Remote Sensing of Environment 323 (2025) 114740

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Mapping tree species fractions in temperate mixed forests using Sentinel-2 time series and synthetically mixed training data

David Klehr [a,*], Johannes Stoffels [b], Andreas Hill [c], Vu-Dong Pham [d], Sebastian van der Linden [d], David Frantz [a]

[a] Geoinformatics - Spatial Data Science, Trier University, Universitätsring 15, 54296 Trier, Germany
[b] Environmental Remote Sensing and Geoinformatics, Trier University, Universitätsring 15, 54286 Trier, Germany
[c] Forest Planning and Inventory, Forest Service Rhineland Palatinate (Germany), Rhein-Mosel-Str. 9, 56281 Emmelshausen, Germany
[d] Earth Observation and Geoinformation Science Lab, Institute of Geography and Geology, University of Greifswald, Friedrich-Ludwig-Jahn-Str. 16, 17489 Greifswald, Germany

ARTICLE INFO

Editor Name: Dr. Marie Weiss

Keywords:
Sentinel-2
Tree species
Forest inventory
Time series reconstruction
Synthetic training data
Regression-based unmixing
Neural network regression

ABSTRACT

Monitoring and mapping of forest stands, including tree species composition can support forest protection and management. Sentinel-2 imagery provide a viable data source due to their high spectral, temporal, and spatial resolution. However, especially in temperate mixed forests challenges with tree species classification persist, mainly due to the high mixing ratio of tree species, which cannot be fully resolved even with the 10 m resolution of Sentinel-2 data. Additional challenges are associated with the commonly low number of reference data for rare tree species, resulting in low classification accuracy for these species. This study proposes an approach to map sub-pixel tree species fractions in mixed temperate forests by combining dense annual multi-spectral Sentinel-2 time series to target differences between species in phenological relevant periods with synthetically mixed training data. This allows for a limited number of pure training samples per tree species, which serves as basis for randomized linear mixing to compute a synthetic spectral library. An artificial neural network is trained for regression for tree species fractions per pixel. To enhance model robustness and stabilize predictions, we implemented this library generation and artificial neural network regression as an ensemble approach. We effectively mapped tree species fractions for the federal state of Rhineland-Palatinate, Germany, with Mean Absolute Errors of 2.76% to 16.05% and R^2 values up to 0.92 – when validated against forest planning data. We show that the data augmentation through synthetic mixing allows for a sample size as small as 30 pure pixels per class, to sufficiently distinguish nine tree species and one 'other species' class, hence substantially increasing the operational potential for deployment when reference data for rare species are limited – while simultaneously generating accurate and information-rich tree species distributions over large areas of mixed forest.

1. Introduction

Forest information layers are vital for understanding the effects of dynamic changes in climate, culture and economy on the multifaceted role of forest ecosystems (European Environment Agency, 2007). Precise and current data regarding the small-scale spatial distinction between coniferous and broadleaf and individual tree species composition are essential for fostering sustainable forest management and serve as pivotal components within forest monitoring programs (Fassnacht et al., 2024; Franklin, 2001; Hermosilla et al., 2024). Traditionally, these information on forest stand information have been derived from field inventories, but it is impossible to sample an entire area with full spatial coverage (Bolyn et al., 2018), and data availability is not guaranteed as it depends on factors such as accessibility, remoteness, and forest ownership (Hemmerling et al., 2021).

Given that traditional forest inventory methods are both resource-intensive and time-consuming, the utilization of remote sensing-based mapping approaches becomes an appealing strategy to complement and optimize extensive forest inventories (Boyd and Danson, 2005; Fassnacht et al., 2024). Through recurring observations, remote sensing

* Corresponding author.
E-mail addresses: klehr@uni-trier.de (D. Klehr), stoffels@uni-trier.de (J. Stoffels), Andreas.Hill@wald-rlp.de (A. Hill), vudong.pham@uni-greifswald.de (V.-D. Pham), sebastian.linden@uni-greifswald.de (S. van der Linden), david.frantz@uni-trier.de (D. Frantz).

https://doi.org/10.1016/j.rse.2025.114740
Received 18 October 2024; Received in revised form 26 March 2025; Accepted 31 March 2025
Available online 7 April 2025

offers opportunities to monitor forest stands more efficiently, comprehensively and at shorter intervals (Kangas et al., 2018; Yin et al., 2017) and thus can support forest management. However, field inventories still remain a fundamental method for operational forest management and are also needed for remote sensing based approaches in terms of training data, and map accuracy assessment or reference data.

Medium spatial resolution satellite imagery is particularly useful in providing detailed information for large or inaccessible areas (Fassnacht et al., 2024; Immitzer et al., 2012), while using minimal financial resources when using open data (White et al., 2016). Interest in tree species classification with satellite images has constantly increased in the last decade (Axelsson et al., 2021; Fassnacht et al., 2016; Kollert et al., 2021; Lechner et al., 2022; Welle et al., 2022). Although it is still remaining a challenging task, because approaches are often not transferable to larger areas and a gap exists between inventory information needs, i.e. species composition, and remote sensing capabilities, i.e. mapping dominant species, genus or forest type (Fassnacht et al., 2024).

Some operational forest monitoring programs and tree species classification approaches have been successfully implemented with Landsat data by employing dense time series and cloud-free image compositing (Hermosilla et al., 2024; Hermosilla et al., 2015; White et al., 2014; Wulder et al., 2022), the data have been successfully utilized to map dominant tree species on a large extend (Hermosilla et al., 2022; Shang et al., 2020; Thompson et al., 2015). Since their launch in 2015 and 2017 (and 2024), Sentinel-2 imagery further improved our monitoring capabilities due to their even higher spatial, spectral, and temporal resolution. The spatial resolution of 10–20 m with 10 spectral bands in the VIS-NIR-SWIR range and the revisit rate of 5 days under cloud-free conditions has already been demonstrated improvements in the research field of tree species classification (Grabska et al., 2019; Hemmerling et al., 2021; Immitzer et al., 2019; Nasiri et al., 2023), with dense time series approaches becoming state-of-the-art (Hemmerling et al., 2021). This is motivated by the observation that changing reflectance properties in different phenological phases can be used to distinguish vegetation types (Grabska et al., 2019). Klosterman and Richardson (2017) describe the influence of vegetation development on VIS reflectance mainly driven by budburst and leaf expansion in spring, and leaf coloring and abscission in autumn. Respectively, observations during relevant phases in the phenological development is crucial (Hill et al., 2010). However, gaps in the time series during the important periods can still occur due to cloud cover. Grabska et al. (2020) demonstrate how best-available-pixel composites of relevant dates can be used in combination with temporal metrics of Sentinel-2 imagery for tree species classification. The approach is developed further by utilizing temporal metrics of multiple years in relevant phenological phases (Grabska-Szwagrzyk et al., 2024). A feasible alternative is to use the complete, sometimes gap-filled, time series to improve the differentiation in phenologically relevant periods (Axelsson et al., 2021; Blickensdörfer et al., 2024; Hemmerling et al., 2021; Lechner et al., 2022; Nasiri et al., 2023; Schwieder et al., 2018; Zeng et al., 2020).

Despite recent advances in the large-scale mapping of tree species, the substantial local variability in tree species and forest structure at the stand level still poses a substantial challenge for currently employed classification procedures. This is especially true when considering that many temperate forests are characterized by a high mixing ratio, even at Sentinel-2's 10 m resolution. This challenge is further aggravated when aiming beyond the most common tree species (Blickensdörfer et al., 2024).

Hermosilla et al. (2022) proposed an approach for computing class membership probabilities and providing them as an indicator of the classification's attribution confidence. This approach enables the assignment of tree species assemblages to individual pixels. However, as this approach is a quasi-side product of the classifier, it is also bound by modern classification frameworks' biggest drawback, i.e., they usually require a huge number of training points. This becomes especially problematic for rare tree species as these are usually underrepresented in current reference databases (Blickensdörfer et al., 2024; Hemmerling et al., 2021) due to them being developed with focus on terrestrial forest management and not specifically designed for remote sensing's needs. Bolyn et al. (2022) proposed an approach using a convolutional neural network to map tree species proportions, targeting the issue of high tree species mixing ratio directly. While this approach has already yielded reliable results, it has not yet fully exploited the potential offered by time series.

While originally being developed as a machine learning-based alternative to traditional spectral mixture analysis (e.g., MESMA (Roberts et al., 1998), WASMA (Somers and Asner, 2014)), Okujeni et al.'s (2013) approach of regression-based unmixing using synthetically mixed training data presents several beneficial traits that might be instrumental for the requirement of mapping per-pixel mixtures of tree species.

Recent studies show, that the approach can be extended to multi-spectral (Cooper et al., 2020) and multi-temporal (Okujeni et al., 2021; Schug et al., 2018) or even time series-based data (Pham et al., 2024; Suess et al., 2018), but not using a dense gap-filled time series yet. It was applied to differentiate various mixtures of different land cover classes (Borges et al., 2022; Okujeni et al., 2018; Schug et al., 2024; Stanimirova et al., 2022) and to distinguish between different leaf types in forests like needleleaf and broadleaf forest (Cooper et al., 2020; Okujeni et al., 2021) and different forest cover type fractions (Bao et al., 2024). Pham et al. (2024) advanced the approach by employing artificial neural network (ANN) regression, a technique that offers benefits such as multi-class regression, which facilitates predictions across multiple categories, and an optimized multi-target loss function (Pham et al., 2024).

This approach is essentially turning a classification task into a regression approach with the considerable advantage of being based on a data augmentation step to synthetically generate a large amount of training data from only a few confirmed endmember locations. The necessity for a substantial quantity of training data, despite a limited availability of samples for certain classes can help addressing the requirements of mapping tree species in mixed forests, particularly when including rare tree species for which reference data are not abundantly available in traditional forest inventory databases. The approach of regression-based unmixing using synthetically mixed training data has not been evaluated using a dense gap-filled time series yet. However, this approach offers the potential to differentiate tree species by requiring only a small number of reference data to utilize the spectral and phenological differences between tree species and predict tree species fractions in mixed forests.

Consequently, the overarching goal of this study is to evaluate the potential of mapping sub-pixel tree species mixtures in mixed temperate forests at 10 m spatial resolution based on limited reference information for its subsequent usage in operational forest management.

For this purpose, the following research questions were formulated:

(1) How accurate are tree species mixture maps derived from a Sentinel-2 time series and machine learning regression with synthetic mixing data augmentation?
(2) What quantity of reference data is required to achieve sufficient accuracy, and still enable operational deployment by a forest survey?

2. Materials

2.1. Study area

Rhineland-Palatinate (RLP), a federal state of Germany, covers an area of about 19,850 km^2 (Fig. 1). With a combined forest cover exceeding 8080 km^2, or 41% of the state's total area, Rhineland-Palatinate is one of the most densely forested states in Germany (Federal Ministry of Food and Agriculture, 2014; Thünen-Institute,

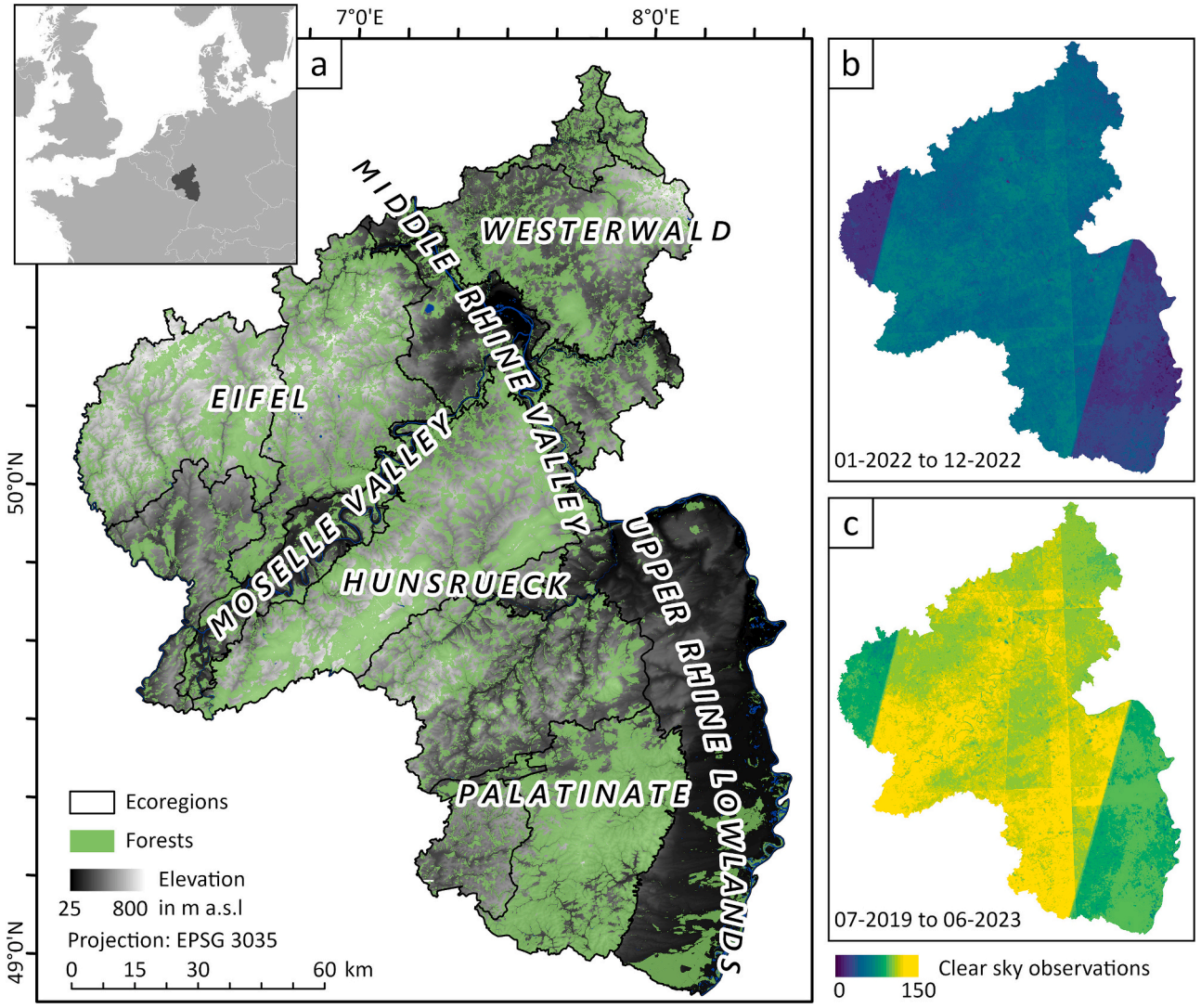

Fig. 1. Study area, i.e., Rhineland-Palatinate, Germany; *a)* Forests and Forest ecoregions after Gauer et al., 2005. (Data source: State Office for Surveying and Geographic Information Rhineland-Palatinate, 2024); *b)* Sentinel-2 Clear sky observations in the study area in 2022; *c)* Sentienel-2 Clear sky observations (07–2019 to 06–2023) used for the time series reconstruction.

