

Data homogeneity impact in tree species classification based on Sentinel-2 multitemporal data case study in central Sweden

Giovanni D'Amico, Mats Nilsson, Arvid Axelsson & Gherardo Chirici

To cite this article: Giovanni D'Amico, Mats Nilsson, Arvid Axelsson & Gherardo Chirici (2024) Data homogeneity impact in tree species classification based on Sentinel-2 multitemporal data case study in central Sweden, International Journal of Remote Sensing, 45:15, 5050-5075, DOI: [10.1080/01431161.2024.2371082](https://doi.org/10.1080/01431161.2024.2371082)

To link to this article: <https://doi.org/10.1080/01431161.2024.2371082>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 10 Jul 2024.



[Submit your article to this journal](#)



Article views: 211



[View related articles](#)



[View Crossmark data](#)

Data homogeneity impact in tree species classification based on Sentinel-2 multitemporal data case study in central Sweden

Giovanni D'Amico^{a,b}, Mats Nilsson^{id}^c, Arvid Axelsson^{id}^c and Gherardo Chirici^{a,d}

^aDepartment of Agriculture, Food, Environment and Forestry, University of Florence, Florence, Italy; ^bCREA Research Centre for Forestry and Wood, Arezzo, Italy; ^cDepartment of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden; ^dFondazione per il Futuro delle Città, Florence, Italy

ABSTRACT



Spatial information on forest composition is invaluable for achieving scientific, ecological, and management objectives and for monitoring multiple changes in forest ecosystems. The increased flow of optical satellite data provides new opportunities to improve tree species mapping. However, the accuracy of such maps is affected by training data, and in particular on the homogeneity of individual classes. Thus, we evaluated the effect of data homogeneity in tree species classification. We performed tree species classification by considering different ways to partition data into tree species classes. The class sets considered were (i) only mixed coniferous and mixed deciduous forest classes, (ii) single-species classes, (iii) single-species, mixed coniferous and mixed deciduous classes, and (iv) single-species, mixed coniferous and mixed deciduous classes and a true mixed class. Using data from the Swedish National Forest Inventory, we varied the threshold that defined dominating species. Tree species were classified for a study area in central Sweden using Sentinel-2 data and two classification approaches: Bayesian inference and random forest (RF). Images were selected by class separability and the most informative images based on variable selection with RF. The most informative images tended to be selected by both methods. However, in forests with tree species of similar spectral behaviour, image selection on the basis of class separability was found to be more reliable. More accurate classification results were achieved as the number of classes decreased and the threshold of plot purity increased. The Bayesian classification approach of only mixed coniferous and mixed deciduous classes gave the highest OA, always greater than 90%. When discriminating between pure plots of Birch (*Betula spp.*), Spruce (*Picea abies*), Scots pine (*Pinus sylvestris*) and Lodgepole pine (*Pinus contorta*), the best OA values were 84% for Bayesian and 80% for RF. In more complicated scenarios, RF resulted in higher overall accuracies (OA).


ARTICLE HISTORY

Received 9 June 2023
Accepted 14 June 2024

KEYWORDS

Bayesian inference; tree species; classification; Sentinel-2; NFI

CONTACT Mats Nilsson  mats.nilsson@slu.se  Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå 901 83, Sweden

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/01431161.2024.2371082>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

1. Introduction

Sustainable forest management and conservation require up-to-date and accurate tree species maps. Having reliable tree species information is critically valuable for a variety of scientific and operational forestry implementations, such as planning (Chiarucci and Piovesan 2020), biodiversity monitoring (Parisi et al. 2023; Wallis et al. 2017; Wang and Gamon 2019), climate change effect in adaptation and mitigation (Hof, Dymond, and Mladenoff 2017; Wazen et al. 2018), and assessments of terrestrial carbon stocks and fluxes (Vangi et al. 2023; Xiao et al. 2019). In addition, the essential role of sustainable forest management and forests is mentioned in the Sustainable Development Goals (SDGs) set by the United Nations in the 2030 Agenda specifically of SDG 15 (life on land) (Gregersen et al. 2017).

National Forest Inventories (NFIs) generally provide aggregated data such as precise estimates of tree species proportions on regional to national levels, but they lack detailed spatially explicit information required for the monitoring and management of natural resources. Moreover, large forest areas, mainly in mountainous regions, are often remote and difficult to access (Young et al. 2017), necessitating time-consuming and costly NFI surveys. In contrast, remote sensing technology can easily obtain full-coverage forest information over large areas even in dense and inaccessible forests. As a result, remote sensing-based approaches that provide spatialized and updatable information on forest cover distribution are increasing in demand (Kangas et al. 2018; Lister et al. 2020).

The constant progress in remote sensing technologies supports the development of species recognition and classification methods. These data are usually gathered from an airborne or spaceborne platform due to the opportunity to easily cover large areas. Primarily data from multispectral (e.g. Axelsson et al. 2021; Hemmerling, Pflugmacher, and Hostert 2021; Immitzer et al. 2019; Sameen, Pradhan, and Aziz 2018), hyperspectral (e.g. Fricker et al. 2019; Vangi et al. 2021), LiDAR (Light Detection and ranging) data (e.g. Lindberg, Holmgren, and Olsson 2021; Michałowska and Rapiński 2021; Waser, Ginzler, and Rehush 2017), or a combination of these are used (e.g. Mäyrä et al. 2021; Plakman et al. 2020). However, large area tree species mapping is primarily still based on multi-spectral satellite data (Breidenbach et al. 2020).

The increasing availability of open-access optical satellite data provides new opportunities for the use of multi-temporal methodologies, under the assumption that changes in reflectance caused by the phenological cycle may increase spectral separability among tree species (Persson, Lindberg, and Reese 2018; Sheeren et al. 2016). In particular, the short revisit period of Sentinel-2 (i.e. 2–3 days in Sweden) allows monitoring of seasonal variations with high temporal resolution and consequently observing phenological differences between tree species, improving vegetation classification results (Fassnacht et al. 2016; Immitzer et al. 2019). However, in large-area land cover analyses with multi-temporal imagery, several critical issues may occur. Images are affected by different atmospheric conditions, such as haze, clouds and cloud shadows. These can affect the separability of tree species or even cover parts of the image, causing data to be missing. The illumination conditions, that is greatly affected by sun angle, also cause shadows to appear different between images. Combined with topography, whole areas can be affected. Moreover, class separability is also strongly influenced by reference data and their homogeneity. In particular, on the one hand,

the prevalence of few forest species allows their consistent spectral behaviours to be identified. On the other hand, the widespread diffusion of mixed forests results in the complexity to identify distinctive spectral signatures.

When a large number of images are available for classification, an appropriate selection among them can improve the classification accuracy. Modern classification approaches are able to use the vast hyperdimensional space created by several bands of all the available images. In particular, machine learning approaches, such as convolutional neural networks, support vector machines, or random forest (RF), have shown accurate classification results even for computationally complex procedures. However, even a traditional classification approach, such as the maximum likelihood classification used in a Bayesian framework, has provided promising results that need further investigation (Axelsson et al. 2021; Cardille and Fortin 2016; Crowley et al. 2019).

The objective of this study was to evaluate the effect of data homogeneity in tree species classification by testing and evaluating two approaches. Bayesian and RF classification approaches were applied to the Swedish boreal forest using Sentinel-2 data. The effect of data homogeneity was evaluated by dividing field data into classes according to four different schemes with increasing complexity. Initially, only coniferous and deciduous forest classes were considered, then single-species classes, and finally both single-species and mixed classes were used.

2. Materials

2.1. Study area

The study area is located in mid-Sweden (62° 37' 18.66" N, 17° 2' 40.63" E) and is defined by two adjacent Sentinel-2 tiles (33VWK and 33XWK) for a total of approximately 100 km × 200 km of which about 14,000 km² on land (Figure 1). The study area contains various types of landscapes, both coastal and inland areas dominated by forest or agriculture. The maximum altitude is 577 m a.s.l. and annual precipitation ranges from 500 to 700 mm. The vegetation period normally begins in mid-May and ends in mid-October. The forest is mainly managed for wood production and dominated by coniferous stands of Norway spruce (*Picea abies* Karst.) with the second most frequent species being Scots pine (*Pinus sylvestris* L.), but Lodgepole pine (*Pinus contorta* Dougl. ex Loud.) and areas with broadleaf forests (mainly Birch, *Betula* spp.) are also present.

2.2. NFI field plots

The Swedish NFI is carried out as two independent annual field inventories covering all of Sweden (Fridman et al. 2014). Each field sample consists either of temporary or permanent plots, all of which are located in rectangular shaped clusters with side lengths of 300 m to 1500 m. Five geographic strata are used with a sampling intensity that decreases towards the north of the country. On both temporary and permanent plots, trees with a stem diameter less than 4 cm are measured on two 0.5 m radius plots, and trees with a stem diameter between 4 cm and 10 cm are measured on a 3.5 m radius plot. If the stem diameter is 10 cm or more, the trees are measured on plots with a 7 m or 10 m radius for temporary and permanent plots, respectively.

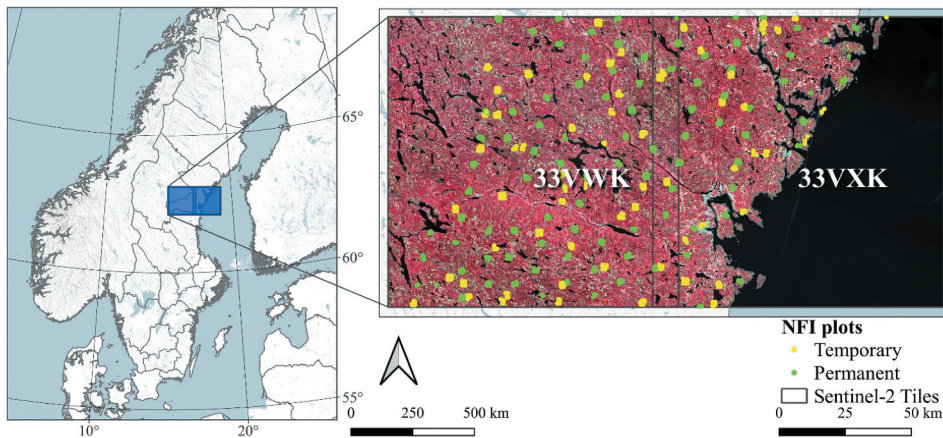


Figure 1. Mid-Sweden study area defined by the extent of two contiguous Sentinel-2 tiles, and the locations of the NFI plots.

