ecology and Conservation* 25, e01441.

Appuhn, K. (2010). *A Forest on the Sea: Environmental Expertise in Renaissance Venice.* Johns Hopkins University Press.

Araújo, M. B. and M. New (2007). "Ensemble forecasting of species distributions". *Trends in Ecology & Evolution* 22.1, 42–47. DOI: 10.1016/j.tree.2006.09.010.

Araza, A., M. Herold, S. De Bruin, P. Ciais, D. A. Gibbs, N. Harris, M. Santoro, J.-P. Wigneron, H. Yang, N. Málaga, et al. (2023). "Past decade above-ground biomass change comparisons from four multi-temporal global maps". *International Journal of Applied Earth Observation and Geoinformation* 118, 103274.

Archambeau, J., P. Ruiz-Benito, S. Ratcliffe, T. Fréjaville, A. Changenet, J. M. M. Castañeda, A. Lehtonen, J. Dahlgren, M. A. Zavala, and M. B. Garzón (2020). "Similar patterns of background mortality across Europe are mostly driven by drought in European beech and a combination of drought and competition in Scots pine". *Agricultural and Forest Meteorology* 280, 107772. DOI: https://doi.org/10.1016/j.agrformet.2019.107772.

Arias, P. A., R. Fu, C. Vera, and M. Rojas (2015). "A correlated shortening of the North and South American monsoon seasons in the past few decades". *Climate dynamics* 45.11, 3183–3203.

Ashcroft, M. B. (2010). "Identifying refugia from climate change". *Journal of biogeography* 37.8, 1407–1413. DOI: https://doi.org/10.1111/j.1365-2699.2010.02300.x.

Astola, H., T. Häme, L. Sirro, M. Molinier, and J. Kilpi (2019). "Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region". *Remote Sensing of Environment* 223, 257–273. DOI: https://doi.org/10.1016/j.rse.2019.01.019.

Avitabile, V., R. Pilli, M. Migliavacca, G. Duveiller, A. Camia, V. Blujdea, R. Adolt, I. Alberdi, S. Barreiro, S. Bender, D. Borota, M. Bosela, O. Bouriaud, J. Breidenbach, I. Cañellas, J. Čavlović, A. Colin, L. Di Cosmo, J. Donis, C. Fischer, A. Freudenschuss, J. Fridman, P. Gasparini, T. Gschwantner, L. Hernández, K. Korhonen, G. Kulbokas, V. Kvist, N. Latte, A. Lazdins, P. Lejeune, K. Makovskis, G. Marin, J. Maslo, A. Michorczyk, M. Mionskowski, F. Morneau, M. Myszkowski, K. Nagy, M. Nilsson, T. Nord-Larsen, D. Pantic, J. Perin, J. Redmond, M. Rizzo, V. Šebeň, M. Skudnik, A. Snorrason, R. Sroga, T. Stoyanov, A. Svensson, A. Talarczyk, S. Teeuwen, E. Thürig,

J. Uva, and S. Mubareka (2024). "Harmonised statistics and maps of forest biomass and increment in Europe". *Scientific Data* 11.1, 274. DOI: 10.1038/s41597-023-02868-8.

Axelsson, A., E. Lindberg, H. Reese, and H. Olsson (2021). "Tree species classification using Sentinel-2 imagery and Bayesian inference". *International Journal of Applied Earth Observation and Geoinformation* 100, 102318.

Ayres, M. P. and M. J. Lombardero (2000). "Assessing the consequences of global change for forest disturbance from herbivores and pathogens". *Science of the Total Environment* 262.3, 263–286.

Babiy, I., S. Im, and V. Kharuk (2022). "Estimating Aboveground Forest Biomass Using Radar Methods". *Contemporary Problems of Ecology* 15.5, 433–448.

Bałazy, R. (2020). "Forest dieback process in the Polish mountains in the past and nowadays-literature review on selected topics". *Folia Forestalia Polonica. Series A. Forestry* 62.3, 184–198. DOI: https://doi.org/10.2478/ffp-2020-0018.

Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R. Evans, et al. (2001). "FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities". *Bulletin of the American Meteorological Society* 82.11, 2415–2434.

Baldocchi, D. and J. Penuelas (2019). "The physics and ecology of mining carbon dioxide from the atmosphere by ecosystems". *Global Change Biology* 25.4, 1191–1197.

Ballantyne, A. á., C. á. Alden, J. á. Miller, P. á. Tans, and J. White (2012). "Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years". *Nature* 488.7409, 70–72.

Banskota, A., N. Kayastha, M. J. Falkowski, M. A. Wulder, R. E. Froese, and J. C. White (2014). "Forest monitoring using Landsat time series data: A review". *Canadian Journal of Remote Sensing* 40.5, 362–384.

Bao, Z., J. Zhang, G. Wang, T. Guan, J. Jin, Y. Liu, M. Li, and T. Ma (2021). "The sensitivity of vegetation cover to climate change in multiple climatic zones using machine learning algorithms". *Ecological indicators* 124, 107443.

Barbati, A., M. Marchetti, G. Chirici, and P. Corona (2014). "European Forest Types and Forest Europe SFM indicators: Tools for monitoring progress on forest biodiversity conservation". *Forest Ecology and Management* 321, 145–157. DOI: https://doi.org/10.1016/j.foreco.2013.07.004.

Barbati, A., P. Corona, and M. Marchetti (2007). "A forest typology for monitoring sustainable forest management: the case of European forest types". *Plant Biosystems* 141.1, 93–103.

Barlow, J., E. Berenguer, R. Carmenta, and F. França (2020). "Clarifying Amazonia's burning crisis". *Global Change Biology* 26.2, 319–321.

Beigaitė, R., H. Tang, A. Bryn, O. Skarpaas, F. Stordal, J. W. Bjerke, and I. Žliobaitė (2022). "Identifying climate thresholds for dominant natural vegetation types at the global scale using machine learning: Average climate versus extremes". *Global Change Biology* 28.11, 3557–3579. DOI: 10.1111/gcb.16110.

Bellman, R. and R. Kalaba (1957). "Dynamic programming and statistical communication theory". *Proceedings of the National Academy of Sciences* 43.8, 749–751.

Berner, L. T. and S. J. Goetz (2022). "Satellite observations document trends consistent with a boreal forest biome shift". *Global change biology* 28.10, 3275–3292.

Berner, L. T., R. Massey, P. Jantz, B. C. Forbes, M. Macias-Fauria, I. Myers-Smith, T. Kumpula, G. Gauthier, L. Andreu-Hayles, B. V. Gaglioti, et al. (2020). "Summer warming explains widespread but not uniform greening in the Arctic tundra biome". *Nature Communications* 11.1, 1–12.

Bischl, B., M. Lang, L. Kotthoff, J. Schiffner, J. Richter, E. Studerus, G. Casalicchio, and Z. M. Jones (2016). "mlr: Machine Learning in R". *Journal of Machine Learning Research* 17.170, 1–5.

Bivand, R., T. Keitt, and B. Rowlingson (2021). *rgdal: Bindings for the 'Geospatial' Data Abstraction Library*. R package version 1.5-23.

Bonannella, **C.**, G. Chirici, D. Travaglini, M. Pecchi, E. Vangi, G. D'amico, and F. Giannetti (2022a). "Characterization of wildfires and harvesting forest disturbances and recovery using Landsat time series: a case study in Mediterranean forests in central Italy". *Fire* 5.3, 68. DOI: 10.3390/fire5030068.

Bonannella, **C.**, T. Hengl, J. Heisig, L. Leal Parente, M. Wright, M. Herold, and S. de Bruin (2022b). *Presence-Absence Points for Tree Species Distribution Modelling for Europe*. Version 0.3. Zenodo. DOI: 10.5281/zenodo.6516590.

– (2022c). *Supplementary material for "Forest tree species distribution for Europe 2000 - 2020: mapping potential and realized distributions using spatiotemporal Machine Learning"*. Version 0.1. DOI: 10.5281/zenodo.6516728.

Bonannella, **C.**, T. Hengl, J. Heisig, L. Parente, M. N. Wright, M. Herold, and S. de Bruin (2022d). "Forest tree species distribution for Europe 2000–2020: mapping potential and realized distributions using spatiotemporal machine learning". *PeerJ*. DOI: https://doi.org/10.7717/peerj.13728.

Bonannella, **C.**, T. Hengl, L. Parente, and S. de Bruin (2023). "Biomes of the world under climate change scenarios: increasing aridity and higher temperatures lead to significant shifts in natural vegetation". *PeerJ* 11, e15593. DOI: https://doi.org/10.7717/peerj.15593.

Bonannella, **C.**, L. Parente, S. de Bruin, and M. Herold (2024a). "Multi-decadal trend analysis and forest disturbance assessment of European tree species: concerning signs of

a subtle shift". *Forest Ecology and Management* 554, 121652. DOI: https://doi.org/10.1016/j.foreco.2023.121652.

Bonannella, **C.**, L. Parente, B. Lerink, S. de Bruin, and M. Herold (2024b). "Impact of NFI coordinate precision on high resolution tree species classification". *In preparation.*

Bontemps, J.-D., O. Bouriaud, C. Vega, and L. Bouriaud (2022). "Offering the appetite for the monitoring of European forests a diversified diet". *Annals of Forest Science* 79.1, 1–9.

Bossard, M., J. Feranec, J. Otahel, et al. (2000). *CORINE land cover technical guide: Addendum 2000*. Vol. 40. European Environment Agency Copenhagen.

Breidenbach, J., D. Ellison, H. Petersson, K. T. Korhonen, H. M. Henttonen, J. Wallerman, J. Fridman, T. Gobakken, R. Astrup, and E. Næsset (2022). "Harvested area did not increase abruptly—how advancements in satellite-based mapping led to erroneous conclusions". *Annals of forest science* 79.1, 1–9.

Breidenbach, J., A. Granhus, G. Hylen, R. Eriksen, and R. Astrup (2020). "A century of National Forest Inventory in Norway–informing past, present, and future decisions". *Forest ecosystems* 7, 1–19.

Breidenbach, J., L. T. Waser, M. Debella-Gilo, J. Schumacher, J. Rahlf, M. Hauglin, S. Puliti, and R. Astrup (2021). "National mapping and estimation of forest area by dominant tree species using Sentinel-2 data". *Canadian Journal of Forest Research* 51.3, 365–379.

Breiman, L. (2001a). "Random forests". *Machine learning* 45.1, 5–32.

– (2001b). "Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)". *Statistical Science* 16.3, 199–231. DOI: 10.1214/ss/1009213726.

Brier, G. W. (1950). "Verification of forecasts expressed in terms of probability". *Monthly weather review* 78.1, 1–3.

Brockerhoff, E. G., L. Barbaro, B. Castagneyrol, D. I. Forrester, B. Gardiner, J. R. González-Olabarria, P. O. Lyver, N. Meurisse, A. Oxbrough, H. Taki, et al. (2017). *Forest biodiversity, ecosystem functioning and the provision of ecosystem services*. DOI: https://dx.doi.org/10.1007/s10531-017-1453-2.

Brown, K. E., F. A. Bhuiyan, and D. A. Talbert (2020). "Uncertainty Quantification in Multimodal Ensembles of Deep Learners". In: *The Thirty-Third International Flairs Conference*.

Bucklin, D. N., M. Basille, A. M. Benscoter, L. A. Brandt, F. J. Mazzotti, S. S. Romanach, C. Speroterra, and J. I. Watling (2015). "Comparing species distribution models constructed with different subsets of environmental predictors". *Diversity and distributions* 21.1, 23–35.

Bussotti, F. and M. Pollastrini (2017). "Observing climate change impacts on European forests: what works and what does not in ongoing long-term monitoring networks". *Frontiers in plant science* 8, 629.

Büttner, G., C. Steenmans, M. Bossard, J. Feranec, and J. Kolár (1998). "The European CORINE land cover database". *International Archives of Photogrammetry and Remote Sensing* 32, 633–638.

Calderón-Loor, M., M. Hadjikakou, and B. A. Bryan (2021). "High-resolution wall-to-wall land-cover mapping and land change assessment for Australia from 1985 to 2015". *Remote Sensing of Environment* 252, 112–148. DOI: https://doi.org/10.1016/j.rse.2020.112148.

Castaldi, F., S. Chabrillat, A. Don, and B. van Wesemael (2019). "Soil organic carbon mapping using LUCAS topsoil database and Sentinel-2 data: An approach to reduce soil moisture and crop residue effects". *Remote Sensing* 11.18, 2121.

Caudullo, G., E. Welk, and J. San-Miguel-Ayanz (2017). "Chorological maps for the main European woody species". *Data in brief* 12, 662–666. DOI: https://doi.org/10.1016/j.dib.2017.05.007.

Cazzolla Gatti, R., P. B. Reich, J. G. Gamarra, T. Crowther, C. Hui, A. Morera, J.-F. Bastin, S. De-Miguel, G.-J. Nabuurs, J.-C. Svenning, et al. (2022). "The number of tree species on Earth". *Proceedings of the National Academy of Sciences* 119.6, e2115329119.

CBD, U. (2011). "COP 10 decision X/2". *Strategic plan for biodiversity* 2020.

Ceccherini, G., G. Duveiller, G. Grassi, G. Lemoine, V. Avitabile, R. Pilli, and A. Cescatti (2020). "Abrupt increase in harvested forest area over Europe after 2015". *Nature* 583.7814, 72–77.

– (2022). "Potentials and limitations of NFIs and remote sensing in the assessment of harvest rates: a reply to Breidenbach et al." *Annals of Forest Science* 79.1, 1–7.

Chakraborty, D., N. Móricz, E. Rasztovits, L. Dobor, and S. Schueler (2021). "Provisioning forest and conservation science with high-resolution maps of potential distribution of major European tree species under climate change". *Annals of Forest Science* 78.2, 1–18.

Cheddadi, R., M. B. Araújo, L. Maiorano, M. Edwards, A. Guisan, M. Carré, M. Chevalier, and P. B. Pearman (2016). "Temperature range shifts for three European tree species over the last 10,000 years". *Frontiers in plant science* 7, 1581. DOI: https://doi.org/10.3389/fpls.2016.01581.

Chefaoui, R. M. and J. M. Lobo (2008). "Assessing the effects of pseudo-absences on predictive distribution model performance". *Ecological modelling* 210.4, 478–486.

Chen, C., W. J. Riley, I. C. Prentice, and T. F. Keenan (2022). "CO2 fertilization of terrestrial photosynthesis inferred from site to global scales". *Proceedings of the National Academy of Sciences* 119.10, e2115627119.

Chen, M., C. R. Vernon, N. T. Graham, M. Hejazi, M. Huang, Y. Cheng, and K. Calvin (2020). "Global land use for 2015–2100 at 0.05 resolution under diverse socioeconomic and climate scenarios". *Scientific Data* 7.1, 1–11. DOI: https://doi.org/10.1038/s41597-020-00669-x.

Chirici, G., F. Giannetti, D. Travaglini, S. Nocentini, S. Francini, G. D'Amico, E. Calvo, D. Fasolini, M. Broll, F. Maistrelli, et al. (2019). "Forest damage inventory after the" Vaia" storm in Italy". *Forest@* 16, 3–9. DOI: https://dx.doi.org/10.3832/efor3070-016.

Choe, H., J. Chi, and J. H. Thorne (2021). "Mapping Potential Plant Species Richness over Large Areas with Deep Learning, MODIS, and Species Distribution Models". *Remote Sensing* 13.13, 2490.

Claverie, M., J. Ju, J. G. Masek, J. L. Dungan, E. F. Vermote, J.-C. Roger, S. V. Skakun, and C. Justice (2018a). "The Harmonized Landsat and Sentinel-2 surface reflectance data set". *Remote sensing of environment* 219, 145–161.

– (2018b). "The Harmonized Landsat and Sentinel-2 surface reflectance data set". *Remote sensing of environment* 219, 145–161.

Cleland, E. E., I. Chuine, A. Menzel, H. A. Mooney, and M. D. Schwartz (2007). "Shifting plant phenology in response to global change". *Trends in ecology & evolution* 22.7, 357–365.

Colbrook, M. J., V. Antun, and A. C. Hansen (2022). "The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem". *Proceedings of the National Academy of Sciences* 119.12, e2107151119.

Coops, N. C., P. Tompalski, T. R. Goodbody, A. Achim, and C. Mulverhill (2023). "Framework for near real-time forest inventory using multi source remote sensing data". *Forestry* 96.1, 1–19.

