



**University of
Zurich**^{UZH}

Large-scale individual tree classification using airborne laser scanning in temperate mixed forests

GEO 610 Master's Thesis

Author

Aline Bornand

15-710-056

Supervised by

Dr. Felix Morsdorf

Dr. Meinrad Abegg (meinrad.abegg@wsl.ch)

Faculty representative

Prof. Dr. Michael Schaepman

31.07.2020

Department of Geography, University of Zurich

Abstract

Remotely sensed information on tree species composition over vast spatial extents is required for various applications, such as mapping and preserving biodiversity or sustainable forest management. Automatic individual tree detection based on airborne laser scanning (ALS) allows tree species recognition on the level of individual trees. Most species classification studies that go beyond the distinction of conifers and broadleaf trees are limited to smaller areas due to reference data availability. This thesis aimed to implement and evaluate the potential of a large-scale tree classification on individual tree level in a diverse temperate forest environment. Our analysis was based on multi-seasonal (leaf-off and leaf-on) ALS acquisitions, supplemented by multispectral aerial imagery, covering an area of 1400 km² in northern Switzerland. We connected field-based species labels from an existing operational forest inventory with individual tree segments extracted from ALS data to receive a reference dataset without any manual delineation or matching steps. This automated creation of reliable reference data was found to represent one of the main challenges, introducing uncertainties into subsequent classification results. We trained and tested a series of object-based Random Forest (RF) machine learning models based on a diverse set of features (ALS point distribution, intensity, return number or multispectral). The integration of features derived from leaf-off and leaf-on ALS acquisitions significantly improved classification accuracy by 6% compared to using only leaf-off. Through applying a recursive feature elimination (RFE) approach, features based on ALS intensity and return number were found to be more important than features related to tree shape, point distribution and spectral features. Despite the use of imperfect reference data in a large-scale context, we were able to discriminate European beech, Norway Spruce and silver fir from the remaining broadleaved and coniferous species with an overall accuracy of 62%, based solely on ALS data. These species predictions can potentially be aggregated to forest composition maps in order to estimate species-specific forest attributes.

Contents

Abstract	I
List of Figures	V
List of Tables	VII
List of Abbreviations	IX
1 Introduction	1
2 Materials	5
2.1 Study Area	5
2.2 ALS Data	6
2.3 Aerial Imagery	6
2.4 Forest Inventory Data	7
3 Methods	9
3.1 Individual tree detection	10
3.2 Creation of reference data: matching ALS-trees to inventory-trees	10
3.3 Evaluation of individual tree detection and matching	11
3.4 Feature extraction	11
3.4.1 Separability of tree species	11
3.4.2 ALS based features	13
3.4.3 Image features	13
3.5 Classification procedure	14
3.5.1 Class grouping	15
3.5.2 Choice of classifier	15
3.5.3 Managing class imbalances	16
3.5.4 Feature selection	16
3.5.5 Training and validation sets	18
3.5.6 Accuracy assessment	18
3.6 Comparison of input datasets	18
3.7 Comparison of individual features	20
3.8 Simplified classification model for predictive mapping	20
4 Results	21
4.1 Individual tree matching	21

4.2	Tree height accuracy	22
4.3	Classification results	23
4.3.1	Relevance of input datasets	23
4.3.2	Individual feature importance	25
4.3.3	Predictive mapping	27
5	Discussion	33
5.1	Compatibility of inventory data and ALS tree segments	34
5.2	Combining two ALS acquisitions and aerial imagery	35
5.2.1	Leaf-off and leaf-on ALS datasets	35
5.2.2	RGBI orthophotos	35
5.3	Evaluation of ALS features	36
5.4	Predictive mapping	38
5.5	Outlook	39
6	Conclusion	41
	References	43
A	Appendix - Additional Figures	51
	Acknowledgements	55
	Personal Declaration	57

List of Figures

2.1	Overview of the study area	5
3.1	Processing and classification workflow	9
3.2	Examples of leaf-on acquisitions of individual trees	12
3.3	Examples of leaf-off acquisitions of individual trees	12
3.4	Multi-class confusion matrix	18
3.5	Results of recursive feature elimination	19
4.1	Example reference tree polygons	21
4.2	Positional accuracy	22
4.3	Tree height distribution	22
4.4	Scatterplot and linear regression of matched individual tree heights.	23
4.5	Height difference as boxplots	24
4.6	Overall feature importance	26
4.7	Class-specific feature importances	29
4.8	Simplified classification applied to a test area near Bremgarten	30
4.9	Detail of Figure 4.8 including train tracks, northeast	31
4.10	Detail of Figure 4.8, southwest	31
5.1	Processing and classification workflow with annotations	33
5.2	Boxplots of exemplary features	37
A.1	Tree segment point density for leaf-on and leaf-off	51
A.2	Tree segment point density per species class	51
A.3	Class-specific feature importances	52
A.4	Correlation matrix, 100 features	53
A.5	Correlation matrix, 48 features	54
A.6	Principal component plot of simplified classification	54

List of Tables

2.1	Summary of the laser scanner and flight specifications	6
2.2	Overview of tree species classes	8
3.1	Overview of ALS derived features	14
3.2	Overview of features derived from orthophotos	15
4.1	Overview of classification datasets	24
4.2	Comparison of classification accuracies among the different datasets	25
4.3	P-values for the McNemar's test	25
4.4	Confusion matrix for ALS-RGBI classification	26
4.5	Confusion Matrix for simplified classification	27

List of Abbreviations

2D	Two-dimensional
3D	Three-dimensional
ALS	Airborne Laser Scanning
AWI	Aargauer Waldinventur (Forest Inventory of Canton of Aargau)
BRF	Balanced Random Forest
CHM	Canopy Height Model
CNN	Convolutional Neural Network
DBH	Diameter at Breast Height
DTM	Digital Terrain Model
GPS	Global Positioning System
IHS	Intensity, Hue, Saturation
ITC	Individual Tree Crown
ITD	Individual Tree Detection
LiDAR	Light Detection And Ranging
m.a.s.l.	Meters Above Sea Level
NDVI	Normalized Difference Vegetation Index
NFI (LFI)	National Forest Inventory (Landesforstinventar)
NIR	Near-infrared
OA	Overall Accuracy
PA	Producer's Accuracy
PC(A)	Principal Component (Analysis)
RF	Random Forest
RFE	Recursive Feature Elimination
RGBI	Red-Green-Blue-Infrared
RMSE	Root Mean Square Error
SVM	Support Vector Machines
UA	User's Accuracy
UAV	Unmanned Aerial Vehicle