In total, over 10,000 plots are surveyed annually by the NFI in the whole Sweden. A sample of 985 NFI plots was selected within the study area. Of these, 469 were temporary plots with a 7 m radius and 516 were permanent plots with a 10 m radius. Each tree that was measured was also positioned relative to the plot centre. Many forest variables were measured or estimated for each plot, including basal-area weighted mean tree height, basal-area weighted mean stem diameter, basal area, stem volume, and above ground tree biomass. The coordinates for the centre of the plots are positioned using a global navigation satellite system (GNSS) with a horizontal accuracy of about 1 m for most of the permanent plots and a horizontal accuracy of about 5 m for most of the temporary plots (Nilsson et al. 2017; Persson et al. 2017). Field plot coordinates were transformed from SWEREF 99TM (EPSG 3006) to WGS84/UTM zone 33 N (EPSG 32,633) since the latter was used in the remote sensing data.

Dominating tree species or species group was determined for each field plot depending on basal area (BA) proportion. If a species or group of them (Birch and other Broadleaf, Norway Spruce, Scots pine, and Lodgepole pine) constituted a larger BA proportion than a certain threshold, it was assigned as dominant of the field plot. When no species' basal area proportion was higher than a certain threshold, the plot was assigned to 1. Mixed Coniferous, when the sum of the BA proportions of conifers exceeded the threshold; 2. Mixed Broadleaf, when the sum of the BA proportions of the broadleaves exceeded the threshold, and 3. Mixed otherwise (Table 1). The basal area proportion thresholds used in this study to assess the influence of plot homogeneity in classification were 70%, 80%, 90%, and 100%.

The study was carried out considering four classification schemes: class set 1 consisting of single species, mixed domination and no domination (Birch and other Broadleaf, Norway Spruce, Scots pine, Lodgepole pine, Mixed Conifer, Mixed Broadleaf, Mixed), class set 2 consisting of single species and mixed domination (Birch and other Broadleaf, Norway Spruce, Scots pine, Lodgepole pine, Mixed Conifer, Mixed Broadleaf), class set 3 consisting of single species only (Birch and other Broadleaf, Norway Spruce,

Table 1. Number of plots by class at different proportions of basal area. Values within parentheses are the number of original plots and the other values are plots available after selection.

Prop%	Birch	Lodgepole P	Scots P	NSpruce	Mixed Broad	Mixed Conif	Mixed
70	(36) 24	(46) 20	(273) 133	(319) 182	(19) 13	(138) 74	(154) 63
80	(27) 12	(42) 14	(235) 104	(249) 128	(11) 5	(183) 101	(238) 96
90	(17) 9	(36) 13	(170) 96	(177) 94	(5) 3	(216) 108	(364) 156
100	(8) 4	(30) 14	(118) 53	(99) 49	(4) 3	(178) 99	(548) 347

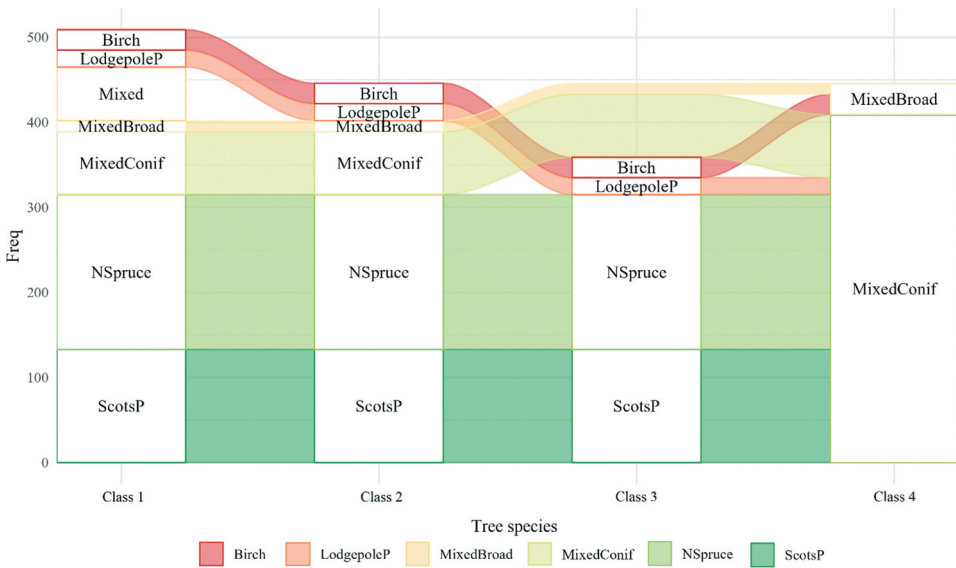


Figure 2. Plot frequency by tree species in the four class sets.

Scots pine, Lodgepole pine), and class set 4 consisting of mixed domination only (Broadleaf, Conifer). The partitioning of field plots into classes is visualized in Figure 2.

2.3. Remote-sensing data

In this study, imagery from the Sentinel-2 (S2) mission was used for tree species classification. The two S2 satellites feature a wide-swath width of 290 km, a Multi-Spectral Imaging (MSI) sensor with 13 high spectral resolution bands, and a varying spatial resolution of 10 m, 20 m, and 60 m depending on band (Drusch et al. 2012). The study area is covered by two S2 tiles, each with a fixed size of 100 km × 100 km (Figure 1). All images with the processing level of L2A (i.e. atmospherically corrected surface reflectance images) from the period of 2018 to 2019, with a cloud cover of less than 15%, were downloaded from ESA Copernicus Open Access Hub. As the winter months are characterized by low sun angles, few daylight hours, and snow, only images acquired in the period between May and November of each year were used in the study. In total, 45 images, obtained by merging the 33VWK and 33VXK tiles (Figure 1), were collected during the period. Among these, 16 images, or 32%, were manually

identified as free of clouds and cloud shadows. Among the spectral bands available from the S2 sensor, we selected the nine visible and near-infrared bands with a spatial resolution of 10 m and 20 m (B2: Blue, B3: Green, B4: Red, B5: Red-Edge1, B6: Red-Edge2, B7: Red-Edge3, B8A: NIR, B11: SWIR1, B12: SWIR2). All 10 m bands were also provided at 20 m resampled by the average value (Pignatale 2022). Thus, to ensure homogeneous information quality for all bands investigated, 20 m resolution images were used. Thus, the spatial resolution of the NIR and SWIR bands, which are more sensitive in discriminating between vegetation classes, is preserved (Alonso, Picos, and Armesto 2021). Spectral data for each image were extracted for the NFI plots as a weighted average of pixels according to their proportion in the circular plot.

2.3.1. Remote sensing variables

Spectrally similar tree species in combination with a 20 m spatial resolution may reduce mapping accuracy in some areas with complex forest structures, often characterized by mixed and diverse forests. Adding vegetation indices as a complement to the spectral bands could better represent the tree coverage characteristics and thus improve map accuracy (Khan et al. 2023). Therefore, we increased the S2 variables and then selected the most representative ones as follows.

We combined the nine S2 bands in pairs, to calculate a set of 36 normalized differential indices n_c , mathematically defined as

$$n_c = \frac{S_i - S_j}{S_i + S_j}, \quad (1)$$

where $i \neq j$, correspond to c -combinations (S_i, S_j) of the set composed of S2 bands.

Then, the data of different S2 band combinations were standardized as

$$n_{st_c} = \frac{(n_c - \mu_c)}{\sigma_c}, \quad (2)$$

where μ_c and σ_c are the mean and standard deviation of the c th S2 indices n_c (D'Amico et al. 2021b).

3. Methods

The procedures herein presented were performed in the statistical software R 4.2.2 (<https://www.r-project.org>., accessed on: 1 May 2023).

3.1. Classification methods

3.1.1. Bayesian approach

Maximum Likelihood (ML) is a method that has commonly been used to classify pixels in satellite images based on spectral data. All pixels are assigned a class based on the relative likelihood of that pixel as part of the probability density function of each class (Hagner and Reese 2007). The algorithm provides a classification by using a decision rule based on an expression describing probabilities for different classes using conditional probabilities (Swain 1978).

In this study, we used the decision rule that k should be selected to maximize

$$\ln(P(\omega_k)) + \sum_{t=1}^n \ln\left(p\left(X_t, \vec{\theta}_{k,t}\right)\right), \quad (3)$$

where $P(\omega_k)$ is the prior probability of class k that was estimated by field sample proportions and $p\left(X_t, \vec{\theta}_{k,t}\right)$ is the likelihood of observing X at time t with the distribution parameters $\vec{\theta}_{k,t}$ (Axelsson et al. 2021). With this decision rule, we assumed that observations (i.e. satellite images) were conditionally independent given a tree species. The Bayesian classification was performed by fitting a normal distribution to field data, although any likelihood function can be used (Gorte and Stein 1998).

3.1.2. Random forest

For comparison, the classification results were compared with those obtained with the RF algorithm. This method was taken as a reference for benchmarking due to extensive and well-documented applications within forestry spatial prediction using remotely sensed data (Chirici et al. 2020). Random Forest is a decision tree algorithm and nowadays is among the most popular ensemble methods to predict forest variables (Belgiu and Drăguț 2016). In RF, a set of de-correlated regression trees (*ntree*) are generated. The results from these regression trees are aggregated to produce predictions without overfitting the data (Breiman 2001). To build and grow trees, RF uses a randomly chosen subset of predictors at each splitting node (*mtry*), and trees are grown without the need for pruning. RF uses a procedure called out-of-bag samples (OOB) to grow trees. In the OOB procedure, each tree of the maximum size is built independently, based on bootstrap samples from the training dataset (i.e. two-thirds of the data), leaving randomly the remaining one-third of the samples. Following the OOB sample procedure, the prediction error (OBB error) for each of the individual trees can be estimated as:

$$OOB_{error} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (4)$$

where \hat{y}_i is the predicted output of an OOB sample, y_i is the actual output and n is the total number of OOB sample units. In this study, we used random forest model growth through the ranger package (Wright, Wager, and Probst 2023), optimizing the model by searching for the combination of *ntree* and *mtry* that minimized the OOB error (Belgiu and Drăguț 2016; Chirici et al. 2020).

In comparison to other machine-learning techniques, RF may minimize output variance and the overfitting issue, improving model stability and accuracy (Breiman 2001). In order to compare the results obtained from the two classification methods, the reference conditions were uniformed by implementing the two classification methods with the same S2 images (section 3.2.1.).