Corona, P., L. Fattorini, S. Franceschi, G. Scrinzi, and C. Torresan (2014). "Estimation of standing wood volume in forest compartments by exploiting airborne laser scanning information: model-based, design-based, and hybrid perspectives". *Canadian Journal of Forest Research* 44.11, 1303–1311.

Crimmins, S. M., S. Z. Dobrowski, J. A. Greenberg, J. T. Abatzoglou, and A. R. Mynsberge (2011). "Changes in climatic water balance drive downhill shifts in plant species' optimum elevations". *Science* 331.6015, 324–327. DOI: https://doi.org/10.1126/science.1199040.

d'Andrimont, R., A. Verhegghen, M. Meroni, G. Lemoine, P. Strobl, B. Eiselt, M. Yordanov, L. Martinez-Sanchez, and M. van der Velde (2021). "LUCAS Copernicus 2018: Earth-observation-relevant in situ data on land cover and use throughout the European Union". *Earth System Science Data* 13.3, 1119–1133. DOI: 10.5194/essd-13-1119-2021.

d'Andrimont, R., M. Yordanov, L. Martinez-Sanchez, B. Eiselt, A. Palmieri, P. Dominici, J. Gallego, H. I. Reuter, C. Joebges, G. Lemoine, et al. (2020). "Harmonised LUCAS

in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union". *Scientific Data* 7.1, 1–15.

Dale, V. H., L. A. Joyce, S. Mcnulty, P. Ronald, and P. Matthew (2001). "Climate Change and Forest Disturbances". *BioScience* 51.9, 723–734.

Dalponte, M., H. O. Ørka, T. Gobakken, D. Gianelle, and E. Næsset (2012). "Tree species classification in boreal forests with hyperspectral data". *IEEE Transactions on Geoscience and Remote Sensing* 51.5, 2632–2645.

Dave, R., C. Saint-Laurent, L. Murray, G. Antunes Daldegan, R. Brouwer, C. A. de Mattos Scaramuzza, L. Raes, S. Simonit, M. Catapan, G. García Contreras, et al. (2018). "Second Bonn challenge progress report". *Application of the Barometer in* 2019.

De Jong, R., S. de Bruin, A. de Wit, M. E. Schaepman, and D. L. Dent (2011). "Analysis of monotonic greening and browning trends from global NDVI time-series". *Remote Sensing of Environment* 115.2, 692–702.

De Sy, V. (2016). *Remote Sensing of Land Use and Carbon Losses Following Tropical Deforestation*. DOI: 10.18174/380263.

Deneu, B., M. Servajean, P. Bonnet, C. Botella, F. Munoz, and A. Joly (2021). "Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment". *PLoS computational biology* 17.4, e1008856.

Deur, M., M. Gašparović, and I. Balenović (2020). "Tree species classification in mixed deciduous forests using very high spatial resolution satellite imagery and machine learning methods". *Remote Sensing* 12.23, 3926.

Dirkse, G. M., W. Daamen, and C. Schuiling (2001). *Toelichting bossenkaart*. Tech. rep. Alterra.

Domke, G. M., S. N. Oswalt, B. F. Walters, and R. S. Morin (2020). "Tree planting has the potential to increase carbon sequestration capacity of forests in the United States". *Proceedings of the National Academy of Sciences* 117.40, 24649–24651.

Dottori, F., P. Salamon, A. Bianchi, L. Alfieri, F. A. Hirpa, and L. Feyen (2016). "Development and evaluation of a framework for global flood hazard mapping". *Advances in Water Resources* 94, 87–102. DOI: https://doi.org/10.1016/j.advwatres.2016.05.002.

Dow, K. and T. E. Downing (2016). *The Atlas of climate change: mapping the world's greatest challenge*. San Francisco: Univ of California Press, 128.

Dowle, M. and A. Srinivasan (2021). *data.table: Extension of 'data.frame'*. R package version 1.14.0.

Draper, F. C., F. R. Costa, G. Arellano, O. L. Phillips, A. Duque, M. J. Macía, H. Ter Steege, G. P. Asner, E. Berenguer, J. Schietti, et al. (2021). "Amazon tree dominance across forest strata". *Nature ecology & evolution* 5.6, 757–767. DOI: 10.1038/s41559-021-01418-y.

Dubayah, R., M. Hofton, J. Blair, J. Armston, H. Tang, and S. Luthcke (2020). "GEDI L2A elevation and height metrics data global footprint level V001". *NASA EOSDIS Land Processes DAAC*.

Dufresne, J.-L., M.-A. Foujols, S. Denvil, A. Caubel, O. Marti, O. Aumont, Y. Balkanski, S. Bekki, H. Bellenger, R. Benshila, et al. (2013). "Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5". *Climate dynamics* 40, 2123–2165.

Dullinger, S., A. Gattringer, W. Thuiller, D. Moser, N. E. Zimmermann, A. Guisan, W. Willner, C. Plutzar, M. Leitner, T. Mang, et al. (2012). "Extinction debt of high-mountain plants under twenty-first-century climate change". *Nature climate change* 2.8, 619–622. DOI: https://doi.org/10.1038/nclimate1514.

Dwyer, J. L., D. P. Roy, B. Sauer, C. B. Jenkerson, H. K. Zhang, and L. Lymburner (2018). "Analysis ready data: enabling analysis of the Landsat archive". *Remote Sensing* 10.9, 1363.

Dyderski, M. K., S. Paź, L. E. Frelich, and A. M. Jagodziński (2017). "How much does climate change threaten European forest tree species distributions?" *Global Change Biology* 24.3, 1150–1163. DOI: https://doi.org/10.1111/gcb.13925.

– (2018). "How much does climate change threaten European forest tree species distributions?" *Global Change Biology* 24.3, 1150–1163. DOI: 10.1111/gcb.13925.

Easdale, M. H., O. Bruzzone, P. Mapfumo, and P. Tittonell (2018). "Phases or regimes? R evisiting NDVI trends as proxies for land degradation". *Land Degradation & Development* 29.3, 433–445.

Eckenwalder, J. E. (2009). *Conifers of the world: the complete reference*. Timber press.

Einzmann, K., C. Atzberger, N. Pinnel, C. Glas, S. Böck, R. Seitz, and M. Immitzer (2021). "Early detection of spruce vitality loss with hyperspectral data: Results of an experimental study in Bavaria, Germany". *Remote Sensing of Environment* 266, 112676.

Elith, J. and C. H. Graham (2009). "Do They? How Do They? Why Do They Differ? On Finding Reasons for Differing Performances of Species Distribution Models". *Ecography* 32.1, 66–77.

ESA (2017). *Land Cover CCI Product User Guide Version 2. Tech. Rep. Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2$_2$.0.pdf*.

Estel, S., T. Kuemmerle, C. Levers, M. Baumann, and P. Hostert (2016). "Mapping cropland-use intensity across Europe using MODIS NDVI time series". *Environmental Research Letters* 11.2, 024015.

European Commission (2021a). *European Green Deal – Research innovation call*. Publications Office of the European Union. DOI: doi/10.2777/33415.

European Commission (2021b). *New EU Forest Strategy for 2030. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*.

– (2022). *Proposal for a Regulation of the European Parliament and of the Council on Nature Restoration*. European Commission Document.

– (2023). *Proposal for a Regulation of the European Parliament and of the Council on a Monitoring Framework for Resilient European Forests*. European Commission Document.

European Parliament (2007). "Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an infrastructure for spatial information in the European Community (INSPIRE)". *Official Journal of the European Union* 50, 1–14.

EUROSTAT (2017). *Land Cover/Use Statistics (LUCAS) Database*. (Visited on 2020).

Ewald, J. (2012). "Vegetation databases provide a close-up on altitudinal tree species distribution in the Bavarian Alps". *Vegetation databases for the 21st century.–Biodiversity & Ecology* 4, 41–48.

Fang, J. and M. J. Lechowicz (2006). "Climatic limits for the present distribution of beech species in the world". *Journal of Biogeography* 33.10, 1804–1819. DOI: https://doi.org/10.1111/j.1365-2699.2006.01533.x.

FAO (2004). *Global forest resources assessment up date 2005—terms and definitions (final version)*.

– (2022). *The State of the World's Forests 2022. Forest pathways for green recovery and building inclusive, resilient and sustainable economies*. Rome, FAO.

Fassnacht, F. E., H. Latifi, K. Stereńczak, A. Modzelewska, M. Lefsky, L. T. Waser, C. Straub, and A. Ghosh (2016). "Review of studies on tree species classification from remotely sensed data". *Remote sensing of environment* 186, 64–87.

Fassnacht, F. E., J. C. White, M. A. Wulder, and E. Næsset (2024). "Remote sensing in forestry: Current challenges, considerations and directions". *Forestry: An International Journal of Forest Research* 97.1, 11–37.

Fawcett, D., J. Bennie, and K. Anderson (2021). "Monitoring spring phenology of individual tree crowns using drone-acquired NDVI data". *Remote Sensing in Ecology and Conservation* 7.2, 227–244.

Feeley, K. J., M. R. Silman, M. B. Bush, W. Farfan, K. G. Cabrera, Y. Malhi, P. Meir, N. S. Revilla, M. N. R. Quisiyupanqui, and S. Saatchi (2011). "Upslope migration of Andean trees". *Journal of Biogeography* 38.4, 783–791. DOI: https://doi.org/10.1111/j.1365-2699.2010.02444.x.

Fei, S., J. M. Desprez, K. M. Potter, I. Jo, J. A. Knott, and C. M. Oswalt (2017). "Divergence of species responses to climate change". *Science Advances* 3.5, e1603055. DOI: https://doi.org/10.1126/sciadv.1603055.

Ferchichi, A., A. B. Abbes, V. Barra, and I. R. Farah (2022). "Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches: A systematic literature review". *Ecological Informatics* 68, 101552.

Ferri, C., J. Hernández-Orallo, and R. Modroiu (2009). "An experimental comparison of performance measures for classification". *Pattern Recognition Letters* 30.1, 27–38. DOI: https://doi.org/10.1016/j.patrec.2008.08.010.

Fick, S. E. and R. J. Hijmans (2017). "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas". *International journal of climatology* 37.12, 4302–4315.

Fidler, F., Y. E. Chee, B. C. Wintle, M. A. Burgman, M. A. McCarthy, and A. Gordon (2017). "Metaresearch for Evaluating Reproducibility in Ecology and Evolution". *BioScience* 67.3, 282–289. DOI: 10.1093/biosci/biw159.

Filazzola, A., S. F. Matter, and J. Roland (2020). "Inclusion of trophic interactions increases the vulnerability of an alpine butterfly species to climate change". *Global change biology* 26.5, 2867–2877.

Fix, E. and J. L. Hodges (1989). "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties". *International Statistical Review / Revue Internationale de Statistique* 57.3, 238. DOI: 10.2307/1403797.

Fois, M., A. Cuena-Lombraña, G. Fenu, and G. Bacchetta (2018). "Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions". *Ecological Modelling* 385, 124–132.

Foley, J. A., G. P. Asner, M. H. Costa, M. T. Coe, R. DeFries, H. K. Gibbs, E. A. Howard, S. Olson, J. Patz, N. Ramankutty, et al. (2007). "Amazonia revealed: forest degradation and loss of ecosystem goods and services in the Amazon Basin". *Frontiers in Ecology and the Environment* 5.1, 25–32.

FOREST EUROPE (2020). "State of Europe's Forests 2020."

Forzieri, G., M. Girardello, G. Ceccherini, J. Spinoni, L. Feyen, H. Hartmann, P. S. Beck, G. Camps-Valls, G. Chirici, A. Mauri, et al. (2021). "Emergent vulnerability to climate-driven disturbances in European forests". *Nature communications* 12.1, 1081. DOI: https://doi.org/10.1038/s41467-021-21399-7.

Forzieri, G., M. Pecchi, M. Girardello, A. Mauri, M. Klaus, C. Nikolov, M. Rüetschi, B. Gardiner, J. Tomaštík, D. Small, et al. (2020). "A spatially explicit database of wind disturbances in European forests over the period 2000–2018". *Earth System Science Data* 12.1, 257–276.

Fourcade, Y., A. G. Besnard, and J. Secondi (2018). "Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics". *Global Ecology and Biogeography* 27.2, 245–256.

Franklin, J. (1995). "Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients". *Progress in physical geography* 19.4, 474–499.

– (2010). *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge: Cambridge University Press. DOI: DOI:10.1017/CBO9780511810602.

Fricker, G. A., J. D. Ventura, J. A. Wolf, M. P. North, F. W. Davis, and J. Franklin (2019). "A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery". *Remote Sensing* 11.19, 2326.

Friedman, J. H. (2002). "Stochastic gradient boosting". *Computational Statistics & Data Analysis* 38.4. Nonlinear Methods and Data Mining, 367–378. DOI: https://doi.org/10.1016/S0167-9473(01)00065-2.

Friend, A. D., W. Lucht, T. T. Rademacher, R. Keribin, R. Betts, P. Cadule, P. Ciais, D. B. Clark, R. Dankers, P. D. Falloon, et al. (2014). "Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2". *Proceedings of the National Academy of Sciences* 111.9, 3280–3285.

Fujimori, S., T. Hasegawa, A. Ito, K. Takahashi, and T. Masui (2018). "Gridded emissions and land-use data for 2005–2100 under diverse socioeconomic and climate mitigation scenarios". *Scientific data* 5.1, 1–13. DOI: https://doi.org/10.1038/sdata.2018.210.

Gamfeldt, L., T. Snäll, R. Bagchi, M. Jonsson, L. Gustafsson, P. Kjellander, M. C. Ruiz-Jaen, M. Fröberg, J. Stendahl, C. D. Philipson, et al. (2013). "Higher levels of multiple ecosystem services are found in forests with more tree species". *Nature communications* 4.1, 1340. DOI: https://doi.org/10.1038/ncomms2328.

Ganatsas, P. (2010). "Forest characteristics of Black pine ecosystems and restoration of burned stands". *Νέες προσεγγίσεις στην αποκατάσταση δασών μαύρης πεύκης*, 139.

Ganteaume, A., A. Camia, M. Jappiot, J. San-Miguel-Ayanz, M. Long-Fournel, and C. Lampin (2013). "A review of the main driving factors of forest fire ignition over Europe". *Environmental management* 51, 651–662.

Gao, B.-C. (1996). "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space". *Remote sensing of environment* 58.3, 257–266.

Gao, T., J. Zhu, X. Zheng, G. Shang, L. Huang, and S. Wu (2015). "Mapping Spatial Distribution of Larch Plantations from Multi-Seasonal Landsat-8 OLI Imagery and Multi-Scale Textures Using Random Forests". *Remote Sensing* 7.2, 1702–1720. DOI: 10.3390/rs70201702.

Gatti, L. V., L. S. Basso, J. B. Miller, M. Gloor, L. Gatti Domingues, H. L. Cassol, G. Tejada, L. E. Aragão, C. Nobre, W. Peters, et al. (2021). "Amazonia as a carbon source linked to deforestation and climate change". *Nature* 595.7867, 388–393.

Gelfand, A. E. and S. Shirota (2021). "The role of odds ratios in joint species distribution modeling". *Environmental and Ecological Statistics* 28.2, 287–302.

Gessler, A., M. Schaub, A. Bose, V. Trotsiuk, R. Valbuena, G. Chirici, and N. Buchmann (2024). "Finding the balance between open access to forest data while safeguarding the integrity of National Forest Inventory-derived information". *New Phytologist*.

Gibson, J., G. Moisen, T. Frescino, and T. C. Edwards (2014). "Using publicly available forest inventory data in climate-based models of tree species distribution: examining effects of true versus altered location coordinates". *Ecosystems* 17, 43–53.

Gillis, M. D., A. Omule, and T. Brierley (2005). "Monitoring Canada's forests: the national forest inventory". *The Forestry Chronicle* 81.2, 214–221.

Giorgetta, M. A., J. Jungclaus, C. H. Reick, S. Legutke, J. Bader, M. Böttinger, V. Brovkin, T. Crueger, M. Esch, K. Fieg, et al. (2013). "Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5". *Journal of Advances in Modeling Earth Systems* 5.3, 572–597.

Giresse, P., J. Maley, and A. Chepstow-Lusty (2023). "A focus on the last 1000 years of natural environmental changes in the tropical rainforests of West and Central Africa. Can we detect anthropogenic disturbances?" *Global and Planetary Change* 220.February 2022, 103995. DOI: 10.1016/j.gloplacha.2022.103995.