1 Introduction

Information on tree species composition and spatial distribution in forest environments is required by a wide variety of applications, ranging from biodiversity assessment, monitoring and conservation (Puumalainen et al., 2003) to sustainable management strategies (Vauhkonen et al., 2014). To give an example, the spatial abundance of Norway spruce has been identified as an important predictor for susceptibility to bark beetle outbreaks (Kärvement al., 2014). In forest management, spatially explicit information on tree species is used as input for species-specific growth and yield models (Vauhkonen et al., 2014) or any species-specific allometric model for attributes like stem diameter, volume or productivity (Korpela and Tokola, 2006; Ørka et al., 2013). Compared to conventional approaches based on labour-intensive field investigations, mapping and monitoring forest species composition using remote sensing would save time, money, and support the analysis of species composition over vast spatial extents (McRoberts and Tomppo, 2007; Hyypä et al., 2008). Over the past three decades, a variety of spaceborne and airborne technologies have been utilised for this purpose (Fassnacht et al., 2016). This involves distinguishing tree species based on satellite multispectral images (Zhang and Liu, 2013; Persson et al., 2018), aerial multi- or hyperspectral images (Waser et al., 2011; Dalponte et al., 2013; Fricker et al., 2019), or unmanned aerial vehicles (UAV) (Franklin and Ahmed, 2018). While spectral signatures can reflect biophysical, biochemical or even physiological attributes, it is possible to retrieve structural properties of trees from airborne laser scanning (ALS) systems, equipped with light detection and ranging (LiDAR) devices (Hyypä et al., 2008). ALS is of particular value for forest characterisation, due to its ability to penetrate the canopy, providing detailed three-dimensional information on vegetation structure (Nilsson, 1996). Small-footprint ALS systems are therefore used at operational level today in order to retrieve inventory parameters relevant for forest management (Wulder et al., 2013; White et al., 2016; Kankare et al., 2017). In regard to tree species recognition, combining spectral and structural remote sensing data has been shown to generally improve classification results (Ørka et al., 2012; Kandare et al., 2017; Parkan and Tuia, 2019).

Another advantage of using ALS is that individual trees can be automatically detected and extracted from LiDAR point clouds. Individual tree detection (ITD) represents a bottleneck in the retrieval of forest parameters and has thus motivated an active research community (Kaartinen et al., 2012). While ITD methods with varying degrees of complexity are being applied today, large-scale tree detection is often based on a canopy height model (CHM). This approach decreases the size of datasets, computational time and the demands on hardware, but also reduces information regarding the understorey (Duncanson et al., 2014). Across most ITD methods, detection accuracy is generally higher for conifer trees, while the closed canopy of broadleaf forests is more challenging to separate (Hastings et al., 2020).

When ITD is applied successfully, species membership can be predicted on individual tree level

through the use of an object-based classification. Many studies on tree level species recognition have reported higher classification accuracies compared to area-based approaches (Ørka et al., 2013; Dalponte et al., 2013; Torabzadeh et al., 2019). Also, tree level data can be easily linked to the tree's biophysical properties and it can be used directly for practical forest management purposes. As input for an object-based classification of tree species, explanatory features are usually calculated for each reference tree in order to capture the differences between species. While features derived from spectral imaging data include reflectance values of selected bands and vegetation indices (Fassnacht et al., 2014), a considerable amount of descriptive features have also been derived from ALS data (Ørka et al., 2009; Lin and Hyyppä, 2016; Shii et al., 2018). ALS-based features can be broadly categorised into spatial point pattern statistics (e.g. height distribution and density of points), radiometric statistics derived from echo intensity, and features utilising the number of returns generated by one laser pulse. Depending on the kind of ITD approach, it may be possible to derive further external and internal tree shape characteristics (Parkan and Tuia, 2019). Due to availability constraints, relatively few studies incorporate temporal changes as additional features. Nevertheless, the integration of LiDAR data acquired under leaf-on and leaf-off conditions has been proven useful for tree species classification in previous studies (Kim et al., 2009; Ørka et al., 2010; Shii et al., 2018).

To find the most effective set of features for classification, iterative feature selection approaches are often applied, because they also provide importance estimates (Fassnacht et al., 2014; Ko et al., 2016; Parkan and Tuia, 2019). As the target classes are generally known and predefined in tree species classification, supervised learning algorithms are by far the most commonly used due to their superior performance (Ghosh et al., 2014; Marrs and Ni-Meister, 2019). Thanks to their ability to handle non-linear classification problems and to their ease of use, Random Forests (RF) and Support Vector Machines (SVM) are currently the most popular and have been shown to have similar performances (Ørka et al., 2012; Dalponte et al., 2012; Ghosh et al., 2014).

The methodologies described above have been tested on numerous experimental sites using state-of-the-art input data and reliable ground-truth (Fassnacht et al., 2016). For example, Parkan and Tuia (2019) received overall accuracies between 84% and 95% using various combinations of hyperspectral imaging and ALS feature sets to classify nine tree species in temperate forests in Switzerland. However, such studies are generally restricted to small areas or limited sample sizes, as labelling a large number of training samples is a tedious process, often the result of repeated visual interpretation and/or field observations.

Depending on the specific application, ecologists or forestry practitioners may require tree species information across landscapes or even countries. Next to availability of suitable remote sensing data, feasibility of such large-scale tree species recognition depends on the automation of reference data creation, which should ideally exploit existing inventory data. In recent years, field plot information from National Forest Inventories (NFI) or smaller management inventories has been increasingly used in combination with ALS data. This has been shown to be effective for enhancing the estimation of forest inventory attributes (Tomppo et al., 2008; Maltamo et al., 2009; Wulder et al., 2012) and is applied operationally for regional forest management inventories in Finland, Norway and Sweden (Barrett et al., 2016; Kangas et al., 2018). However, large-scale tree species recognition, especially on tree level, is not yet performed operationally (White et al., 2016).

Nonetheless, individual tree level information is expected to play an important role in next generation's forest mapping systems, especially in conifer-dominated boreal forests (Kankare et al., 2017).

In temperate forest regions, large-scale or even country-wide tree classification attempts with high spatial resolution have, so far, mainly been focused on the distinction between broadleaved and coniferous trees. For instance, Schumacher and Nord-Larsen (2014) performed a classification for whole of Denmark (tree segment level), Waser et al. (2017) for whole of Switzerland (spatial resolution of 3 m) and Krzystek et al. (2020) for the combined area of Bavarian Forest National Park and Šumava National Park (tree segment level). They all applied a combination of multispectral aerial imagery and ALS.

Operational large-scale ALS acquisitions with suitable point density for forest characterisation are becoming more commonplace in Central Europe. Combined with existing inventory sources, such as the NFI, this can potentially be utilised for large-scale individual tree classification of multiple species. However, such attempts are still associated with many uncertainties. Focusing mainly on ALS data, the goal of this thesis is to implement and evaluate the potential of a large-scale tree classification on individual tree level in a diverse temperate forest environment. Our analysis is based on leaf-off and leaf-on ALS acquisitions, supplemented by RGBI aerial imagery from the same year, covering an area of 1400 km². We connect field-based species labels from an existing operational forest inventory with tree segments extracted from ALS data to receive a reference dataset without any manual delineation or matching steps. To evaluate the potential of multi-temporal ALS data, we assess classification accuracy based on each input dataset as well as based on their combination. We also evaluate feature importance to assess which features are suitable for use in a large-scale application with imperfect reference data. Finally, we apply a classification model to a test area for qualitative assessment. In order to locate possibilities and limitations associated with this approach, we explored the following objectives:

1. What is the potential of national forest inventory data to serve as reference for labelling ALS-derived individual tree segments?
2. What is the added value of combining ALS leaf-off and leaf-on acquisitions and aerial imagery for tree species classification?
3. Which features, based on ALS point distribution, intensity, return number or multispectral bands, are the most important for predicting tree species in a large-scale context?