3.2. Image and feature selection

To achieve a high classification accuracy, classification algorithms require remote sensing data of high quality. For the Bayesian classification, image selection to ensure the greatest separability between class pairs can be used (Axelsson et al. 2021). However, since image selection is the same as feature selection, a selection among images should be suited to

the classification method that is used. Therefore, classification results based on images selected by using Z-value and an RF-based feature selection approach were compared.

3.2.1. Z-values

The images were sorted according to the class separation level within them. For each pair of classes, a Mahalanobis distance (Mahalanobis 1936) of the difference between the class population averages was calculated as

$$Z = \left(\hat{\mu}_a - \hat{\mu}_b \right)^T \left(\frac{\hat{\Sigma}_a}{m_a} + \frac{\hat{\Sigma}_b}{m_b} \right)^{-1} \left(\hat{\mu}_a - \hat{\mu}_b \right), \quad (5)$$

where a and b represent the different classes, $\hat{\mu}$ is the sample mean vector, $\hat{\Sigma}$ is the sample variance-covariance matrix, and m is the sample size.

Based on the Z-values, 16 images were selected, corresponding to the cloud-free image proportion (section 2.3). To select these 16 images, two images with the largest Z-value for each class pair were selected. This resulted in different sets of images selected for different class sets. After this, additional images were chosen by the highest sum of Z-values over class pairs, to reach a total of 16 images.

3.3. Feature selection

3.3.1. Bayesian approach

For each image, we performed a recursive feature elimination (RFE) (Guyon et al. 2002) to choose a good feature subset for tree species classification among the S2 bands (9) and indices (36). The wrapper-type algorithm identifies good combinations of features by generating a set of classification models and iteratively removing features that do not improve classification accuracy. The RFE was conducted using a naïve Bayes algorithm (nb-RFE) as prediction model, specifically, the nbFuncs method from the caret R package (Kuhn 2008) was used. The feature elimination method employs backward selection, meaning that the RFE search process starts with the full feature set and proceeds to iteratively remove features that do not contribute to or are detrimental to classification accuracy, until a good combination of features is found (Ramezan 2022). A 10-fold cross-validation resampling method was employed in the iterative development of the RFE model, providing performance estimates that incorporate variation due to feature selection. Thus, the model with the highest accuracy provides a good subset of metrics for tree species classification.

3.3.2. Random forest

For the RF classification model, a feature selection was performed using the Variable Selection Using Random Forest (VSURF) R-package (Genuer, Poggi, and Tuleau-Malot 2015). The algorithm works in two steps: 1. variables are ranked according to importance using RF, and irrelevant ones are eliminated; 2. it performs feature selection to generate two separate subsets for interpretation and prediction purposes, respectively. The subset considered is the prediction subset, which includes the response-related variables. The data set containing S2 variables (bands and indices) from all available (45) images and the

information classes for each plot served as input to VSURF. After the variable selection, all images were classified using RF and ranked according to their OA.

3.4. NFI plot selection

To limit the influence of ground vegetation in young forests with limited canopy cover, plots with heights below 5 m were excluded from the analyses. In addition, plots affected by cuttings in the time period between the field inventory and the image acquisition were identified and removed based on the NDVI of the six S2 images with the best Z-values (Section 3.2.) (Table 1).

3.5. Accuracy assessment

Leave-one-out cross-validation was conducted to assess the accuracy of the classification models. The primary measure of accuracy was overall accuracy, which is defined as the number of correctly classified items divided by the total number of items. The accuracy of the Bayesian inference method was also evaluated by varying the number of images used for classification. More specifically, classification was made using the best available images that had been chosen by Z-values (Section 3.2). All the possible combinations were tested by progressively adding images from the 16 selected. Applying the classification method to a single image corresponds to ordinary ML classification.

4. Results

4.1. Bayesian approach

4.1.1. Image selection

For each class set (Figure 2) and BA threshold, we selected 16 images out of 45 for the classification according to their class separability (Z value). Two of the images were selected in all cases (30 May 2018 and 1 July 2018), and five images were selected in 80% of the cases (30 May 2018, 1 July 2018, 16 July 2018, 7 October, and 14 October 2018). When considering each class set, 7 out of the 16 selected images were the same for all BA thresholds in class 1 and 2, eight for class 3, and six for class 4 (Table 2). As shown in Table 2, the best images for discriminating between all classes (class set 1) are in 2018: 10, 25, and 30 May, 1 and 16 July, 7 and 14 October, and when excluding plots in mixed coniferous and deciduous forests (class set 2), the best images are 10, 25, and 30 May 2018, 1 July, 7 and 14 October and for 17 May 2019. When considering only single-species classes (class set 3), the best images are on 10 and 30 May 2018, 26 Jun, 1, 16, 24 July, 7 Octoberth, and 22nd, and when only using the two classes coniferous and broadleaf trees (class set 4), the best images are 30 May, 26 June, 1, 16, and 24 July, and 3 August 2018.

4.1.2. Feature selection

For each of the 16 selected images used, a naïve Bayes classification model was produced from the RFE method. Features were selected from the 36 indices and nine original S2 bands. Depending on the class set, a different number of features were selected. On

Table 2. The best images selected by class set and proportion of basal area based on Z-value.

Rank	Class 1			Class 2			Class 3			Class 4				
	70	80	90	70	80	90	70	80	90	70	80	90	100	
1	2018-05-25	2018-05-25	2018-07-04	2018-09-02	2018-07-01	2018-05-25	2018-07-04	2018-05-30	2018-07-01	2018-06-26	2018-07-16	2018-07-01	2018-06-26	2019-06-06
2	2018-07-01	2018-07-06	2019-09-04	2018-05-30	2018-06-24	2018-05-05	2018-07-01	2018-09-02	2018-07-16	2018-07-16	2018-06-26	2018-07-16	2018-07-16	2018-07-06
3	2018-10-22	2018-10-14	2018-10-07	2018-10-07	2018-10-07	2018-10-07	2019-06-11	2019-05-17	2019-08-28	2019-05-17	2018-05-25	2018-06-26	2018-07-11	2018-07-16
4	2018-07-24	2018-07-24	2018-10-14	2018-05-30	2018-05-05	2018-09-02	2019-09-04	2018-10-24	2018-08-03	2019-05-17	2018-05-10	2019-06-06	2019-06-16	2018-10-07
5	2018-05-30	2018-07-11	2018-07-11	2018-05-25	2018-07-01	2018-07-01	2019-09-19	2018-08-15	2018-09-19	2019-07-26	2018-05-05	2018-05-17	2018-07-16	2018-07-01
6	2018-10-19	2018-07-01	2018-06-26	2018-10-19	2018-07-16	2018-07-16	2018-08-03	2018-07-01	2018-05-10	2018-05-30	2018-09-19	2018-07-11	2018-09-02	2018-07-11
7	2018-10-14	2018-10-22	2018-10-19	2018-05-10	2018-07-24	2018-07-24	2018-05-30	2018-07-16	2018-10-07	2018-05-10	2018-10-07	2019-06-16	2018-08-03	2018-06-26
8	2018-06-24	2018-10-07	2018-05-10	2018-10-14	2019-05-17	2018-10-07	2018-10-14	2018-05-25	2018-10-14	2018-10-07	2019-06-16	2018-10-24	2018-07-01	2018-10-14
9	2018-10-07	2018-05-10	2018-05-30	2019-05-17	2018-05-10	2018-05-10	2019-06-06	2018-09-19	2019-05-17	2019-06-16	2018-07-24	2018-07-01	2019-06-06	2019-07-26
10	2018-05-05	2018-05-30	2018-07-16	2018-10-24	2018-10-14	2018-05-30	2018-06-26	2018-05-10	2018-10-19	2018-06-26	2018-07-16	2018-05-25	2018-05-25	2019-07-19
11	2018-07-16	2018-05-05	2018-07-01	2018-08-15	2018-10-22	2018-05-25	2018-10-14	2018-05-25	2018-10-14	2018-10-24	2018-10-24	2018-07-24	2018-07-24	2018-05-30
12	2019-05-17	2019-06-06	2019-06-11	2018-07-01	2018-08-03	2019-08-28	2019-06-16	2019-10-07	2018-07-01	2019-09-04	2018-10-22	2018-06-24	2018-10-22	2018-05-05
13	2018-05-10	2018-09-02	2019-09-19	2018-09-19	2018-09-19	2019-05-17	2018-10-07	2018-06-26	2018-07-24	2018-07-24	2018-07-24	2018-10-19	2018-05-30	2018-05-12
14	2018-08-03	2018-07-16	2018-08-03	2019-10-07	2018-05-25	2019-07-26	2018-05-10	2018-05-05	2018-07-16	2018-10-22	2018-07-01	2019-07-26	2018-06-04	2018-08-03
15	2018-09-19	2019-08-28	2018-06-06	2018-10-19	2018-10-19	2019-05-17	2018-10-07	2018-09-02	2018-10-19	2018-10-07	2018-10-14	2019-09-04	2018-10-14	2018-10-22
16	2018-09-02	2019-05-17	2018-05-25	2018-05-05	2018-07-04	2019-06-16	2019-07-26	2019-08-15	2018-07-04	2018-10-24	2018-07-11	2018-07-06	2018-10-24	2018-07-24

average, 28 metrics were selected for the classification of class set 1. For the class set 2, 28 features were selected on average, for the class set 3, 20 features were selected on average, and for class set 4, 15 features were selected on average. On average, the number of metrics decreased with decreasing number of target classes. In fact, when discriminating between coniferous and deciduous trees, with a basal area proportion threshold of 100%, only six metrics were selected (Figure 3).

Looking at the individual metrics that were selected, those most frequently chosen were: SWIR1, SWIR2, and RedEdge2, followed by the normalized differential indices of RedEdge1 and SWIR1, and Green and Red, followed by the original NIR and RedEdge3. As for metrics selected per class set, results are presented in Figure 4. Here, we see that SWIR1 appears to be the most frequently chosen feature for classification when looking at class sets that include single tree species dominated classes (class sets 1, 2, and 3), while for classification of deciduous and coniferous mixed classes (class set 4), NIR appears to be the most commonly used.

4.1.3. Classification accuracy

The OA increased as more images were used for classification, a pattern that held true for all class sets. Moreover, the highest OA values were observed as the complexity of class sets decreased and plot homogeneity (BA threshold) increased. When including plots with all single species classes and mixed classes (class set 1), OA values averaged 0.5. The highest OA for classification using class set 1 was 60%, and it was obtained with the combination of seven images and plots with a basal area threshold of 90% (Table 3). The average OA values using two images are the lowest, ranging from 31% to 36%, varying with the species homogeneity of the plots (Figure 5).