Gobeyn, S., A. M. Mouton, A. F. Cord, A. Kaim, M. Volk, and P. L. Goethals (2019). "Evolutionary algorithms for species distribution modelling: A review in the context of machine learning". *Ecological Modelling* 392.June 2018, 179–195. DOI: 10.1016/j.ecolmodel.2018.11.013.

Godsoe, W., J. Franklin, and F. G. Blanchet (2017). "Effects of biotic interactions on modeled species' distribution can be masked by environmental gradients". *Ecology and evolution* 7.2, 654–664.

Gomes, C., H. Nocairi, M. Thomas, F. Ibanez, J.-F. Collin, and G. Saporta (2012). "Stacking prediction for a binary outcome". In: *Compstat 2012*. Limassol, Cyprus, 271–282.

Gonzalez, P., R. P. Neilson, J. M. Lenihan, and R. J. Drapek (2010). "Global patterns in the vulnerability of ecosystems to vegetation shifts due to climate change". *Global Ecology and Biogeography* 19.6, 755–768. DOI: https://doi.org/10.1111/j.1466-8238.2010.00558.x.

Goodbody, T. R., N. C. Coops, J. E. Luther, P. Tompalski, C. Mulverhill, C. Frizzle, R. Fournier, S. Furze, and S. Herniman (2021). "Airborne laser scanning for quantifying criteria and indicators of sustainable forest management in Canada". *Canadian Journal of Forest Research* 51.7, 972–985.

Goodbody, T. R., N. C. Coops, M. Queinnec, J. C. White, P. Tompalski, A. T. Hudak, D. Auty, R. Valbuena, A. LeBoeuf, I. Sinclair, et al. (2023). "sgsR: a structurally guided sampling toolbox for LiDAR-based forest inventories". *Forestry*, cpac055.

Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone". *Remote Sensing of Environment* 202, 18–27. DOI: 10.1016/J.RSE.2017.06.031.

Götmark, F. and C. Kiffer (2014). "Regeneration of oaks (*Quercus robur*/*Q. petraea*) and three other tree species during long-term succession after catastrophic disturbance (windthrow)". *Plant Ecology* 215, 1067–1080. DOI: https://doi.org/10.1007/s11258-014-0365-4.

Gottschalk, T. K., B. Aue, S. Hotes, and K. Ekschmitt (2011). "Influence of grain size on species-habitat models". *Ecological Modelling* 222.18, 3403–3412. DOI: 10.1016/j.ecolmodel.2011.07.008.

Grabska, E., D. Frantz, and K. Ostapowicz (2020). "Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians". *Remote Sensing of Environment* 251, 112103.

Gregoire, T. G. and H. T. Valentine (2007). *Sampling strategies for natural resources and the environment*. CRC Press.

Gschwantner, T., I. Alberdi, S. Bauwens, S. Bender, D. Borota, M. Bosela, O. Bouriaud, J. Breidenbach, J. Donis, C. Fischer, et al. (2022). "Growing stock monitoring by European National Forest Inventories: Historical origins, current methods and harmonisation". *Forest Ecology and Management* 505, 119868.

Guisan, A. and W. Thuiller (2005). "Predicting species distribution: Offering more than simple habitat models". *Ecology Letters* 8.9, 993–1009. DOI: 10.1111/j.1461-0248.2005.00792.x.

Guisan, A., R. Tingley, J. B. Baumgartner, I. Naujokaitis-Lewis, P. R. Sutcliffe, A. I. Tulloch, T. J. Regan, L. Brotons, E. Mcdonald-Madden, C. Mantyka-Pringle, T. G. Martin, J. R. Rhodes, R. Maggini, S. A. Setterfield, J. Elith, M. W. Schwartz, B. A. Wintle, O. Broennimann, M. Austin, S. Ferrier, M. R. Kearney, H. P. Possingham, and Y. M. Buckley (2013). "Predicting species distributions for conservation decisions". *Ecology Letters* 16.12, 1424–1435. DOI: 10.1111/ele.12189.

Hall, M. A. and G. Holmes (2003). "Benchmarking attribute selection techniques for discrete class data mining". *IEEE Transactions on Knowledge and Data engineering* 15.6, 1437–1447.

Hamedianfar, A., C. Mohamedou, A. Kangas, and J. Vauhkonen (2022). "Deep learning for forest inventory and planning: a critical review on the remote sensing approaches so far and prospects for further applications". *Forestry* 95.4, 451–465.

Hampton, S. E., C. A. Strasser, J. J. Tewksbury, W. K. Gram, A. E. Budden, A. L. Batcheller, C. S. Duke, and J. H. Porter (2013). "Big data and the future of ecology". *Frontiers in Ecology and the Environment* 11.3, 156–162. DOI: 10.1890/120103.

Hamrouni, Y., E. Paillassa, V. Chéret, C. Monteil, and D. Sheeren (2021). "From local to global: A transfer learning-based approach for mapping poplar plantations at national scale using Sentinel-2". *ISPRS Journal of Photogrammetry and Remote Sensing* 171, 76–100.

Hanberry, B. B. and M. H. Hansen (2015). "Latitudinal range shifts of tree species in the United States across multi-decadal time scales". *Basic and Applied Ecology* 16.3, 231–238. DOI: https://doi.org/10.1016/j.baae.2015.02.002.

Hanewinkel, M., D. A. Cullmann, M.-J. Schelhaas, G.-J. Nabuurs, and N. E. Zimmermann (2013). "Climate change may cause severe loss in the economic value of European forest land". *Nature climate change* 3.3, 203–207. DOI: https://doi.org/10.1038/nclimate1687.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013). "High-Resolution Global Maps of 21st-Century Forest Cover Change". *Science (New York, N.Y.)* 850.November, 2011–2014. DOI: 10.1126/science.1244693. eprint: 1011.1669v3.

Hao, T., J. Elith, G. Guillera-Arroita, and J. J. Lahoz-Monfort (2019). "A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD". *Diversity and Distributions* 25.5, 839–852.

Hao, T., J. Elith, J. J. Lahoz-Monfort, and G. Guillera-Arroita (2020). "Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models". *Ecography* 43.4, 549–558.

Harris, C. R., K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant (2020). "Array programming with NumPy". *Nature* 585, 357–362. DOI: 10.1038/s41586-020-2649-2.

Harris, N. L., D. A. Gibbs, A. Baccini, R. A. Birdsey, S. De Bruin, M. Farina, L. Fatoyinbo, M. C. Hansen, M. Herold, R. A. Houghton, et al. (2021). "Global maps of twenty-first century forest carbon fluxes". *Nature Climate Change* 11.3, 234–240.

Harrison, S. (2017). "BIOME 6000 DB classified plotfile version 1". DOI: https://doi.org/10.17864/1947.99.

Hastie, T., J. Qian, and K. Tay (2016). *An Introduction to glmnet*.

Hausfather, Z., K. Marvel, G. A. Schmidt, J. W. Nielsen-Gammon, and M. Zelinka (2022). "Climate simulations: recognize the 'hot model' problem". *Nature* 605.7908, 26–29. DOI: https://doi.org/10.1038/d41586-022-01192-2.

Hayhoe, K., J. Edmonds, R. Kopp, A. LeGrande, B. Sanderson, M. Wehner, and D. Wuebbles (2017). "Climate models, scenarios, and projections". In: *Climate Science Special Report: Fourth National Climate Assessment*. Vol. 1. US Global Change Research Program, Washington, DC, 133–160.

He, K. S., J. Zhang, and Q. Zhang (2009). "Linking variability in species composition and MODIS NDVI based on beta diversity measurements". *acta oecologica* 35.1, 14–21.

Hefley, T. J. and M. B. Hooten (2016). "Hierarchical species distribution models". *Current Landscape Ecology Reports* 1.2, 87–97.

Heisig, J. and T. Hengl (2020). *Harmonized Tree Species Occurrence Points for Europe*. URL: https://zenodo.org/record/4061816, Dataset Version: 0.2. DOI: 10.5281/ZENODO.4061816. (Visited on 2021).

Hengl, T., L. Leal Parente, J. Krizan, and C. Bonannella (2020). *Continental Europe Digital Terrain Model at 30 m resolution based on GEDI, ICESat-2, AW3D, GLO-30, EUDEM, MERIT DEM and background layers*. Version v0.3. DOI: 10.5281/zenodo.4724549.

Hengl, T., M. G. Walsh, J. Sanderman, I. Wheeler, S. P. Harrison, and I. C. Prentice (2018). "Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential". *PeerJ* 6. DOI: https://doi.org/10.7717/peerj.5457.

Hermosilla, T., A. Bastyr, N. C. Coops, J. C. White, and M. A. Wulder (2022). "Mapping the presence and distribution of tree species in Canada's forested ecosystems". *Remote Sensing of Environment* 282, 113276.

Hermosilla, T., M. A. Wulder, J. C. White, N. C. Coops, G. W. Hobart, and L. B. Campbell (2016). "Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring". *International Journal of Digital Earth* 9.11, 1035–1054.

Hickler, T., K. Vohland, J. Feehan, P. A. Miller, B. Smith, L. Costa, T. Giesecke, S. Fronzek, T. R. Carter, W. Cramer, I. Kühn, and M. T. Sykes (2012). "Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model". *Global Ecology and Biogeography* 21.1, 50–63.

Hierro, J. L., J. L. Maron, and R. M. Callaway (2005). "A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range". *Journal of ecology* 93.1, 5–15.

Higgins, S. I., R. Buitenwerf, and G. R. Moncrieff (2016). "Defining functional biomes and monitoring their change globally". *Global Change Biology* 22.11, 3583–3593.

Higgins, S. I., T. Conradi, and E. Muhoko (2023). "Shifts in vegetation activity of terrestrial ecosystems attributable to climate trends". *Nature Geoscience*, 1–7.

Higgins, S. I., R. B. O'Hara, O. Bykova, M. D. Cramer, I. Chuine, E.-M. Gerstner, T. Hickler, X. Morin, M. R. Kearney, G. F. Midgley, et al. (2012). "A physiological analogy of the niche for projecting the potential distribution of plants". *Journal of Biogeography* 39.12, 2132–2145.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005). "Very high resolution interpolated climate surfaces for global land areas". *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25.15, 1965–1978.

Hijmans, R. J. (2021). *terra: Spatial Data Analysis*. R package version 1.4-0.

Himes, A., K. Puettmann, and B. Muraca (2020). "Trade-offs between ecosystem services along gradients of tree species diversity and values". *Ecosystem services* 44, 101133. DOI: https://doi.org/10.1016/j.ecoser.2020.101133.

Hlásny, T., S. Zimová, K. Merganičová, P. Štěpánek, R. Modlinger, and M. Turčáni (2021a). "Devastating outbreak of bark beetles in the Czech Republic: Drivers, impacts, and management implications". *Forest Ecology and Management* 490, 119075. DOI: https://doi.org/10.1016/j.foreco.2021.119075.

Hlásny, T., L. König, P. Krokene, M. Lindner, C. Montagné-Huck, J. Müller, H. Qin, K. F. Raffa, M.-J. Schelhaas, M. Svoboda, et al. (2021b). "Bark beetle outbreaks in Europe: state of knowledge and ways forward for management". *Current Forestry Reports* 7, 138–165. DOI: https://doi.org/10.1007/s40725-021-00142-x.

Hof, A. R., C. C. Dymond, and D. J. Mladenoff (2017). "Climate change mitigation through adaptation: the effectiveness of forest diversification by novel tree planting regimes". *Ecosphere* 8.11, e01981.

Hoffer, R. (1984). "Remote sensing to measure the distribution and structure of vegetation". *The Role of Terrestrial Vegetation in the Global Carbon Cycle: Measurement by Remote Sensing*, 131–59.

Hoogen, J. van den, N. Robmann, D. Routh, T. Lauber, N. van Tiel, O. Danylo, and T. W. Crowther (2021). "A geospatial mapping pipeline for ecologists". *BioRxiv*. DOI: 10.1101/2021.07.07.451145.

Hosseinzadeh, M., A. Wachal, H. Khamfroush, and D. E. Lucani (2021). "Optimal accuracy-time trade-off for deep learning services in edge computing systems". In: *ICC 2021-IEEE International Conference on Communications*. IEEE, 1–6.

Huang, S., L. Tang, J. P. Hupy, Y. Wang, and G. Shao (2021). "A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing". *Journal of Forestry Research* 32.1, 1–6.

Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira (2002). "Overview of the radiometric and biophysical performance of the MODIS vegetation indices". *Remote sensing of environment* 83.1-2, 195–213.

Huete, A. R. (1988). "A soil-adjusted vegetation index (SAVI)". *Remote sensing of environment* 25.3, 295–309.

Hunter, J. D. (2007). "Matplotlib: A 2D graphics environment". *Computing in science & engineering* 9.3, 90–95.

Hurtt, G. C., L. Chini, R. Sahajpal, S. Frolking, B. L. Bodirsky, K. Calvin, J. C. Doelman, J. Fisk, S. Fujimori, K. Klein Goldewijk, et al. (2020). "Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6". *Geoscientific Model Development* 13.11, 5425–5464. DOI: https://doi.org/10.5194/gmd-13-5425-2020.

ICP Forests (2010). *Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests*. UNECE ICP Forests Programme Co-ordinating Centre. Hamburg.

Ij, H. (2018). "Statistics versus machine learning". *Nat Methods* 15.4, 233.

IPCC (2014). "Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]", 1–151.

– (2021). "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen". *Cambridge University Press* In Press.

– (2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK and New York, USA: Cambridge University Press.

Isajev, V., B. Fady, H. Semerci, and V. Andonovski (2004). *EUFORGEN Technical Guidelines for genetic conservation and use for European Black pine (*Pinus nigra*)*. Bioversity International.

Iturbide, M., J. Bedia, and J. M. Gutiérrez (2018a). "Background sampling and transferability of species distribution model ensembles under climate change". *Global and Planetary Change* 166.March, 19–29. DOI: 10.1016/j.gloplacha.2018.03.008.

– (2018b). "Tackling Uncertainties of Species Distribution Model Projections with Package mopa." *R Journal* 10.1.

Jaime, L., E. Batllori, J. Margalef-Marrase, M. Á. P. Navarro, and F. Lloret (2019). "Scots pine (*Pinus sylvestris* L.) mortality is explained by the climatic suitability of both host tree and bark beetle populations". *Forest Ecology and Management* 448, 119–129. DOI: https://doi.org/10.1016/j.foreco.2019.05.070.

Jiang, Z., A. R. Huete, K. Didan, and T. Miura (2008). "Development of a two-band enhanced vegetation index without a blue band". *Remote sensing of Environment* 112.10, 3833–3845.

Jiménez-Valverde, A., A. T. Peterson, J. Soberón, J. Overton, P. Aragón, and J. M. Lobo (2011). "Use of niche models in invasive species risk assessments". *Biological invasions* 13.12, 2785–2797.

Jonášová, M., E. Vávrová, and P. Cudlín (2010). "Western Carpathian mountain spruce forest after a windthrow: Natural regeneration in cleared and uncleared areas". *Forest Ecology and Management* 259.6, 1127–1134. DOI: https://doi.org/10.1016/j.foreco.2009.12.027.

Jonckheere, I., R. Hamilton, J. Michel, and E. Donegan (2024). *Good practices in sample-based area estimation*. Rome. DOI: 10.4060/cc9276en.

Jones, B. and B. C. O'Neill (2016). "Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways". *Environmental Research Letters* 11.8, 084003. DOI: https://doi.org/10.1088/1748-9326/11/8/084003.

Jones, E. W. (1959). "*Quercus* L." *Journal of Ecology* 47.1, 169–222.

Jump, A. S. and J. Peñuelas (2005). "Running to stand still: adaptation and the response of plants to rapid climate change". *Ecology letters* 8.9, 1010–1020. DOI: https://doi.org/10.1111/j.1461-0248.2005.00796.x.

Kaplan, J. O., K. M. Krumhardt, and N. Zimmermann (2009). "The prehistoric and preindustrial deforestation of Europe". *Quaternary science reviews* 28.27-28, 3016–3034. DOI: https://doi.org/10.1016/j.quascirev.2009.09.028.

Karger, D. N., O. Conrad, J. Böhner, T. Kawohl, H. Kreft, R. W. Soria-Auza, N. E. Zimmermann, H. P. Linder, and M. Kessler (2017). "Climatologies at high resolution for the earth's land surface areas". *Scientific data* 4.1, 1–20.