2 Materials

2.1 Study Area

The study area consists of the entire Canton of Aargau. With its 1403 km² it is the 10th largest canton of Switzerland, with altitude ranging from 261 to 908 m.a.s.l.. 35 % of the area are covered by forest (49070 ha or 490 km²). The area comprises the biogeographical regions Table Jura (17%) and Swiss Plateau (83%) and is mostly characterised by semi-natural temperate mixed forest, of which 6% are unmanaged forest reserves. The most common species are European beech (*Fagus sylvatica*), Norway spruce (*Picea abies*), Sycamore Maple (*Acer pseudoplatanus*) and Silver fir (*Abies alba*) (cf. Table 2.2). Table Jura and forest reserves are generally more dominated by broadleaf trees as compared to the remaining, more intensively managed forest regions in Canton of Aargau (Departement Bau Verkehr und Umwelt, 2018).

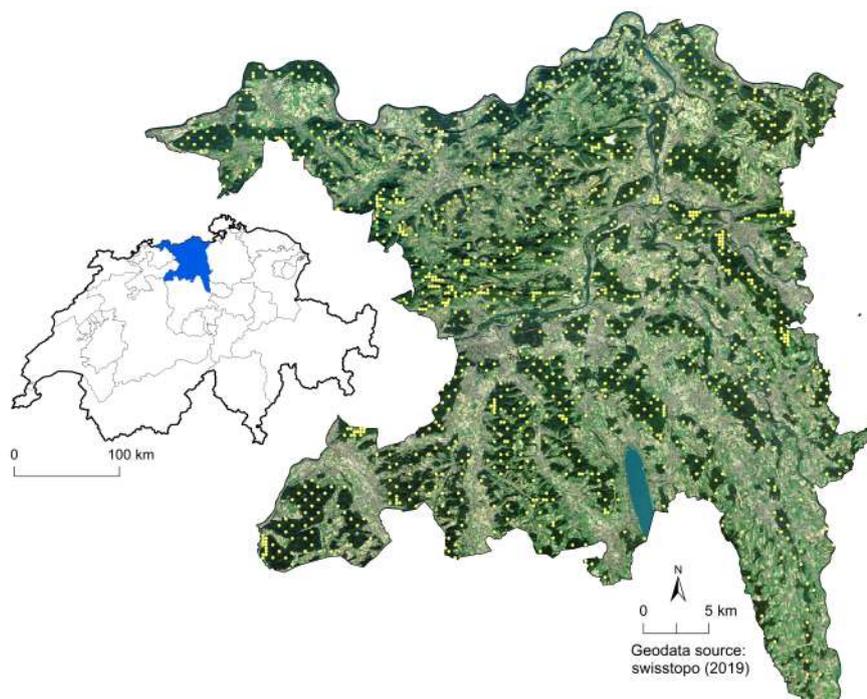


FIGURE 2.1: Overview of the study area Canton of Aargau (right) and its location in Switzerland (left). The yellow dots indicate the approximate location of inventory sampling plots used in this study.

2.2 ALS Data

ALS data for the Canton of Aargau was acquired during March and April 2014 under defoliated conditions (leaf-off) and during June and July 2014 under foliated conditions (leaf-on). The sensor specifications are summarised in Table 2.1. The full waveform ALS data was acquired and processed by Milan Geoservice GmbH (Kamenz, Germany). The processing steps involved the extraction of laser returns from full waveform data, transformation of the point cloud into the Swiss CH-1903 (LV03) Cartesian coordinate system, flight strip adjustment as well as filtering and classification of the point cloud into ground and vegetation points. A digital terrain model (DTM) was also generated by the contractor using TerraScan software (TerraScan v014, TerraSolid, Helsinki, Finland).

The further derivation of a CHM and terrain-corrected heights is described by Leiterer et al. (2015). The final dataset contains a three-dimensional point cloud composed of planimetric coordinates (x and y), ellipsoidal heights (z), terrain-corrected heights (z_{AG}), echo type (i.e., first, intermediate, last echo) and range normalised intensity values for every echo.

TABLE 2.1: Summary of the laser scanner and flight specifications for both leaf-on and leaf-off acquisitions.

ALS parameter	Leaf-off	Leaf-on
Acquisition date	March/April 2014	June/July 2014
ALS sensor	Riegl LMS-Q680i	
Operating platform	Airplane	
Area of coverage [km ²]	~ 1400	
Mean operating altitude above ground [m]	600	700
Scanning method	Rotating multi-facet mirror	
Pulse detection method	Full-waveform processing	
Pulse length [ns]	~ 4	
Sampling interval [ns]	1	
Max scan angle [deg]	22	
Mean point density [pts/m ²]	15	30
Mean pulse density [pls/m ²]	~ 11	~ 11
Pulse footprint [cm]	30	35
Laser wavelength [nm]	1550	
Scan rate [Hz]	120	
Pulse repetition frequency [kHz]	300	
Beam divergence [mrad]	0.5	
Angular step width [deg]	0.0176	

2.3 Aerial Imagery

Orthophotos were included as an additional data source. These were acquired on the 16th, 17th, and 18th of July 2014 (leaf-on season) using an UltraCamEagle digital camera. The images were acquired as part of a yearly campaign on behalf of the Cantonal authorities and cover the entire Canton of Aargau. They were orthorectified using a digital surface model created by stereoscopic image matching. Featuring a spatial resolution of 0.25 m, they contain the 4 spectral bands red,

green, blue and near infrared (RGBI).

2.4 Forest Inventory Data

In addition to the Swiss National Forest Inventory (NFI), the Canton of Aargau has its own Forest Inventory (Aargauer Waldinventur, AWI), which has so far been conducted in 2005 and 2016. The sample plots are located on grid points of a 707x707 m grid containing four times as many plots as the NFI base grid. Within the scope of the AWI 2016, around 16'000 individual trees were registered and measured on approximately 1300 sample plots (Departement Bau Verkehr und Umwelt, 2018).

The AWI uses the same inventory methodology as the NFI (described in Fischer and Traub (2019)). Sampling plots consist of concentric circles around a known location. Within a 200 m² circle, every tree with a diameter at breast height (DBH) larger than 12 cm is recorded as a tally tree, and within a 500 m² circle, every tree with a DBH larger than 36 cm is recorded. Unlike DBH, tree height is only measured for a randomly selected subset of trees.

After filtering the AWI dataset for living, standing trees with a height measurement and a complete geographic location, we received 3609 reference trees, as shown in Table 2.2. Available data associated with every tree includes: plot and tree number, height, species, DBH, basal area, tree status (living, dead, standing etc.), canopy layer, crown length and approximate tree age. The location is given in azimuth and distance from the plot centre point, whose coordinates are measured by GPS with an accuracy of ~ 1 m.

TABLE 2.2: Number of trees of each species present in the reference dataset and their class grouping for the simplified classification and main species classification.