2024), with 46.1% and 25.6% of forest areas being owned by municipalities and state, and 26.7% by private entities, respectively.

The forest area is divided into deciduous (62.7%) and coniferous (34.1%) tree types, while the remaining 3.2% of the area comprises temporarily unforested areas (e.g. due to calamities). The most frequently occurring tree species are European Beech (*Fagus sylvatica*), Sessile Oak (*Quercus petraea*) and Pedunculate Oak (*Quercus robur*), Norway Spruce (*Piecea abies*), Scots Pine (*Pinus sylvestris*), and Douglas Fir (*Pseudotsuga menziesii*) (in the subsequent text, common Oak and pedunculate Oak are collectively referred to as 'Oak'). These tree species collectively cover approximately 23.1%, 21.5%, 14.9%, 9.2%, 7% (i.e. 75.7% in total) of the forested area respectively (Thünen-Institute, 2024). However, more rare tree species can be locally important, like silver Fir (*Abies alba, 0.7%*), European Larch (*Larix decidua, 2.3%*), sycamore Maple (*Acer pseudoplatanus, 2.5%*), black Alder (*Alnus glutinosa, 1.3%*) and Birch (*Betula pubescens* and *pendula*, 3.5%). A list of tree species used for fraction prediction can be found in Table 1. In the following text, the tree species and groups listed are referred to as "tree species", despite the potential for taxonomic inaccuracy.

The low mountain ranges of the Eifel, the Hunsrück, and parts of the Palatinate exhibit diverse forest growth conditions owing to their cool and precipitation-rich mountain climate, which contrasts with the warm and dry continental climates prevalent in the Middle Rhine and Moselle Valleys, as well as the Upper Rhine Lowlands. Reflecting the variability in climate and natural surroundings, the study area is partitioned into 16 ecoregions, of which are 14 of relevant size, with unique forest growth conditions (Gauer et al., 2005). The weather conditions in Rhineland-Palatinate during the years under investigation were characterized by high average annual temperatures of 11.0 °C to 11.2 °C in 2020, 2022 and 2023, making them the three warmest years on record. Precipitation sums of 701 mm was recorded in 2022, the year of prediction in this study, which were relatively dry compared to the long-term average. (Competence Center for Climate Change Impacts - Rhineland-Palatinate, 2024).

2.2. Reference data

2.2.1. Forest planning data

The training (Section 3.1) and validation data (Section 3.5) used in this study were extracted from a statewide forest planning dataset that was provided by the State Forest Service of Rhineland-Palatinate and comprised stand-level information of state, communal and private forests. The data comprises information about the stocking tree species such as covering area, species proportion, timber volume and growth

Table 1
Tree species groups and auxiliary classes, with number of training samples overall, and for covered forest ecological region.

Name	Description	No. of samples	No. of ecoregions
Beech	Pure pixels of species European Beech (*Fagus sylvatica)*	57	14
Oak	Trees of species Sessile Oak (*Quercus petraea*) and Pedunculate Oak (*Quercus robur*)	58	12
Maple	Pure pixels of species sycamore Maple (*Acer pseudoplatanus*)	49	11
Alder	Pure pixels of species black Alder (*Alnus glutinosa*)	32	9
Other deciduous trees	Pure pixels of genus Poplar (*Populus*), cherry tree (*Prunus*), Birch (*Betula*), Ash (*Fraxinus*), Robinia (*Robinia*), and sweet chestnut species (*Castanea sativa*)	57	11
Spruce	Pure pixels of species Norway Spruce (*Piecea abies)*	49	10
Douglas Fir	Pure pixels of species Douglas Fir (*Pseudotsuga menziesii*)	55	9
Pine	Pure pixels of species Scots Pine (*Pinus sylvestris*)	45	11
Larch	Pure pixels of species European Larch (*Larix decidua*), Japanese Larch (*Larix kaempferi*)	41	10
Fir	Pure pixels of species European silver Fir (*Abies alba*)	31	8
Ground	Pure pixels of bare soil, grass or low shrub	32	12
Shadow	Pure Pixel of shadowed ground	30	10

rates (i.e., inventory data) as well as the amount of wood to be harvested, or planting actions (i.e., planning data).

The inventory data were gathered by forest experts who visit and examine each forest stand of the forest enterprise in an exhaustive field survey where the size of a stand can vary between 0.5 ha and 20 ha. Within this survey procedure, precise geographic location and extent of the forest stands are stored in a GIS database. A multitude of silvicultural attributes are then assigned to each forest stand geometry, which are not further georeferenced within the stand. This can been seen as a major challenge for its usage in remote sensing applications (Stoffels et al., 2015). Each prominent tree species with an overall covering area above 0.1 ha is stored as an individual database entry recording means of its stock layer (top layer, intermediate layer or ground layer), age, development phase, covering area (i.e., aggregated crown projection area), stock density and yield level. A renewal-interval of ten years is mandatory but not synchronized among the forest enterprises, such that the reference date can vary between regions.

The assessment of the inventory data during the field survey is realized by visual expert judgement that is complemented by stand-wise angle count sampling and height measurements in order to validate basal areas, stock densities, tree species covering areas and proportions as well as yield levels. However, the application, number and locations of these measurements are not predefined by a statistical sample-design but subjectively chosen by the expert to cover the stand variability. In case the expert conducts angle count-sampling, the software provides an interface where the angle count-sampling results (i.e. number of observed tree species at the chosen sample location) are attached to the respective tree-record in the database. Basal areas, stock densities, tree species covering areas and proportions are then derived directly from the angle count sampling results, and not from the visual expert judgement. Validation is also supported in the provided planning-software by means of digital orthophoto interpretation (covering area measurements) and LiDAR-derived Canopy Height Models (height measurements).

Due to this time-expensive full-census survey procedure, the gathered information on the forest stand- and enterprise level is considered very precise, although no quantitative accuracy metrics can be provided in general. In the case of recording rare tree species, this full-census survey can be superior to sample-based procedures due to insufficient sample-sizes that prevent a detection of rare events.

2.2.2. National forest inventory

This study utilizes information derived from the latest national forest inventory (NFI) from 2021 and 2022. The NFI involves the collection of data on various forest and tree parameters (e.g., species, height, diameter at breast height, etc.) in a network of almost 80,000 sampling points distributed throughout Germany. The Thünen-Institute uses this sample to estimate forest properties for both the individual federal states and entire Germany (Federal Ministry of Food and Agriculture, 2022). In this study, the state-wide estimation of the tree species-specific areas of the forest in RLP were employed for the purpose of cross-comparing area estimates (Thünen-Institute, 2024).

2.2.3. Digital orthophotos

We used openly available 20 cm R/G/B digital orthophotos for visual inspection of the study area of the years 2022 and 2023 provided by the State Office for Surveying and Geographic Information Rhineland-Palatinate. The acquisition dates vary across the study area, but the whole area is covered with a mosaic of images of May 05, 2022, June 06, 2022, June 18, 2022, July 02, 2022, July 03, 2022, July 17, 2022, May 25, 2023, June 01, 2023, June 04, 2023, June 06, 2023, June 15, 2023, August 06, 2023, August 08, 2023, August 09, 2023.

2.3. Sentinel-2 time series

This study utilized all available Sentinel-2 A/B Level 1C images with cloud coverage below 70% from July 2019 to June 2023 to reconstruct a gap-filled time series for 2022. All images were processed to level 2 using the Framework for Operational Radiometric Correction for Environmental monitoring (FORCE) version 3.7.12 (Frantz, 2019). The processing includes cloud and cloud shadow masking using a modified version of the Fmask algorithm (Frantz et al., 2018). Radiometric correction includes a radiative transfer based atmospheric correction (Frantz et al., 2016a), topographic correction using an enhanced C-correction (Buchner et al., 2020; Kobayashi and Sanga-Ngoie, 2008), adjacency effect correction, as well as bi-directional reflectance function correction (Roy et al., 2017). The ImproPhe algorithm was employed to adjust the spatial resolution of the 20 m bands of Sentinel-2 to 10 m (Frantz et al., 2016b). The data were co-registered using Landsat 8 to ensure high multitemporal geolocation stability (Rufin et al., 2021; Yan et al., 2018). During level 2 processing, the images were reprojected to EPSG:3035 and were divided into rectangular tiles (30 km × 30 km), representing a data cube structure for efficient higher-level processing.

The availability of data per pixel is not solely dependent on cloud coverage. Approximately 75% of the study area lies within the overlapping footprints of orbits 8 and 108 (Fig. 1*b*). For the entire target year 2022 between 20 and 40 clear sky observations can be used in the overlap region, while areas with single coverage have only between 10 and 15 observations in the northwest and 15 to 20 observations in the southeast (Fig. 1*b*). With the applied time series reconstruction approach (Section 3.2), the number of used clear sky observations used for reconstruction was increased to 90 to 150 in the overlap region and between 45 and 60 in the shoulder regions (Fig. 1*c*).

3. Methods

The employed workflow consists of five major parts (Fig. 2), which are described in detail in the following subsections, i.e. training data collection (3.1), Sentinel-2 processing and feature engineering for 2022 (3.2), tree species mixture modeling (3.3), inference of tree species mixtures and variability (3.4), and validation in forest stands (3.5). In addition to the validation, a model deviation product was computed and

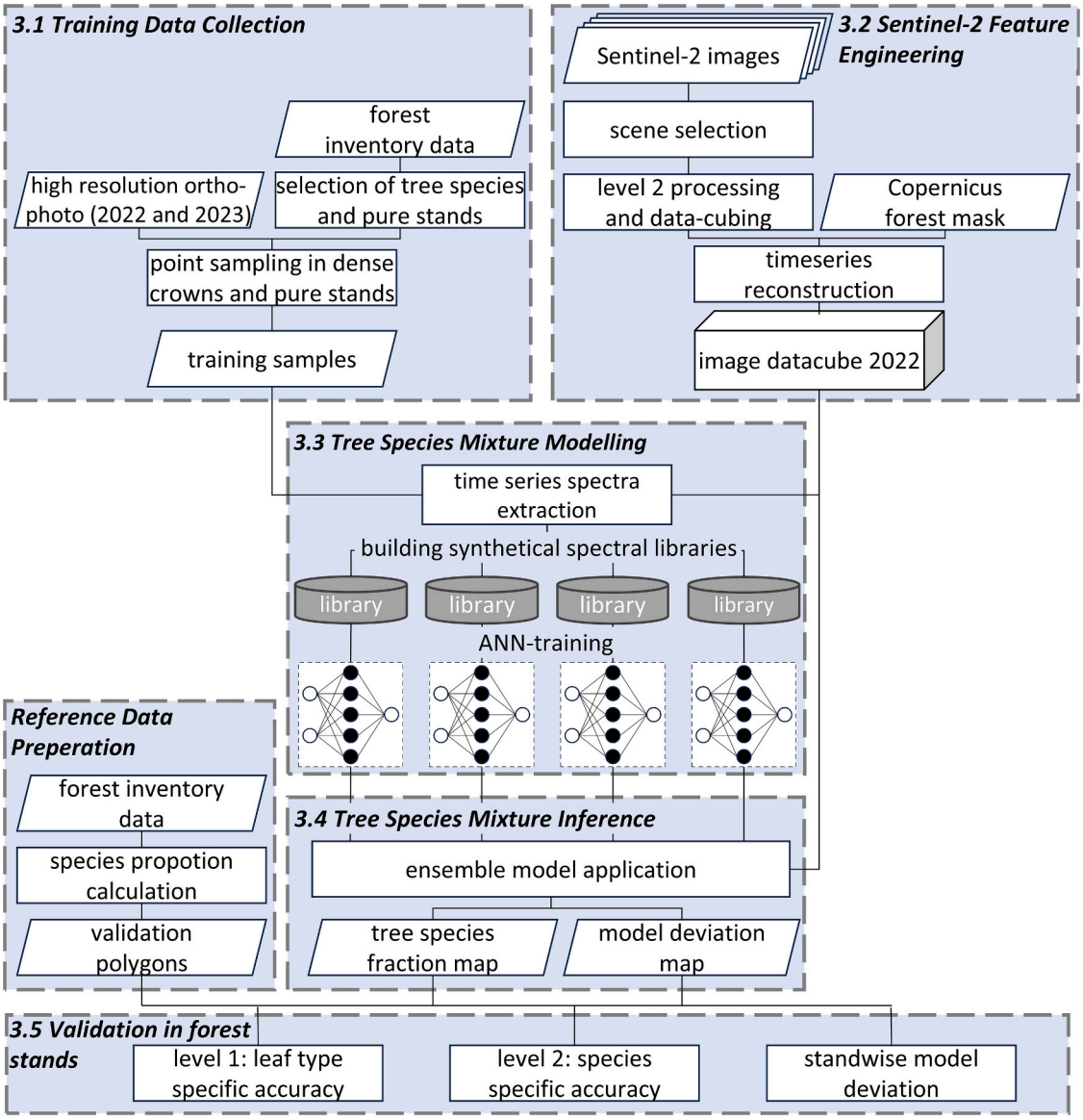

Fig. 2. Workflow for the tree species mapping approach consisting of training data and satellite image pre-processing, model training and application, and validation.

further analyzed.

3.1. Training data collection

In this study we target nine tree species, one joint class for other deciduous trees, and two background classes, shadow and ground (Table 1). We collected training data in all ecoregions, where pure stands were present, to cover different phenological conditions, i.e. variations in sprouting time, leaf senescence, or varying reflectance characteristics of the same tree species within the study area. This was crucial to cover within-class variance in the training data for better generalization capabilities. To achieve this, the forest planning data described in Section 2.2.1 were used to locate pure stands together with spatially high resolution digital orthophotos from 2022 and 2023 to manually select pure pixels of highest quality for each target class. The orthophotos were used to ensure that only dense crowns (i.e., gap closure) in homogeneous stands were sampled to avoid influences of understory trees (potentially other tree species) and ground surface (soil or vegetation other than trees). We selected pure stands with a minimum size of 0.2 ha to ensure a reasonable distance of multiple Sentinel pixels to the stands' border. For the class of other deciduous trees (other DT), pure stands of the following rare species were selected: Poplar, cherry tree, Birch, Ash, Robinia, and sweet chestnut. To provide suitable spatial distribution of the data, we covered all forest-ecoregions defined by Gauer et al., 2005, when pure stands were present for each tree species and background class. In addition, the sample collection also accounts for varying topographic conditions, including differences in altitude and aspect ground and shadow samples were selected in areas of bare soil and vegetation other than trees, and in shadowed forest borders respectively. We deliberately picked relatively few, but high-quality pixels, to demonstrate the effectiveness of the proposed method with respect to training data limitations, i.e., aiming at a number of training sites that would be realistic for a state service to provide on a regular basis.