Karger, D. N., B. Dabaghchian, S. Lange, W. Thuiller, N. E. Zimmermann, and C. H. Graham (2020). *"High resolution climate data for Europe"*. DOI: http://dx.doi.org/10.16904/envidat.150.

Kattge, J., G. Bönisch, S. Díaz, S. Lavorel, I. C. Prentice, P. Leadley, S. Tautenhahn, G. D. Werner, T. Aakala, M. Abedi, et al. (2020). "TRY plant trait database–enhanced coverage and open access". *Global change biology* 26.1, 119–188.

Kautz, M., F. J. Peter, L. Harms, S. Kammen, and H. Delb (2023). "Patterns, drivers and detectability of infestation symptoms following attacks by the European spruce bark beetle". *Journal of Pest Science* 96.1, 403–414. DOI: https://doi.org/10.1007/s10340-022-01490-8.

Keenan, R. J. (2015). "Climate change impacts and adaptation in forest management: a review". *Annals of Forest Science* 72.2, 145–167. DOI: 10.1007/s13595-014-0446-5.

Keith, D. A., J. R. Ferrer, E. Nicholson, M. J. Bishop, B. A. Polidoro, E. Ramirez-Llodra, M. G. Tozer, J. L. Nel, R. Mac Nally, E. J. Gregr, et al. (2020). "The IUCN global ecosystem typology v1. 01: Descriptive profiles for biomes and ecosystem functional groups".

Kennedy, R. E., Z. Yang, and W. B. Cohen (2010). "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms". *Remote Sensing of Environment* 114.12, 2897–2910.

Key, C. H. and N. C. Benson (1999). "The Normalized Burn Ratio (NBR): A Landsat TM radiometric measure of burn severity". *United States Geological Survey, Northern Rocky Mountain Science Center: Bozeman, MT, USA*.

Klaus, M., A. Holsten, P. Hostert, and J. P. Kropp (2011). "Integrated methodology to assess windthrow impacts on forest stands under climate change". *Forest Ecology and Management* 261.11, 1799–1810. DOI: https://doi.org/10.1016/j.foreco.2011.02.002.

Kolář, T., P. Čermák, M. Trnka, T. Žid, and M. Rybníček (2017). "Temporal changes in the climate sensitivity of Norway spruce and European beech along an elevation gradient in Central Europe". *Agricultural and Forest Meteorology* 239, 24–33. DOI: https://doi.org/10.1016/j.agrformet.2017.02.028.

Krakau, U.-K., M. Liesebach, T. Aronen, M.-A. Lelu-Walter, and V. Schneck (2013). "Scots pine (*Pinus sylvestris* L.)" In: *Forest tree breeding in Europe: Current state-of-the-art and perspectives*. Springer, 267–323. DOI: https://doi.org/10.1007/978-94-007-6146-9_6.

Krause, A., P. Papastefanou, K. Gregor, L. S. Layritz, C. S. Zang, A. Buras, X. Li, J. Xiao, and A. Rammig (2022). "Quantifying the impacts of land cover change on gross primary productivity globally". *Scientific Reports* 12.1, 1–10. DOI: https://doi.org/10.1038/s41598-022-23120-0.

Kronauer, H. (2007). "Ruhe nach dem Sturm? Schäden durch Orkan Kyrill". *AFZ-Der Wald* 62.5, 250–251.

Kronauer, H. (2000). "Schwere Sturmschäden in Baden-Württemberg: Lothar stellt Wiebke in den Schatten". *AFZ/Der*.

Kuhn, M. (2021). *caret: Classification and Regression Training*. R package version 6.0-88.

Lakshminarayanan, B., A. Pritzel, and C. Blundell (2016). "Simple and scalable predictive uncertainty estimation using deep ensembles". *arXiv preprint arXiv:1612.01474*.

Lang, N., W. Jetz, K. Schindler, and J. D. Wegner (2023). "A high-resolution canopy height model of the Earth". *Nature Ecology & Evolution* 7.11, 1778–1789.

Lasslop, G., S. Hantson, S. P. Harrison, D. Bachelet, C. Burton, M. Forkel, M. Forrest, F. Li, J. R. Melton, C. Yue, et al. (2020). "Global ecosystems and fire: Multi-model

assessment of fire-induced tree-cover and carbon storage reduction". *Global Change Biology* 26.9, 5027–5041.

Latałowa, M. and W. O. van der Knaap (2006). "Late Quaternary expansion of Norway spruce (*Picea abies* L.) Karst. in Europe according to pollen data". *Quaternary Science Reviews* 25.21-22, 2780–2805. DOI: https://doi.org/10.1016/j.quascirev.2006.06.007.

Latifi, H., F. E. Fassnacht, F. Hartig, C. Berger, J. Hernández, P. Corvalán, and B. Koch (2015). "Stratified aboveground forest biomass estimation by remote sensing data". *International Journal of Applied Earth Observation and Geoinformation* 38, 229–241.

Latifi, H., A. Nothdurft, C. Straub, and B. Koch (2012). "Modelling stratified forest attributes using optical/LiDAR features in a central European landscape". *International Journal of Digital Earth* 5.2, 106–132.

Lawton, J. H. (1998). "Daily, G.C. (Ed.). 1997. Nature's services. Societal dependence on natural ecosystems. Island Press, Washington, DC. 392 pp. ISBN 1-55963-475-8 hbk), 1 55963 476 6 (soft cover". In: *Animal Conservation forum*. Vol. 1. 1. Cambridge University Press, 75–76. DOI: 10.1017/s1367943098221123.

Leeper, T. J. (2017). *aws.s3: AWS S3 Client Package*. R package version 0.3.8.

Lefebvre, D., A. G. Williams, G. J. Kirk, J. Burgess, J. Meersmans, M. R. Silman, F. Román-Dañobeytia, J. Farfan, P. Smith, et al. (2021). "Assessing the carbon capture potential of a reforestation project". *Scientific reports* 11.1, 1–10.

Lenoir, J., J.-C. Gégout, P. A. Marquet, P. de Ruffray, and H. Brisse (2008). "A significant upward shift in plant species optimum elevation during the 20th century". *Science* 320.5884, 1768–1771. DOI: https://doi.org/10.1111/ecog.00967.

Lenoir, J. and J.-C. Svenning (2015). "Climate-related range shifts–a global multidimensional synthesis and new research directions". *Ecography* 38.1, 15–28.

Li, L., M. A. Friedl, Q. Xin, J. Gray, Y. Pan, and S. Frolking (2014). "Mapping crop cycles in China using MODIS-EVI time series". *Remote Sensing* 6.3, 2473–2493.

Lindenmayer, D. B., P. J. Burton, and J. F. Franklin (2012). *Salvage logging and its ecological consequences*. Island Press.

Lindgren, A., Z. Lu, Q. Zhang, and G. Hugelius (2021). "Reconstructing past global vegetation with random forest machine learning, sacrificing the dynamic response for robust results". *Journal of Advances in Modeling Earth Systems* 13.2, e2020MS002200.

Lindner, M., M. Maroschek, S. Netherer, A. Kremer, A. Barbati, J. Garcia-Gonzalo, R. Seidl, S. Delzon, P. Corona, M. Kolström, et al. (2010). "Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems". *Forest ecology and management* 259.4, 698–709. DOI: https://doi.org/10.1016/j.foreco.2009.09.023.

Lister, A. J., H. Andersen, T. Frescino, D. Gatziolis, S. Healey, L. S. Heath, G. C. Liknes, R. McRoberts, G. G. Moisen, M. Nelson, et al. (2020). "Use of remote sensing data to

improve the efficiency of national forest inventories: a case study from the United States national forest inventory". *Forests* 11.12, 1364.

Liu, Z., C. Peng, T. Work, J.-N. Candau, A. DesRochers, and D. Kneeshaw (2018). "Application of machine-learning methods in forest ecology: recent progress and future challenges". *Environmental Reviews* 26.4, 339–350.

Lobo, J. M., A. Jiménez-Valverde, and J. Hortal (2010). "The uncertain nature of absences and their importance in species distribution modelling". *Ecography* 33.1, 103–114. DOI: 10.1111/j.1600-0587.2009.06039.x.

Łoś, H., G. S. Mendes, D. Cordeiro, N. Grosso, H. Costa, P. Benevides, and M. Caetano (2021). "Evaluation of Xgboost and Lgbm Performance in Tree Species Classification with Sentinel-2 Data". In: *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 5803–5806.

Lovejoy, T. E. and C. Nobre (2018). *Amazon tipping point*.

Lucas-Borja, M. E., J. T. Van Stan, M. Heydari, R. Omidipour, F. Rocha, P. A. Plaza-Alvarez, D. A. Zema, and M. Muñoz-Rojas (2021). "Post-fire restoration with contour-felled log debris increases early recruitment of Spanish black pine (*Pinus nigra* Arn. ssp. *salzmannii*) in Mediterranean forests". *Restoration Ecology* 29.4, e13338. DOI: https://doi.org/10.1111/rec.13338.

Madonsela, S., M. A. Cho, A. Ramoelo, and O. Mutanga (2017). "Remote sensing of species diversity using Landsat 8 spectral variables". *ISPRS Journal of Photogrammetry and Remote Sensing* 133, 116–127. DOI: https://doi.org/10.1016/j.isprsjprs.2017.10.008.

Maes, J., A. G. Bruzón, J. I. Barredo, S. Vallecillo, P. Vogt, I. M. Rivero, and F. Santos-Martín (2023). "Accounting for forest condition in Europe based on an international statistical standard". *Nature Communications* 14.1, 3723. DOI: https://doi.org/10.1038/s41467-023-39434-0.

Maharjan, S. K., F. J. Sterck, N. Raes, Y. Zhao, and L. Poorter (2023). "Climate change induced elevational range shifts of Himalayan tree species". *Biotropica* 55.1, 53–69.

Mahecha, M. D., F. Gans, G. Brandt, R. Christiansen, S. E. Cornell, N. Fomferra, G. Kraemer, J. Peters, P. Bodesheim, G. Camps-Valls, et al. (2020). "Earth system data cubes unravel global multivariate dynamics". *Earth System Dynamics* 11.1, 201–234. DOI: https://doi.org/10.5194/esd-11-201-2020.

Mäkinen, H., P. Nöjd, and S. Helama (2022). "Recent unexpected decline of forest growth in North Finland: examining tree-ring, climatic and reproduction data". *Silva Fennica* 56.4. DOI: https://doi.org/10.14214/sf.10769.

Maksic, J., I. Venancio, M. Shimizu, C. Chiessi, P. Piacsek, G. Sampaio, F. W. Cruz, and F. Alexandre (2022). "Brazilian biomes distribution: Past and future". *Palaeogeography, Palaeoclimatology, Palaeoecology* 585, 110717.

Málaga, N., S. De Bruin, R. E. McRoberts, A. A. Olivos, R. de la Cruz Paiva, P. D. Montesinos, D. R. Suarez, and M. Herold (2022). "Precision of subnational forest AGB estimates within the Peruvian Amazonia using a global biomass map". *International Journal of Applied Earth Observation and Geoinformation* 115, 103102.

Malhi, Y., S. Adu-Bredu, R. A. Asare, S. L. Lewis, and P. Mayaux (2013). "African rainforests: past, present and future". *Philosophical Transactions of the Royal Society B: Biological Sciences* 368.1625, 20120312.

Manzoor, S. A., G. Griffiths, and M. Lukac (2018). "Species distribution model transferability and model grain size-finer may not always be better". *Scientific Reports* 8.1, 1–9. DOI: 10.1038/s41598-018-25437-1.

Marconi, S., B. G. Weinstein, S. Zou, S. A. Bohlman, A. Zare, A. Singh, D. Stewart, I. Harmon, A. Steinkraus, and E. P. White (2022). "Continental-scale hyperspectral tree species classification in the United States National Ecological Observatory Network". *Remote Sensing of Environment* 282, 113264.

Marengo, J. A. and J. C. Espinoza (2016). "Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts". *International Journal of Climatology* 36.3, 1033–1050.

Markus, T., T. Neumann, A. Martino, W. Abdalati, K. Brunt, B. Csatho, S. Farrell, H. Fricker, A. Gardner, D. Harding, et al. (2017). "The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation". *Remote sensing of environment* 190, 260–273.

Martinez del Castillo, E., C. S. Zang, A. Buras, A. Hacket-Pain, J. Esper, R. Serrano-Notivoli, C. Hartl, R. Weigel, S. Klesse, V. Resco de Dios, et al. (2022). "Climate-change-driven growth decline of European beech forests". *Communications biology* 5.1, 1–9.

Martínez-Vilalta, J., B. C. López, N. Adell, L. Badiella, and M. Ninyerola (2008). "Twentieth century increase of Scots pine radial growth in NE Spain shows strong climate interactions". *Global Change Biology* 14.12, 2868–2881. DOI: https://doi.org/10.1111/j.1365-2486.2008.01685.x.

Mathes, T., D. Seidel, and P. Annighöfer (2023). "Response to extreme events: do morphological differences affect the ability of beech (*Fagus sylvatica* L.) to resist drought stress?" *Forestry* 96.3, 355–371. DOI: https://doi.org/10.1093/forestry/cpac056.

Mauri, A., M. Girardello, G. Forzieri, F. Manca, P. S. Beck, A. Cescatti, and G. Strona (2023). "Assisted tree migration can reduce but not avert the decline of forest ecosystem services in Europe". *Global Environmental Change* 80, 102676. DOI: https://doi.org/10.1016/j.gloenvcha.2023.102676.

Mauri, A., M. Girardello, G. Strona, P. S. Beck, G. Forzieri, G. Caudullo, F. Manca, and A. Cescatti (2022). "EU-Trees4F, a dataset on the future distribution of European tree species". *Scientific data* 9.1, 1–12.

Mauri, A., G. Strona, and J. San-Miguel-Ayanz (2017). "EU-Forest, a high-resolution tree occurrence dataset for Europe". *Scientific Data* 4.1, 160123. DOI: 10.1038/sdata.2016.123.

Mayer, A. L. and G. N. Cameron (2003). "Consideration of grain and extent in landscape studies of terrestrial vertebrate ecology". *Landscape and Urban Planning* 65.4, 201–217. DOI: 10.1016/S0169-2046(03)00057-4.

McKinney, W. et al. (2010). "Data structures for statistical computing in python". In: *Proceedings of the 9th Python in Science Conference*. Vol. 445. Austin, TX, 51–56.

McRoberts, R. E. (2006). "A model-based approach to estimating forest area". *Remote Sensing of Environment* 103.1, 56–66.

McRoberts, R. E., T. Gobakken, and E. Næsset (2012). "Post-stratified estimation of forest area and growing stock volume using lidar-based stratifications". *Remote Sensing of Environment* 125, 157–166.

McRoberts, R. E., G. R. Holden, M. D. Nelson, G. C. Liknes, W. K. Moser, A. J. Lister, S. L. King, E. B. LaPoint, J. W. Coulston, W. B. Smith, et al. (2005). "Estimating and circumventing the effects of perturbing and swapping inventory plot locations". *Journal of Forestry* 103.6, 275–279.

McRoberts, R. E., E. Næsset, S. Saatchi, and S. Quegan (2022). "Statistically rigorous, model-based inferences from maps". *Remote Sensing of Environment* 279, 113028.

McRoberts, R. E., E. O. Tomppo, and E. Næsset (2010). "Advances and emerging issues in national forest inventories". *Scandinavian Journal of Forest Research* 25.4, 368–381.

Mehra, A., P. Tripathy, A. Faridi, and A. Chinmay (2019). "Ensemble Learning Approach to Improve Existing Models". *International Journal of Innovative Science and Research Technology* 4 (12).

Menezes-Silva, P. E., L. Loram-Lourenço, R. D. F. B. Alves, L. F. Sousa, S. E. d. S. Almeida, and F. S. Farnese (2019). "Different ways to die in a changing world: Consequences of climate change for tree species performance and survival through an ecophysiological perspective". *Ecology and evolution* 9.20, 11979–11999. DOI: https://doi.org/10.1002/ece3.5663.

Mennis, J. (2006). "Exploring the influence of ENSO on African vegetation variability using multidimensional map algebra". *GIScience & Remote Sensing* 43.4, 352–376.

Meyer, H. and E. Pebesma (2021). "Predicting into unknown space? Estimating the area of applicability of spatial prediction models". *Methods in Ecology and Evolution* 12.9, 1620–1633. DOI: https://doi.org/10.1111/2041-210X.13650.