Simplified classes	Main species classes	Species names	Total trees per class
spruce	spruce	<i>Picea abies</i> (712)	712
fir	fir	<i>Abies alba</i> (413)	413
misc. conifer	pine	<i>Pinus sylvestris</i> (96) <i>Pinus nigra</i> (7) <i>Pinus strobus</i> (5)	108
	-	<i>Larix decidua</i> (52) <i>Larix kaempferi</i> (4) <i>Pseudotsuga menziesii</i> (36) <i>Taxus baccata</i> (25) <i>Thuja spec.</i> (38) other conifers (2)	157
beech	beech	<i>Fagus sylvatica</i> (1231)	1231
misc. broadleaf	maple	<i>Acer pseudoplatanus</i> (139) <i>Acer platanoides</i> (7) <i>Acer campestre</i> (3)	149
	ash	<i>Fraxinus excelsior</i> (212)	212
	oak	<i>Quercus petraea</i> (122) <i>Quercus robur</i> (69) <i>Quercus rubra</i> (27) <i>Quercus pubescens</i> (2)	220
	-	<i>Alnus glutinosa</i> (12) <i>Alnus incana</i> (3) <i>Betula pendula</i> (4) <i>Castanea sativa</i> (1) <i>Carpinus betulus</i> (34) <i>Juglans regia</i> (4) <i>Populus nigra s.l.</i> (5) <i>Populus tremula</i> (1) <i>Populus spec.</i> (1) <i>Prunus avium</i> (15) <i>Robinia pseudoacacia</i> (12) <i>Salix alba</i> (3) <i>Salix spec.</i> (2) <i>Sorbus torminalis</i> (1) <i>Tilia cordata</i> (25) <i>Tilia platyphyllos</i> (31) <i>Ulmus glabra</i> (11) <i>Aesculus hippocastanum</i> (2) <i>Platanus spec.</i> (2)	169

3 Methods

Subsequent sections follow the workflow illustrated in Figure 3.1. Individual trees were extracted from ALS data, then matched and validated using inventory data. Features were computed from the resulting labelled segments, evaluated and used as input for tree species classification.

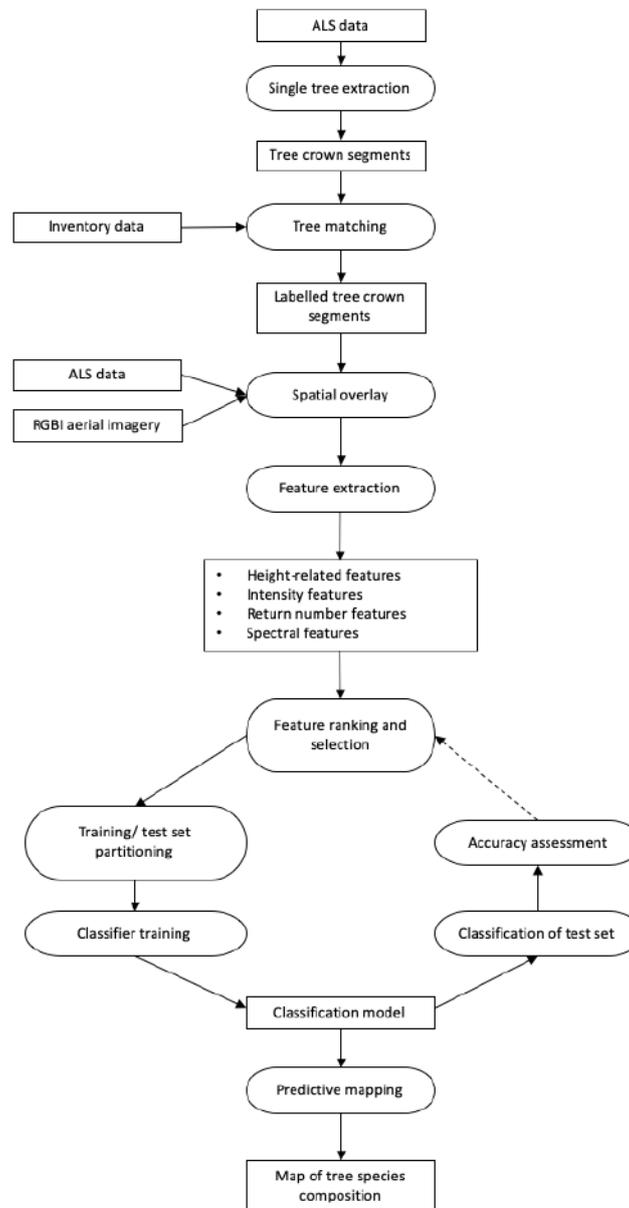


FIGURE 3.1: Processing and classification workflow. Squares represent products, while round boxes symbolise processing steps.

3.1 Individual tree detection

Individual tree detection was performed for the entire area of Canton of Aargau based on terrain-corrected ALS data, with points belonging to buildings set to a height of zero. Beforehand, an area-based Random Forest classifier was applied to discern the forest types broadleaved and coniferous. When validated with leaf-off colour infrared images, this binary classification yielded an overall accuracy of 75-85%.

As a basis for tree extraction, the CHM was smoothed using a Gaussian filter to reduce data pits according to the approach of Duncanson et al. (2014). The following steps apply functions included in the Digital Forestry Toolbox by M.Parkan (2018), which are based on methods originally presented by e.g. Q. Chen et al. (2006). Local maxima were detected using a tree top marker function with a search radius parameter determined by dominant forest type in the region. Using the CHM and the local maxima coordinates, a marker controlled watershed segmentation was computed, resulting in Individual Tree Crown (ITC) segments. ALS returns located within each segment were assigned the respective segment ID. Segments higher than 3 m and containing more than 25 points were considered as trees.

3.2 Creation of reference data: matching ALS-trees to inventory-trees

In order to receive a dataset that is amenable to validation, tree segments detected by airborne laser scanning are individually paired to trees measured by field inventory. To achieve this, a computer vision based feature matching algorithm was applied, built on the function *matchFeatures*, which is already implemented in MATLAB (2014). For every inventory tree, the algorithm searches for the corresponding ALS segment (ITC), based on tree positions (x,y) and height (h) of both the ITC and the inventory trees. Search radius was restricted to 5 m, as 99% of ITCs possess a crown diameter smaller than that. Ultimately, the reference tree is matched to the ITC, which is closest in the dimensions x,y and h. This individual tree matching process is associated with the following assumptions (or simplifications), which need to be taken into account when using resulting data for further analysis and interpretation:

- Each ITC is assumed to contain only one individual tree, when in reality it could also contain several undetected smaller trees or one tree could be split into multiple segments.
- The tree peak is situated vertically above the corresponding tree stem.
- The geographical location of the reference tree is assumed to correctly represent ground truth.
- The tree height measured by the field inventory crew is comparable to the height detected by ALS.
- No major disturbances such as wood harvesting or windfall occurred between the years 2014 and 2016.
- The sampled trees did not grow substantially between 2014 and 2016.

3.3 Evaluation of individual tree detection and matching

Tree detection and matching quality is evaluated by comparing position and height of reference trees and matched ITCs.

The positional error for a tree is defined as the Euclidean distance between the detected tree position and the reference tree position (Yin and Wang, 2016). The mean positional error, e_{pos} , is used as a measure of tree detection accuracy:

$$e_{pos} = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - x_{ref,i})^2 + (y_i - y_{ref,i})^2}, \quad (3.1)$$

where (x_i, y_i) is the location of the i th ITC and $(x_{ref,i}, y_{ref,i})$ is the location of the corresponding reference tree. A lower e_{pos} value is associated with a generally more accurate tree detection result. The tree height estimated by ALS is compared to the reference height via a scatterplot and an analysis of resulting linear regression. R^2 and root mean squared error ($RMSE$) are computed. The latter reports the deviation of estimated parameters from their true values. As the most extreme outliers are most certainly due to mismatches (i.e. not the same tree), an additional robust linear regression is fitted to decrease the weight of those points on the regression fit. Robust regression is done by iterated re-weighted least squares (IRLS). This results in each point's the weight being adjusted in an iterative process (Huber, 2011).