The number of samples per tree species, the used tree species for each class, and their coverage within the forest ecological regions is shown in Table 1. Training samples could not be found for every class in every ecoregion, especially samples for Fir and Alder could only be identified in a limited number of regions.

3.2. Sentinel-2 feature engineering

We used bands 2–4 (visible light), 5–7 (red edge), 8 and 8a (near-infrared) and 11–12 (short wave infrared) as input features for later model training.

The ANN regression approach, as well as the synthetic mixing, technically requires an equidistant and gap-free time series, as well as a dense temporal quantification is needed to retain spectral-temporal differences and distinctions of phenological development between tree classes (Blickensdörfer et al., 2024; Hemmerling et al., 2021; Schwieder et al., 2016). To achieve this we used the time series reconstruction method of Bolton et al. (2020), who proposed using weighted splines of observations from multiple years. The strength of this method relies on consistent –forest phenology between years to overcome large data gaps such as those found in single orbit regions (Fig. 1*b-c*). We adopted this method while making minor adjustments, which we describe below. We implemented the reconstruction algorithm as a user-defined R-function in the higher-level module of FORCE (the code can be accessed on GitHub https://github.com/davidfrantz/force-udf/tree/main/rstats/ts/spline-reconstruction). Prior to the pixelwise reconstruction, an outlier detection was performed to decrease artifacts resulting from omission errors in the cloud detection (Frantz, 2019). The reconstruction was then calculated for the target year 2022, including observations from the two previous years 2021 and 2020 - if they were sufficiently similar.

The similarity was determined by calculating the Euclidean distance between the spline function of 2022 using observations from 07/2021 to 06/2023 and the spline functions of 2021 (using observations from 07/2020 to 06/2022) and 2020 (using observations from 07/2019 to 06/2021). The inclusion of observations in six months before and after the target year ensured reliable spline calculation in winter months. The Euclidean distance resulted in weights for the observations for 2020 and 2021. If the Euclidean distance of the spline function of one year was greater than the distance to the average time series of 2022, the observations of the respective year were disregarded. The final reconstruction was then calculated using a combined dataset of weighted observations.

To illustrate this approach, Fig. 3*a-c)* displays the result of the spline reconstruction of a pixel located in a non-overlapping Sentinel-2 orbit, as well as the lack of sufficient observations in fall 2022, which could be successfully reconstructed with help of observations from 2020 and 2021. Fig. 3*d-f)* demonstrates that if the time series are not sufficiently similar, for example due to bark beetle-induced disturbances, the reconstruction will not include observations from previous years.

The maximum allowable weight parameter for previous years has been set to 0.2, thus past years only influence the current year's prediction noticeably if there is a big data gap in a specific part of the current year. We used a cubic smoothing spline using the *stats* package (R Core Team, 2021) in *R* with a smoothing factor of 0.5. The reconstruction was calculated for all pixels within the forest mask deduced by the Copernicus high resolution layer forest type (Copernicus - Land Monitoring Service, 2024).

The reconstruction was calculated for the year 2022 at 10-day temporal resolution for all 10 spectral bands from March to November, hence effectively using Sentinel-2 data from 2019 to 2023.

3.3. Tree species mixture modeling

For modeling tree species fractions, we employed multi-output ANN regression, with training data based on synthetic mixing of endmembers. Firstly, a synthetic training dataset was generated, comparable with the synthetic spectral library introduced by Okujeni et al., 2013. For this purpose, the time series of the target year for March to November for every interpolated Sentinel-2 band were extracted at the locations of the training sample. This resulted in feature vectors of 280 elements (10 spectral bands for 28 time steps) for every sample point serving as an endmember. Using these pure endmembers, 256,000 training vectors (v)

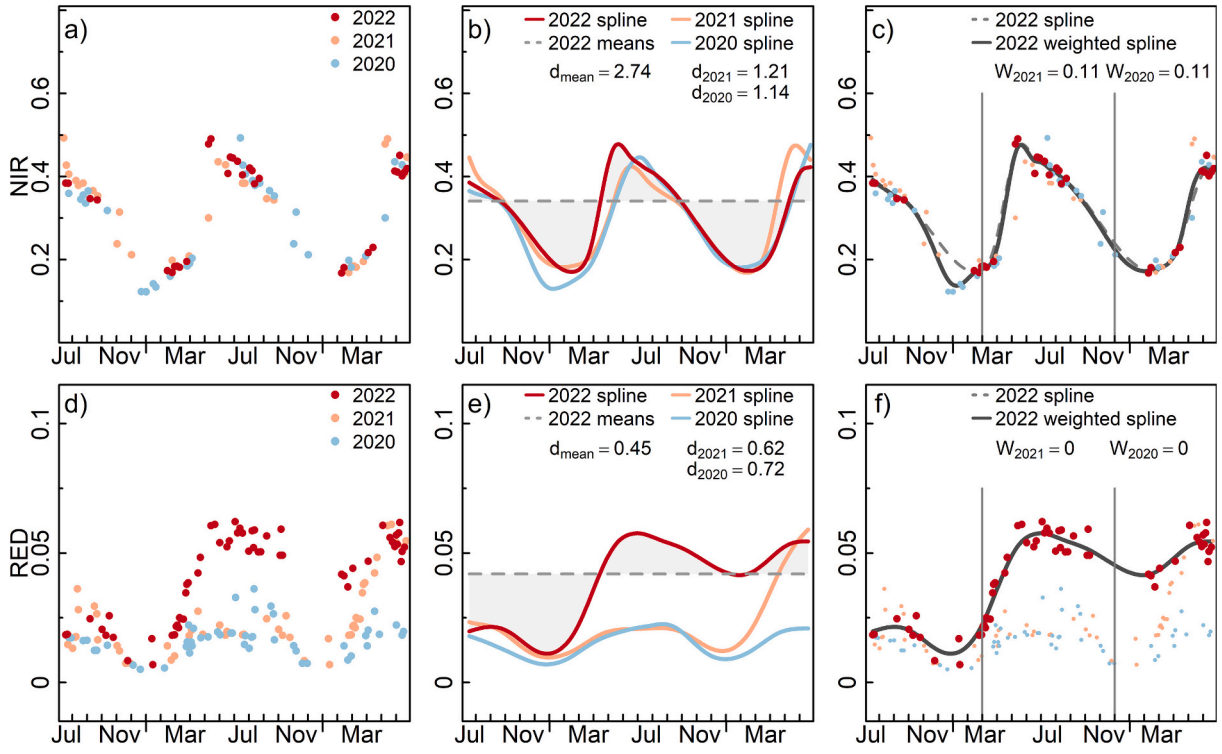

Fig. 3. Spline-based time series reconstruction for two example pixels. (*a-c*): A healthy Beech pixel in a Sentinel-2 single-orbit area; (*d-f*): A pixel of pure Spruce forest affected by a bark beetle attack in the year 2022 (*a, d*) all observations of 2020, 2021, and 2022, each year with observations six months before and after for a successful spline calculation; (*b, e*) spline functions of 2020, 2021, 2022 and mean function of 2022 with corresponding euclidean distance values (d_{year}) to the 2022 spline; (*c,f*): weighted spline function with observations from 2020 to 2022 and the weights (W_{year}) of the corresponding years (weight of 2022 equals one).

were synthetically mixed by linear combination of endmembers (e_j) and fractions (f_j) in a random manner, with the condition of summing up to 1, with a mixing complexity of up to $m = 3$, using the formula:

$$v_i = \sum_{j=1}^{m} f_j * e_j \quad (1)$$

The calculation contained one to three randomly selected endmembers with a distinct likelihood of each mixing complexity. Synthetic data augmentation by linear combination has been shown effective to improve model training (Zhang et al., 2017). The fractions were selected at random with the condition that they are positive and that their sum equals one. Pure endmembers were also included in the spectral library, i.e. a mixing complexity of one. Within-class mixing was permitted as well, hence also generating synthetically mixed feature vectors from endmembers of the same tree species, i.e., to account for regional variability, e.g. in terms of phenology. Two matrices stored the resulting synthetic training data and a 12-node vector of endmember class fractions for each data point, i.e., the feature matrix, and the predictor matrix.

Secondly, ANN models were trained. For this study, we used a Keras model architecture with the TensorFlow package version 2.10.0 (TensorFlow, 2023) and Python version 3.9.10 (Python Software Foundation, 2022). The network architecture is fully connected and employs five hidden layers, each with 128 nodes. The feature vector used was a synthetic mixed spectrum, containing 280 nodes, to predict nine tree species, other deciduous trees, ground, and shadow fraction, stored in the predictor matrix. The training process was performed over 250 epochs using a learning rate of 0.001 and a learning rate decay of 0.5 every epoch, with an Adam optimizer. The used batch size of 256 synthetic vectors resulted in 1000 training iterations per epoch. The Mean Absolute Error (MAE) was used as the loss function for training, which calculates the difference between all predicted and actual fraction vectors within one training iteration. An average MAE was also calculated per epoch to evaluate performance and progress. The ANN training process was based on the approach developed by (Pham et al., 2024).

One consequence of using the synthetic mixing approach with a high degree of randomness is that the results of individual models exhibit variability. Thus, we employed an ensemble approach, as recommended by Okujeni et al. (2018). Hence, the steps for constructing synthetic training data and training a multi-output ANN were executed ten times, resulting in individual trained models to account for inter-model uncertainty.

3.4. Tree species mixture inference and variability

The trained ANN models were applied to the corresponding reconstructed Sentinel-2 time series data-cubes. The annual prediction was performed by pixel-wise application of the $n = 10$ models, followed by averaging the different predictions ($x_{1..n}$), to account for inter-model uncertainty:

$$AVG_{spec} = \frac{1}{n} * \sum_{i=1}^{n} x_{spec,i} \quad (2)$$

For efficient large-scale calculation, we used 30 by 30 km tiles defined by FORCE, to implement parallel processing in Python. This ensemble prediction resulted in a 12-band raster containing the average tree species, ground and shadow fractions for the entire study area's forest extent. The inference of the tree species mixture maps was completed with L1-normalization of the average prediction to ensure that the sum of fractions for every pixel equals 100%:

$$AVG'_{spec} = AVG_{spec} \Big/ \sum_{j=1}^{12} AVG_j \quad (3)$$

This implementation additionally allowed an analysis of the mean absolute deviation (MAD) per tree class between all individual models, which was calculated as:

$$MAD_{spec} = \frac{1}{n} * \sum_{i=1}^{n} \left| x_{spec,i} - \bar{x}_{spec} \right| \quad (4)$$

where $x_{1..n}$ are the individual model predictions and $\bar{x}$ is the average prediction of n models. The resulting deviation per class was sampled using two different strategies: a pixel-based approach and a polygon-based approach, averaging the MAD values using the validation polygons, and depending on the species. The polygon-based approach allows for an examination of the relationship between the prediction error of the ensemble prediction and the MAD within each validation polygon. In addition, we have calculated and sampled the deviation between five ensemble approaches to analyze the effect of ensemble application.

3.5. Validation in forest stands

The average fraction raster was validated using information about tree species proportions within the polygon-based forest planning data described in Section 2.2.1. To ensure dense crown conditions in validation data, we selected polygons that contain fully grown trees, i.e. trees of stage dimensioning or maturation and data vintage newer than 2017. The units were required to exclusively contain trained tree species. Additionally, individual polygons must exceed 1 ha. The proportion of species within the forest unit was determined by the area of each tree species present. To further increase the reliability of the validation data, some validation polygons were excluded due to oddly shaped geometries unfeasible for validation (e.g. slim and tubular or separated subpolygons) or inconsistencies in the recorded information detectable with the orthophoto (e.g. disturbance or high crown sparseness). The final validation dataset comprised 2621 polygons, distributed across the entire study area in forest owned by state or municipalities with high spatial distribution per species where possible (Table 2).

Due to the data design of the reference dataset, fraction validation on pixel level was not possible. Polygon-based validation is a well-established method for evaluating fraction maps (Cooper et al., 2020; Okujeni et al., 2021), wherein the average predicted fraction of all pixels within the polygon of one class - referred as *predicted proportion* - is compared with the reference proportion. In accordance with the data design, the validation does not consist of a map assessment for the entire predicted map, but rather, it is a validation for the utilization of this method within forest stands. Due to the nature of the reference data and their collection in the field for the purpose of forest management, additional processing of the inferred tree species fraction maps was necessary before validation. The forest unit polygons only contain the proportion of tree species relative to the forested area, i.e., non-forested and shadowed areas are unaccounted for. As such, statistical validation of the ground and shadow fraction was not possible within the polygons. To account for Spruce stands affected by bark beetle outbreaks, we implemented a forest loss mask after the approach of (Dutrieux et al., 2021; Stoffels et al., 2024) to remove afflicted stands for validation. We used the Root Mean Squared Error (RMSE), and the MAE as error statistics (Willmott and Matsuura, 2005), as well as the R^2, slope and intercept from ordinary least squares regression. The validation design was divided into two levels. The first level assessed the broad-leafed and needle-leafed tree type fractions only (i.e., the sum of corresponding species, Table 2). The second level was carried out at individual tree species. Two distinct approaches were employed to examine the second level error metrics. The first approach utilized all validation polygons for error metric calculation, thereby representing the overall performance of each tree species. The second approach employed only polygons where an individual tree species occurred in the validation data, i.e., excluding all validation polygons with a proportion of 0%. We deemed this distinction necessary for various reasons. The map may be evaluated with reference to how effectively the fractional estimate of the species

Table 2
Numbers of validation polygons considered where tree species are present.

	deciduous species					coniferous species				
	Beech	Oak	Maple	Alder	Other DT	Spruce	D. Fir	Pine	Larch	Fir
Level 1: Number of samples	2121					2024				
Level 2: Number of samples	1541	1006	266	201	619	1072	931	644	613	72
Number of ecoregions	14	14	13	10	14	13	13	13	14	7

was generated where the corresponding tree species knowingly exists. It is also relevant to assess the model's overall performance, i.e., including whether species' absence was accurately predicted, or commission errors are frequent. As our validation dataset cannot be balanced across the entire fractional data range across all the tree species, both assessments need to be considered, especially with regard to the frequency of occurrence of given tree species, i.e., in a similar notion as evaluating producer's and user's accuracy in a classification.