Mezei, P., R. Jakuš, J. Pennerstorfer, M. Havašová, J. Škvarenina, J. Ferenčík, J. Slivinskỳ, S. Bičárová, D. Bilčík, M. Blaženec, et al. (2017). "Storms, temperature maxima and the Eurasian spruce bark beetle *Ips typographus*—An infernal trio in Norway spruce forests

of the Central European High Tatra Mountains". *Agricultural and Forest Meteorology* 242, 85–95. DOI: https://doi.org/10.1016/j.agrformet.2017.04.004.

Mitchell, S. (2013). "Wind as a natural disturbance agent in forests: a synthesis". *Forestry* 86.2, 147–157. DOI: https://doi.org/10.1093/forestry/cps058.

Moncrieff, G. R., W. J. Bond, and S. I. Higgins (2016). "Revising the biome concept for understanding and predicting global change impacts". *Journal of Biogeography* 43.5, 863–873.

Moncrieff, G. R., T. Hickler, and S. I. Higgins (2015). "Intercontinental divergence in the climate envelope of major plant biomes". *Global Ecology and Biogeography* 24.3, 324–334.

Morin, X., D. Viner, and I. Chuine (2008). "Tree species range shifts at a continental scale: new predictive insights from a process-based model". *Journal of Ecology* 96.4, 784–794.

Moris, J. V., R. Berretti, A. Bono, R. Sino, G. Minotta, M. Garbarino, R. Motta, G. Vacchiano, J. Maringer, M. Conedera, et al. (2023). "Resprouting in European beech confers resilience to high-frequency fire". *Forestry* 96.3, 372–386. DOI: https://doi.org/10.1093/forestry/cpac018.

Mountrakis, G., J. Im, and C. Ogole (2011). "Support vector machines in remote sensing: A review". *ISPRS journal of photogrammetry and remote sensing* 66.3, 247–259.

Mucina, L. (2019). "Biome: evolution of a crucial ecological and biogeographical concept". *New Phytologist* 222.1, 97–114.

Nabuurs, G.-J., W. Daamen, G. Dirkse, J. Paasman, P. Kuikman, and A. Verhagen (2003). *Present readiness of, and white spots in the Dutch national system for greenhouse gas reporting of the land use, land-use change and forestry sector (LULUCF)*. Tech. rep. Alterra.

Nabuurs, G.-J., N. Harris, D. Sheil, M. Palahi, G. Chirici, M. Boissière, C. Fay, J. Reiche, and R. Valbuena (2022). "Glasgow forest declaration needs new modes of data ownership". *Nature Climate Change*, 1–3.

Næsset, E. (2002). "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data". *Remote sensing of environment* 80.1, 88–99.

Næsset, E., R. E. McRoberts, A. Pekkarinen, S. Saatchi, M. Santoro, Ø. D. Trier, E. Zahabu, and T. Gobakken (2020). "Use of local and global maps of forest canopy height and aboveground biomass to enhance local estimates of biomass in miombo woodlands in Tanzania". *International Journal of Applied Earth Observation and Geoinformation* 93, 102138.

Nandy, S., R. Singh, S. Ghosh, T. Watham, S. P. S. Kushwaha, A. S. Kumar, and V. K. Dadhwal (2017). "Neural network-based modelling for forest biomass assessment". *Carbon Management* 8.4, 305–317. DOI: 10.1080/17583004.2017.1357402.

Nave, L. E., B. F. Walters, K. Hofmeister, C. H. Perry, U. Mishra, G. M. Domke, and C. Swanston (2019). "The role of reforestation in carbon sequestration". *New Forests* 50.1, 115–137.

Neale, R. B., C. Chen, A. Gettelman, P. H. Lauritzen, S. Park, D. L. Williamson, A. J. Conley, R. Garcia, D. Kinnison, J. F. Lamarque, D. Marsh, M. Mills, A. K. Smith, S. Tilmes, F. Vitt, H. Morrison, P. Cameron Smith, W. D. Collins, M. J. Iacono, R. C. Easter, S. J. Ghan, X. Liu, P. J. Rasch, and M. A. Taylor (2010). *Description of the NCAR Community Atmosphere Model (CAM5.0)*.

Nelder, J. A. and R. W. M. Wedderburn (1972). "Generalized Linear Models". *Journal of the Royal Statistical Society. Series A (General)* 135.3, 370–384.

Nesha, M. K., M. Herold, V. De Sy, A. E. Duchelle, C. Martius, A. Branthomme, M. Garzuglia, O. Jonsson, and A. Pekkarinen (2021). "An assessment of data sources, data quality and changes in national forest monitoring capacities in the Global Forest Resources Assessment 2005–2020". *Environmental Research Letters* 16.5, 054029.

Neumann, M., V. Mues, A. Moreno, H. Hasenauer, and R. Seidl (2017). "Climate variability drives recent tree mortality in Europe". *Global Change Biology* 23.11, 4788–4797. DOI: https://doi.org/10.1111/gcb.13724.

Nezami, S., E. Khoramshahi, O. Nevalainen, I. Pölönen, and E. Honkavaara (2020). "Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks". *Remote Sensing* 12.7, 1070.

Nolan, C., J. T. Overpeck, J. R. Allen, P. M. Anderson, J. L. Betancourt, H. A. Binney, S. Brewer, M. B. Bush, B. M. Chase, R. Cheddadi, et al. (2018). "Past and future global transformation of terrestrial ecosystems under climate change". *Science* 361.6405, 920–923. DOI: https://doi.org/10.1126/science.aan5360.

O'Lear, S., ed. (2020). *A Research Agenda for Environmental Geopolitics*. Elgar Research Agendas. Edward Elgar Publishing.

Olaya, V. (2009). "Chapter 6 Basic Land-Surface Parameters". In: *Geomorphometry*. Ed. by T. Hengl and H. I. Reuter. Vol. 33. Developments in Soil Science. Elsevier, 141–169. DOI: 10.1016/S0166-2481(08)00006-8.

Olofsson, P., G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder (2014). "Good practices for estimating area and assessing accuracy of land change". *Remote sensing of Environment* 148, 42–57.

Olsson, H. (2015). "Forestry remote sensing in Sweden". In: *Proceedings of the EARSeL Symposium; June*, 15–18.

Ouaknine, A., T. Kattenborn, E. Laliberté, and D. Rolnick (2023). "OpenForest: A data catalogue for machine learning in forest monitoring". *arXiv preprint arXiv:2311.00277*.

Päivinen, R., R. Astrup, R. A. Birdsey, J. Breidenbach, J. Fridman, A. Kangas, P. E. Kauppi, M. Köhl, K. T. Korhonen, V. K. Johannsen, et al. (2023). "Ensure forest-data integrity for climate change studies". *Nature Climate Change*, 1–2.

Pardos, M., M. Del Río, H. Pretzsch, H. Jactel, K. Bielak, F. Bravo, G. Brazaitis, E. Defossez, M. Engel, K. Godvod, et al. (2021). "The greater resilience of mixed forests to drought mainly depends on their composition: Analysis along a climate gradient across Europe". *Forest Ecology and Management* 481, 118687. DOI: https://doi.org/10.1016/j.foreco.2020.118687.

Patacca, M., M. Lindner, M. E. Lucas-Borja, T. Cordonnier, G. Fidej, B. Gardiner, Y. Hauf, G. Jasinevičius, S. Labonne, E. Linkevičius, et al. (2023). "Significant increase in natural disturbance impacts on European forests since 1950". *Global Change Biology* 29.5, 1359–1376. DOI: https://doi.org/10.1111/gcb.16531.

Pearson, R. G. and T. P. Dawson (2003). "Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?" *Global ecology and biogeography* 12.5, 361–371.

Pearson, R. G., W. Thuiller, M. B. Araújo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T. P. Dawson, and D. C. Lees (2006). "Model-based uncertainty in species range prediction". *Journal of biogeography* 33.10, 1704–1711.

Pebesma, E. (2018). "Simple Features for R: Standardized Support for Spatial Vector Data". *The R Journal* 10.1, 439–446. DOI: 10.32614/RJ-2018-009.

Pebesma, E. J. and R. S. Bivand (2005). "Classes and methods for spatial data in R". *R News* 5.2, 9–13.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. (2011). "Scikit-learn: Machine learning in Python". *Journal of machine learning research* 12.Oct, 2825–2830.

Pekel, J.-F., A. Cottam, N. Gorelick, and A. S. Belward (2016). "High-resolution mapping of global surface water and its long-term changes". *Nature* 540.7633, 418–422.

Peltola, H., M.-L. Nykänen, and S. Kellomäki (1997). "Model computations on the critical combination of snow loading and windspeed for snow damage of Scots pine, Norway spruce and Birch sp. at stand edge". *Forest Ecology and Management* 95.3, 229–241. DOI: https://doi.org/10.1016/S0378-1127(97)00037-6.

Pérez Chaves, P., K. Ruokolainen, and H. Tuomisto (2018). "Using remote sensing to model tree species distribution in Peruvian lowland Amazonia". *Biotropica* 50.5, 758–767. DOI: 10.1111/btp.12597.

Pietrzykowski, M. and B. Woś (2021). "The impact of climate change on forest tree species dieback and changes in their distribution". *Climate Change and the Microbiome: Sustenance of the Ecosphere*, 447–460. DOI: https://doi.org/10.1007/978-3-030-76863-8_23.

Popkin, G. (2021). "Germany's trees are dying. A fierce debate has broken out over how to respond". *Science* 374 (6572). DOI: 10.1126/science.acx9735.

Porfirio, L. L., R. M. Harris, E. C. Lefroy, S. Hugh, S. F. Gould, G. Lee, N. L. Bindoff, and B. Mackey (2014). "Improving the use of species distribution models in conservation planning and management under climate change". *PLoS ONE* 9.11, 1–21. DOI: 10.1371/journal.pone.0113749.

Potapov, P., M. C. Hansen, I. Kommareddy, A. Kommareddy, S. Turubanova, A. Pickens, B. Adusei, A. Tyukavina, and Q. Ying (2020). "Landsat analysis ready data for global land cover and land cover change mapping". *Remote Sensing* 12.3, 426.

Potapov, P., M. C. Hansen, A. Pickens, A. Hernandez-Serna, A. Tyukavina, S. Turubanova, V. Zalles, X. Li, A. Khan, F. Stolle, et al. (2022). "The global 2000-2020 land cover and land use change dataset derived from the Landsat archive: first results". *Frontiers in Remote Sensing* 3, 856903.

Potapov, P., X. Li, A. Hernandez-Serna, A. Tyukavina, M. C. Hansen, A. Kommareddy, A. Pickens, S. Turubanova, H. Tang, C. E. Silva, et al. (2021). "Mapping global forest canopy height through integration of GEDI and Landsat data". *Remote Sensing of Environment* 253, 112165.

Prates-Clark, C. D. C., S. S. Saatchi, and D. Agosti (2008). "Predicting geographical distribution models of high-value timber trees in the Amazon Basin using remotely sensed data". *Ecological Modelling* 211.3-4, 309–323. DOI: 10.1016/j.ecolmodel.2007.09.024.

Prentice, I. C. and T. Webb III (1998). "BIOME 6000: reconstructing global mid-Holocene vegetation patterns from palaeoecological records". *Journal of Biogeography* 25.6, 997–1005.

Pretzsch, H., D. I. Forrester, and T. Rötzer (2015). "Representation of species mixing in forest growth models. A review and perspective". *Ecological Modelling* 313, 276–292.

Probst, P., A.-L. Boulesteix, and B. Bischl (2019). "Tunability: Importance of hyperparameters of machine learning algorithms". *The Journal of Machine Learning Research* 20.1, 1934–1965.

Qi, J., A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian (1994). "A modified soil adjusted vegetation index". *Remote sensing of environment* 48.2, 119–126.

Qiao, F., Z. Song, Y. Bao, Y. Song, Q. Shu, C. Huang, and W. Zhao (2013). "Development and evaluation of an Earth System Model with surface gravity waves". *Journal of Geophysical Research: Oceans* 118.9, 4514–4524.

Qiao, H., X. Feng, L. E. Escobar, A. T. Peterson, J. Soberón, G. Zhu, and M. Papeş (2019). "An evaluation of transferability of ecological niche models". *Ecography* 42.3, 521–534.

Quinlan, J. R. (1986). "Induction of decision trees". *Machine Learning* 1.1, 81–106. DOI: 10.1007/BF00116251.

R Core Team (2021a). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

– (2021b). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raczko, E. and B. Zagajewski (2017). "Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images". *European Journal of Remote Sensing* 50.1, 144–154.

Radeloff, V. C., D. P. Roy, M. A. Wulder, M. Anderson, B. Cook, C. J. Crawford, M. Friedl, F. Gao, N. Gorelick, M. Hansen, et al. (2024). "Need and vision for global medium-resolution Landsat and Sentinel-2 data products". *Remote Sensing of Environment* 300, 113918.

Rees, W. G., A. Hofgaard, S. Boudreau, D. M. Cairns, K. Harper, S. Mamet, I. Mathisen, Z. Swirad, and O. Tutubalina (2020). "Is subarctic forest advance able to keep pace with climate change?" *Global Change Biology* 26.7, 3965–3977.

Reiche, J., A. Mullissa, B. Slagter, Y. Gou, N.-E. Tsendbazar, C. Odongo-Braun, A. Vollrath, M. J. Weisse, F. Stolle, A. Pickens, et al. (2021). "Forest disturbance alerts for the Congo Basin using Sentinel-1". *Environmental Research Letters* 16.2, 024005.

Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, et al. (2019). "Deep learning and process understanding for data-driven Earth system science". *Nature* 566.7743, 195–204. DOI: https://doi.org/10.1038/s41586-019-0912-1.

Rigling, A., C. Bigler, B. Eilmann, E. Feldmeyer-Christe, U. Gimmi, C. Ginzler, U. Graf, P. Mayer, G. Vacchiano, P. Weber, et al. (2013). "Driving factors of a vegetation shift from Scots pine to pubescent oak in dry Alpine forests". *Global Change Biology* 19.1, 229–240. DOI: https://doi.org/10.1111/gcb.12038.

Rigo, D. de, T. Houston Durrant, G. Caudullo, and J. I. Barredo (2016). "European forests: an ecological overview". In: *European Atlas of Forest Tree Species*. Publication Office of the European Union, Luxembourg, 24–31.

Ripley, B. and W. Venables (2017). *nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models*. R package version 7.3-12. URL: http://CRAN.R-project.org/package=nnet.

Roberts, D. R., V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, et al. (2017). "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure". *Ecography* 40.8, 913–929. DOI: https://doi.org/10.1111/ecog.02881.

Rodrigo, A., J. Retana, and F. X. Picó (2004). "Direct regeneration is not the only response of Mediterranean forests to large fires". *Ecology* 85.3, 716–729.

Rolf, E., K. Klemmer, C. Robinson, and H. Kerner (2024). "Mission Critical–Satellite Data is a Distinct Modality in Machine Learning". *arXiv preprint arXiv:2402.01444*.

Rolinski, S., C. Müller, J. Heinke, I. Weindl, A. Biewald, B. L. Bodirsky, A. Bondeau, E. R. Boons-Prins, A. F. Bouwman, P. A. Leffelaar, et al. (2018). "Modeling vegetation and carbon dynamics of managed grasslands at the global scale with LPJmL 3.6". *Geoscientific Model Development* 11.1, 429–451. DOI: 10.5194/gmd-11-429-2018.

Romeiro, J. M. N., T. Eid, C. Antón-Fernández, A. Kangas, and E. Trømborg (2022). "Natural disturbances risks in European Boreal and Temperate forests and their links to climate change–A review of modelling approaches". *Forest Ecology and Management* 509, 120071. DOI: https://doi.org/10.1016/j.foreco.2022.120071.

Rong, X. and M. X. Rong (2014). *Package 'deepnet'*.

Rounsevell, M., M. Fischer, M. Torre-Marin Rando, and A. Mader (2018). *The IPBES regional assessment report on biodiversity and ecosystem services for Europe and Central Asia*. Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity ...

Ruiz-Benito, P., G. Vacchiano, E. R. Lines, C. P. Reyer, S. Ratcliffe, X. Morin, F. Hartig, A. Mäkelä, R. Yousefpour, J. E. Chaves, et al. (2020). "Available and missing data to model impact of climate change on European forests". *Ecological Modelling* 416, 108870.

Sampaio, G., C. Nobre, M. H. Costa, P. Satyamurty, B. S. Soares-Filho, and M. Cardoso (2007). "Regional climate change over eastern Amazonia caused by pasture and soybean cropland expansion". *Geophysical Research Letters* 34.17.