3.4 Feature extraction

As input for an object-based, supervised classification of tree species, explanatory features (also called metrics, descriptors, criteria or variables) are commonly calculated for each reference tree (Ørka et al., 2009; Kim et al., 2009; Zhang and Liu, 2013; Shii et al., 2018; Torabzadeh et al., 2019; Parkan and Tuia, 2019). Feature extraction can be seen as an attempt to capture tree species characteristics ideally enabling an unambiguous differentiation of classes.

3.4.1 Separability of tree species

The patterns of ALS point clouds of individual trees (as illustrated in Figures 3.2 and 3.3) are influenced by structural characteristics, which affect the laser beam by locally modifying the opacity of space (Vauhkonen et al., 2014; Parkan and Tuia, 2019). Many trees exhibit species-specific differences in their structure. The vertical distribution of echoes is influenced by crown shape as well as crown density, which may be subject to seasonal changes. Multiple echoes of comparably low intensity are generated when the laser beam is only partially reflected e.g. by small branches, whereas large and opaque structures such as the stem or main branches may cause single returns with higher intensity. Coniferous trees are expected to show small to no seasonal changes regarding these structural properties. Conversely, the examples in Figures 3.2 and 3.3 illustrate that especially for beech, the dense summer foliage results in comparably high intensity values in the upper tree crown, while fewer points are situated inside the crown and on the forest floor. Whereas during leaf-off season, the point density inside the tree crown and the understorey is

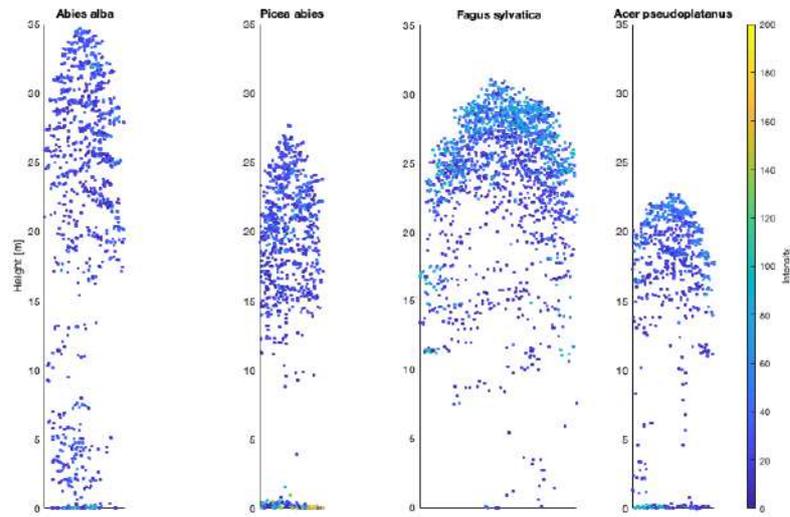


FIGURE 3.2: Examples of leaf-on acquisitions of individual trees. The colour represents laser return intensity.

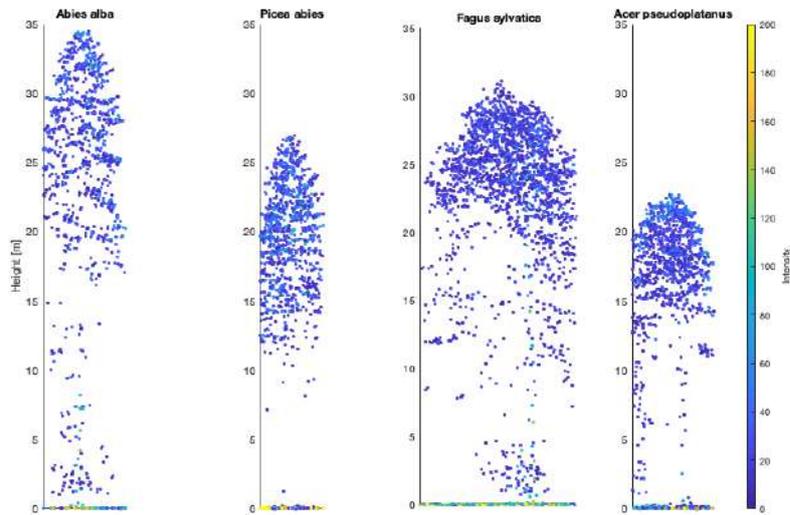


FIGURE 3.3: Examples of leaf-off acquisitions of individual trees. The colour represents laser return intensity.

generally higher. Features related to crown density and seasonal differences thereof are, thus, expected to be important in the classification of tree species.

Spectral reflectance characteristics have also been widely used for tree classification (e.g. Dalponte et al., 2012; Heinzel and Koch, 2012), but they play a secondary role in this study, as the available imagery is limited to four spectral bands.

Furthermore, differences in structural and reflectance characteristics within a class due to local environmental conditions, phenological phase, genetic differences or damage must be taken into account when classifying.

3.4.2 ALS based features

ALS based features were calculated (in MATLAB (2014)) for every sample tree. These are created by the spatial intersection of the 2D crown polygon (ITC) and the terrain corrected 3D point cloud in both leaf-on and leaf-off conditions. From each of these individual tree point clouds, three categories of features were retrieved: 1. features describing the spatial structure and distribution of LiDAR returns (height-related features), 2. features based on return intensity, and 3. features based on the return number (echo category). All of these ALS based features are summarised in Table 3.1.

For most canopy related features, a height threshold of three meters (according to the Swiss NFI threshold for forest) was applied in order to exclude echoes belonging to the understory and forest floor. This notably applies to the first category, where a set of commonly used height metrics was estimated, including statistical summary measures (*max*, *min*, *median*, σ , c_v , γ_1 , κ) and percentiles (Donoghue et al., 2007; Ørka et al., 2009; Latifi et al., 2012; Zhang and Liu, 2013; Torabzadeh et al., 2019). Also, to estimate vertical distribution, tree height was divided into 10 bins of equal height and relative frequency bins were calculated as the percentage of echoes per vertical bin (Korpela et al., 2010).

Similarly to the height metrics, statistical measures and percentiles were retrieved from the intensity values (Donoghue et al., 2007; Kim et al., 2009; Ørka et al., 2012; Shii et al., 2018). In addition, mean values were calculated of only first and last return intensities. While the percentiles are solely based on intensity values and not on height bins as in Torabzadeh et al. (2019), intensity was summarised inside the highest 3 m of the tree crown for all as well as only for first returns.

Multiple returns are often generated by vegetation, where portions of the LiDAR pulse penetrate deeper through gaps in the canopy, reflecting off branches and leaves at various heights or from the ground. This typically leads to 3 to 6 recorded returns per emitted pulse. Hence, additional features were extracted as the proportion of echoes in the different echo categories (fraction of pulse types); single, intermediate, and last of many to the total number of returns (Ørka et al., 2012; Sasaki et al., 2012).

Finally, differences between the features extracted from leaf-on and corresponding leaf-off data were calculated and set as additional features, resulting in approximately 180 ALS-derived features in total.

3.4.3 Image features

Image features are calculated based on the spatial intersection of the LiDAR derived crown polygons and RGBI orthophotos using *R statistics version 3.6.1* (R Core Team, 2019). Visual inspection indicates a strong variation in the spatial congruence of ITC polygons and sunlit treetops (cf. Figure 4.1).