3.6. Analysis of training sample size effect

To analyze what quantity of training data is required to achieve sufficient prediction accuracy, we performed an ablation experiment by systematically degrading the number of training points. In an iterative manner, we used a subsample of the available pure pixels within the training dataset (see Section 3.1) by randomly selecting 1, 2, 4, 6, […], or 34 pure pixels for each tree class. This resulted in 18 degraded training datasets. For degraded training dataset, the workflow of spectral library generation, ANN training, and fraction inference (described

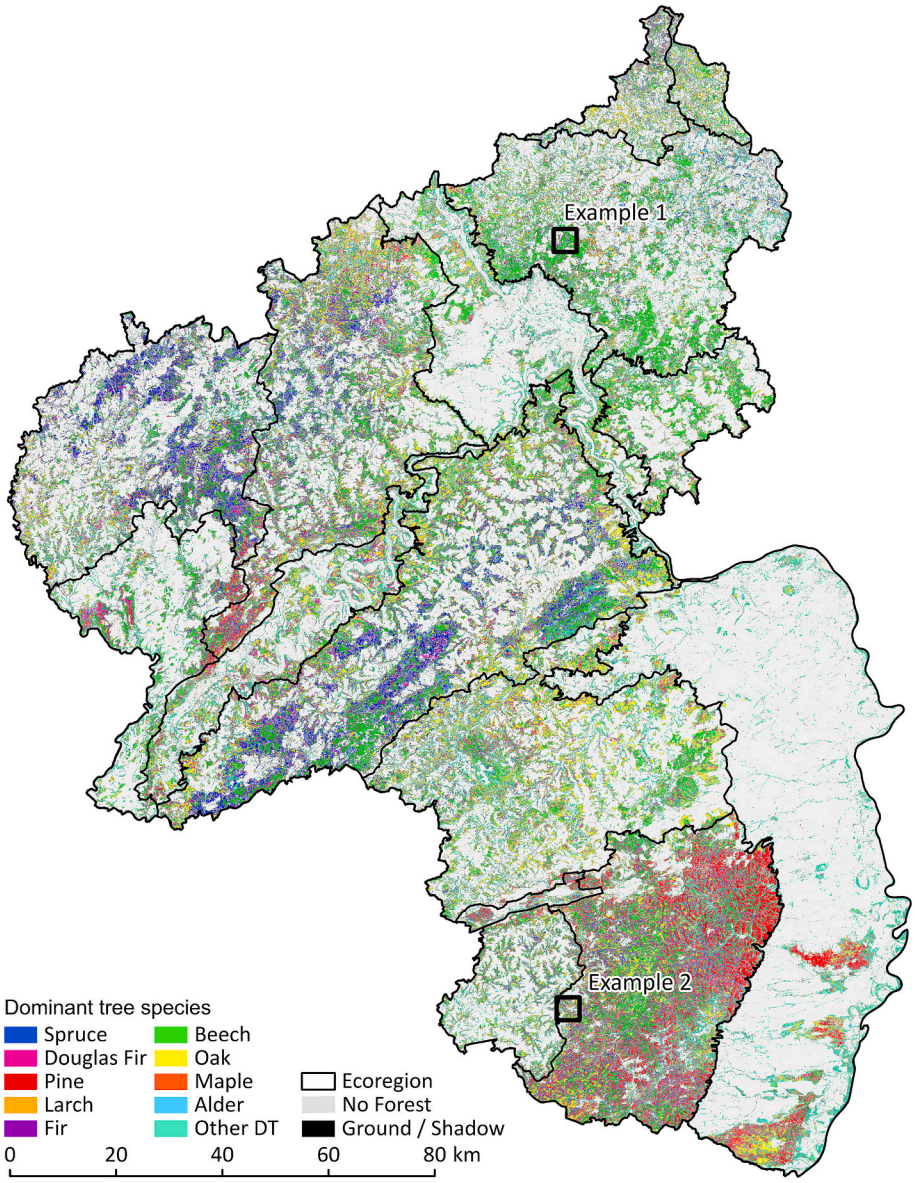

Fig. 4. Tree species fraction map of Rhineland-Palatinate, Germany. The dominant class and its fraction are visualized using the HSV-color space with a species specific hue, saturation as fraction of the dominant tree species, and value as the sum background and shadow fraction. A detailed description of the color coding process is described in the supplemental. The two example areas are visualized in detail in Fig. 5.

in Sections 3.3 and 3.4) was repeated. The resulting predictions were validated as described in Section 3.5 to investigate the method's performance in relation to an increasing sample size.

4. Results

4.1. Tree type and tree species mixture maps

Fig. 4 shows the resulting tree species fraction map of the study area. Details about the visualization approach using the HSV-color space are explained in the supplementary document (Section 1). Maps of individual tree species fractions of the entire study area can be found in the supplementary document too (Figs. A4 to A15). The tree species proportions per forest ecological region and whole Rhineland-Palatinate of dense crown forest pixels are shown in Table 3. The difference between the official number of total area of forest of RLP by (Landesforsten, 2012) and the forest area proposed in this analysis is attributable to the use of the forest mask of Copernicus, which uses a different definition of forest, in combination with the applied forest loss mask, which accounts for an area of 673.96 km^2.

Fig. 5 presents representative examples for two validation sites (locations in Fig. 4), showing the results of the tree type (level 1 – RGB color coded) and tree species maps (level 2 – HSV color coded), also showing the reference and predicted proportion within the polygon in the bar chart. Fig. 5*a-c)* illustrates a heterogeneous forest comprising Beech, Spruce, and Larch. The stand proportion of Spruce is underestimated by 8%, while the stand proportion of Larch is overestimated by 3%, indicated by the bar chart comparison. Fig. 5*d-f)* depicts a heterogeneous stand comprising Beech, Oak, and Pine. The predicted distribution of tree species shows a slight underestimation of the proportion of Pine and a slight overestimation of Beech. A detailed description of this two example polygons can be found in the supplementary (Section 2).

We have evaluated the predicted area of each tree species with the state-wide projection of the NFI dataset described in Section 2.2.2 (Fig. 6). Our approach slightly underestimates the total predicted area of the prevalent tree species Beech and Oak, while rare species like Alder and Larch are overestimated. Spruce is underestimated in relation to the NFI assessment.

4.2. Validation in forest stand

4.2.1. Level one: Leaf type proportions in forest stands

The validation of the leaf type mixture maps resulted in R^2 values of 0.91 for both forest types (Fig. 7). The slopes of the models are close to 1, with a slight tendency for underestimation of tree type fractions for high reference proportions. The values of MAE and RMSE are 8.91% and 12.71% respectively for both leaf types. The distributions of the residuals indicate a slight overestimation for deciduous species for low proportions and a slight underestimation for coniferous species for high proportions.

4.2.2. Level two: Tree species proportion in forest stands

At the second level of validation, we compared the tree species mixture map with species-specific validation data in forest stand polygons. We used two different approaches, representing the overall model performance and the presence-only performance. Fig. 8 shows the scatterplots and regression models of each individual tree species, while the respective error metrics and R^2 can be found in Table 4. The regression models for deciduous tree species achieve R^2 values ranging from 0.41 to 0.83. All deciduous species exhibit a tendency to be underestimated with increasing validation proportions, indicated by slopes between 0.98 and 1.19 (apart from other DT class). It is noted that the model can only underestimate the real value, as prediction and references values are limited between 0% and 100%. The best error statistics for presence-only performance are observed for Alder with MAE of 11.64%, and other DT with RMSE of 16.87%, while best error statistics for overall performance are observed for Alder, with MAE and RMSE of 3.17% and 6.75% respectively. Beech exhibits the highest errors for overall performance, with MAE of 9.12% and RMSE of 14.21%. With regard to presence-only performance, Oak exhibits the highest error values, with MAE of 14.61% and RMSE of 20.52%.

For coniferous species, the underestimation with increasing validation proportions is again noticeable, due to the 100% limit. Slopes are close to, or higher than one, with values between 0.97 and 1.27, except for the Fir species, which has a slope of 0.42. Nevertheless, the R^2 values indicate reliable predictions for Spruce, Douglas Fir, Pine, and Larch, with values between 0.66 and 0.92. The predictions of Fir in terms of overall performance appears to be more unreliable, with an R^2 of 0.27. In contrast, the model for presence-only performance achieves an R^2 of 0.66, which is considerably higher. It is important to exercise caution when interpreting the results for the Fir species, given the high number of validation polygons that do not contain this species (2549) and the low number of validation polygons with presence (72). However, Fir exhibits the best error statistics for overall performance, with MAE and RMSE of 2.76% and 5.61%, respectively and presence-only performance, with MAE and RMSE of 7.49% and 12.18%. Douglas Fir exhibits the highest error statistics. Notably, the MAE (16.05%) and RMSE (20.85%) in the presence-only performance are higher than those observed for the other species.

Table 3
Tree species proportion [%] per forest ecological region. 40: Sauerland; 41: Bergisches Land; 43: Niederrheinische Bucht; 44: Nordwesteifel; 45: Osteifel; 46: Mittelrheintal; 47: Westerwald; 48: Taunus; 65: Oberrheinisches Tiefland und Rhein-Main Ebene; 66: Hunsrück; 67: Moseltal; 68: Gutland; 69: Saarländisch-Pfälzisches Muschelkalkgebiet; 70: Saar-Nahe Bergland; 71: Westricher Moorniederung; 72: Pfälzerwald.

Ecoregion	Forest [km^2]	Beech	Oak	Maple	Alder	Other DT	Spruce	Do. Fir	Pine	Larch	Fir
RLP	7272.81	22.73	18.48	6.2	5.16	16.39	9.31	5.76	8.95	4.44	2.57
40	69.27	21.12	24.46	10.86	11.84	16.79	2.73	3.38	2.21	4.44	2.17
41	98.99	19.66	23.8	7.4	11.53	15.75	6.56	5.27	2.96	3.17	3.89
43	0.20	4.19	6.96	14.82	2.97	40.56	2.66	4.35	6.14	13.92	3.42
44	579.46	18.6	14.1	4.29	5.25	12.27	26	8.02	3.53	3.31	4.64
45	850.86	23.21	20.47	7.53	4.73	13.2	10.77	8.18	5.2	3.83	2.88
46	282.28	18.22	19.92	10.33	3.08	31.36	1.25	2.5	3.39	7.93	2.03
47	657.98	33.4	13.6	10.67	8.7	17.2	4.11	2.61	2.25	5.59	1.88
48	213.6	39.67	19.8	6.15	3.65	14.4	3.18	3.96	2.9	4.82	1.46
65	429.79	7.33	14.89	5.72	3.88	36.58	1.34	1.8	16.33	10.06	2.07
66	1275.21	24.48	19.08	4.97	4.09	12.95	18.75	6.16	3.3	3.52	2.71
67	277.59	17.86	21.88	8.74	2.07	26.67	2.28	7.26	5.13	5.32	2.8
68	218.28	29.27	13.2	5.67	3.47	18.46	7.43	6.52	8.94	5.44	1.62
69	169.52	19.89	20.37	8.53	5.6	21.8	6.58	4.23	5.27	5.44	2.3
70	808.35	21.38	32.3	7.41	3.09	20.82	3.05	4.34	3.12	3.11	1.39
71	32.90	13.05	13.64	2.42	8.46	14.14	10.78	5.8	25.07	4.93	1.72
72	1304.25	22.16	12.4	2.63	7.03	7.8	6.32	7.49	27.79	3.41	2.98

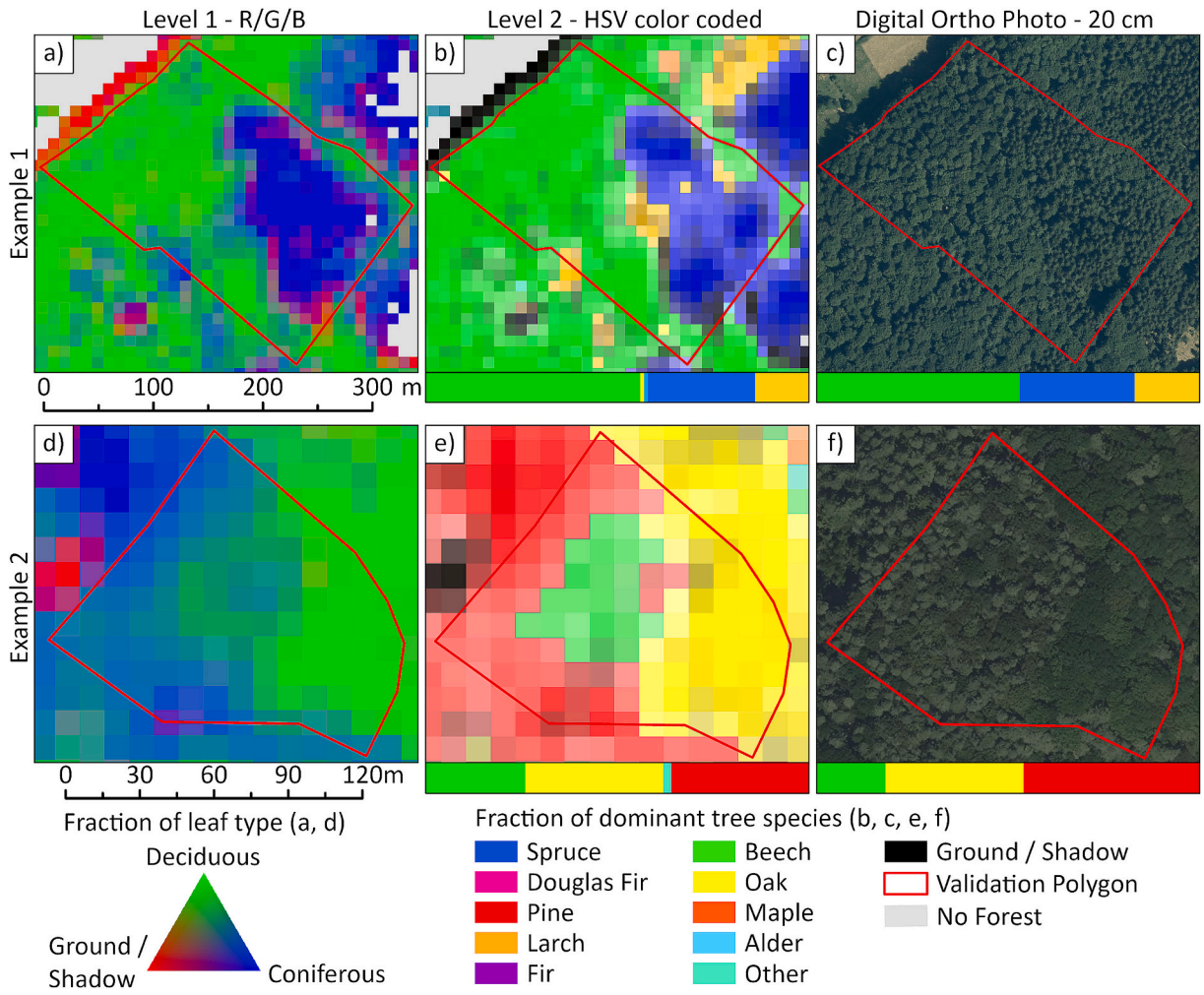

Fig. 5. Tree species fraction map at two example validation polygons; *a)* and *d)* show the level 1 product of leaf type and ground fractions R/G/B color-coded; *b)* and *e)* show the HSV color-coded level 2 tree species fraction map. Hue is defined as color for the dominant tree species, saturation is defined as fraction of the dominant tree species, and value is defined as the sum background and shadow fraction (detailed description in the supplemental). The bar chart here refers to the predicted proportions in the validation polygons; *c)* and *f)* show the validation polygons and tree species distribution for validation based on the official forest information as a bar chart with the digital orthophoto; Orthophoto data (*c)* 04.06.2023; *f)* 02.07.2022): ©GeoBasis-DE / LVermGeoRP (2024), dl-de/by-2-0, http://www.lvermgeo.rlp.de