San-Miguel-Ayanz, J., D. de Rigo, G. Caudullo, T. Houston Durrant, and A. Mauri (2016). "European Atlas of Forest Tree Species".

San-Miguel-Ayanz, J., E. Schulte, G. Schmuck, A. Camia, P. Strobl, G. Liberta, C. Giovando, R. Boca, F. Sedano, P. Kempeneers, et al. (2012). "Comprehensive monitoring of wildfires in Europe: the European forest fire information system (EFFIS)". In: *Approaches to managing disaster-Assessing hazards, emergencies and disaster impacts*. IntechOpen. DOI: https://doi.org/10.5772/28441.

Sanderson, B. M., R. Knutti, and P. Caldwell (2015). "A representative democracy to reduce interdependency in a multimodel ensemble". *Journal of Climate* 28.13, 5171–5194.

Schadauer, K., R. Astrup, J. Breidenbach, J. Fridman, S. Gräber, M. Köhl, K. T. Korhonen, V. K. Johannsen, F. Morneau, R. Päivinen, et al. (2024). "Access to exact National Forest Inventory plot locations must be carefully evaluated". *The New phytologist*.

Schelhaas, M.-J., G.-J. Nabuurs, and A. Schuck (2003). "Natural disturbances in the European forests in the 19th and 20th centuries". *Global Change Biology* 9.11, 1620–1633. DOI: https://doi.org/10.1046/j.1365-2486.2003.00684.x.

Schelhaas, M.-J., S. Teeuwen, J. Oldenburger, G. Beerkens, G. Velema, J. Kremers, B. Lerink, M. Paulo, H. Schoonderwoerd, W. Daamen, et al. (2022). *Zevende Nederlandse*

Bosinventarisatie: Methoden en resultaten. Tech. rep. Wettelijke Onderzoekstaken Natuur & Milieu.

Schepaschenko, D., J. Chave, O. L. Phillips, S. L. Lewis, S. J. Davies, M. Réjou-Méchain, P. Sist, K. Scipal, C. Perger, B. Herault, et al. (2019). "The Forest Observation System, building a global reference dataset for remote sensing of forest biomass". *Scientific Data* 6.1, 198.

Scherrer, D., D. Ascoli, M. Conedera, C. Fischer, J. Maringer, B. Moser, P. S. Nikolova, A. Rigling, and T. Wohlgemuth (2022). "Canopy disturbances catalyse tree species shifts in Swiss forests". *Ecosystems* 25.1, 199–214. DOI: https://doi.org/10.1007/s10021-021-00649-1.

Schloss, A., D. Kicklighter, J. Kaduk, U. Wittenberg, and T. P. O. T. P. N. M. Intercomparison (1999). "Comparing global models of terrestrial net primary productivity (NPP): comparison of NPP to climate and the Normalized Difference Vegetation Index (NDVI)". *Global Change Biology* 5.S1, 25–34.

Schmidt, M., M. Hanewinkel, G. Kändler, E. Kublin, and U. Kohnle (2010). "An inventory-based approach for modeling single-tree storm damage—experiences with the winter storm of 1999 in southwestern Germany". *Canadian journal of forest research* 40.8, 1636–1652. DOI: https://doi.org/10.1139/X10-099.

Schramm, M., E. Pebesma, M. Milenković, L. Foresta, J. Dries, A. Jacob, W. Wagner, M. Mohr, M. Neteler, M. Kadunc, et al. (2021). "The openeo api–harmonising the use of earth observation cloud services using virtual data cube functionalities". *Remote Sensing* 13.6, 1125.

Schratz, P., J. Muenchow, E. Iturritxa, J. Richter, and A. Brenning (2019). "Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data". *Ecological Modelling* 406, 109–120. DOI: https://doi.org/10.1016/j.ecolmodel.2019.06.002.

Seddon, N. (2022). "Harnessing the potential of nature-based solutions for mitigating and adapting to climate change". *Science* 376.6600, 1410–1416.

Sedmáková, D., R. Sedmák, M. Bosela, M. Ježík, M. Blaženec, T. Hlásny, and R. Marušák (2019). "Growth-climate responses indicate shifts in the competitive ability of European beech and Norway spruce under recent climate warming in East-Central Europe". *Dendrochronologia* 54, 37–48. DOI: https://doi.org/10.1016/j.dendro.2019.02.001.

Seidl, R. and K. Blennow (2012). "Pervasive growth reduction in Norway spruce forests following wind disturbance". *PLoS One* 7.3, e33301. DOI: https://doi.org/10.1371/journal.pone.0033301.

Seidl, R., M.-J. Schelhaas, and M. J. Lexer (2011). "Unraveling the drivers of intensifying forest disturbance regimes in Europe". *Global Change Biology* 17.9, 2842–2852.

Seidl, R., D. Thom, M. Kautz, D. Martin-Benito, M. Peltoniemi, G. Vacchiano, J. Wild, D. Ascoli, M. Petr, J. Honkaniemi, et al. (2017). "Forest disturbances under climate change". *Nature climate change* 7.6, 395–402. DOI: https://doi.org/10.1111/j.1365-2486.2011.02452.x.

Senay, S. D., S. P. Worner, and T. Ikeda (2013). "Novel Three-Step Pseudo-Absence Selection Technique for Improved Species Distribution Modelling". *PLoS ONE* 8.8. DOI: 10.1371/journal.pone.0071218.

Senf, C., D. Pflugmacher, Y. Zhiqiang, J. Sebald, J. Knorn, M. Neumann, P. Hostert, and R. Seidl (2018). "Canopy mortality has doubled in Europe's temperate forests over the last three decades". *Nature Communications* 9.1, 1–8.

Senf, C., J. Sebald, and R. Seidl (2021). "Increasing canopy mortality affects the future demographic structure of Europe's forests". *One Earth* 4.5, 749–755.

Senf, C. and R. Seidl (2021a). "Mapping the forest disturbance regimes of Europe". *Nature Sustainability* 4.1, 63–70. DOI: https://doi.org/10.1038/s41893-020-00609-y.

– (2021b). "Storm and fire disturbances in Europe: Distribution and trends". *Global Change Biology* 27.15, 3605–3619. DOI: https://doi.org/10.1111/gcb.15679.

Shabani, F., L. Kumar, and M. Ahmadi (2018). "Assessing accuracy methods of species distribution models: AUC, Specificity, Sensitivity and the True Skill Statistic". *Global Journal of Human Social Science* 18.1, 6–18.

Shannon, C. E. (1948). "A mathematical theory of communication". *The Bell system technical journal* 27.3, 379–423.

Shen, X. and L. Cao (2017). "Tree-species classification in subtropical forests using airborne hyperspectral and LiDAR data". *Remote Sensing* 9.11, 1180.

Shi, X., Y. D. Wong, M. Z.-F. Li, C. Palanisamy, and C. Chai (2019). "A feature learning approach based on XGBoost for driving assessment and risk prediction". *Accident Analysis & Prevention* 129, 170–179. DOI: https://doi.org/10.1016/j.aap.2019.05.005.

Sirén, A. P., C. S. Sutherland, A. V. Karmalkar, M. J. Duveneck, and T. L. Morelli (2022). "Forecasting species distributions: Correlation does not equal causation". *Diversity and Distributions*.

Smith, S. L., H. B. O'Neill, K. Isaksen, J. Noetzli, and V. E. Romanovsky (2022). "The changing thermal state of permafrost". *Nature Reviews Earth & Environment* 3.1, 10–23.

Soleymani, R., E. Granger, and G. Fumera (2020). "F-measure curves: A tool to visualize classifier performance under imbalance". *Pattern Recognition* 100, 107146.

Sommerfeld, A., C. Senf, B. Buma, A. W. D'Amato, T. Després, I. Díaz-Hormazábal, S. Fraver, L. E. Frelich, Á. G. Gutiérrez, S. J. Hart, et al. (2018). "Patterns and drivers of recent disturbances across the temperate forest biome". *Nature communications* 9.1, 4355. DOI: https://doi.org/10.1038/s41467-018-06788-9.

Sothe, C., C. De Almeida, M. Schimalski, L. La Rosa, J. Castro, R. Feitosa, M. Dalponte, C. Lima, V. Liesenberg, G. Miyoshi, et al. (2020). "Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data". *GIScience & Remote Sensing* 57.3, 369–394.

STAC Community (2024). *Spatio Temporal Asset Catalog (STAC) Specification*. `https://github.com/radiantearth/stac-spec`. Accessed: 14 February 2024.

Ståhl, G., S. Saarela, S. Schnell, S. Holm, J. Breidenbach, S. P. Healey, P. L. Patterson, S. Magnussen, E. Næsset, R. E. McRoberts, et al. (2016). "Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation". *Forest Ecosystems* 3.1, 1–11.

Stevens, N., C. E. Lehmann, B. P. Murphy, and G. Durigan (2017). "Savanna woody encroachment is widespread across three continents". *Global change biology* 23.1, 235–244.

Storch, T., H.-P. Honold, S. Chabrillat, M. Habermeyer, P. Tucker, M. Brell, A. Ohndorf, K. Wirth, M. Betz, M. Kuchler, et al. (2023). "The EnMAP imaging spectroscopy mission towards operations". *Remote Sensing of Environment* 294, 113632.

Strickland, G. E. I., J. E. Luther, J. C. White, M. A. Wulder, G. E. I. Strickland, J. E. Luther, J. C. White, and M. A. Wulder (2020). "Extending Estimates of Tree and Tree Species Presence-Absence through Space and Time Using Landsat Composites". *Canadian Journal of Remote Sensing* 46.5, 567–584. DOI: `10.1080/07038992.2020.1811083`.

Strobl, P. (2020). "The new Copernicus digital elevation model". *GSICS Quarterly* 14.1, 17–18.

Strona, G., A. Mauri, J. A. Veech, G. Seufert, J. San-Miguel Ayanz, and S. Fattorini (2016). "Far from naturalness: How much does spatial ecological structure of European tree assemblages depart from potential natural vegetation?" *Plos One* 11.12, e0165178. DOI: `https://doi.org/10.1371/journal.pone.0165178`.

Su, Y., B. Gabrielle, and D. Makowski (2021). "The impact of climate change on the productivity of conservation agriculture". *Nature Climate Change* 11.7, 628–633. DOI: `https://doi.org/10.1038/s41558-021-01075-w`.

Susmel, L. (1994). *I rovereti di pianura della Serenissima*. Cleup.

Svenning, J.-C., M. C. Fitzpatrick, S. Normand, C. H. Graham, P. B. Pearman, L. R. Iverson, and F. Skov (2010). "Geography, topography, and history affect realized-to-potential tree species richness patterns in Europe". *Ecography* 33.6, 1070–1080. DOI: `https://doi.org/10.1111/j.1600-0587.2010.06301.x`.

Svenning, J.-C. and F. Skov (2004). "Limited filling of the potential range in European tree species". *Ecology Letters* 7.7, 565–573. DOI: https://doi.org/10.1111/j.1461-0248.2004.00614.x.

– (2007). "Ice age legacies in the geographical distribution of tree species richness in Europe". *Global Ecology and Biogeography* 16.2, 234–245. DOI: https://doi.org/10.1111/j.1466-8238.2006.00280.x.

Tadono, T., H. Nagai, H. Ishida, F. Oda, S. Naito, K. Minakawa, and H. Iwamoto (2016). "Generation of the 30 m-mesh global digital surface model by ALOS PRISM". *The international archives of the photogrammetry, remote sensing and spatial information sciences* 41, 157–162.

Temesgen, H., D. Affleck, K. Poudel, A. Gray, and J. Sessions (2015). "A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models". *Scandinavian Journal of Forest Research* 30.4, 326–335.

Therneau, T. M. and E. J. Atkinson (2011). "An Introduction to Recursive Partitioning Using the RPART Routines". *Mayo clinic* 61, 33.

Thuiller, W., C. Albert, M. B. Araújo, P. M. Berry, M. Cabeza, A. Guisan, T. Hickler, G. F. Midgley, J. Paterson, F. M. Schurr, et al. (2008). "Predicting global change impacts on plant species' distributions: future challenges". *Perspectives in plant ecology, evolution and systematics* 9.3-4, 137–152. DOI: https://doi.org/10.1016/j.ppees.2007.09.004.

Thuiller, W., L. Brotons, M. B. Araújo, and S. Lavorel (2004). "Effects of restricting environmental range of data to project current and future species distributions". *Ecography* 27.2, 165–172.

Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso". *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, 267–288.

Tilman, D., R. M. May, C. L. Lehman, and M. A. Nowak (1994). "Habitat destruction and the extinction debt". *Nature* 371.6492, 65–66. DOI: https://doi.org/10.1038/371065a0.

Tinner, W., D. Colombaroli, O. Heiri, P. D. Henne, M. Steinacher, J. Untenecker, E. Vescovi, J. R. Allen, G. Carraro, M. Conedera, et al. (2013). "The past ecology of *Abies alba* provides new perspectives on future responses of silver fir forests to global warming". *Ecological Monographs* 83.4, 419–439. DOI: https://doi.org/10.1890/12-2231.1.

Tompalski, P., N. C. Coops, J. C. White, and M. A. Wulder (2014). "Simulating the impacts of error in species and height upon tree volume derived from airborne laser scanning data". *Forest ecology and management* 327, 167–177.

Tomppo, E., T. Gschwantner, M. Lawrence, R. E. McRoberts, K. Gabler, K. Schadauer, C. Vidal, A. Lanz, G. Ståhl, E. Cienciala, et al. (2010). "National forest inventories". *Pathways for Common Reporting. European Science Foundation* 1, 541–553.

Trisos, C. H., I. O. Adelkan, E. Totin, A. Ayanlade, J. Efitre, A. Gemeda, K. Kalaba, C. Lennard, C. Masao, Y. Mgaya, G. Ngaruiya, D. Olago, N. P. Simpson, and S. Zakieldeen (2022). "Africa". In: *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama]*. Cambridge University Press. Chap. 9, 1285–1455. DOI: 10.1017/9781009325844.011.

Tucker, C. J. (1979). "Red and photographic infrared linear combinations for monitoring vegetation". *Remote sensing of Environment* 8.2, 127–150.

Tuia, D., C. Persello, and L. Bruzzone (2016). "Domain adaptation for the classification of remote sensing data: An overview of recent advances". *IEEE geoscience and remote sensing magazine* 4.2, 41–57.

Turubanova, S., P. Potapov, M. C. Hansen, X. Li, A. Tyukavina, A. H. Pickens, A. Hernandez-Serna, A. P. Arranz, J. Guerra-Hernandez, C. Senf, et al. (2023). "Tree canopy extent and height change in Europe, 2001–2021, quantified using Landsat data archive". *Remote Sensing of Environment* 298, 113797.

UNFCCC (2018). "Paris Climate Change Conference - November 2015". In: *Proceedings of the COP 21*. Paris, France: United Nations Framework Convention on Climate Change.

United States Department of Agriculture (USDA) (2007). *A History of the Forest Survey in the United States: 1830–2004*. Tech. rep. FS-877. Forest Service.

Valavi, R., G. Guillera-Arroita, J. J. Lahoz-Monfort, and J. Elith (2021). "Predictive performance of presence-only species distribution models: a benchmark study with reproducible code". *Ecological Monographs* 0.0, 1–27. DOI: 10.1002/ecm.1486.

Valinger, E., G. Kempe, and J. Fridman (2014). "Forest management and forest state in southern Sweden before and after the impact of storm Gudrun in the winter of 2005". *Scandinavian Journal of Forest Research* 29.5, 466–472. DOI: https://doi.org/10.1080/02827581.2014.927528.

Van Rijsbergen, C. J. (1979). *Information Retrieval*. 2nd. Butterworth-Heinemann.

Van Rossum, G. and F. L. Drake (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, et al. (2011). "The representative concentration pathways: an overview". *Climatic change* 109.1, 5–31.

Venter, Z. S. and M. A. Sydenham (2021). "Continental-scale land cover mapping at 10 m resolution over Europe (ELC10)". *Remote Sensing* 13.12, 2301.

Verkerk, P., R. Costanza, L. Hetemäki, I. Kubiszewski, P. Leskinen, G. Nabuurs, J. Potočnik, and M. Palahí (2020). "Climate-smart forestry: the missing link". *Forest Policy and Economics* 115, 102164.