Exploiting only sunlit pixels for feature calculation is highly recommended, as shadowed parts frequently represent lower parts of the canopy and generally suffer from a low signal-to-noise ratio (Leckie et al., 2005; Trier et al., 2018). Therefore, a shadow mask is applied to exclude such pixels from the analysis. This mask is created using image intensity from IHS - transformation, a procedure that transforms RGB bands into an intensity, hue, saturation (IHS) colour system

TABLE 3.1: Overview of ALS derived features. **Abbreviations:** σ : Standard deviation, c_v : Coefficient of variation, κ : Kurtosis, γ_1 : Skewness.

Category	Feature abbreviations	Feature descriptions
Height-related	maxh, minh, medh, stdh, cvh, kurth, skewh	height statistics (of points with $h > 3$ m): <i>max, min, median, $\sigma, c_v, \kappa, \gamma_1$</i>
	hp10 ... hp90, hp99	10 th to 90 th and 99 th height percentiles
	relh, reldia	<i>max height / mean height</i> and <i>crown diameter / max height</i>
	iqrh	interquartile range of point heights
	rfd1 ... rfd10	relative frequency bins: nr. of points per 10% height bin / total nr. of points
	foge	Fraction of ground echoes (lower than 3 m)
Intensity	maxi, mini, medi, stdi, cvi, kurti, skewi	intensity statistics (of points with $h > 3$ m): <i>max, min, median, $\sigma, c_v, \kappa, \gamma_1$</i>
	meaif, medif, meail, medil	mean and median of first and last return intensities (of points with $h > 3$ m)
	ip10 ... ip90, ip99	10 th to 90 th and 99 th intensity percentiles
	int_top, int_topf	cumulative intensity for the top part of the canopy (maximum tree height – 3m) of all and first returns
Return number	fofe, fome, fole, fose	fraction of first/intermediate/last/single returns
	medop	median opacity (return number / number of returns)

(Waser et al., 2011). Intensity corresponds to brightness, saturation to colour purity, and hue to the dominant colour of the pixel. According to visual inspection, an intensity threshold of < 0.14 achieves to exclude most shaded areas while still retaining shade differences inside tree crowns. Besides the four original bands of the RGBI orthoimages, additional features were derived, an overview of which is given in Table 3.2. Since the sample trees are distributed over more than 1000 km², varying illumination and topography can affect their reflectance properties. Relative band values (the individual band divided by the sum of all bands) were computed as additional spectral variables, because they provide an approach to reduce such effects (Waser et al., 2011; Ørka et al., 2012).

The three bands resulting from the IHS - transformation were also included as features (Waser et al., 2011; Heinzl and Koch, 2012; Schumacher and Nord-Larsen, 2014).

Additionally, the normalised difference vegetation index (NDVI) was computed. The NDVI is the most common vegetation index and describes a relation between the near infrared and the red light band. Lastly, four features were derived from principal component analysis (PCA) of the (possibly correlated) original bands (Hill and Thomson, 2005).

From the above described raster layers, the pixel values inside each ITC image segment were aggregated by the median and, in some cases, the coefficient of variation. This allowed to receive 27 object-based spectral features in total.

3.5 Classification procedure

The following steps were conducted using the package *caret* (Kuhn, 2008) within *R* (R Core Team, 2019), which allows for a single consistent environment for training machine learning algorithms and tuning their associated parameters.

TABLE 3.2: Overview of features derived from orthophotos, which contain the image bands red (R), green (G), blue (B) and near-infrared (NIR).

Category	Feature abbreviations	Feature descriptions
Original bands	med_red, med_green, med_blue, med_IR	1. band: R, 2. band: G, 3. band: B, 4. band: NIR
Relative band values	med_red_ratio, med_green_ratio, med_blue_ratio, med_IR_ratio	$R/(R + G + B + NIR)$; $G/(R + G + B + NIR)$; $B/(R + G + B + NIR)$; $NIR/(R + G + B + NIR)$
IHS of RGB	med_I_rgb, med_H_rgb, med_S_rgb	transforms red, green, and blue values into intensity, hue, and saturation
NDVI	med_ndvi	$NDVI = (NIR - R)/(NIR + R)$
PCs 1-4	med_PC1 etc., cv_PC1 etc.	first four principal components of RGBI bands

3.5.1 Class grouping

Mixed temperate forests are usually characterised by a few dominant tree species while many other species only amount to a small number of occurrences. Consequently, it is more challenging to obtain a large enough reference sample of the latter in order to enable adequate classification. In practice, this often means that one can either only classify the dominant species and exclude the rest, or one can perform an exhaustive classification where the rare species are summarised into miscellaneous classes. Such mixed classes decrease model stability and may lead to lower accuracies, as the model is not able to consistently characterise them due to their heterogeneity. However, if rare species are simply ignored in a wall-to-wall classification scenario, they would inevitably be (wrongly) classified as one of the majority species, thus generating a biased impression of the species distribution. Moreover, the proportion of dominant versus non-dominant species is often of interest for biodiversity-related questions. For these reasons, we decided to build and use two models with different class partitions (cf. Table 2.2): The first class grouping ("main species classes") includes only the main tree species, where more than 100 reference trees are available. European beech (*Fagus sylvatica*), Norway spruce (*Picea abies*) and European silver fir (*Abies alba*) each constitute an individual species class, while the maple species (*Acer pseudo-platanus*, *Acer platanoides*, *Acer campestre*), oak species (*Quercus petraea*, *Quercus robur*, *Quercus rubra*) and pine species (*Pinus sylvestris*, *Pinus nigra*, *Pinus strobus*) are grouped at the genus taxonomic level. This partition is used during comparison of input datasets and evaluation of feature importance. The second class grouping ("simplified classes") includes all present trees, but only beech, spruce and silver fir are represented as an individual class, while all of the other species are summarised into either "miscellaneous broadleaf" or "miscellaneous conifer", respectively. This setup is used when classifying test areas for qualitative validation.

3.5.2 Choice of classifier

Recent studies in tree species classification that use large sets of mixed input variables (spectral, texture, geometric, indices), often prefer non-parametric machine learning methods like Random Forest (RF), support vector machines or neural networks (e.g. Dalponte et al., 2012; Fassnacht et

al., 2016; Marrs and Ni-Meister, 2019). Due to a large-scale dataset with imperfect reference data in the present study, high classification accuracies are not to be expected even with an optimal classifier. The characteristic of intuitive derivation of accuracy and variable importance therefore served as key criteria for choice of classifier. Consequently, Random Forest was chosen because of its inbuilt capacity to compute and rank predictor variable importance, its low sensitivity to overfitting, and its ease of implementation (Breiman, 2001; Cutler et al., 2007).

The RF algorithm is based on ensembles of decision trees, where each individual tree is built on a random sample from the dataset with replacement (Bootstrap Aggregation or “bagging”) and a random subset of features. This approach increases variation in the trees grown and prevents strongly correlated trees from occurring, if there are very prominent predictors. Once a forest consists of a (determined) number of trees, class prediction for a new observation is given by the majority consensus based on the classification result of each tree. This aggregate model is more stable and less susceptible to overfitting than a single decision tree. There are two parameters influencing the performance of RF: the number of decision trees (*ntree*) and the number of randomly considered variables at each split in the tree building process (*mtry*). For *ntree*, the default value (500) was selected since values larger than the default are known to have little influence on the overall classification accuracy (Pal, 2005; Duro et al., 2012; Naidoo et al., 2012), but increase computational cost. *Mtry* was optimised using a 10-fold cross-validation.