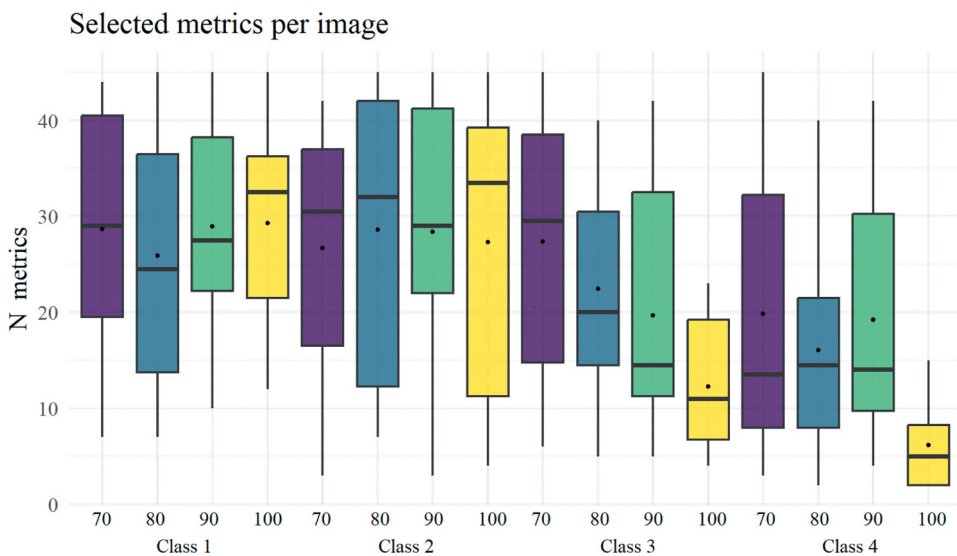


Figure 3. Boxplot of number of metrics selected in the 16 images per class set and basal area proportion threshold. Black dots represent the arithmetic mean while black center lines represent the median.

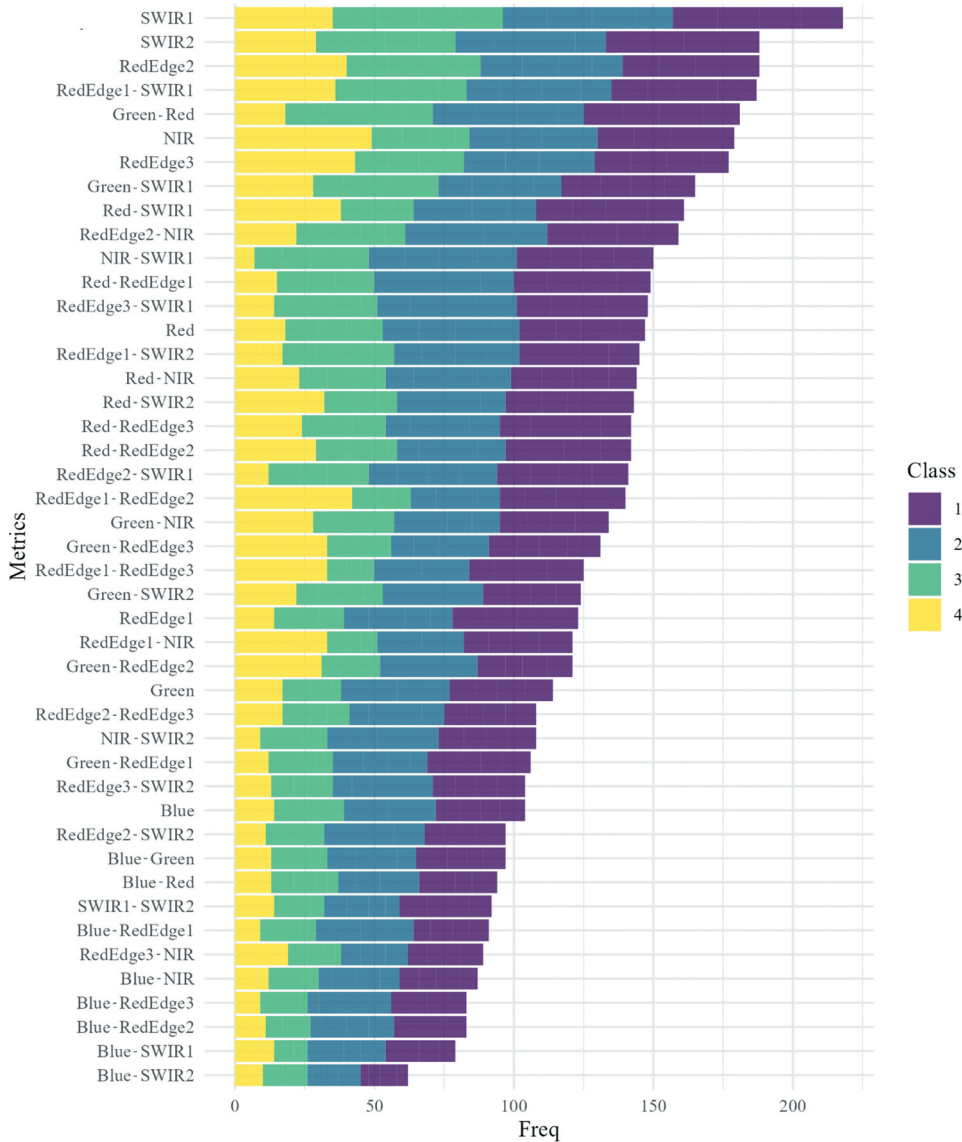


Figure 4. Number of times different metrics were selected in the 16 images per class set.

In terms of OA results slightly higher than class set 1 were obtained by removing the class with mixed forests from the analysis (class set 2). The OA values, using all 16 images, were 56% for both a basal area threshold of 70% and 100%. While the best OA values of 61% was obtained using a BA threshold equal to 100% with two combinations with both eight and nine images (Table 4). Using only two images, the average OA value is between 36% and 43%, depending on the species homogeneity of the plots (Figure 5).

A significant improvement in OA results was obtained by removing mixed classes (class set 3). With pure classes and combinations of at least 13 images, the average OA results 75%. The best OA values of 84% result with four combination images (considering 4, 6, 7

Table 3. Best confusion matrix (with OA and No. of images) per basal area proportion considering class set 1. Birch and other Broadleaf, Norway Spruce, Scots pine, Lodgepole pine, Mixed Conifer, Mixed Broadleaf, Mixed.

		Reference							
		Mixed	Mixed Broadleaf	Mixed Conifer	Birch	Lodgepole Pine	Scots Pine	Norway Spruce	
Prediction	70								
	Mixed	19	1	4	4	1	5	15	
	Mixed Broadleaf	14	9	0	5	0	2	6	
	Mixed Conifer	7	0	26	1	1	12	20	
	Birch	10	3	2	14	0	1	2	
	Lodgepole Pine	0	0	6	0	13	19	9	
	Scots Pine	5	0	14	0	3	77	17	
	Norway Spruce	8	0	22	0	2	17	113	
	OA = 53.24%; 7, and 9 images								
	80								
	Mixed	51	2	3	3	0	8	16	
	Mixed Broadleaf	6	0	0	1	0	1	0	
	Mixed Conifer	7	0	58	0	1	33	30	
	Birch	11	2	0	8	0	0	0	
	Lodgepole Pine	0	0	5	0	9	4	1	
	Scots Pine	9	1	17	0	3	51	10	
Norway Spruce	12	0	17	0	1	7	70		
OA = 54.14%; 8 images									
90									
Mixed	95	1	10	2	0	3	13		
Mixed Broadleaf	0	0	0	0	0	0	0		
Mixed Conifer	27	0	62	0	2	26	17		
Birch	12	2	0	7	0	0	0		
Lodgepole Pine	0	0	5	0	9	4	0		
Scots Pine	9	0	11	0	1	59	7		
Norway Spruce	13	0	20	0	1	4	57		
OA = 60.33%; 7 images									
100									
Mixed	199	1	6	1	0	2	3		
Mixed Broadleaf	2	0	0	2	0	0	0		
Mixed Conifer	45	0	48	0	2	9	11		
Birch	8	2	0	1	0	0	0		
Lodgepole Pine	7	0	7	0	9	2	0		
Scots Pine	36	0	15	0	3	37	5		
Norway Spruce	50	0	23	0	0	3	30		
OA = 57.64%; 6, and 7 images									

and 9 image combinations) and tree species pure plots (BA = 100%). With 80% and 90% G, the highest OA of 81% results with combinations of 5–6 and 7 images, respectively (Table 5). While using two images, the average OA varies between 49% and 62% when looking at different plot homogeneity (Figure 5).

When classifying only the two classes of coniferous and broadleaf species (class set 4), an average OA of 99% was achieved with at least 10 images used for classification and tree species pure plots (BA = 100%). With BA = 90%, the OA was 97% with at least 10 images, while with BA = 80%, the OA was 95% with at least 14 images. With BA = 70%, the average OA was 90% using at least 11 images. Considering plots with BA = 100%, the maximum OA achieved for any combination of images was more than 99%. The best combinations of 4–10 images with BA = 90% ensured an OA of 99%. With 10 images, the maximum OAs