4.3. Ensemble deviation

The analysis of the ensemble deviation using an example polygon is illustrated and described in detail in the supplementary (Section 3). Fig. 9*a)* depicts the distribution of the MAD in dependence on the average of the predicted fractions using the two sampling strategies described in Section 3.4. The polygon-based view demonstrates a slight yet consistent rise in the deviation with rising fractions, reaching a peak (MAD of 5.64% in average) at approximately 40% average fraction, after which there is a slight yet continuous decline. In contrast, the pixel-based metrics shows a rapid increase in deviation with an increasing fraction, reaching a maximum deviation between 10% and 20% average fraction (MAD of 8.66% in average). Subsequently, the distribution aligns with the polygon-based metrics, exhibiting a strong resemblance from approximately 40% average fraction, with a slight dispersion of the MAD values. The distribution indicates that single pixels with high deviation can occur with predictions between 15% and 30% for the specific tree species. These high MAD values of individual pixels are smoothed by the averaging in the polygon-based approach, which compensates for this effect. The third boxplot class in plot *a)* represents the deviation between five ensemble prediction. A notable reduction in the deviation can be observed in comparison to the MAD within a single ensemble model.

Fig. 9*b)* illustrates the correlation between the prediction error and MAD by emphasizing species fraction predictions with a high MAD in the scatterplot of predicted versus reference proportion. The Pearson correlation coefficient of 0.72 indicates a strong positive correlation. Most of these predictions fall within the range of 25 to 75%, as illustrated in plot *a)*. However, they are particularly prevalent in samples with a high prediction error.

4.4. Effect of training sample size

The development of the ANN's model performance with increasing pure species training sample size can be observed in the error statistics graphs in Fig. 10. A swift improvement in MAE and RMSE is evident when increasing the sample size to ten pure samples for some species, including Beech, Oak, Maple, Douglas Fir and Pine (e.g. for Beech overall MAE from 22.82% to 12.53%). However, when the sample size is further increased, the improvement appears to stagnate for Beech, Oak, Maple and Pine with a sample size between 20 and 30. Some species start with comparatively low error values, and show only small reductions, such as Alder, Larch, Fir and the other DT class (e.g. other DT overall MAE from 9.13% to 6.90%). A minimum is reached at a sample

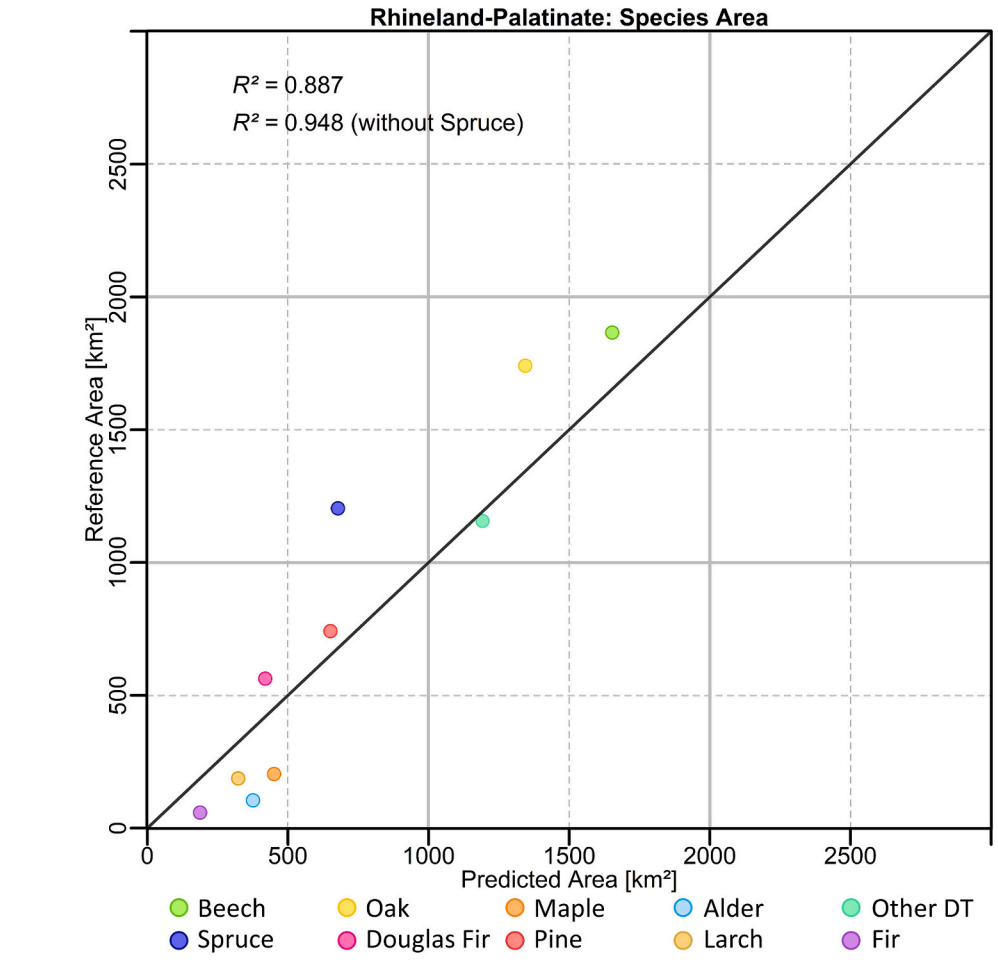

Fig. 6. Validation of predicted tree species area with expected area based on tree species proportions based on the NFI (Thünen-Institute).

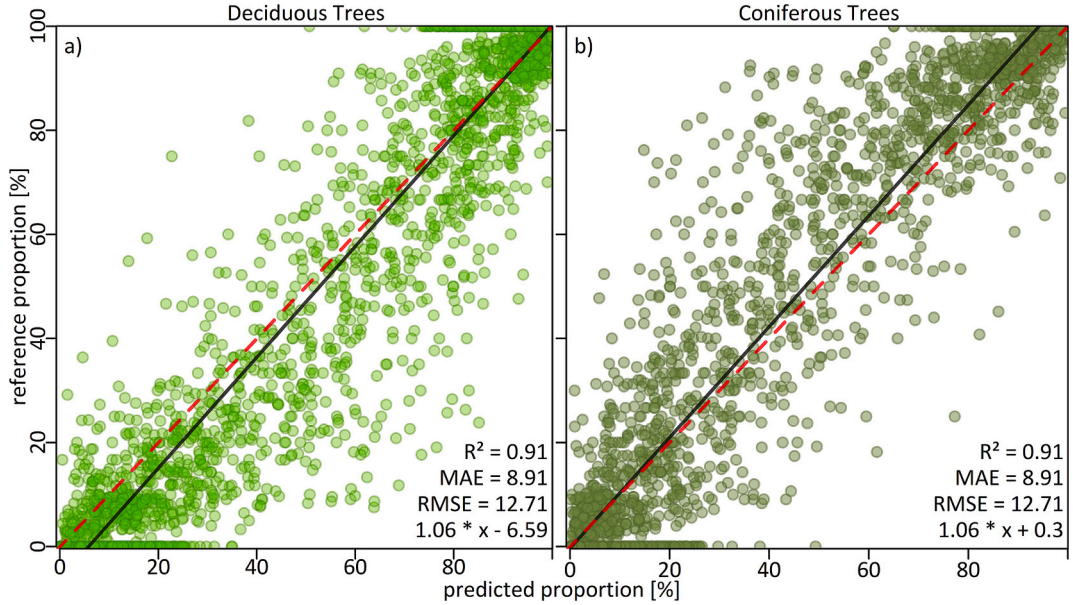

Fig. 7. Scatterplots of predicted proportions vs reference proportions for leaf types for deciduous trees (*a*) and coniferous trees (*b*); black line = regression line, red dashed line = 1:1 line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

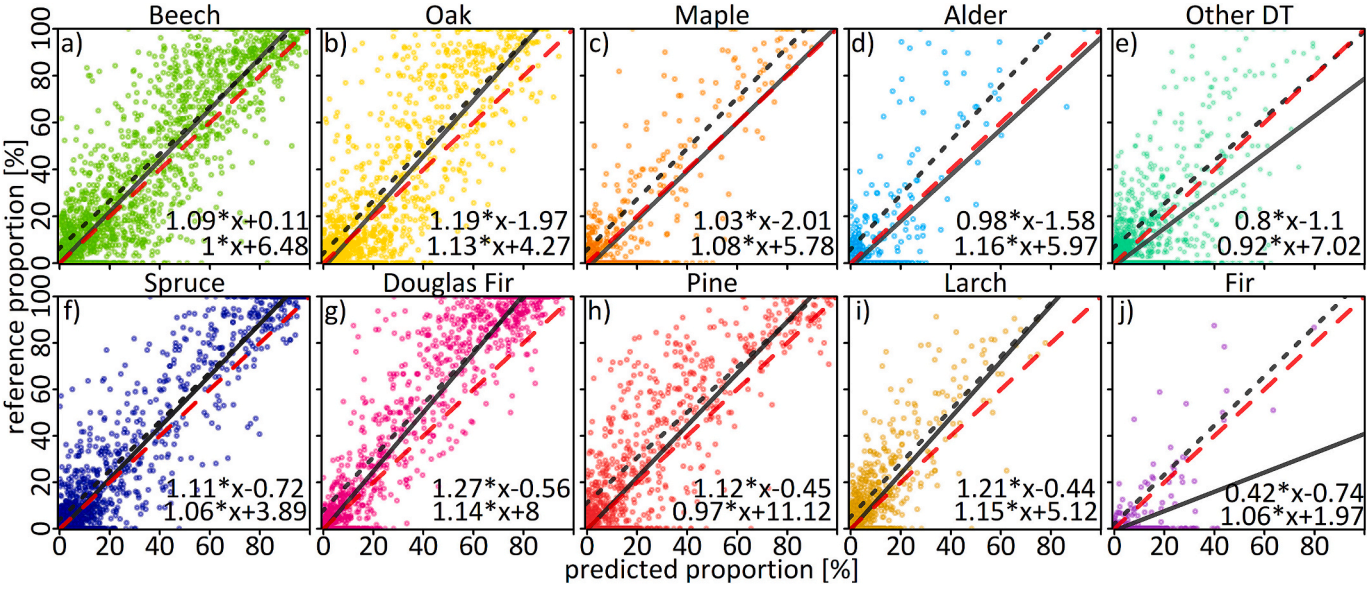

Fig. 8. Scatterplots of predicted proportions vs reference proportions for individual tree species for deciduous tree species and for coniferous tree species; the black solid line represents the validation using all validation polygons (i.e. overall error model) and belongs to the top formula; the black dotted line represents validation using only validation polygons where the species is actually present (i.e. presence-only error model) and belongs to the second formula; the red dashed line represents the 1:1 relationship. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Error metrics and R^2 of the predicted proportions in the validation polygons for all tree species. Overall metrics include all available validation polygons and presence-only metrics only include polygons, where the respective species is present.

	MAE [%]		RMSE [%]		R^2	
	overall	presence-only	overall	presence-only	overall	presence-only
Beech	9.12	12.8	14.21	17.44	0.83	0.76
Oak	8.01	14.61	13.88	20.52	0.77	0.7
Maple	3.83	11.78	7.69	16.89	0.72	0.78
Alder	3.17	11.64	6.75	18.5	0.51	0.61
Other DT	6.79	11.72	11.43	16.87	0.41	0.49
Spruce	5.41	9.73	9.6	13.9	0.9	0.88
Do. Fir	6.89	16.05	12.78	20.85	0.92	0.84
Pine	5.33	13.97	10.68	19.28	0.79	0.73
Larch	3.18	8.28	6.67	12.42	0.69	0.71
Fir	2.76	7.49	5.61	12.18	0.27	0.66

size of 20, after which no further reduction can be detected.

Douglas Fir and Spruce continue improving when the sample size is increased to 30. Nevertheless, for all species, the effects of sample size become increasingly less pronounced as the number of training samples increases, e.g. Larch overall RMSE from 8.65% with 24 samples to 7.09% with 34 samples.

The development of the R^2 is similar across all species, with the exception of Alder, other DT class, and Fir in the overall metrics. Once again, a notable improvement is observed with the increase in sample size to ten (e.g. Oak from 0.10 to 0.59). The improvement for Beech, Oak, Maple, and Pine is considerably more rapid, while Larch is influenced by a reduction towards a sample size of 6, followed by the improvement. Regardless, all the aforementioned species are characterized by a saturation of R^2 between a sample size of 20 and 30. The development is comparable to a root function, but with a positive exponent.

The development of the overall performance in R^2 for Alder, other DT, and Fir and the presence-only performance in R^2 for other DT appear to be noisy and more linear, with continued improvement to a sample size of 22 and stabilization between 24 and 34 samples. Also, the development of the R^2 of Fir in the presence only performance is influenced by strong fluctuations. However, the development of the overall performance in R^2 for Fir requires caution due to the relatively low number of available validation polygons in comparison to the high number of polygons of absence, which results in only minor improvements in overall performance.