3.5.3 Managing class imbalances

As shown in Table 2.2, the distribution of tree species in the study area is unbalanced, leading to classes of strongly varying sizes. In supervised classification algorithms including RF, which are constructed to minimise the overall error rate, such class imbalances commonly lead to over-estimation of majority classes (C. Chen et al., 2004). If this imbalance is not taken into account, models become overly adapted to the most frequent species classes and do not generalise well to other species classes. Balanced Random Forests (BRF) and weighted Random Forests are two approaches addressing this challenge and can be easily embedded into the RF algorithm (C. Chen et al., 2004). BRF has previously been applied in tree classification by others (e.g. Ørka et al., 2012) and is implemented in the R package *randomForest* (Liaw and Wiener, 2002). A BRF can either be achieved by artificially enlarging the minority classes (upsampling) or by downsampling the larger classes. We applied the downsampling approach, whereby a subset the size of the smallest class is randomly drawn from each species-class (with replacement) to grow each tree in Random Forest. Overall accuracy metrics of classification are usually lower after downsampling, but there are less false positives in majority classes and more trees predicted as minority classes. The predicted class proportions are expected to lie closer to reality, whereas the individual predictions are less likely to be correct.

3.5.4 Feature selection

Although a large number of input features is generally less challenging for Random Forest as it is for parametric classifiers (Breiman, 2001), it is recommended to apply a form of feature reduction to decrease feature space dimensionality. Feature reduction aims to produce a compact

classification model by removing redundant features. It thereby enhances generalisation and reduces computational cost (Fassnacht et al., 2014; Perez-Riverol et al., 2017). As a trade-off between efficiency and accuracy, the choice of the number of features used in the final model is often subjective and depends on the aim of the study (Ko et al., 2016). Various approaches to feature reduction exist and have been compared in recent years, including methods specifically aimed at tree species classification (e.g. Fassnacht et al., 2014; Waser, 2012; Marrs and Ni-Meister, 2019). These can be subdivided into two categories: Feature selection methods select a subset of the original predictor variables, whereas feature extraction methods calculate new predictor variables that typically summarise the content of several original predictors (Fassnacht et al., 2016). Only the former allow a meaningful interpretation of the selected predictors and were therefore applied in this thesis. Furthermore, feature selection methods also represent a tool for the comparison and ranking of features, enabling deduction of information about a) what type of features from what kind of data source is most valuable in tree species classification and b) about the intrinsic differences among tree species classes in regard to the used data sources. These methods thus indicate which feature types from what data sources are worth extracting for future classifications.

The majority of feature selection approaches can be placed in the categories "filter methods" and "wrapper methods". While filter methods are mostly based on uni- or multivariate correlation, wrapper methods iteratively select features based on classifier performance. For an efficient feature selection workflow it is recommended to apply methods of both categories (Perez-Riverol et al., 2017). A frequently used and versatile wrapper for Random Forest is the backward selection algorithm Recursive Feature Elimination (RFE) (Pullanagari et al., 2018), which is implemented in the R package *caret*. This technique begins by building a model on the entire set of predictors and computing an importance score for each predictor. The least important predictor(s) are removed, the model is re-built, and importance scores are recomputed. In practice, the number of predictor subsets to evaluate as well as each subset's size can be specified in order to reduce the number of models to compute. The subset size optimising the performance criteria is used to select the predictors based on their importance ranking. The optimal subset can then be used to train the final model (Kuhn and Johnson, 2019). During the process of elimination, we implemented 10-fold cross validation to optimise variable selection and to ascertain the standard deviation of the importance estimates. When applied with Random Forest, RFE uses the (mean) decrease in accuracy, which is achieved by randomly permuting values of each feature in the out-of-bag observations, to estimate the feature importance scores. As output, RFE provides these importance scores for each class (and the averaged overall importance) and accuracy metrics for each model built during the process as well as the optimal subset of features.

Many of the features used in this study are correlated because the same features were extracted from both leaf-on as well as leaf-off ALS datasets and additional features based on their differences. As a consequence, the RFE algorithm might select several features, which are all important but also strongly correlated. In order to mitigate this effect and receive a sleek model that can be applied to large areas, a simple Pearson correlation matrix (Perez-Riverol et al., 2017) is used as a multivariate filter to remove some of the most correlated features.

3.5.5 Training and validation sets

Although the RF algorithm already provides an accuracy estimate based on the out-of-bag samples, validation with an independent test set is still recommended if the number of available reference trees allows it (Fassnacht et al., 2016). Since the overall class distribution should be preserved in both training and test set, the data was partitioned using random sampling within each class. Except where declared otherwise, all models in this thesis were trained on 60 % and assessed using the remaining 40% of the data. The reference data consists of a comparably large number (>3000) of sample trees which are regularly distributed over an forest area of 490 km². Training and test set are therefore expected to possess a similar distribution in terms of, for instance, tree height and segment quality.

3.5.6 Accuracy assessment

The primary mean of error assessment for multi-class classification is the interpretation of the confusion matrix (or error matrix), which is a cross-tabulation of the classification predictions against the reference dataset (cf. Figure 3.4). User's accuracy (*UA*) and producer's accuracy (*PA*) were calculated from the confusion matrix to describe performance by class (e.g. Yin and Wang, 2016). *UA* is the number of true positives divided by all positive predictions per class, while *PA* is the number of true positives divided by the actual number of trees in each class. Additionally, the generic scores overall accuracy (*OA*) and kappa coefficient (κ) are provided (Cohen, 1960). These should not be interpreted independently, as they only partially summarise classification performance (Yin and Wang, 2016; Parkan and Tuia, 2019).

		Reference			
		C. 1	C. 2	...	C. N
Prediction	C. 1	$n_{1,1}$	$n_{2,1}$...	$n_{N,1}$
	C. 2	$n_{1,2}$	$n_{2,2}$...	$n_{N,2}$

	C. N	$n_{1,N}$	$n_{2,N}$...	$n_{N,N}$

FIGURE 3.4: Multi-class confusion matrix. **Abbreviations:** C: class, N: total number of classes, n_{obs} : number of observations, $n_{i,j}$: the number of times a tree of class i was assigned to class j .

3.6 Comparison of input datasets

The relevance of the different data sources was assessed by comparing the classification performance achieved with several sets of features that are listed in Table 4.1. The seven "main species classes" (cf. Table 2.2) were used as response classes because the ability to differentiate between individual tree species is the main model characteristic to be tested. Also, homogeneous classes

generally lead to more stable classification results. In order to accurately evaluate the effect of different input datasets on model outcome, other factors such as the number of features as well as the partition of samples into training and test set need to be held constant. Therefore, RFE (cf. Chapter 3.5.4) is performed to reduce the amount of features in each dataset to approximately the same number. For random forest there is often a plateau of good performance for larger subset sizes, usually because unimportant variables are infrequently used in splits and do not significantly affect performance (Kuhn and Johnson, 2019). Figure 3.5 illustrates that this behaviour also applies to the present datasets. Based on this, it was decided to reduce each dataset to approximately 30 features. Another reason for this number is that only 27 features were extracted from the RGBI-orthophotos. On the same training and test sets (60%, 40%), an otherwise identical classification was performed with each reduced feature set using 10-fold cross-validated Random Forest with automatic downsampling (cf. Chapter 3.5). The results are given in Table 4.2.