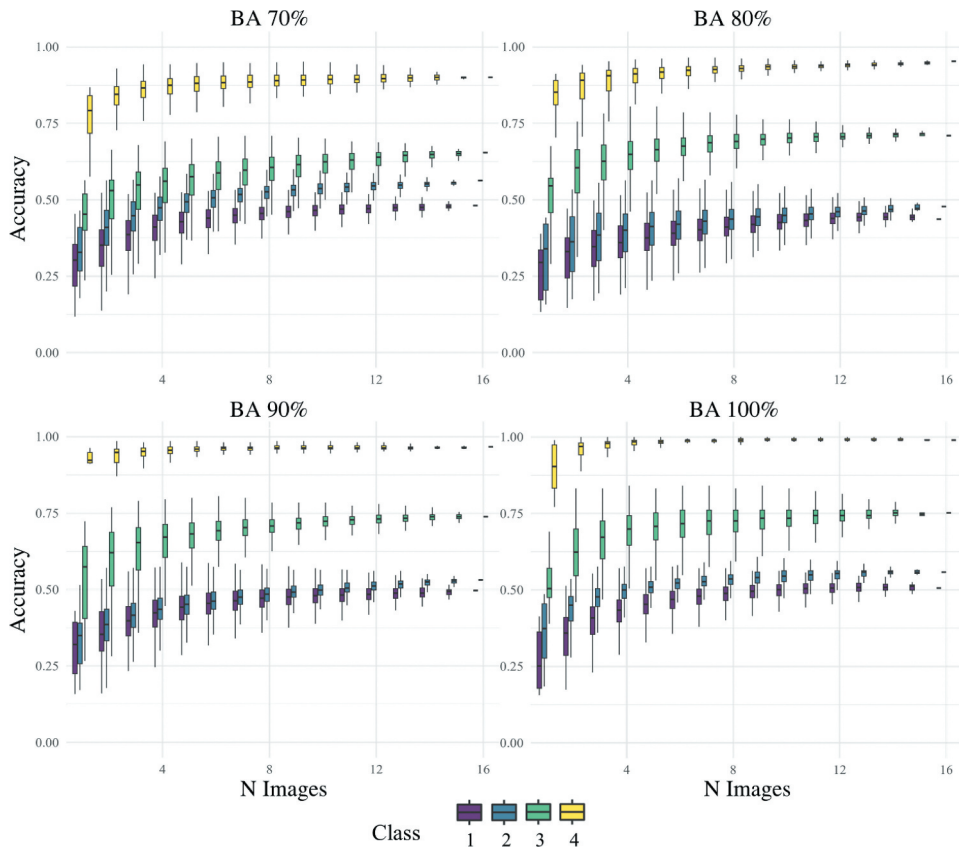


Figure 5. Overall accuracy boxplots considering every possible combination of images used for classification per class set and basal area proportion.

achieved were 97% for BA = 80% and 95% for BA = 70% (Table 6). The worst OA values stood at 49% using two images, while on average, the OAs using two images were above 84% (Figure 5).

4.2. Random forest

4.2.1. Image and variable selection

In the VSURF approach, an average of 12 images are selected for all class sets and different BAs. While considering each class set separately, the selected images are 13 for class set 1 (all tree species classes), 15 for class set 2 (excluding mixed coniferous and deciduous forests), 11 for class set 3 (pure forests only), and 8 for class set 4 in which conifers and broadleaf trees are discriminated. The 1 July 2018 image was always selected, the 14 October 2018 image was selected in 80% of the cases, while in more than 60% of cases were selected 2018: 30 May and 7, 19, and 22 October (Figure 6).

The VSURF algorithm mainly provides the most important variables for prediction. Analysing the frequency of variables by dates, it turns out that the images with the most variables were 2018: 14 October, and 1 July, followed by 30 May,

Table 4. Best confusion matrix (with OA and No. of images) per basal area proportion considering class set 2. Birch and other Broadleaf, Norway Spruce, Scots pine, Lodgepole pine, Mixed Conifer, Mixed Broadleaf.

		Reference					
		Mixed Broadleaf	Mixed Conifer	Birch	Lodgepole Pine	Scots Pine	Norway Spruce
Prediction	70						
	Mixed Broadleaf	6	2	5	0	1	6
	Mixed Conifer	0	25	1	1	15	20
	Birch	7	3	17	0	0	6
	Lodgepole Pine	0	8	0	15	17	6
	Scots Pine	0	16	0	3	81	24
	Norway Spruce	1	21	0	0	22	131
	OA = 60.00%; 9 images						
	80						
	Mixed Broadleaf	1	0	1	0	1	1
	Mixed Conifer	0	51	1	1	36	28
	Birch	4	1	12	0	2	4
	Lodgepole Pine	0	6	0	9	8	4
	Scots Pine	0	23	0	3	67	17
	Norway Spruce	1	25	0	1	9	95
	OA = 57.04%; 6, and 8 images						
	90						
	Mixed Broadleaf	0	0	0	0	0	0
	Mixed Conifer	0	58	0	1	26	18
	Birch	3	0	6	0	0	1
Lodgepole Pine	0	16	0	7	6	6	
Scots Pine	0	8	1	2	57	8	
Norway Spruce	0	20	0	2	3	54	
OA = 60.10%; 4, and 5 images							
100							
Mixed Broadleaf	0	0	3	0	0	0	
Mixed Conifer	0	52	0	0	13	10	
Birch	2	0	1	0	0	0	
Lodgepole Pine	0	2	0	12	2	0	
Scots Pine	0	14	0	2	36	4	
Norway Spruce	1	31	0	0	2	35	
OA = 61.26%; 8 images							

7 October, and 16 June 2019 (Figure 6). The most selected indices were the normalized differences between Green and Red, RedEdge2 and NIR, Green and Red, Red and SWIR1, and Green and SWIR1, while the most selected band was SWIR1.

4.3. Bayesian classifier versus random forest

The two classifiers produced comparable results in terms of best OA at varying BA proportion and class set, with limited differences (Figure 7). Specifically, RF tended

Table 5. Best confusion matrix (with OA and No. of images) per basal area proportion considering class set 3. Birch and other Broadleaf, Norway Spruce, Scots pine, Lodgepole pine.

		Reference				
		Birch	Lodgepole Pine	Scots Pine	Norway Spruce	
Prediction	70					
	Birch	17	0	3	6	
	Lodgepole Pine	1	10	20	17	
	Scots Pine	1	4	81	24	
	Norway Spruce	1	3	17	125	
	OA = 70.61%; 7, and 8 images					
	80					
	Birch	13	0	2	2	
	Lodgepole Pine	0	7	9	4	
	Scots Pine	0	3	87	17	
	Norway Spruce	0	1	11	106	
	OA = 81.30%; 5, and 6 images					
	90					
	Birch	10	0	0	3	
	Lodgepole Pine	0	3	4	1	
	Scots Pine	0	4	65	12	
Norway Spruce	1	3	9	80		
OA = 81.03%; 7 images						
100						
Birch	2	0	0	0		
Lodgepole Pine	0	8	4	1		
Scots Pine	0	2	42	3		
Norway Spruce	2	0	6	43		
OA = 84.07%; 4, 7, 8, and 9 images						

to achieve the greatest OA with mixed classes (class sets 1 and 2), while the Bayesian approach had comparable results for pure classes (class set 3) and was better in the classification of coniferous and broadleaf species (class set 4). Considering the proportions of BA, with tree species pure plot (BA = 100%), except for class set 2, the Bayesian method always achieved a greater OA than RF. Otherwise, with the lowest plot homogeneity considered (BA = 70%), the OA achieved by RF was greater for all class sets.

5. Discussion

The present study evaluated the use of Sentinel-2 time series for tree species classification, considering several progressively more homogeneous classes and comparing Bayesian inference and RF in a test area located in the central part of Sweden. The use of multi-temporal data provided greater accuracy than data from

Table 6. Best confusion matrix (with OA and No. of images) per basal area proportion considering class set 4. Broadleaf, Conifer.

		Reference	
		Mixed Broadleaf	Mixed Conifer
Prediction	70		
	Mixed Broadleaf	31	20
	Mixed Conifer	2	402
	OA = 95.16%; 9 images		
	80		
	Mixed Broadleaf	19	9
	Mixed Conifer	2	311
	OA = 96.77%; 6, 7, 9, and 11 images		
	90		
	Mixed Broadleaf	16	2
	Mixed Conifer	1	252
	OA = 98.89%; 5, and 6 images		
	100		
	Mixed Broadleaf	6	0
Mixed Conifer	0	191	
OA = 100%; 2-14 images			

a single date image, as seen in previous studies (Axelsson et al. 2021; Persson, Lindberg, and Reese 2018). It also offers the advantage of producing and integrating new data quickly and economically at various scales, thus facilitating decision-making (Hussain et al. 2023). In tree species mapping, environmental conditions, such as cloud cover and atmospheric and topographical effects are relevant issues. These problems are overcome i. partly by the short revisit time of twin Sentinel-2 satellites, ii. by the proposed method (Axelsson et al. 2021), in which the same images are not required for the entire study area.

The use of Z-value for image selection was confirmed to improve the overall accuracy of classification, identifying mainly cloud-free images with good radiometric quality. Indeed, the higher cloudiness in the study area during the year 2019 led to fewer images selected from that year. Overall, out of 34 images selected at least once, only 11 were acquired in 2019. Considering the four-class sets of class combinations and the selected images, we see that, for class sets 1 and 2, containing mixed classes, the most selected images were the same even with varying species homogeneity. Images from 5, 10, 25, and 30 May, 1 and 16 July, 7 and 14 October 2018 and 17 May 2019 were selected from at least three of the four class sets considered. Data from the month of May, with five images (four from 2018 and one from 2019), appear to be the most occurring and therefore the most effective in discriminating between mixed species. As reported by Persson et al. (2018), late spring is particularly effective for discriminating species because of the greater phenological variation among them due to differences between species in the leaf-out period.

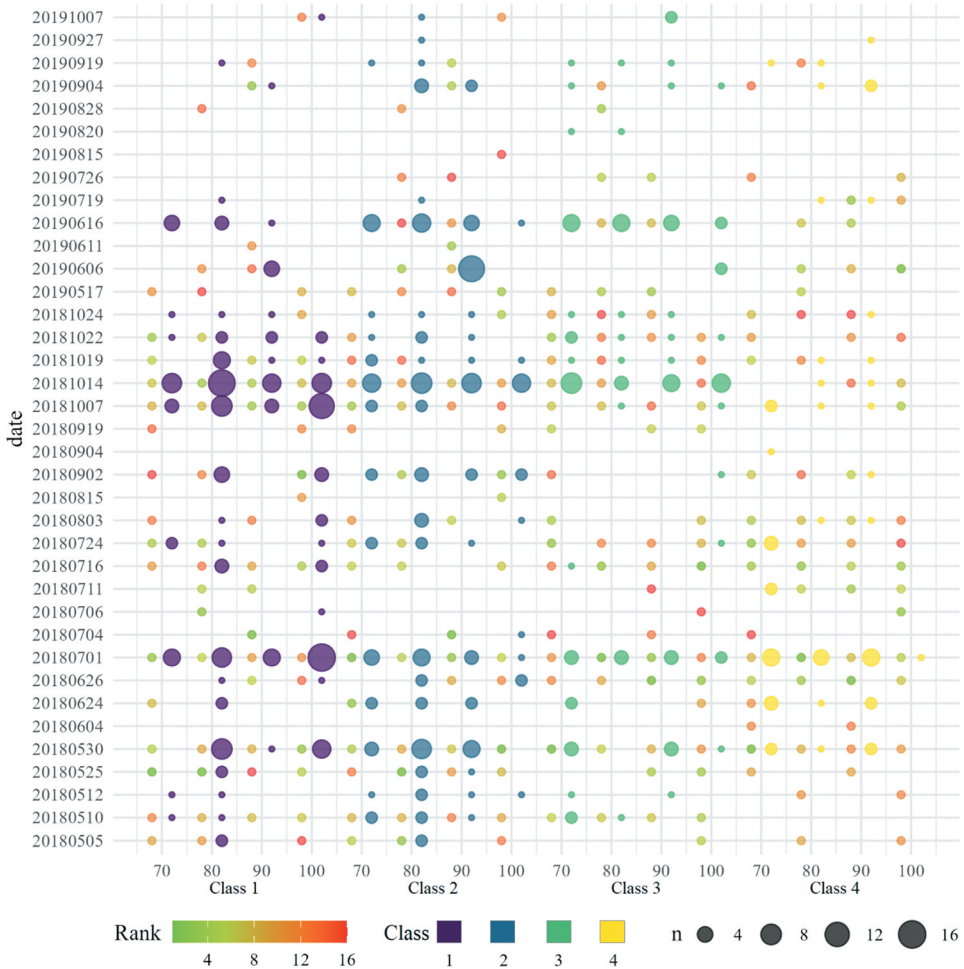


Figure 6. Comparison of image selection per class set and BA threshold with ranking (from Z-values, Table 2) and VSURF colored by class. The VSURF point size varies according to the number of indices (n) selected for each image.