5. Discussion

5.1. Tree species fraction accuracy

The combination of a regression approach with synthetically mixed training data from a densely reconstructed multispectral Sentinel-2 time series has enabled the prediction of different tree species within mixed pixels, and hence mixed forest stands. The validation demonstrates the ability to effectively predict the proportion in forest stands of needle-leafed and broadleaved species, as well as proportions of nine tree species and one remainder class. The effectiveness of separating deciduous and coniferous forest (level 1 in this study) was already reported in previous studies using regression-based unmixing, although they focused on the distinction between different vegetation types outside of forests (Cooper et al., 2020; Okujeni et al., 2021). Nevertheless, although different classes were considered, as well as a different validation design was used (pixel-based instead of polygon-based), the achieved error metrics are comparable to our level 1 assessment (e.g., MAE of 7.8% to 8.8% (Okujeni et al., 2021)), which provides a first qualitative indication that our approach performs in a similar way as published research, with the advancement of distinguishing between different tree species.

The validation of the individual tree species must be considered in a differentiated manner. To validate the metrics, two different approaches have been employed: an evaluation of the metrics in the overall context and in the presence-only context. In general our results show the typical patterns of underestimating higher reference proportions, which is evident in many studies that estimate sub-pixel fractions with machine learning-based approaches (Cooper et al., 2020; Guerschman et al., 2015; Kowalski et al., 2022; Okujeni et al., 2021; Okujeni et al., 2018; Pham et al., 2024; Schug et al., 2020; Senf et al., 2020). This effect may

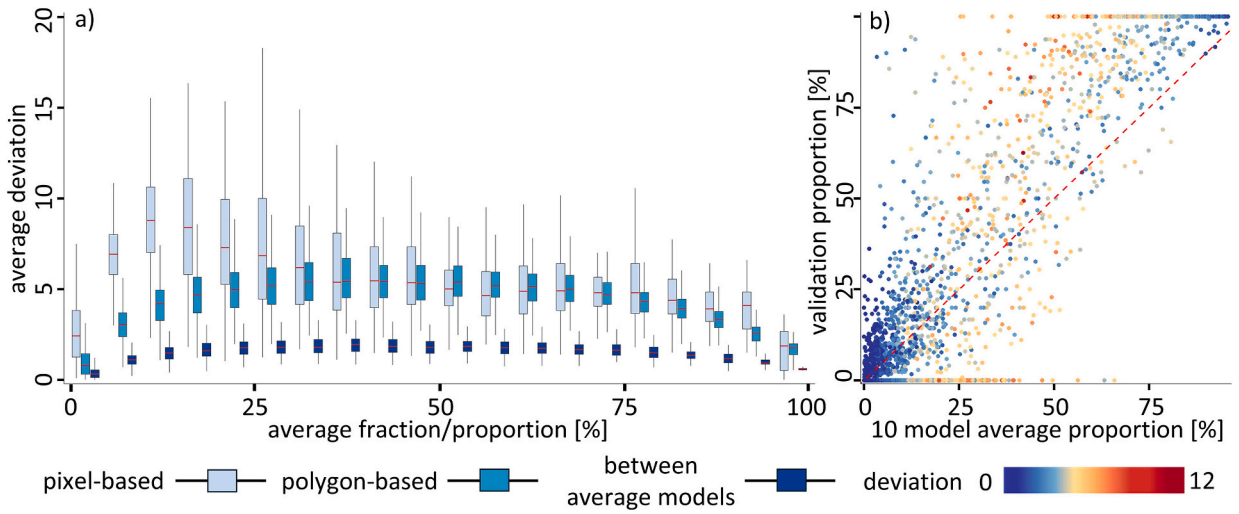

Fig. 9. Distribution of model's deviation and the correlation of deviation to the prediction error; *a)* Distribution of the average deviation dependent on the predicted fraction by an averaged model for individual pixels and within validation polygons, and the deviation between multiple averaged models; *b)* Distribution of the deviation within polygons for predicted average proportions of all tree species against the validation proportion.

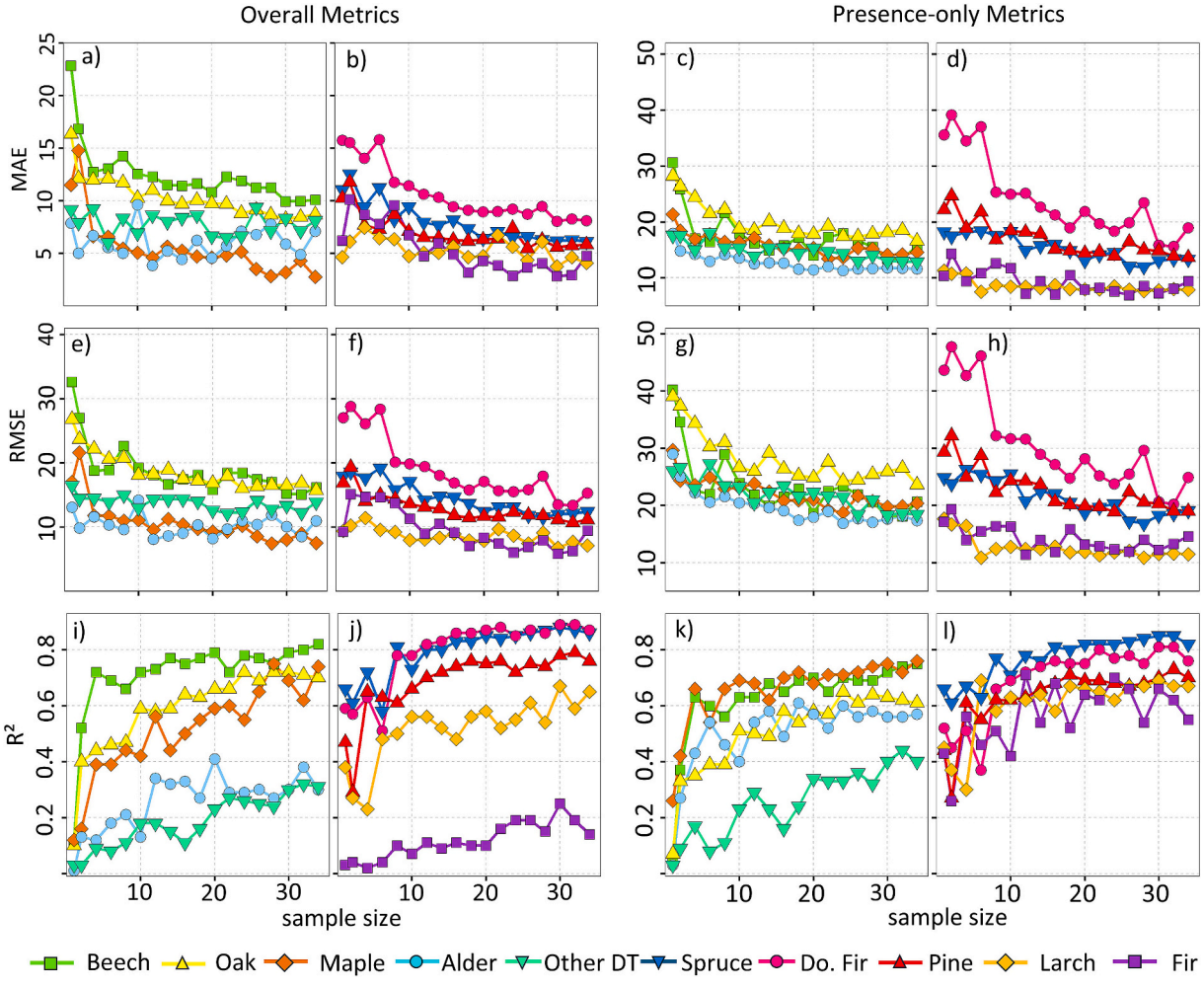

Fig. 10. Development of model's error statistic with increasing sample size for each tree species class, ground and shadow; (*a-d*): MAE, (*e-h*): RMSE, (*i-l*): R^2. Two validation sets are shown, i.e., using all validation polygons (overall metrics – column one and two); and using only validation polygons of species' presence (presence only metrics – column three and four). 614 Deciduous (columns one and three) and coniferous (columns two and four) species are shown separately for better differentiation.

be amplified in this study due to specific properties of reference data. Near the edge of polygons, parts of the neighboring stand may extend into the validation polygon, which are not accounted for in the ground truth dataset. For all polygons with species borders, this effect may lead to a proportional reduction of the predicted present tree species. Parts of the stand may be excluded by the shape and neglected for validation, or parts of neighboring stands are included by the polygon and reduce the proportion of tree species of validation data.

In order to validate the resulting species proportions, the predicted area of ground and shadow were excluded, thereby reflecting the methodology of reference data collection of sampling crown area and thus maps the proportions of tree species' crowns in the polygon. However, the occurrence of shadow is particularly prevalent in coniferous stands (particularly Spruce, Douglas Fir and Fir) due to the pronounced crown structure of these trees, which promotes the formation of shadows (Lopatin et al., 2019), leading to a partial prediction of shadows on crowns instead of the ground. One potential solution to this issue could be to utilize 3D data, such as airborne laser scanning data, to effectively differentiate between shaded crowns and shaded ground. In instances where the crown is shaded, the predicted shadow fraction could be apportioned between the existing predicted tree species or the predicted dominant tree species. This could also lead to an improvement in the underestimation of coniferous species fractions.

In general, error metrics support the hypothesis that different tree species fractions can be predicted at stand level. Fir, however, appears to be an exception as it is characterized by substantially poorer metrics, especially regarding the overall situation, and indicates an error-prone fraction prediction. One reason for this may be the aforementioned low number of available validation polygons with Fir presence (72) compared to 2549 polygons of Fir absence (proportion of only 2.7%). No pure stands of Fir were part of the validation dataset as all of these needed to be included in the training dataset (Table 1). In addition, the distribution of the validation polygons is highly clustered: 42 of the polygons (58%) are located within one ecoregion, while 16 additional polygons (22%) are all located in another ecoregion. Additionally, the distribution of the training samples for this tree species was not as representative as with the other species. While the samples could be distributed across nine ecoregions, the samples within these could only be collected in a very clustered manner. Nevertheless, the scatterplot and error metrics of presence-only performance suggest that differentiation from other tree species is possible, in principle, with possibly better training and validation conditions.

While Spruce performed well in the polygon-based assessment, the overall predicted area of Spruce was seemingly underestimated in comparison to the state-wide NFI assessment. However, the reference area of Spruce requires careful consideration due to the acquisition time of NFI data, which is subject to potential bias from excessive, and temporally very dynamic forest disturbances caused by bark beetle (*Ips typographus*) since 2018 (Knutzen et al., 2025; Senf and Seidl, 2021).

The forest loss mask applied to our approach in coniferous stands encompasses an area of 673.94 km^2, predominantly comprising Spruce stands with bark beetle damage from 2018 to 2023. Consequently, it can be assumed that parts of the NFI-based Spruce area estimation were already affected by bark beetle calamity during the investigation period of this study. With this in mind, the R^2 between our prediction and the state-wide NFI assessment increases from 0.89 to 0.95 when excluding Spruce.

The validation of this study is based on stand-based forest planning data. Validation data for land-use fraction maps are often generated at a pixel level (Suess et al., 2018), or use conglomerates of a few connected pixels (e.g. Schug et al., 2020), and are often based on expert reference fraction estimation from high-resolution digital orthophoto (Okujeni et al., 2018; Okujeni et al., 2013). A pixel-based validation based on orthophotos is unfeasible, as even for experts, distinguishing nine different tree species and a rest class in high-resolution RBG-IR data is a challenging, and error-prone task to achieve. While introducing its own challenges and limitations, polygon-based validation is also still a common strategy for evaluating fraction maps (Cooper et al., 2020; Okujeni et al., 2021). Consequently, area or stand-based validations represent the optimal available option, but some assumptions and consequences for interpretation need to be considered. The validation of forest stands species' proportions does not have the same meaning and does not necessarily imply similar accuracies at pixel level. The use of NFI plots may be a future option for a small scale assessment, although it cannot provide a validation at the pixel-level as well due to NFI plot sizes that vary angle count sampling dependent (Hill et al., 2018). Drawbacks for validation attempts using NFI data are already described, like e.g. GPS location accuracy, or the estimation of tree species shares in mixed pixels (Blickensdörfer et al., 2024; Bolyn et al., 2022; Freudenberg et al., 2024), in addition to challenges related to infer a bird-eye view-compatible crown area from the individual trees in the plots. Nevertheless, it would be valuable to develop sub-pixel validation strategies point-based NFI data in the future, especially given that the intercomparison of our fraction estimates with state-wide area statistics from NFI are in good agreement.

A benefit of the employed stand-based forest planning data is the actuality that it can provide. We were able to use the most current data from the past years, i.e. from 2017 to 2022. Updates of this dataset are performed annually as new forest stands are surveyed. The surveying follows a ten-year repeat cycle, meaning each forest stand is revisited and updated once every ten years. A potential annual implementation would allow for the generation of a customized current validation data set for each year of prediction, by using the sampled stands of e.g. the last five years.

It is important to note that the used stand-based forest planning database was not generated for use in the field of remote sensing. The information on forest stand areas is based on expert estimation in the field, which represents a subjective estimation. Hence, the validation should not be overinterpreted and rather be seen as a meaningful intercomparison to prove the validity. Despite the assumption of high data precision, it might be possible that prediction results might yield more precise predictions than the estimations in the database. This, however, cannot be answered conclusively, and would require a tailored, and very laborious new reference dataset, which presents an interesting topic for future studies.

In traditional SMA, the estimation accuracy is negatively impacted by high correlation and high similarity between endmember spectra (Somers et al., 2011). This limitation persists when a regression-based approach with synthetically mixed training data augmentation is employed (Okujeni et al., 2013). The application of Sentinel-2 data allows for the implementation of a higher spectral resolution compared to Landsat data. Furthermore, the time series reconstruction method employed with a high temporal resolution enables the targeting of phenological characteristics and the differentiation of the selected tree species. This information has already led to the generation of improved results for tree species classifications (Grabska et al., 2019; Hill et al., 2010; Kollert et al., 2021). The most crucial phenological phases are budburst in spring and leaf senescence (Hill et al., 2010; Klosterman and Richardson, 2017). Both phases can be targeted with the applied reconstruction method, which extends from March to November, even if image availability is limited in a given year. Based on the findings of the study, this has resulted in sufficiently robust differentiation for the nine selected tree species.