Vidal, C., I. Alberdi, J. Redmond, M. Vestman, A. Lanz, and K. Schadauer (2016). "The role of European National Forest Inventories for international forestry reporting". *Annals of Forest Science* 73.4, 793–806.

Von Berg, E. (1995). *Kertomus Suomenmaan metsistä 1858*. Ed. by M. Leikola. Original work published 1858 (in Finnish). Jyväskylä, Finland: Kustannusosakeyhtiö Metsälehti, Gummerus Kirjapaino Oy, 93.

Walthert, L. and E. S. Meier (2017). "Tree species distribution in temperate forests is more influenced by soil than by climate". *Ecology and Evolution* 7.22, 9473–9484.

Wang, H., H. Liu, N. Huang, J. Bi, X. Ma, Z. Ma, Z. Shangguan, H. Zhao, Q. Feng, T. Liang, et al. (2021). *Satellite-derived NDVI underestimates the advancement of alpine vegetation growth over the past three decades*.

Wang, S., Y. Zhang, W. Ju, J. M. Chen, P. Ciais, A. Cescatti, J. Sardans, I. A. Janssens, M. Wu, J. A. Berry, et al. (2020). "Recent global decline of CO2 fertilization effects on vegetation photosynthesis". *Science* 370.6522, 1295–1300.

Waser, L. T., M. Küchler, K. Jütte, and T. Stampfer (2014). "Evaluating the potential of WorldView-2 data to classify tree species and different levels of ash mortality". *Remote Sensing* 6.5, 4515–4545.

Waser, L. T., M. Rüetschi, A. Psomas, D. Small, and N. Rehush (2021). "Mapping dominant leaf type based on combined Sentinel-1/-2 data–Challenges for mountainous countries". *ISPRS Journal of Photogrammetry and Remote Sensing* 180, 209–226.

Watanabe, M., T. Suzuki, R. O'ishi, Y. Komuro, S. Watanabe, S. Emori, T. Takemura, M. Chikira, T. Ogura, M. Sekiguchi, K. Takata, D. Yamazaki, T. Yokohata, T. Nozawa, H. Hasumi, H. Tatebe, and M. Kimoto (2010). "Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity". *Journal of Climate* 23.23, 6312–6335". DOI: `https://doi.org/10.1175/2010JCLI3679.1`.

Weigel, R., J. Gilles, M. Klisz, M. Manthey, and J. Kreyling (2019). "Forest understory vegetation is more related to soil than to climate towards the cold distribution margin of European beech". *Journal of Vegetation Science* 30.4, 746–755.

Welle, T., L. Aschenbrenner, K. Kuonath, S. Kirmaier, and J. Franke (2022). "Mapping dominant tree species of German forests". *Remote Sensing* 14.14, 3330.

Welsink, A.-J., J. Reiche, V. De Sy, S. Carter, B. Slagter, D. R. Suarez, B. Batros, M. Peña-Claros, and M. Herold (2023). "Towards the use of satellite-based tropical forest disturbance alerts to assess selective logging intensities". *Environmental Research Letters* 18.5, 054023.

Wermelinger, B., A. Rigling, D. Schneider Mathis, M. Kenis, and M. M. Gossner (2021). "Climate change effects on trophic interactions of bark beetles in inner alpine Scots pine forests". *Forests* 12.2, 136. DOI: https://doi.org/10.3390/f12020136.

Wessel, M., M. Brandmeier, D. Tiede, R. Seitz, and C. Straub (2018). "Comparison of Different Machine-Learning Algorithms for Tree Species Classification Based on Sentinel Data". In: *PFGK18*.

White, J., M. Wulder, A. Varhola, M. Vastaranta, N. Coops, B. Cook, et al. (2013). *A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach*. Information Report FI-X-010. Victoria, BC: Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre.

White, J. C., M. Wulder, G. Hobart, J. Luther, T. Hermosilla, P. Griffiths, N. Coops, R. Hall, P. Hostert, A. Dyk, et al. (2014). "Pixel-based image compositing for large-area dense time series applications and science". *Canadian Journal of Remote Sensing* 40.3, 192–212.

White, J. C., M. A. Wulder, T. Hermosilla, N. C. Coops, and G. W. Hobart (2017). "A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series". *Remote Sensing of Environment* 194, 303–321.

Whittaker, R. H. and P. L. Marks (1975). "Methods of assessing terrestrial productivty". In: *Primary productivity of the biosphere*. Springer, 55–118.

Wickham, H. (2019). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.4.0.

Wickham, H., R. François, L. Henry, and K. Müller (2021). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.7.

Williams, J. W., S. T. Jackson, and J. E. Kutzbach (2007). "Projected distributions of novel and disappearing climates by 2100 AD". *Proceedings of the National Academy of Sciences* 104.14, 5738–5742.

Wilson, A. M. and W. Jetz (2016). "Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions". *PLoS biology* 14.3, e1002415.

Wise, J. (2023). "COP28: Loss and damage fund a step forward but not enough, say climate experts". *BMJ* 383. DOI: 10.1136/bmj.p2871.

Witjes, M., L. Parente, C. J. van Diemen, T. Hengl, M. Landa, L. Brodskỳ, L. Halounova, J. Križan, L. Antonić, C. M. Ilie, et al. (2022). "A spatiotemporal ensemble machine learning framework for generating land use/land cover time-series maps for Europe (2000–2019) based on LUCAS, CORINE and GLAD Landsat". *PeerJ* 10, e13573.

Witjes, M., L. Parente, J. Križan, T. Hengl, and L. Antonić (2023). "Ecodatacube. eu: Analysis-ready open environmental data cube for Europe". *PeerJ* 11, e15478. DOI: https://doi.org/10.7717/peerj.15478.

Wohlgemuth, T., M. Hanewinkel, R. Seidl, and T. Wohlgemuth (2022). "Wind Disturbances". In: *Disturbance Ecology*. Springer, 173–194. DOI: https://doi.org/10.1007/978-3-030-98756-5_8.

Wolpert, D. H. (1992). "Stacked generalization". *Neural networks* 5.2, 241–259.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, et al. (2008). "Free access to Landsat imagery." *SCIENCE VOL 320: 1011*.

Wright, R. E. (1995). "Logistic regression."

Wulder, M. A., T. Hermosilla, J. C. White, C. W. Bater, G. Hobart, and S. C. Bronson (2024). "Development and implementation of a stand-level satellite-based forest inventory for Canada". *Forestry: An International Journal of Forest Research*, cpad065.

Wulder, M. A., T. R. Loveland, D. P. Roy, C. J. Crawford, J. G. Masek, C. E. Woodcock, R. G. Allen, M. C. Anderson, A. S. Belward, W. B. Cohen, et al. (2019). "Current status of Landsat program, science, and applications". *Remote sensing of environment* 225, 127–147.

Wulder, M. A., D. P. Roy, V. C. Radeloff, T. R. Loveland, M. C. Anderson, D. M. Johnson, S. Healey, Z. Zhu, T. A. Scambos, N. Pahlevan, et al. (2022). "Fifty years of Landsat science and impacts". *Remote Sensing of Environment* 280, 113195.

Xu, H., X. Lian, I. J. Slette, H. Yang, Y. Zhang, A. Chen, and S. Piao (2022). "Rising ecosystem water demand exacerbates the lengthening of tropical dry seasons". *Nature Communications* 13.1. DOI: 10.1038/s41467-022-31826-y.

Yamazaki, D., D. Ikeshima, J. C. Neal, F. O'Loughlin, C. C. Sampson, S. Kanae, and P. D. Bates (2017). "MERIT DEM: A new high-accuracy global digital elevation model and its merit to global hydrodynamic modeling". In: *AGU fall meeting abstracts*. Vol. 2017.

Yang, C., Q. Huang, Z. Li, K. Liu, and F. Hu (2017). "Big Data and cloud computing: innovation opportunities and challenges". *International Journal of Digital Earth* 10.1, 13–53.

Ye, N., J. Morgenroth, C. Xu, and N. Chen (2021). "Indigenous forest classification in New Zealand–A comparison of classifiers and sensors". *International Journal of Applied Earth Observation and Geoinformation* 102, 102395.

Yin, J., P. Gentine, L. Slater, L. Gu, Y. Pokhrel, N. Hanasaki, S. Guo, L. Xiong, and W. Schlenker (2023). "Future socio-ecosystem productivity threatened by compound drought–heatwave events". *Nature Sustainability*, 1–14. DOI: 10.1038/s41893-022-01024-1.

Zaehle, S., S. Sitch, B. Smith, and F. Hatterman (2005). "Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics". *Global Biogeochemical Cycles* 19.3.

Zelinka, M. D., T. A. Myers, D. T. McCoy, S. Po-Chedley, P. M. Caldwell, P. Ceppi, S. A. Klein, and K. E. Taylor (2020). "Causes of higher climate sensitivity in CMIP6 models". *Geophysical Research Letters* 47.1.

Zeng, Y., D. Hao, T. Park, P. Zhu, A. Huete, R. Myneni, Y. Knyazikhin, J. Qi, R. R. Nemani, F. Li, et al. (2023). "Structural complexity biases vegetation greenness measures". *Nature Ecology & Evolution* 7.11, 1790–1798.

Zevallos, J. and W. Lavado-Casimiro (2022). "Climate Change Impact on Peruvian Biomes". *Forests* 13.2, 238.

Zhang, B., L. Zhao, and X. Zhang (2020). "Three-dimensional convolutional neural network model for tree species classification using airborne hyperspectral images". *Remote Sensing of Environment* 247, 111938.

Zhang, C. and Y. Ma (2012). *Ensemble machine learning: methods and applications*. Springer.

Zhang, C., W. Li, and D. Travis (2007). "Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach". *International Journal of Remote Sensing* 28.22, 5103–5122.

Zhang, J. and S. Li (2017). "A Review of Machine Learning Based Species' Distribution Modelling". In: *2017 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII)*. IEEE, 199–206.

Zhang, J., S. E. Nielsen, Y. Chen, D. Georges, Y. Qin, S.-S. Wang, J.-C. Svenning, and W. Thuiller (2017). "Extinction risk of North American seed plants elevated by climate and land-use change". *Journal of Applied Ecology* 54.1, 303–312. DOI: https://doi.org/10.1111/1365-2664.12701.

Zhao, Q., Z. Zhu, H. Zeng, W. Zhao, and R. B. Myneni (2020). "Future greening of the Earth may not be as large as previously predicted". *Agricultural and Forest Meteorology* 292, 108111.

Zhou, Z.-H. (2019). *Ensemble methods: foundations and algorithms*. Chapman and Hall/CRC.

Zhu, K., C. W. Woodall, and J. S. Clark (2012). "Failure to migrate: lack of tree range expansion in response to climate change". *Global Change Biology* 18.3, 1042–1052. DOI: https://doi.org/10.1111/j.1365-2486.2011.02571.x.

Zhu, K., C. W. Woodall, S. Ghosh, A. E. Gelfand, and J. S. Clark (2014). "Dual impacts of climate change: forest migration and turnover through life history". *Global Change Biology* 20.1, 251–264. DOI: https://doi.org/10.1111/gcb.12382.

Zhu, Z., S. Piao, X. Lian, R. B. Myneni, S. Peng, and H. Yang (2017). "Attribution of seasonal leaf area index trends in the northern latitudes with "optimally" integrated ecosystem models". *Global change biology* 23.11, 4798–4813.

Zhu, Z., S. Piao, R. B. Myneni, M. Huang, Z. Zeng, J. G. Canadell, P. Ciais, S. Sitch, P. Friedlingstein, A. Arneth, et al. (2016). "Greening of the Earth and its drivers". *Nature climate change* 6.8, 791–795.

Zhu, Z., M. A. Wulder, D. P. Roy, C. E. Woodcock, M. C. Hansen, V. C. Radeloff, S. P. Healey, C. Schaaf, P. Hostert, P. Strobl, et al. (2019). "Benefits of the free and open Landsat data policy". *Remote Sensing of Environment* 224, 382–385. DOI: 10.1016/j.rse.2019.02.016.

Zurell, D., J. Franklin, C. König, P. J. Bouchet, C. F. Dormann, J. Elith, G. Fandos, X. Feng, G. Guillera-Arroita, A. Guisan, J. J. Lahoz-Monfort, P. J. Leitão, D. S. Park, A. T. Peterson, G. Rapacciuolo, D. R. Schmatz, B. Schröder, J. M. Serra-Diaz, W. Thuiller, K. L. Yates, N. E. Zimmermann, and C. Merow (2020). "A standard protocol for reporting species distribution models". *Ecography* 43.9, 1261–1277. DOI: 10.1111/ecog.04960.

Zweifel, R., C. Pappas, R. L. Peters, F. Babst, D. Balanzategui, D. Basler, A. Bastos, M. Beloiu, N. Buchmann, A. K. Bose, et al. (2023). "Networking the forest infrastructure towards near real-time monitoring–A white paper". *Science of the Total Environment* 872, 162167.

Acknowledgements

Pursuing this PhD has been a journey of profound learning and insight, not only into because of its interdisciplinary nature, spanning multiple and diverse topics, but also into the nature of academia itself. This experience has shaped my understanding of research, its challenges, and its vast potential to contribute to society. As I reflect on this chapter of my academic life, I am grateful for the intellectual growth it has fostered and the clarity it has brought to my career aspirations.

Gaining such insights in this journey was made possible only through a continuous interactions with a diverse community of colleagues-including scholars, practitioners, and professionals from various non-academic fields. Engaging with them has been instrumental in broadening my perspective, refining my approaches to research and overall enriching my academic journey.

First I would like to thank my promotor, Martin Herold. Due to a global pandemic first and long distance later, our interactions were mostly virtual; despite that, your support and vision were never lacking during my whole PhD journey. I deeply appreciated how every occasion we had to discuss new research ideas or opportunities for collaborations was always productive and on point: never I have walked out of a meeting with you without either a deeper understanding of the big picture or a new piece of knowledge, which was invaluable.

I do not know how this thesis would look like without the constant feedback and constructive criticism provided by my co-promotor, Sytze de Bruin. Since our first online interview for the position, where I had a first glimpse of your inquisitive nature, your contributions have been invaluable in challenging me to truly *think* deeply. Whether during the crafting of a paragraph or the planning of an experimental design, your meticulous attention to detail and insistence on internal coherence have greatly improved the way I think about things and the words I use. A more conscious use of language produces clearer thinking after all.

Given the arrangements that gave me the chance to pursue this PhD, it would not be an exaggeration to say that none of this would have been possible without the contribution of my two other supervisors Tomislav Hengl and Leandro Parente. Having both of you

on my team basically meant having the best of both worlds: while Tom would be an endless source of ideas, fueled by an equally inexhaustible passion for science and research, Leandro would just quietly sit next to you and help you make these ideas come true.

I would also like to thank all four of my opponents for taking the time to read and evaluate my thesis, and for their willingness to travel to Wageningen specifically for the defence. Each of you has a very interesting research profile, some more closely related to things I have done over my PhD track and some that deal with equally interesting aspects of the field that I haven't had the time to explore. If possible, I would like to work together with each one of you in the near future.

I want to thank my paranynmphs, Robert and Xuemeng, for their role in my PhD journey. Robert, during one of my sporadic appearances at the GRS group, what started as a casual conversation in front of the coffee machine turned out to be a long-lasting exchange of research notes, impressions and experiences (I will save to the reader my snarky and opinionated remarks on coffee and coffee quality during these interactions though). Like Sytze, you are one of the people I consider as a cornerstone of my time at GRS and you've always made me feel welcome. Xuemeng, since you joined OpenGeoHub, our interactions have spanned a wide array of topics, from science to food, to social customs. Each discussion, regardless of the subject, has been immensely enriching and has broadened my perspectives. I am grateful for the diversity of our conversations: thank you for sharing your insights and enthusiasm throughout these years.

My deepest gratitude goes to OpenGeoHub: many thanks for allowing me to carry out this PhD under your support, and for providing the opportunity to connect with scientists and other enthusiast globally who share my passion for open source software and the principles of reproducible research. Doing a PhD at OpenGeoHub also gave me the pleasure of working alongside a warm and diverse group of people, coming from all kinds of different backgrounds. A special thanks goes to my PhD-twin Martijn, with whom I literally shared every step of the way: from the selection procedure for the PhD position to writing the thesis, from learning how to be an independent researcher to despairing about how little we think we know about anything at all and we should better go doing something else. All of this coupled with the occasional drink, (s)word fights about food, endless boardgames sessions and so on and so forth. Thanks for sharing the ups and downs of this surprising, improbable journey.