Regarding Z-values, pure tree species classes (BA = 100%) generally showed higher values. The homogeneity of classes, in fact, results in reduced noise in predictors that promote greater discrimination among species. Class set 1 contains mixed forests, based on the Z-value, the period starting in May and ending in August was the best for discriminating mixed broadleaf forests, while for mixed coniferous forests, the period from May to October presented the best values.

Considering class set 2, the highest Z-values for discrimination between mixed coniferous and deciduous forests were obtained from images taken in the spring. Among the most difficult classes to discriminate are mixed deciduous and birch-dominated forests, with Z-values tending to be low, especially for a BA threshold lower than 90%, with the selected images distributed throughout the growing season. In class set 3, on the other hand, where mixed classes were excluded, 13 images were selected for at least three class sets out of four. The selected images mainly belonged to the months of May (5th and 30th

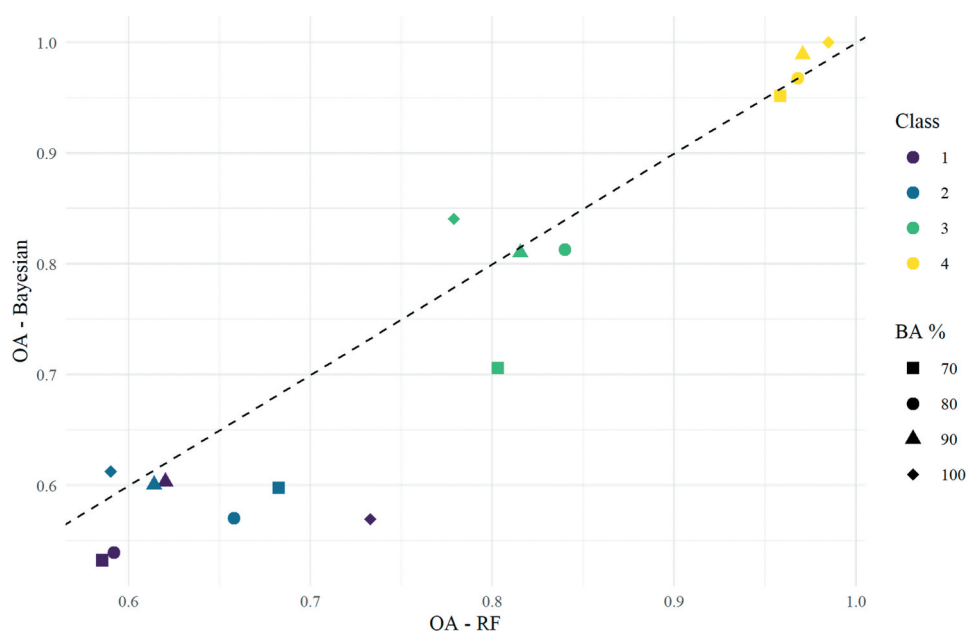


Figure 7. Scatterplot of the best overall accuracy of Bayesian versus RF approach.

–2018 and 17th –2019), July (1st, 16th, 24th –2018) and October (7th, 14th, 19th, 22nd, 24th –2018) (Table 2). Therefore, the best seasons for differentiating pure classes were autumn images from the period in which leaf colour change occurs, and spring images when leaf-out occurs. Considering Z-values, spring and autumn images provided a better basis to discriminate between spruce and the two pine species. In contrast, to distinguish between conifers and broadleaf trees (class set 4), the best images in at least three of the BA thresholds were acquired in the summer, with five images between late June and early August (2018: 26 June, 1, 16, and 24 July, and 3 August) and in the spring (2018: 30 May). This agreed with Lisein et al. (2015), who found that late spring and early summer images are optimal for species discrimination. Accordingly, higher Z-values were obtained between June and July for discriminating between coniferous and broadleaf trees, with the highest values recorded for greater BA thresholds.

The VSURF approach for image selection selected a total of 29 images at least once. Of these, three differed from the Z-value selection (2018: 4 September 2019: 20 August and 27 September). Of the 34 images selected with Z-value, eight were excluded from the VSURF selection (2018: 4 June, 15 August, 19 September, 2019: 17 May, 11 June, 26 July, 15 and 28 August).

The frequent presence of clouds in the study area led to the partial selection of cloud-covered images even with VSURF (11 images). However, cloud-free images were highly selected, with a median of eight images, compared to cloud-covered images (median of three images). Similarly to Z-values image selection, the most informative images with VSURF resulted in 1 July 2018 and 7 and 14 October.

Considering the VSURF image selection for class sets 1 and 2, containing mixed classes, the most selected images were mainly in late May (2018, 30th), July (2018, 1st), and October

(2018, 7th, 14th, and 22nd), with other images selected in September (2018, 2nd) and June (2019, 16th). In class set 3, with only pure classes, except for 1 July 2018 and 16 June 2019 images, the summer period appeared less informative, with four images selected in the second half of October (14th, 19th, 22nd, and 24th). In class set 4, with only mixed conifer and broadleaf classes, even with VSURF selection, the summer period is the most informative with images mainly selected in July, along with some images selected in October.

Although the images selected by the two methods considered generally ranged over the same periods of the year, some points should be noted. First, class separability allows the identification of images (i.e. period) in which, even for species with similar spectral properties, classes can best be distinguished from each other. Image selection with VSURF and other model-based classification methods ensures the collection of predictors that provide more robust models. However, the selection of images with higher spectral separability between classes allows more opportunity to generalize classifications when there are classes with similar spectral properties or unbalanced datasets (Su 2020). For instance, it was found that in half of the cases, the image with the highest Z-value ranking (Table 2) was not included in the VSURF selection.

For national or continental applications, cloud-based platform approaches offer significant opportunities allowing to process vast datasets without extensive local hardware. Particularly promising are the approaches based on pixel-based composite images where best available pixel (BAP) (White et al. 2014) and medoid (Flood 2013; Kennedy et al. 2018) are the most common (Francini et al. 2023).

Our results show that the use of multi-temporal images provided better results than single images. In addition, the OA improved as the mixed classes were excluded from the analysis and as the proportion of plot basal area where a plot was defined as homogeneous increased. Regarding individual species, Birch, Lodgepole pine, Scots pine, and Norway spruce were classified quite accurately with the selected combinations of images at different proportions of BA. However, given the spectral similarity between different conifer species, the classification of only the two classes coniferous and deciduous led to a major increase in accuracy.

These results demonstrate the impact in classification accuracy of using different basal area thresholds to assign plots to a class. Although pure tree species plots (BA = 100%) ensure the best results, given the homogeneity of the spectral response of the pixels, classes are usually assigned considering a lower BA threshold (Bravo-Oviedo et al. 2014). Nevertheless, if mixed classes are excluded (e.g. in class sets 3 and 4), the 80% BA threshold which is commonly used to discriminate between classes gives accuracies that are comparable with the ones obtained using higher BA thresholds. Particularly, limiting in terms of OA is the use of mixed classes. The extensive presence of conifers, already unbalanced in plot proportion (Table 1), has affected the frequent misclassifications. However, the effect of mixed classes on classification should be investigated in other environmental contexts characterized by the presence of other forest species (Oreti et al. 2021).

Comparing the results achieved by the Bayesian approach and the RF approach, chosen as the benchmark for machine learning given its simplicity of implementation and wide popularity, some aspects emerge. The OA generally increases for both the Bayesian approach and the RF approach as the homogeneity of the plot increases, independently of the classes considered and the BA threshold. However, it has been noted that under more challenging conditions such as in cases of low homogeneity in plots (BA = 70%) and with mixed classes (class sets 1 and 2), where spectral properties

tend to be mixed up between classes, RF achieves better results (Nasiri et al. 2023). On the other hand, under the conditions typically investigated, with both the limited presence of mixed classes and high BA thresholds to ensure spectral differences for training classification models, the Bayesian method was particularly effective.

Simultaneously, by decreasing the number of classes considered, the best results were obtained by selecting fewer S2 metrics. In the simplest situation, with only coniferous and broadleaf classes and plots characterized by 100% BA, only six metrics ensured the best results. However, to discriminate between more than two classes, a larger number of metrics, such as to provide greater spectral differentiation between classes, were selected (Vangi et al. 2021). Among the selected metrics, SWIR1, SWIR2, and NIR were the most frequent, followed by the normalized difference SWIR1/Red-Edge1, Red/Green, and NIR/Red-Edge3. In tree species mapping, the highest reflectance values occur mainly in the infrared region as compared to the visible region. On the other hand, canopy reflectance is known to be high in the near-infrared bands, while chlorophyll levels in leaves and their variations result in absorption features that are particularly sensitive in the red edge regions (Fassnacht et al. 2016; Immitzer et al. 2019; Puletti, Chianucci, and Castaldi 2018).