5.2. Ensemble approach and model variability

The use of an ensemble approach has already been demonstrated using synthetic training datasets and it has been proven to enhance the quality of results (Cooper et al., 2020; Okujeni et al., 2021; Okujeni et al., 2018). We have shown that a notable reduction in the deviation between prediction approaches can be achieved, which also contributes to the robustness of the resulting fraction prediction. However, the

implementation of a ten-model ensemble is computationally intensive. In the context of operational deployment, one potential solution for this would be to implement a model convergence functionality like it is already implemented in FORCE when using ensemble predictions, wherein additional models are only added to the ensemble if the previous models are in insufficient agreement. This is performed at the pixel level and has the potential to substantially reduce processing time. Another approach would be to employ a single model with greater capacity, which also could offer significant advantages in terms of computational efficiency and streamlined performance, as it could directly predict mean and variance without requiring ensemble aggregation. Future work could experiment with such models, incorporating training techniques like dropout or regularization to enhance robustness and reduce variability (Srivastava et al., 2014).

As we found a high correlation between MAD and prediction errors, the product of the deviation can be effectively utilized in operational contexts to convey a value of precision to forest managers. In forest areas exhibiting a high prediction deviation for a specific tree species, the predicted fraction may be error prone, necessitating the implementation of additional sampling procedures on site.

5.3. Operationality of the unmixing approach for forest management institutions

The development of the demonstrated method was driven by the idea to design a workflow that may be deployed operationally by forest management authorities. In future work, we deem it necessary to investigate the application of this method to a larger area, e.g., on the national scale. In principle, if enough training samples are available representing different species in different climatic conditions and phenological developments, we hypothesize that our method can be successfully transferred to larger areas. An alternative approach could be to expand the method by generating regionally tuned models in a similar fashion as Hermosilla et al. (2022). In this case, the relatively small number of training samples is advantageous, as it still ensures operability. It is likely that the number of required samples could be reduced per regional model, as a lower intraclass variability is expected.

Tree species classification applications require a much larger sample size per tree species than presented here, which is especially problematic for rare tree species (Blickensdörfer et al., 2024; Grabska et al., 2019; Hemmerling et al., 2021). This can more easily be circumvented by using the synthetic data augmentation method used here - although it cannot be completely prevented (as discussed with Fir). The number of training samples per class is still a crucial factor, but a notable improvement is achieved relatively early with already few training samples that are magnitudes below the requirements of classification architectures (Fig. 9). Furthermore, it is possible to iteratively assess whether a larger number of pure training samples may be necessary (as done in Fig. 9).

Available reference datasets, i.e. forest inventory data, are usually not generated with remote sensing requirements in mind (Kangas et al., 2018). However, given the relatively small number of required samples, the development of an independent remote-sensing-tailored survey of well-suited pure populations distributed over the study area would be a viable and realistic option. Such a database could potentially be implemented by the forest administration to boost operational suitability even further (e.g., by specifically adding pure locations of underrepresented species), especially given that single sufficiently large patches of a pure species would suffice. In principle, it would also be possible to include other rare tree species as individual class, such as birch, although their inclusion in the model needs to be verified with regards to separability between the other classes. If such a database would be implemented, we suggest checking them on an annual basis with DOPs to prevent disturbances and diseases within the trainings samples. Again, this is supported by the small sample size required.

It is further possible to transfer the methodology to other sites, as the Sentinel-2 data and the Copernicus forest mask (Copernicus - Land Monitoring Service, 2024) for Europe are open data. Global or local alternatives for a forest mask need to be considered on different continents. Furthermore, the FORCE framework (Frantz, 2019), with the implemented interpolation method (https://github.com/davidfrantz/force-udf/tree/main/rstats/ts/spline-reconstruction), and the complete workflow used in this study (https://github.com/davidklehr/tree-species-unmixing) are also freely available. It is necessary to identify training samples of pure tree species, potentially through the utilization of local forest surveys or an independent set of temporally consistent pure training samples as previously described. The necessity for high quality validation data may be somewhat challenging, depending on the availability of data on species distribution from forest institutions.

5.4. Outlook

Building upon this study, some key avenues for future research and development are planned. First, the approach will be extended to a national scale, enabling a broader evaluation of its applicability and operational feasibility across diverse regions. This expansion will allow for the integration of the methodology into large-scale forest monitoring systems, enhancing its relevance for national forestry management. Additionally, the use of the NFI as a reference dataset is a promising direction for future work. The NFI offers both opportunities and challenges, as it can provide more detailed and comprehensive forest information. Integrating this data could lead to the prediction of a wider range of tree species, potentially yielding more stable and reliable results and enabling the method to be implemented nationwide. Finally, further refinement of the model architecture is planned. To advance the current ensemble approach, the development of a more complex and robust single model architecture is expected to reduce computational costs while maintaining or improving predictive accuracy. These future directions aim to strengthen the operational capacity of the approach, broaden its applicability, and enhance its robustness for large-scale, real-world forestry applications.

6. Conclusion

The overreaching goal of this study was to assess the potential of mapping sub-pixel tree species mixtures in mixed temperate forests, given the size-limitations of the available reference information. The findings demonstrate that neural network regression in combination with synthetically mixed training data is an effective approach for predicting sub-pixel tree species fractions of nine distinct species and one ‘other species’ class, along with a background and shadow class. By using a densely reconstructed time series of ten Sentinel-2 bands, phenological and spectral characteristics between tree species are targeted directly for differentiation. The resulting tree species fraction maps were validated with tree species proportions in existing mixed forest stand polygons and provide information regarding the composition of tree species in these stands, which can be complex in temperate mixed forests. The maps can provide valuable information for the management, development, and protection of forests. A key advantage of this approach is its minimal requisite number of pure training pixels, in comparison to conventional classification methods. This makes it an easily accessible solution for operational use in forest monitoring and management. The application of an ensemble approach can reduce inter-model deviation, and the resulting per-species deviation maps can provide valuable insights into prediction certainty due to the correlation between prediction error and model deviation.

Funding

This work was supported by the Federal Ministry of Digital and Transport of the Federal Republic of Germany under contract number

FKZ 50EW2408C as part of the ForestPulse project.

CRediT authorship contribution statement

David Klehr: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Johannes Stoffels:** Writing – review & editing, Supervision, Conceptualization. **Andreas Hill:** Writing – review & editing. **Vu-Dong Pham:** Writing – review & editing, Software, Methodology. **Sebastian van der Linden:** Writing – review & editing. **David Frantz:** Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We thank the European Space Agency and the European Commission for freely and openly sharing Sentinel-2 imagery and the high-resolution layer of the forest mask. We thank State Office for Surveying and Geographic Information Rhineland-Palatinate for providing the digital orthophotos and the digital elevation model. We thank Jari Mahler for many helpful ideas and assistance. We thank the four anonymous reviewers for their constructive feedback.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2025.114740.

Data availability

Data can be found in an online repository at https://zenodo.org/records/13756960.

The time series reconstruction FORCE-UDF can be found at https://github.com/davidfrantz/force-udf/tree/main/rstats/ts/spline-reconstruction.

The python Unmixing workflow can be found at https://github.com/davidklehr/tree-species-unmixing.

References

Axelsson, A., Lindberg, E., Reese, H., Olsson, H., 2021. Tree species classification using Sentinel-2 imagery and Bayesian inference. Int. J. Appl. Earth Obs. Geoinf. 100. https://doi.org/10.1016/j.jag.2021.102318.

Bao, S.G., Wang, W.J., Liu, Z., Zhang, H.K., Wang, L., Ma, J., Sun, H., Ba, S., Wang, Y., He, H.S., 2024. Revealing post-megafire spectral and compositional recovery in the Siberian boreal forest using Landsat time series and regression-based unmixing approach. Remote Sens. Environ. 311, 114307. https://doi.org/10.1016/j.rse.2024.114307.

Blickensdörfer, L., Oehmichen, K., Pflugmacher, D., Kleinschmit, B., Hostert, P., 2024. National tree species mapping using Sentinel-1/2 time series and German National Forest Inventory data. Remote Sens. Environ. 304. https://doi.org/10.1016/j.rse.2024.114069.

Bolton, D.K., Gray, J.M., Melaas, E.K., Moon, M., Eklundh, L., Friedl, M.A., 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. Remote Sens. Environ. 240, 111685. https://doi.org/10.1016/j.rse.2020.111685.

Bolyn, C., Michez, A., Gaucher, P., Lejeune, P., Bonnet, S., 2018. Forest mapping and species composition using supervised per pixel classification of Sentinel-2 imagery. Biotechnol. Agron. Soc. Environ. 172–187. https://doi.org/10.25518/1780-4507.16524.

Bolyn, C., Lejeune, P., Michez, A., Latte, N., 2022. Mapping tree species proportions from satellite imagery using spectral–spatial deep learning. Remote Sens. Environ. 280, 113205. https://doi.org/10.1016/j.rse.2022.113205.

Borges, J., Higginbottom, T.P., Cain, B., Gadiye, D.E., Kisingo, A., Jones, M., Symeonakis, E., 2022. Landsat time series reveal forest loss and woody encroachment in the Ngorongoro conservation area, Tanzania. Remote Sens Ecol. Conserv. 8, 808–826. https://doi.org/10.1002/rse2.277.

Boyd, D.S., Danson, F.M., 2005. Satellite remote sensing of forest resources: three decades of research development. Progress Phys. Geogr.: Earth Environ. 29, 1–26. https://doi.org/10.1191/0309133305pp432ra.

Buchner, J., Yin, H., Frantz, D., Kuemmerle, T., Askerov, E., Bakuradze, T., Bleyhl, B., Elizbarashvili, N., Komarova, A., Lewińska, K.E., Rizayeva, A., Sayadyan, H., Tan, B., Tepanosyan, G., Zazanashvili, N., Radeloff, V.C., 2020. Land-cover change in the Caucasus Mountains since 1987 based on the topographic correction of multi-temporal Landsat composites. Remote Sens. Environ. 248. https://doi.org/10.1016/j.rse.2020.111967.

Competence Center for Climate Change Impacts - Rhineland-Palatinate. Witterungsrückblick. https://www.klimawandel-rlp.de/de/daten-und-fakten/witterungsrueckblick/. (Accessed 28 February 2024).

Cooper, S., Okujeni, A., Jänicke, C., Clark, M., van der Linden, S., Hostert, P., 2020. Disentangling fractional vegetation cover: regression-based unmixing of simulated spaceborne imaging spectroscopy data. Remote Sens. Environ. 246. https://doi.org/10.1016/j.rse.2020.111856.

Copernicus - Land Monitoring Service, 2024. High Resolution Layer Forest Type. https://land.copernicus.eu/en/products/high-resolution-layer-forest-type. (Accessed 10 June 2024).

Dutrieux, R., Féret, J.-B., Ose, K., 2021. Mise au point d'une méthode reproductible pour le suivi généralisé des dégâts de scolytes par télédétection satellitaire. In: Office national des forêts, Fontainebleau, France (Ed.), Rendez-vous techniques de l'ONF - n°69-70 télédétection satellitaire.

European Environment Agency, 2007. European Forest Types, Categories and Types for Sustainable Forest Management Reporting and Policy Technical Report No. 9. EEA, Copenhagen, p. 114. https://www.eea.europa.eu/publications/technical_report_2006_9. (Accessed 30 July 2024).

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classification from remotely sensed data. Remote Sens. Environ. 186. https://doi.org/10.1016/j.rse.2016.08.013.

Fassnacht, F.E., White, J.C., Wulder, M.A., Næsset, E., 2024. Remote sensing in forestry: current challenges, considerations and directions. Forestry: Int. J. Forest Res. 97, 11–37. https://doi.org/10.1093/forestry/cpad024.

Federal Ministry of Food and Agriculture, 2014. Der Wald in Deutschland: Ausgewählte Ergebnisse der dritten Bundeswaldinventur. 56. https://www.bmel.de/DE/themen/wald/wald-in-deutschland/bundeswaldinventur.html. (Accessed 28 February 2024).

Federal Ministry of Food and Agriculture, 2022. Bundeswaldinventur: Vermessung des Waldes: Vierte Bundeswaldinventur. https://www.bundeswaldinventur.de/vierte-bundeswaldinventur-2022/vermessung-des-waldes. (Accessed 30 January 2025).

Franklin, S.E., 2001. Remote Sensing for Sustainable Forest Management. CRC Press, Boca Raton, Fla.

Frantz, D., 2019. FORCE—Landsat + Sentinel-2 analysis ready data and beyond. Remote Sens. 11, 1124. https://doi.org/10.3390/rs11091124.

Frantz, D., Roder, A., Stellmes, M., Hill, J., 2016a. An operational radiometric Landsat preprocessing framework for large-area time series applications. IEEE Trans. Geosci. Remote Sensing 54, 3928–3943. https://doi.org/10.1109/TGRS.2016.2530856.

Frantz, D., Stellmes, M., Roder, A., Udelhoven, T., Mader, S., Hill, J., 2016b. Improving the spatial resolution of land surface phenology by fusing medium- and coarse-resolution inputs. IEEE Trans. Geosci. Remote Sensing 54, 4153–4164. https://doi.org/10.1109/TGRS.2016.2537929.

Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the Fmask algorithm for Sentinel-2 images: separating clouds from bright surfaces based on parallax effects. Remote Sens. Environ. 215, 471–481. https://doi.org/10.1016/j.rse.2018.04.046.

Freudenberg, M., Schnell, S., Magdon, P., 2024. A Sentinel-2 machine learning dataset for tree species classification in Germany. Earth Syst. Sci. Data. https://doi.org/10.5194/essd-2024-206.

Gauer, J., Wolff, B., Holzhausen, M., 2005. Waldökologische Naturräume Deutschlands. Mitteilungen des Vereins für forstliche Standortskunde und Forstpflanzenzüchtung.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest stand species mapping using the Sentinel-2 time series. Remote Sens. 11, 1197. https://doi.org/10.3390/rs11101197.

Grabska, E., Frantz, D., Ostapowicz, K., 2020. Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the polish Carpathians. Remote Sens. Environ. 251, 112103. https://doi.org/10.1016/j.rse.2020.112103.

Grabska-Szwagrzyk, E., Tiede, D., Sudmanns, M., Kozak, J., 2024. Map of forest tree species for Poland based on Sentinel-2 data. Earth Syst. Sci. Data 16, 2877–2891. https://doi.org/10.5194/essd-16-2877-2024.

Guerschman, J.P., Scarth, P.F., McVicar, T.R., Renzullo, L.J., Malthus, T.J., Stewart, J.B., Rickards, J.E., Trevithick, R., 2015. Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. Remote Sens. Environ. 161, 12–26. https://doi.org/10.1016/j.rse.2015.01.021.

Hemmerling, J., Pflugmacher, D., Hostert, P., 2021. Mapping temperate forest tree species using dense Sentinel-2 time series. Remote Sens. Environ. 267. https://doi.org/10.1016/j.rse.2021.112743.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sens. Environ. 158, 220–234. https://doi.org/10.1016/j.rse.2014.11.005.