To the GRS colleagues, although I'm not always around, I want to say how much I appreciated our limited but meaningful small talks. Among others, I would like to thank Alvaro, Anne, Anne-Juul, Antoinette, Arnan, Bart, Benjamin, Chenglong, Dainius, Daniela, Diego, Federico, Frida, Gerlinde, Gonzalo, Ioannis, Jens, Johannes Balling, Johannes Reiche, Julia, Karimon, Kirsten, Laura, Layla, Magdalena, Marc, Mathieu, Milutin, Nandika, Na, Natalia, Panpan, Ron, Sabina, Sarah, Sebastian, Sietse, Truus, Ximena and anybody that I might not have mentioned here.

My life as a PhD student in the Netherlands has been good, although not how I imagined it. I would not be who I am today without my years in the Wageningen community, especially during the restrictive times imposed by a global pandemic. During my time in Wageningen, I primarily lived in "Plantsoen," an old office building converted into student housing. It was there that I met many of the people who introduced me to the life in the city, the Dutch culture and many other ones. These people remain important to me today, even though distance prevents us from seeing each other as often; I would particularly like to remember Bea, Caro, Django, Kazem, Manolis, Pepe and Simen.

I would like to thank my friends and family for their support throughout this journey. Their steady presence and encouragement have been a quiet backdrop to my studies, providing the space I needed to focus on my PhD. I appreciate the understanding and patience they have shown along the way.

Lastly, I am deeply grateful to Ida, who has been an unexpected but extraordinary companion during the final months of my PhD journey. Our time together has been a whirlwind of new experiences and mutual growth, profoundly enriching this phase of my life.

About the author

Carmelo Bonannella was born on November 9, 1993 in Lipari, Italy, an island of the Aeolian archipelago. He grew up there, spending most of his childhood hopping from one island to the other. His early years were characterized by reading books, penning short stories and exploring the surrounding landscape, rich in history and natural wonders. Despite this seemingly natural inclination for arts and literature, he completed high school with majors in mathematics and physics, ultimately opting for a career in science. In 2012, he relocated to Sicily to pursue a Bachelor's degree in Materials Engineering at the University of Messina. His deep-rooted passion for the environment accompanied him, nurtured by his experiences in the wild yet delicate landscapes of his islands, his idea being to combine science and living materials like plants to restore polluted or destroyed ecosystems.

Carmelo quickly realized that his ideas didn't find fertile ground within that university department: courses on soil and water bioengineering were notably absent from Italian faculties at that time. He then decided to switch gears in 2013, enrolling in a Bachelor's program in Forestry at the Mediterranea University of Reggio Calabria. During these formative years, he rekindled his passion for biogeography, ecology, and complex systems, subjects he had initially encountered in both his academic studies and leisure readings. His discovery of a rich and unbroken legacy of scholarly work in silviculture, illustrating the paradigm shift from viewing forests merely as timber sources to promoting sustainable management, was largely attributed to the guidance and expertise of his now-retired BSc thesis advisor, Giuliano Menguzzato. Encouraged at that time by his advisor's suggestion, he continued his academic journey by enrolling in a Master's program in Forestry at the University of Florence in 2016, recognized as Italy's oldest formal forestry school. The fieldwork for his Bachelor's thesis brought an important realization: his efforts, if confined to a local area as specific as his islands, would be limited in scope. Aspiring to impact

broader regions and devise comprehensive strategies for forest management, during his Master's studies he tried to learn as much as possible about GIS and remote sensing forestry applications. He became quite interested in forest inventories and national forest inventory surveys as well, still today the state of the art tools for understanding and monitoring forest resources. Leveraging his strong motivation and quick learning skills, he managed to win a part-time position as a GIS technician in the university's remote sensing laboratory during his studies. During this time, he began his Master's thesis, captivated by the University of British Columbia's Remote Sensing faculty's work in applying remote sensing for forest monitoring in Canadian forests. Under the supervision of Gherardo Chirici, he first studied and then mastered these methods, adapting them for application in Central Italy, his goal understanding the spectral signature and the subsequent recovery patterns of common forest disturbances in Mediterranean forests. His Master's thesis work was also what secured him a PhD position in 2019 at the University of Florence, feeling that the work done for his thesis was only the tip of the iceberg and that he could have done more and better given time and funds.

Nine months, a scientific publication and the beginning of a global pandemic later, Carmelo felt constrained in his research pursuits. This led him to make a bold decision: he left his PhD program at the University of Florence to seek new opportunities abroad. In July 2020, he relocated to Wageningen, joining the OpenGeoHub Foundation as a Research Assistant. Concurrently, he embarked on a new PhD journey at the Laboratory of Geo-Information Science and Remote Sensing at Wageningen University & Research. As an external PhD candidate at a start-up led by two individuals equally passionate about a greener planet, he found an ideal blend of experiences. This unique position allowed him to merge the scientific rigor and stimulating environment of the university with the opportunity to rapidly reach people and implement his ideas in practical, real-world scenarios typically encountered in the corporate sphere. Given the particular focus of OpenGeoHub on open source and open science, these principles had a big influence on Carmelo's way of conducting science: while the concepts were not new for him, he now embraced them throughly, trying to make everything he produced easily accessible to everyone and fully reproducible. Carmelo's PhD research, titled "Spatiotemporal modeling of vegetation dynamics in a changing environment: combining Earth observation and machine learning", which you are now reading, is part of an ongoing Horizon Europe research project "Open-Earth-Monitor Cyberinfrastructure" and of the completed GeoHarmonizer research project. During his PhD research, Carmelo has led several successful international collaborations, showcasing his work at various international conferences and symposia. His research has garnered significant recognition, to the extent that prominent international organizations, including the European Forest Institute (EFI), The Nature Conservancy (TNC), and the World Resource Institute (WRI), have already begun utilizing some of his research outputs.

Carmelo's current research interests still lie in biogeography, forest ecology and forest resources management, especially with the approaching threat posed by climate change. He is especially interested in combining traditional ground measurements with state of the art machine learning techniques and remote sensing data application to develop strategies for a greener future. Following his PhD defence, Carmelo plans to continue as a postdoctoral researcher in forest resources management at the OpenGeoHub Foundation, particularly to continue the development of a reforestation planner for European forests for the "Open-Earth-Monitor Cyberinfrastructure" research project and in tight collaboration with the European Forest Institute.

Peer-reviewed Journal Publications

Arsevska, E., T. Hengl, D. A. Singleton, P.-J. M. Noble, C. Caminade, O. A. Eneanya, P. H. Jones, J. M. Medlock, K. M. Hansford, **C. Bonannella**, and A. D. Radford (2024). "Risk factors for tick attachment in companion animals in Great Britain: a spatiotemporal analysis covering 2014 – 2021". *Parasites Vectors* 17.1, 29. DOI: 10.1186/s13071-023-06094-4.

Bonannella, **C.**, G. Chirici, D. Travaglini, M. Pecchi, E. Vangi, G. D'amico, and F. Giannetti (2022a). "Characterization of wildfires and harvesting forest disturbances and recovery using Landsat time series: a case study in Mediterranean forests in central Italy". *Fire* 5.3, 68. DOI: 10.3390/fire5030068.

Bonannella, **C.**, T. Hengl, J. Heisig, L. Parente, M. N. Wright, M. Herold, and S. de Bruin (2022r). "Forest tree species distribution for Europe 2000–2020: mapping potential and realized distributions using spatiotemporal machine learning". *PeerJ*. DOI: https://doi.org/10.7717/peerj.13728.

Bonannella, **C.**, T. Hengl, L. Parente, and S. de Bruin (2023b). "Biomes of the world under climate change scenarios: increasing aridity and higher temperatures lead to significant shifts in natural vegetation". *PeerJ* 11, e15593. DOI: https://doi.org/10.7717/peerj.15593.

Bonannella, **C.**, L. Parente, S. de Bruin, and M. Herold (2024). "Multi-decadal trend analysis and forest disturbance assessment of European tree species: concerning signs of a subtle shift". *Forest Ecology and Management* 554, 121652. DOI: https://doi.org/10.1016/j.foreco.2023.121652.

Da Re, D., G. Marini, C. Bonannella, F. Laurini, M. Manica, N. Anicic, A. Albieri, P. Angelini, D. Arnoldi, M. Blaha, F. Bertola, B. Caputo, C. De Liberato, A. della Torre, E. Flacio, A. Franceschini, F. Gradoni, P. Kadriaj, V. Lencioni, I. Del Lesto, F. La Russa, R. P. Lia, F. Montarsi, D. Otranto, G. L'Ambert, A. Rizzoli, P. Rombolà, F. Romiti, G. Stancher, A. Torina, E. Velo, C. Virgilito, F. Zandonai, and R. Rosà (2023a). "Inferring

the seasonal dynamics and abundance of an invasive species using a spatio-temporal stacked machine learning model". *Ecography*. In review. DOI: 10.32942/X2NG70.

Da Re, D., G. Marini, C. Bonannella, F. Laurini, M. Manica, N. Anicic, A. Albieri, P. Angelini, D. Arnoldi, M. Blaha, F. Bertola, B. Caputo, C. De Liberato, A. della Torre, E. Flacio, A. Franceschini, F. Gradoni, P. Kadriaj, V. Lencioni, I. Del Lesto, F. La Russa, R. P. Lia, F. Montarsi, D. Otranto, G. L'Ambert, A. Rizzoli, P. Rombolà, F. Romiti, G. Stancher, A. Torina, E. Velo, C. Virgilito, F. Zandonai, and R. Rosà (2023b). "VectAbundance: a spatio-temporal database of vector observations". *Scientific Data*. In review. DOI: 10.32942/X2S60T.

Hengl, T., P. Sorenson, L. Parente, K. Cornish, J. Battigelli, **C. Bonannella**, M. Gorzelak, and K. Nichols (2023b). "Assessment of soil organic carbon stocks in Alberta using 2-scale sampling and 3D predictive soil mapping". *FACETS* 8, 1–17.

Holler, S., D. Kübler, O. Conrad, O. Schmitz, **C. Bonannella**, T. Hengl, J. Böhner, S. Günter, and M. Lippe (2024). "Quo vadis, smallholder forest landscape? An introduction to the LPB-RAP model". *PLoS ONE* 19.2, e0297439. DOI: 10.1371/journal.pone.0297439.

Other Scientific Publications

Bonannella, **C.**, T. Hengl, J. Heisig, L. Leal Parente, M. Wright, M. Herold, and S. de Bruin (2022b). *Potential and realized distribution at 30m for Aleppo pine (Pinus halepensis) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6962632.

– (2022c). *Potential and realized distribution at 30m for Austrian pine (Pinus nigra) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6964167.

– (2022d). *Potential and realized distribution at 30m for Common hazel (Corylus avellana) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6951687.

– (2022e). *Potential and realized distribution at 30m for Goat willow (Salix caprea) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6967538.

– (2022f). *Potential and realized distribution at 30m for Holm oak (Quercus ilex) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6967263.

– (2022g). *Potential and realized distribution at 30m for Norway spruce (Picea abies) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6961187.

– (2022h). *Potential and realized distribution at 30m for Olive tree (Olea europaea) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6960111.

– (2022i). *Potential and realized distribution at 30m for Scots pine (Pinus sylvestris) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6966804.

– (2022j). *Potential and realized distribution at 30m for Silver fir (Abies alba) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6953790.

– (2022k). *Potential and realized distribution at 30m for Stone pine (Pinus pinea) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6966748.

– (2022l). *Potential and realized distribution at 30m for Sweet cherry (Prunus avium) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6966847.

– (2022m). *Potential and realized distribution at 30m for Sweet chestnut (Castanea sativa) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6951628.

– (2022n). *Potential and realized distribution at 30m for the European beech (Fagus sylvatica) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6956944.

– (2022o). *Potential and realized distribution at 30m for Turkey oak (Quercus cerris) in Europe for 2000 - 2020*. Version 3. Zenodo. DOI: 10.5281/zenodo.6967119.

– (2022p). *Presence-Absence Points for Tree Species Distribution Modelling for Europe*. Version 0.3. Zenodo. DOI: 10.5281/zenodo.6516590.

– (2022q). *Supplementary material for "Forest tree species distribution for Europe 2000 - 2020: mapping potential and realized distributions using spatiotemporal Machine Learning"*. Version 0.1. DOI: 10.5281/zenodo.6516728.

Bonannella, **C.**, T. Hengl, L. Leal Parente, and S. de Bruin (2023a). *Current and future global distribution of potential biomes under climate change scenarios*. Version 2. Zenodo. DOI: 10.5281/zenodo.7822868.

Bonannella, **C.**, L. Leal Parente, S. de Bruin, and M. Herold (2023c). *Prevalent trends in realized probability of occurrence of main European forest tree species for 2000–2020*. Version 1. Zenodo. DOI: 10.5281/zenodo.8273144.

Hengl, T., E. Arsevska, and **C. Bonannella** (2023a). *Predicted occurrence probability for ticks in Great Britain (2014 to 2021) at 1 km spatial resolution*. Version v1. Zenodo. DOI: 10.5281/zenodo.7625175.

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Writing of project proposal (4.5 ECTS)

- Forest tree species distribution modeling and prediction in a spatiotemporal framework

Post-graduate courses (6.5 ECTS)

- OpenGeoHub Summer School; OpenGeoHub Foundation (2020)
- Forest Management and Biodiversity across Europe; PE&RC (2023)
- Uncertainty propagation in spatial environmental modeling; PE&RC (2022)

Deficiency, refresh, brush-up courses (0.9 ECTS)

- Geo for Good 2020; Google (2020)
- Workshop on Big Data and Artificial Intelligence for Earth Observation; European Commission (2020)

Competence strengthening / skills courses (5.6 ECTS)

- Project and Time Management; WUR-WGS (2021)
- Writing grant proposals; WUR-WGS (2021)
- How to use art and design in Science; Springer Nature (2022)
- Scientific Publishing; WUR-WGS (2022)
- Career orientation; WUR-WGS (2022)

Scientific integrity/ethics in science activities (0.6 ECTS)

- Scientific integrity; WUR-WGS (2021)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.8 ECTS)

- PE&RC First Years (2021)
- PE&RC Midterm (2022)
- PE&RC Day (2022)

Discussion groups / local seminars / other scientific meetings (5.1 ECTS)

- OpenGeoHub Summer School; OpenGeoHub Foundation (2021–2022)
- BIOMAC webinars; VU Amsterdam (2021)
- Open-Earth-Monitor Cyberinfrastructure project launch; Open-Earth-Monitor Cyberinfrastructure consortium (2022)
- SUPERB & IUFRO Monthly Forest Restoration Talks; SUPERB consortium, IUFRO (2023–2024)

International symposia, workshops and conferences (13.2 ECTS)

- OpenDataScience Europe (ODSE) workshop; oral presentation; Wageningen, The Netherlands (2021)
- Living Planet Symposium; poster presentation; Bonn, Germany (2022)
- ODSE workshop; oral presentation; Prague, Czech Republic (2022)
- ForestSAT; oral presentation; Berlin, Germany (2022)
- Land & Carbon Lab Global Summit; oral presentation; Brussels, Belgium (2023)
- Open-Earth-Monitor Cyberinfrastructure Global Workshop; oral presentation; Bolzen, Italy (2023)

Societally relevant exposure (0.9 ECTS)

- Right tree, right place: redesigning Europe's forests with AI; blog article (2022)
- Predicting world's vegetation under climate change scenarios: maps of forest areas most threatened by climate change; blog article (2023)
- From the Open-Earth-Monitor Global Workshop 2023: key takeaways and what to expect in 2024; blog article (2023)

Lecturing / supervision of practical's / tutorials (1.2 ECTS)

- Spatiotemporal ensemble machine learning in R with examples: distribution of *Fagus sylvatica* (2021)
- Forest species distribution modeling and spatial planning in R (2022)
- Covariate data for species distribution modeling (2023)

Supervision of MSc students (3 ECTS)

- Modeling distribution of African Boswellia (Frankincense) in the Sahel region in a climate change scenario (2022)

This research received funding from Grant Agreement Connecting Europe Facility (CEF) Telecom project 2018-EU-IA-0095 by the European Union ("GeoHarmonizer") and from the European Commission Horizon Europe research project "Open-Earth-Monitor Cyber-infrastructure" (grant agreement No. 101059548). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Financial support from Wageningen University and Research for printing this thesis is gratefully acknowledged.

Cover design by Carmelo Bonannella and Beatriz Sarabia

Printed by ProefschriftMaken