The critical metrics identified in this study predominantly originate from the bands provided at 20 m resolution. According to the findings reported by Alonso et al. (2021) in a study on forest classification in northern Spain, the use of S2 imagery with a 10-metre resolution yielded improved accuracy solely for land uses of anthropogenic origin. Nonetheless, with recent advancements in image downscaling techniques (Wu et al. 2023), and the capability to conduct analyses on cloud computing platforms that offer unrestricted access to comprehensive satellite data repositories, it might be advantageous to employ higher-resolution imagery even for studies focused on mapping tree species.

Among the classification algorithms available and used with remote sensing data, Bayesian inference has beneficial properties. Although any machine learning method can provide pixel-wise probabilities for each tree species in each image, to incorporate new available images into the classification results, implementation of ex-novo models is required. On the other hand, the Bayesian inference method used in this study makes it easy to update the classification with new images as they become available, representing an approach to automate classification from remotely sensed data. The Z-value provides information on the quality of the new images, selecting those to be used to improve the previous classification. Implementation of the presented approach might prove suitable as a method of continuous mapping of large area forests.

6. Conclusion

In this study, the data homogeneity impact in forest trees species classification was analysed. Based on Sentinel-2 images available in a central Sweden study area, we mapped forest species, considering four class sets of species aggregations and different single-species proportions of basal area as the definition of a homogenous plot. To do this, a Bayesian inference was sequentially applied accumulating the likelihoods to calculate the a posteriori probability from a stream of remote sensing images. With the aim of testing, the same images were then used with an RF model. Based on the combinations of the 16 best selected images, we calculated the classification overall accuracy. The main results were that when using only the two class sets of coniferous

and broadleaf trees and basal area of 100% with at least 10 images the OA was equal to 99% and the OA was always greater than 90% even with different proportions of basal area. All things being equal, the accuracy of RF is slightly lower. Excluding mixed classes, then with individual tree species classes (class set 3), the highest OA of 81% for both BA = 80% and 90% and OA = 84% for BA = 100% were obtained with combinations of 6 or 7 images. While in the more challenging situations with mixed classes and lower basal area thresholds (BA = 70% and 80%), RF provided slightly higher accuracies. The Bayesian approach, within our study area, proves efficient for tree species classification. The mapping of forest species supports the development of policies for forest conservation, fostering sustainable forest management critical to achieving the SDGs.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was financially supported by MISTRA Digital Forest. The authors also acknowledge the Swedish National Forest Inventory for providing the field data used in the study.

ORCID

Mats Nilsson  <http://orcid.org/0000-0001-7394-6305>

Arvid Axelsson  <http://orcid.org/0000-0001-6542-2528>

Data availability statement

The field data used was provided by the Swedish National Forest Inventory.

References

- Alonso, L., J. Picos, and J. Armesto. 2021. "Forest Land Cover Mapping at a Regional Scale Using Multi-Temporal Sentinel-2 Imagery and RF Models." *Remote Sensing* 13 (12): 2237. <https://doi.org/10.3390/rs13122237>.
- Axelsson, A., E. Lindberg, H. Reese, and H. Olsson. 2021. "Tree Species Classification Using Sentinel-2 Imagery and Bayesian Inference." *International Journal of Applied Earth Observation and Geoinformation* 100 : 102318. <https://doi.org/10.1016/j.jag.2021.102318>.
- Belgiu, M., and L. Drăguț. 2016. "Random Forest in Remote Sensing: A Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114:24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Bravo-Oviedo, A., H. Pretzsch, C. Ammer, E. Andenmatten, A. Barbat, S. Barreiro, P. Brang, et al. 2014. "European Mixed Forests: Definition and Research Perspectives." *Forest Systems* 23 (3): 518–533. <https://doi.org/10.5424/fs/2014233-06256>.
- Breidenbach, J., L. T. Waser, M. Debella-Gilo, J. Schumacher, J. Rahlf, M. Hauglin, and R. Astrup. 2020. "National Mapping and Estimation of Forest Area by Dominant Tree Species Using Sentinel-2 Data." *Canadian Journal of Forest Research* 51 (3): 365–379. <https://doi.org/10.1139/cjfr-2020-0170>.
- Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. <https://doi.org/10.1023/A:1010933404324>.

- Cardille, J. A., and J. A. Fortin. 2016. "Bayesian Updating of Land-Cover Estimates in a Data-Rich Environment." *Remote Sensing of Environment* 186:234–249. <https://doi.org/10.1016/j.rse.2016.08.021>.
- Chiarucci, A., and G. Piovesan. 2020. "Need for a Global Map of Forest Naturalness for a Sustainable Future." *Conservation Biology* 34 (2): 368–372. <https://doi.org/10.1111/cobi.13408>.
- Chirici, G., F. Giannetti, R. E. McRoberts, D. Travaglini, M. Pecchi, F. Maselli, M. Chiesi, and P. Corona. 2020. "Wall-To-Wall Spatial Prediction of Growing Stock Volume Based on Italian National Forest Inventory Plots and Remotely Sensed Data." *International Journal of Applied Earth Observation and Geoinformation* 84:101959. <https://doi.org/10.1016/j.jag.2019.101959>.
- Crowley, M. A., J. A. Cardille, J. C. White, and M. A. Wulder. 2019. "Multi-Sensor, Multi-Scale, Bayesian Data Synthesis for Mapping Within-Year Wildfire Progression." *Remote Sensing Letters* 10 (3): 302–311. <https://doi.org/10.1080/2150704X.2018.1536300>.
- D'Amico, G., S. Francini, F. Giannetti, E. Vangi, D. Travaglini, F. Chianucci, W. Mattioli, et al. 2021b. "A Deep Learning Approach for Automatic Mapping of Poplar Plantations Using Sentinel-2 Imagery." *GIScience & Remote Sensing* 58 (8): 1352–1368. <https://doi.org/10.1080/15481603.2021.1988427>.
- Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, and P. Bargellini. 2012. "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." *Remote Sensing of Environment* 120:25–36. <https://doi.org/10.1016/j.rse.2011.11.026>.
- Fassnacht, F. E., H. Latifi, K. Stereńczak, A. Modzelewska, M. Lefsky, L. T. Waser, C. Straub, and A. Ghosh. 2016. "Review of Studies on Tree Species Classification from Remotely Sensed Data." *Remote Sensing of Environment* 186:64–87. <https://doi.org/10.1016/j.rse.2016.08.013>.
- Flood, N. 2013. "Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (A Multi-Dimensional Median)." *Remote Sensing* 5 (12): 6481–6500. <https://doi.org/10.3390/rs5126481>.
- Francini, S., T. Hermosilla, N. C. Coops, M. A. Wulder, J. C. White, and G. Chirici. 2023. "An Assessment Approach for Pixel-Based Image Composites." *ISPRS Journal of Photogrammetry and Remote Sensing* 202:1–12. <https://doi.org/10.1016/j.isprsjprs.2023.06.002>.
- Fricke, G. A., J. D. Ventura, J. A. Wolf, M. P. North, F. W. Davis, and J. Franklin. 2019. "A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery." *Remote Sensing* 11 (19): 2326. <https://doi.org/10.3390/rs11192326>.
- Fridman, J., S. Holm, M. Nilsson, P. Nilsson, A. H. Ringvall, and G. Ståhl. 2014. "Adapting National Forest Inventories to Changing Requirements – the Case of the Swedish National Forest Inventory at the Turn of the 20th Century." *Silva Fennica* 48 (3): 1095. <http://dx.doi.org/10.14214/sf.1095>.
- Genuer, R., J. M. Poggi, and C. Tuleau-Malot. 2015. "VSURF: An R Package for Variable Selection Using Random Forests." *The R Journal* 7 (2): 19–33. <https://doi.org/10.32614/RJ-2015-018>.
- Gorte, B., and A. Stein. 1998. "Bayesian Classification and Class Area Estimation of Satellite Images Using Stratification." *IEEE Transactions on Geoscience and Remote Sensing* 36 (3): 803–812. <https://doi.org/10.1109/36.673673>.
- Gregersen H., H. El Lakany, and J. Blaser. 2017. "Forests for sustainable development: a process approach to forest sector contributions to the UN 2030 Agenda for Sustainable Development." *International Forestry Review: Shifting global development discourses — implications for forests and livelihoods* 19 (1): 10–23. <https://www.ingentaconnect.com/content/cfa/ifr/2017/00000019/a00101s1;jsessionid=15palhg7ydfpo.x-ic-live-01>.
- Guyon, I., J. Weston, S. Barnhill, and V. Vapnik. 2002. "Gene Selection for Cancer Classification Using Support Vector Machines." *Machine Learning* 46 (1/3): 389–422. <https://doi.org/10.1023/A:1012487302797>.
- Hagner, O., and H. Reese. 2007. "A Method for Calibrated Maximum Likelihood Classification of Forest Types." *Remote Sensing of Environment* 110 (4): 438–444. <https://doi.org/10.1016/j.rse.2006.08.017>.
- Hemmerling, J., D. Pflugmacher, and P. Hostert. 2021. "Mapping Temperate Forest Tree Species Using Dense Sentinel-2 Time Series." *Remote Sensing of Environment* 267:112743. <https://doi.org/10.1016/j.rse.2021.112743>.