Hermosilla, T., Bastyr, A., Coops, N.C., White, J.C., Wulder, M.A., 2022. Mapping the presence and distribution of tree species in Canada's forested ecosystems. Remote Sens. Environ. 282. https://doi.org/10.1016/j.rse.2022.113276.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Bater, C.W., Hobart, G.W., 2024. Characterizing long-term tree species dynamics in Canada's forested ecosystems using annual time series remote sensing data. For. Ecol. Manag. 572, 122313. https://doi.org/10.1016/j.foreco.2024.122313.

Hill, R.A., Wilson, A.K., George, M., Hinsley, S.A., 2010. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. Appl. Veg. Sci. 13, 86–99. https://doi.org/10.1111/j.1654-109X.2009.01053.x.

Hill, A., Buddenbaum, H., Mandallaz, D., 2018. Combining canopy height and tree species map information for large-scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes. Eur. J. Forest Res. 137, 489–505. https://doi.org/10.1007/s10342-018-1118-z.

Immitzer, M., Atzberger, C., Koukal, T., 2012. Tree species classification with random Forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sens. 4, 2661–2693. https://doi.org/10.3390/rs4092661.

Immitzer, M., Neuwirth, M., Böck, S., Brenner, H., Vuolo, F., Atzberger, C., 2019. Optimal input features for tree species classification in Central Europe based on multi-temporal Sentinel-2 data. Remote Sens. 11, 2599. https://doi.org/10.3390/rs11222599.

Kangas, A., Astrup, R., Breidenbach, J., Fridman, J., Gobakken, T., Korhonen, K.T., Maltamo, M., Nilsson, M., Nord-Larsen, T., Næsset, E., Olsson, H., 2018. Remote sensing and forest inventories in Nordic countries – roadmap for the future. Scand. J. For. Res. 33. https://doi.org/10.1080/02827581.2017.1416666.

Klosterman, S., Richardson, A.D., 2017. Observing spring and fall phenology in a deciduous Forest with aerial drone imagery. Sensors (Basel, Switzerland) 17. https://doi.org/10.3390/s17122852.

Knutzen, F., Averbeck, P., Barrasso, C., Bouwer, L.M., Gardiner, B., Grünzweig, J.M., Hänel, S., Haustein, K., Johannessen, M.R., Kollet, S., Müller, M.M., Pietikäinen, J.-P., Pietras-Couffignal, K., Pinto, J.G., Rechid, D., Rousi, E., Russo, A., Suarez-Gutierrez, L., Veit, S., Wendler, J., Xoplaki, E., Gliksman, D., 2025. Impacts on and damage to European forests from the 2018–2022 heat and drought events. Nat. Hazards Earth Syst. Sci. 25, 77–117. https://doi.org/10.5194/nhess-25-77-2025.

Kobayashi, S., Sanga-Ngoie, K., 2008. The integrated radiometric correction of optical remote sensing imageries. Int. J. Remote Sens. 29, 5957–5985. https://doi.org/10.1080/01431160701881889.

Kollert, A., Bremer, M., Löw, M., Rutzinger, M., 2021. Exploring the potential of land surface phenology and seasonal cloud free composites of one year of Sentinel-2 imagery for tree species mapping in a mountainous region. Int. J. Appl. Earth Obs. Geoinf. 94. https://doi.org/10.1016/j.jag.2020.102208.

Kowalski, K., Okujeni, A., Brell, M., Hostert, P., 2022. Quantifying drought effects in central European grasslands through regression-based unmixing of intra-annual Sentinel-2 time series. Remote Sens. Environ. 268, 112781. https://doi.org/10.1016/j.rse.2021.112781.

Landesforsten, 2012. Der Wald in Rheinland-Pfalz: Ergebnisse der Bundeswaldinventur, p. 2012.

Lechner, M., Dostálová, A., Hollaus, M., Atzberger, C., Immitzer, M., 2022. Combination of Sentinel-1 and Sentinel-2 data for tree species classification in a central European biosphere reserve. Remote Sens. 14. https://doi.org/10.3390/rs14112687.

Lopatin, J., Dolos, K., Kattenborn, T., Fassnacht, F.E., 2019. How canopy shadow affects invasive plant species classification in high spatial resolution remote sensing. Remote Sens Ecol. Conserv. 5, 302–317. https://doi.org/10.1002/rse2.109.

Nasiri, V., Beloiu, M., Asghar Darvishsefat, A., Griess, V.C., Maftei, C., Waser, L.T., 2023. Mapping tree species composition in a Caspian temperate mixed forest based on spectral-temporal metrics and machine learning. Int. J. Appl. Earth Obs. Geoinf. 116, 103154. https://doi.org/10.1016/j.jag.2022.103154.

Okujeni, A., van der Linden, S., Tits, L., Somers, B., Hostert, P., 2013. Support vector regression and synthetically mixed training data for quantifying urban land cover. Remote Sens. Environ. 137, 184–197. https://doi.org/10.1016/j.rse.2013.06.007.

Okujeni, A., Canters, F., Cooper, S.D., Degerickx, J., Heiden, U., Hostert, P., Priem, F., Roberts, D.A., Somers, B., van der Linden, S., 2018. Generalizing machine learning regression models using multi-site spectral libraries for mapping vegetation-impervious-soil fractions across multiple cities. Remote Sens. Environ. 216, 482–496. https://doi.org/10.1016/j.rse.2018.07.011.

Okujeni, A., Jänicke, C., Cooper, S., Frantz, D., Hostert, P., Clark, M., Segl, K., van der Linden, S., 2021. Multi-season unmixing of vegetation class fractions across diverse Californian ecoregions using simulated spaceborne imaging spectroscopy data. Remote Sens. Environ. 264, 112558. https://doi.org/10.1016/j.rse.2021.112558.

Pham, V.-D., Thiel, F., Frantz, D., Okujeni, A., Schug, F., van der Linden, S., 2024. Learning the variations in annual spectral-temporal metrics to enhance the transferability of regression models for land cover fraction monitoring. Remote Sens. Environ. 308. https://doi.org/10.1016/j.rse.2024.114206.

Python Software Foundation, 2022. 3.9.10 Documentation. https://docs.python.org/release/3.9.10/. (Accessed 28 May 2024).

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/. accessed 5 March 2024.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R.O., 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. Remote Sens. Environ. 65, 267–279. https://doi.org/10.1016/S0034-4257(98)00037-6.

Roy, D., Li, Z., Zhang, H., 2017. Adjustment of Sentinel-2 multi-spectral instrument (MSI) red-edge band reflectance to nadir BRDF adjusted reflectance (NBAR) and quantification of red-edge band BRDF effects. Remote Sens. 9, 1325. https://doi.org/10.3390/rs9121325.

Rufin, P., Frantz, D., Yan, L., Hostert, P., 2021. Operational Coregistration of the sentinel-2A/B image archive using multitemporal Landsat spectral averages. IEEE Geosci. Remote Sensing Lett. 18, 712–716. https://doi.org/10.1109/LGRS.2020.2982245.

Schug, F., Okujeni, A., Hauer, J., Hostert, P., Nielsen, J.Ø., van der Linden, S., 2018. Mapping patterns of urban development in Ouagadougou, Burkina Faso, using machine learning regression modeling with bi-seasonal Landsat time series. Remote Sens. Environ. 210, 217–228. https://doi.org/10.1016/j.rse.2018.03.022.

Schug, F., Frantz, D., Okujeni, A., van der Linden, S., Hostert, P., 2020. Mapping urban-rural gradients of settlements and vegetation at national scale using Sentinel-2 spectral-temporal metrics and regression-based unmixing with synthetic training data. Remote Sens. Environ. 246, 111810. https://doi.org/10.1016/j.rse.2020.111810.

Schug, F., Pfoch, K.A., Pham, V.-D., van der Linden, S., Okujeni, A., Frantz, D., Radeloff, V.C., 2024. Land cover fraction mapping across global biomes with Landsat data, spatially generalized regression models and spectral-temporal metrics. Remote Sens. Environ. 311, 114260. https://doi.org/10.1016/j.rse.2024.114260.

Schwieder, M., Leitão, P.J., Da Cunha Bustamante, M.M., Ferreira, L.G., Rabe, A., Hostert, P., 2016. Mapping Brazilian savanna vegetation gradients with Landsat time series. Int. J. Appl. Earth Obs. Geoinf. 52, 361–370. https://doi.org/10.1016/j.jag.2016.06.019.

Schwieder, M., Leitão, P.J., Pinto, J.R.R., Teixeira, A.M.C., Pedroni, F., Sanchez, M., Bustamante, M.M., Hostert, P., 2018. Landsat phenological metrics and their relation to aboveground carbon in the Brazilian savanna. Carbon Balance Manag. 13, 7. https://doi.org/10.1186/s13021-018-0097-1.

Senf, C., Seidl, R., 2021. Mapping the forest disturbance regimes of Europe. Nat Sustain 4, 63–70. https://doi.org/10.1038/s41893-020-00609-y.

Senf, C., Laštovička, J., Okujeni, A., Heurich, M., van der Linden, S., 2020. A generalized regression-based unmixing model for mapping forest cover fractions throughout three decades of Landsat data. Remote Sens. Environ. 240, 111691. https://doi.org/10.1016/j.rse.2020.111691.

Shang, C., Coops, N.C., Wulder, M.A., White, J.C., Hermosilla, T., 2020. Update and spatial extension of strategic forest inventories using time series remote sensing and modeling. Int. J. Appl. Earth Obs. Geoinf. 84, 101956. https://doi.org/10.1016/j.jag.2019.101956.

Somers, B., Asner, G.P., 2014. Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: wavelength adaptive spectral mixture analysis. Int. J. Appl. Earth Obs. Geoinf. 31, 57–66. https://doi.org/10.1016/j.jag.2014.02.006.

Somers, B., Asner, G.P., Tits, L., Coppin, P., 2011. Endmember variability in spectral mixture analysis: a review. Remote Sens. Environ. 115, 1603–1616. https://doi.org/10.1016/j.rse.2011.03.003.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 1929–1958.

Stanimirova, R., Graesser, J., Olofsson, P., Friedl, M.A., 2022. Widespread changes in 21st century vegetation cover in Argentina, Paraguay, and Uruguay. Remote Sens. Environ. 282, 113277. https://doi.org/10.1016/j.rse.2022.113277.

State Office for Surveying and Geographic Information Rhineland-Palatinate, 2024. Open Data. https://lvermgeo.rlp.de/geodaten-geoshop/open-data. (Accessed 12 August 2024).

Stoffels, J., Hill, J., Sachtleber, T., Mader, S., Buddenbaum, H., Stern, O., Langshausen, J., Dietz, J., Ontrup, G., 2015. Satellite-based derivation of high-resolution Forest information layers for operational Forest management. Forests 6, 1982–2013. https://doi.org/10.3390/f6061982.

Stoffels, J., Milles, A., Faust, J., Frantz, D., 2024. Monitoring von Borkenkäferschäden mit Satellitenbildzeitreihen. In: Nationalpark Hunsrück-Hochwald, B. (Ed.), Die Dynamik im Fokus - 10 Jahre Nationalpark Hunsrück Hochwald: die wissenschaftliche Reise geht weiter, pp. 222–231.

Suess, S., van der Linden, S., Okujeni, A., Griffiths, P., Leitão, P.J., Schwieder, M., Hostert, P., 2018. Characterizing 32 years of shrub cover dynamics in southern Portugal using annual Landsat composites and machine learning regression modeling. Remote Sens. Environ. 219, 353–364. https://doi.org/10.1016/j.rse.2018.10.004.

TensorFlow, 2023. API Documentation - TensorFlow v2.10.1. https://www.tensorflow.org/versions/r2.10/api_docs. (Accessed 28 May 2024).

Thompson, S.D., Nelson, T.A., White, J.C., Wulder, M.A., 2015. Mapping dominant tree species over large forested areas using Landsat best-available-pixel image composites. Can. J. Remote. Sens. 41, 203–218. https://doi.org/10.1080/07038992.2015.1065708.

Thünen-Institute, 2024. Bundeswaldinventur Ergebnisdatenbank: Vierte Bundeswaldinventur. https://bwi.info/. (Accessed 11 October 2024).

Welle, T., Aschenbrenner, L., Kuonath, K., Kirmaier, S., Franke, J., 2022. Mapping dominant tree species of German forests. Remote Sens. 14, 3330. https://doi.org/10.3390/rs14143330.

White, J.C., Wulder, M.A., Hobart, G.W., Luther, J.E., Hermosilla, T., Griffiths, P., Coops, N.C., Hall, R.J., Hostert, P., Dyk, A., Guindon, L., 2014. Pixel-based image compositing for large-area dense time series applications and science. Can. J. Remote. Sens. 40, 192–212. https://doi.org/10.1080/07038992.2014.945827.

White, J.C., Coops, N.C., Wulder, M.A., Vastaranta, M., Hilker, T., Tompalski, P., 2016. Remote sensing technologies for enhancing forest inventories: a review. Can. J. Remote. Sens. 42, 619–641. https://doi.org/10.1080/07038992.2016.1207484.

Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. 79–82.

Wulder, M.A., Roy, D.P., Radeloff, V.C., Loveland, T.R., Anderson, M.C., Johnson, D.M., Healey, S., Zhu, Z., Scambos, T.A., Pahlevan, N., Hansen, M., Gorelick, N., Crawford, C.J., Masek, J.G., Hermosilla, T., White, J.C., Belward, A.S., Schaaf, C.,

Woodcock, C.E., Huntington, J.L., Lymburner, L., Hostert, P., Gao, F., Lyapustin, A., Pekel, J.-F., Strobl, P., Cook, B.D., 2022. Fifty years of Landsat science and impacts. Remote Sens. Environ. 280, 113195. https://doi.org/10.1016/j.rse.2022.113195.

Yan, L., Roy, D.P., Li, Z., Zhang, H.K., Huang, H., 2018. Sentinel-2A multi-temporal misregistration characterization and an orbit-based sub-pixel registration methodology. Remote Sens. Environ. 215, 495–506. https://doi.org/10.1016/j.rse.2018.04.021.

Yin, H., Khamzina, A., Pflugmacher, D., Martius, C., 2017. Forest cover mapping in post-soviet Central Asia using multi-resolution remote sensing imagery. Sci. Rep. 7, 1375. https://doi.org/10.1038/s41598-017-01582-x.

Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. Remote Sens. Environ. 237. https://doi.org/10.1016/j.rse.2019.111511.

Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D., 2017. mixup: Beyond Empirical Risk Minimization, p. 13. http://arxiv.org/pdf/1710.09412v2.