- Hof, A. R., C. C. Dymond, and D. J. Mladenoff. 2017. "Climate Change Mitigation Through Adaptation: The Effectiveness of Forest Diversification by Novel Tree Planting Regimes." *Ecosphere* 8 (11): e01981. <https://doi.org/10.1002/ecs2.1981>.
- Hussain, S., A. Raza, H. G. Abdo, M. Mubeen, A. Tariq, and W. Nasim, . . . A. A. Al Dughairi. 2023. "Relation of Land Surface Temperature with Different Vegetation Indices Using Multi-Temporal Remote Sensing Data in Sahiwal Region, Pakistan." *Geoscience Letters* 10 (1): 33. <https://doi.org/10.1186/s40562-023-00287-6>.
- Immitzer, M., M. Neuwirth, S. Böck, H. Brenner, F. Vuolo, and C. Atzberger. 2019. "Optimal Input Features for Tree Species Classification in Central Europe Based on Multi-Temporal Sentinel-2 Data." *Remote Sensing* 11 (22): 2599. <https://doi.org/10.3390/rs11222599>.
- Kangas, A., R. Astrup, J. Breidenbach, J. Fridman, T. Gobakken, K. T. Korhonen, M. Maltamo, et al. 2018. "Remote Sensing and Forest Inventories in Nordic Countries—Roadmap for the Future." *Scandinavian Journal of Forest Research* 33 (4): 397–412. <https://doi.org/10.1080/02827581.2017.1416666>.
- Kennedy, R. E., Z. Yang, N. Gorelick, J. Braaten, L. Cavalcante, W. B. Cohen, and S. Healey. 2018. "Implementation of the LandTrendr Algorithm on Google Earth Engine." *Remote Sensing* 10 (5): 691. <https://doi.org/10.3390/rs10050691>.
- Khan, S. N., A. N. Khan, A. Tariq, L. Lu, N. A. Malik, M. Umair, W. A. Haamleh, and F. H. Zawaideh. 2023. "County-Level Corn Yield Prediction Using Supervised Machine Learning." *European Journal of Remote Sensing* 56 (1): 2253985. <https://doi.org/10.1080/22797254.2023.2253985>.
- Kuhn, M. 2008. "Building Predictive Models in R Using the Caret Package." *Journal of Statistical Software* 28 (5): 1–26. <https://doi.org/10.18637/jss.v028.i05>.
- Lindberg, E., J. Holmgren, and H. Olsson. 2021. "Classification of Tree Species Classes in a Hemi-Boreal Forest from Multispectral Airborne Laser Scanning Data Using a Mini Raster Cell Method." *International Journal of Applied Earth Observation and Geoinformation* 100:102334. <https://doi.org/10.1016/j.jag.2021.102334>.
- Lisein, J., A. Michez, H. Claessens, P. Lejeune, and M. Cristani. 2015. "Discrimination of Deciduous Tree Species from Time Series of Unmanned Aerial System Imagery." *Public Library of Science One* 10 (11): 1–20. <https://doi.org/10.1371/journal.pone.0141006>.
- Lister, A. J., H. Andersen, T. Frescino, D. Gatzliolis, S. Healey, L. S. Heath, and B. T. Wilson. 2020. "Use of Remote Sensing Data to Improve the Efficiency of National Forest Inventories: A Case Study from the United States National Forest Inventory." *Forests* 11 (12): 1364. <https://doi.org/10.3390/f11121364>.
- Mahalanobis, P. C. 1936. "On the Generalized Distance in Statistics." In *Proceedings of the National Institute of Sciences of India*, Calcutta, 49–55. National Institute of Science of India. http://bayes.acs.unt.edu:8083/BayesContent/class/Jon/MiscDocs/1936_Mahalanobis.pdf.
- Mäyrä, J., S. Keski-Saari, S. Kivinen, T. Tanhuanpää, P. Hurskainen, P. Kullberg, and P. Vihervaara. 2021. "Tree Species Classification from Airborne Hyperspectral and LiDAR Data Using 3D Convolutional Neural Networks." *Remote Sensing of Environment* 256:112322. <https://doi.org/10.1016/j.rse.2021.112322>.
- Michałowska, M., and J. Rapiński. 2021. "A Review of Tree Species Classification Based on Airborne LiDAR Data and Applied Classifiers." *Remote Sensing* 13 (3): 353. <https://doi.org/10.3390/rs13030353>.
- Nasiri V., M. Belouï, A. Asghar Darvishsefat, V. C. Griess, C. Maftai, and L. T. Waser. 2023. "Mapping tree species composition in a Caspian temperate mixed forest based on spectral-temporal metrics and machine learning." *International Journal of Applied Earth Observation and Geoinformation* 116:103154. <https://doi.org/10.1016/j.jag.2022.103154>.
- Nilsson, M., K. Nordkvist, J. Jonzén, N. Lindgren, P. Axensten, J. Wallerman, and H. Olsson. 2017. "A Nationwide Forest Attribute Map of Sweden Predicted Using Airborne Laser Scanning Data and Field Data from the National Forest Inventory." *Remote Sensing of Environment* 194:447–454. <https://doi.org/10.1016/j.rse.2016.10.022>.
- Oreti, L., D. Giuliarelli, A. Tomao, and A. Barbati. 2021. "Object Oriented Classification for Mapping Mixed and Pure Forest Stands Using Very-High Resolution Imagery." *Remote Sensing* 13 (13): 2508. <https://doi.org/10.3390/rs13132508>.

- Parisi, F., E. Vangi, S. Francini, G. D'Amico, G. Chirici, M. Marchetti, F. Lombardi, et al. 2023. "Sentinel-2 Time Series Analysis for Monitoring Multi-Taxon Biodiversity in Mountain Beech Forests." *Frontiers in Forest Global Change* 6:1020477. <https://doi.org/10.3389/ffgc.2023.1020477>.
- Persson, H. J., H. Olsson, M. J. Soja, L. M. Ulander, and J. E. Fransson. 2017. "Experiences from Large-Scale Forest Mapping of Sweden Using TanDEM-X Data." *Remote Sensing* 9 (12): 1253. <https://doi.org/10.3390/rs9121253>.
- Persson, M., E. Lindberg, and H. Reese. 2018. "Tree Species Classification with Multi-Temporal Sentinel-2 Data." *Remote Sensing* 10 (11): 1794. <https://doi.org/10.3390/rs10111794>.
- Pignatale, F. C. 2022. "Sen2Cor 2.11.00 Software Release Note." *European Space Agency*. OMPC.TPZG. SRN.003. 4000136252/21/I-BG.
- Plakman, V., T. Janssen, N. Brouwer, and S. Veraverbeke. 2020. "Mapping Species at an Individual-Tree Scale in a Temperate Forest, Using Sentinel-2 Images, Airborne Laser Scanning Data, and Random Forest Classification." *Remote Sensing* 12 (22): 3710. <https://doi.org/10.3390/rs12223710>.
- Puletti, N., F. Chianucci, and C. Castaldi. 2018. "Use of Sentinel-2 for Forest Classification in Mediterranean Environments." *Annals of Silvicultural Research* 42 (1): 32–38. <https://doi.org/10.12899/asr-1463>.
- Ramezan, C. A. 2022. "Transferability of Recursive Feature Elimination (RFE)-Derived Feature Sets for Support Vector Machine Land Cover Classification." *Remote Sensing* 14 (24): 6218. <https://doi.org/10.3390/rs14246218>.
- Sameen, M. I., B. Pradhan, and O. S. Aziz. 2018. "Classification of Very High Resolution Aerial Photos Using Spectral-Spatial Convolutional Neural Networks." *Journal of Sensors* 2018:1–12. <https://doi.org/10.1155/2018/7195432>.
- Sheeren, D., M. Fauvel, V. Josipović, M. Lopes, C. Planque, J. Willm, and J.-F. Dejoux. 2016. "Tree Species Classification in Temperate Forests Using Formosat-2 Satellite Image Time Series." *Remote Sensing* 8 (9): 734. <https://doi.org/10.3390/rs8090734>.
- Su, T. 2020. "Object-Based Feature Selection Using Class-Pair Separability for High-Resolution Image Classification." *International Journal of Remote Sensing* 41 (1): 238–271. <https://doi.org/10.1080/01431161.2019.1641242>.
- Swain, P. H. 1978. "Fundamentals of Pattern Recognition in Remote Sensing." *Remote Sensing; the Quantitative Approach*, edited by P.H. Swain and S.M. Davis, 136–187. New York: McGraw-Hill International Book Company.
- Vangi, E., G. D'Amico, S. Francini, C. Borghi, F. Giannetti, P. Corona, M. Marchetti, et al. 2023. "Large-Scale High Resolution Yearly Modeling of Forest Growing Stock Volume and Above-Ground Carbon Pool." *Environmental Modelling and Software* 159:105580. <https://doi.org/10.1016/j.envsoft.2022.105580>.
- Vangi, E., G. D'Amico, S. Francini, F. Giannetti, B. Lasserre, M. Marchetti, and G. Chirici. 2021. "The New Hyperspectral Satellite PRISMA: Imagery for Forest Types Discrimination." *Sensors* 21 (4): 1182. <https://doi.org/10.3390/s21041182>.
- Wallis, C. I., G. Brehm, D. A. Donoso, K. Fiedler, J. Homeier, D. Paulsch, and J. Bendix. 2017. "Remote Sensing Improves Prediction of Tropical Montane Species Diversity but Performance Differs Among Taxa." *Ecological Indicators* 83:538–549. <https://doi.org/10.1016/j.ecolind.2017.01.022>.
- Wang, R., and J. A. Gamon. 2019. "Remote Sensing of Terrestrial Plant Biodiversity." *Remote Sensing of Environment* 231:111218. <https://doi.org/10.1016/j.rse.2019.111218>.
- Waser, L. T., C. Ginzler, and N. Rehus. 2017. "Wall-To-Wall Tree Type Mapping from Countrywide Airborne Remote Sensing Surveys." *Remote Sensing* 9 (8): 766. <https://doi.org/10.3390/rs9080766>.
- Wazen, N., V. Garavaglia, N. Picard, C. Besacier, and B. Fady. 2018. "Distribution Maps of Twenty-Four Mediterranean and European Ecologically and Economically Important Forest Tree Species Compiled from Historical Data Collections." *Annals of Silvicultural Research* 44 (2): 95–101. <https://doi.org/10.12899/asr-1933>.
- White, J. C., M. A. Wulder, G. W. Hobart, J. E. Luther, T. Hermosilla, P. Griffiths, N. C. Coops, et al. 2014. "Pixel-Based Image Compositing for Large-Area Dense Time Series Applications and Science." *Canadian Journal of Remote Sensing* 40 (3): 192–212. <https://doi.org/10.1080/07038992.2014.945827>.

- Wright, M. N., S. Wager, and P. Probst. 2023. "Package 'Ranger'". Version 0.15.1.
- Wu, J., L. Lin, C. Zhang, T. Li, X. Cheng, and F. Nan. 2023. "Generating Sentinel-2 All-Band 10-M Data by Sharpening 20/60-M Bands: A Hierarchical Fusion Network." *ISPRS Journal of Photogrammetry and Remote Sensing* 196:16–31. <https://doi.org/10.1016/j.isprsjprs.2022.12.017>.
- Xiao, J., F. Chevallier, C. Gomez, L. Guanter, J. A. Hicke, A. R. Huete, and X. Zhang. 2019. "Remote Sensing of the Terrestrial Carbon Cycle: A Review of Advances Over 50 Years." *Remote Sensing of Environment* 233:111383. <https://doi.org/10.1016/j.rse.2019.111383>.
- Young, B., J. Yarie, D. Verbyla, F. Huettmann, K. Herrick, and F. S. Chapin. 2017. "Modeling and Mapping Forest Diversity in the Boreal Forest of Interior Alaska." *Landscape Ecology* 32 (2): 397–413. <https://doi.org/10.1007/s10980-016-0450-2>.