The significance of differences between all possible pairs of the classification results was then determined using a non-parametric McNemar’s chi-squared test, which focuses on the binary distinction between correct and incorrect class allocations (Foody, 2004). A $p = 0.05$ -level of significance was applied.

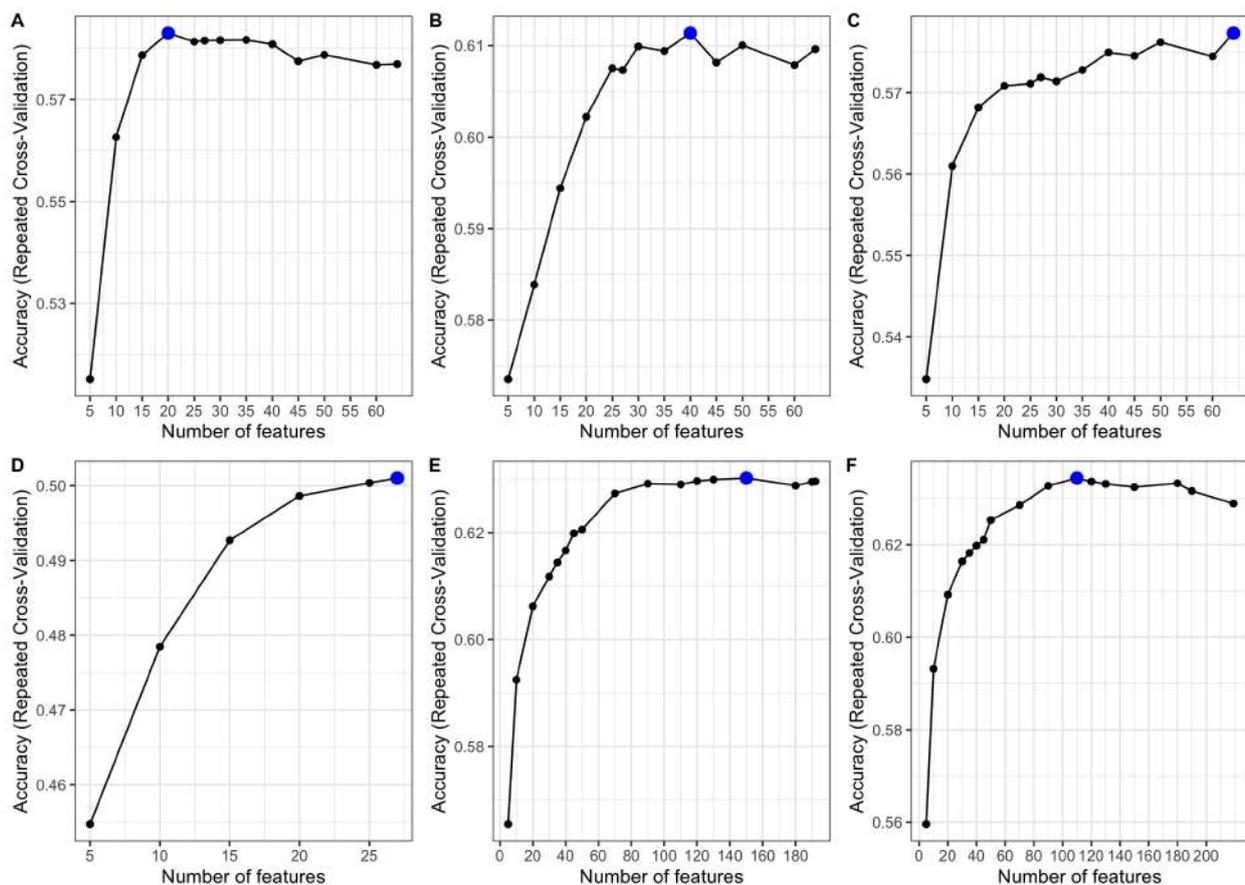


FIGURE 3.5: The results of recursive feature elimination show the relation between the number of features and classification accuracy for each dataset (cf. table 4.1). A: ALS-leafOn, B: ALS-leafOff, C: ALS-diff, D: RGBI, E: ALS-all, F: ALS-RGBI. The given accuracy values were derived using the out-of-bag samples during the RFE iterations. They are therefore not directly comparable to Balanced Random Forest results validated on an independent test set.

3.7 Comparison of individual features

The relevance of individual features (and types of features) was assessed using the feature importance scores produced by RFE (cf. Chapter 3.5.4). Incorporating downsampling (i.e. a Balanced Random Forest approach) into the RFE algorithm would cause unnecessary complexity, so a dataset containing 100 samples from each main tree species class was used as input. The 100 samples were chosen based on minimal height difference between the inventory and ALS measurements, thus reducing the possibility of including mismatches in the classification. We abstained from only using trees with maximal segment quality or number of ALS returns, as the features needed to be evaluated on a dataset of mixed quality, which is the typical scenario for large scale applications (Ko et al., 2016).

To investigate the role and importance of all individual features within the classifier, RFE is performed while applying a 3-times repeated 10-fold cross-validation. The resulting feature importance scores of the 30 RFE runs are aggregated by mean and standard deviation for each class as well as for overall importance.

3.8 Simplified classification model for predictive mapping

Based on the results from dataset and feature evaluation, a simplified classification model was built in order to be applied to a new test area for qualitative validation. For this, a simplified class partition was used, as described in Chapter 3.5.1. Additionally, a binary tree type classification (conifer/broadleaf) was performed for further comparison.

Among the 100 features estimated as most important by RFE (cf. Figure 4.6), about 60% are highly correlated. For the purpose of predictive mapping on a larger scale, the number of features was reduced to 48. This was done using a Pearson correlation matrix, where out of feature groups with an absolute correlation higher than 0.7 only the feature with the highest RFE importance score was selected. Furthermore, only ALS-derived features and no spectral features were used in the predictive mapping model. There are only very few spectral features amongst the most important features (cf. Chapter 5.2) and processing large amounts of aerial imagery would cause an increase in computational cost disproportionate to actual gain in accuracy.

The 48 selected features were used to train a Random Forest model with downsampling and the same parameters as used in Chapter 3.6. The model accuracy was first validated by using 50% of the available reference trees as training and the other 50% as test data. Then, the final classification model was trained on all available reference trees. The selected features were extracted from all tree polygons in the test region and the classification model was applied to predict their class membership.

4 Results

4.1 Individual tree matching

The individual tree extraction algorithm detected 27'316'000 tree segments in the entire Canton of Aargau. Out of 3317 reference trees in the inventory dataset, 3609 were matched to one of the ALS-derived segments (resulting in a detection rate of 0.92). These matched pairs are distributed over 1038 different inventory sample plots, with 3.19 trees per plot on average. Matches with an ALS tree height of more than 50 m were removed from the dataset, as these are outliers mainly caused by power lines, which have not yet been filtered out during the ITC extraction process. Figure 4.1 illustrates the tree positions and crown polygons for four random inventory sample plots. In Figure 4.2, the deviation of ALS-derived from reference tree locations is distributed randomly around (0, 0), which indicates that there is no systematic bias associated with direction. The average distance of the detected tree location (tree tops) from the reference locations (tree stems) is $e_{pos} = 2.17$ m.



FIGURE 4.1: Examples of reference tree positions (red) and their matched ALS tree (yellow) including crown polygon (white) overlaid on aerial orthophotos from 2014.

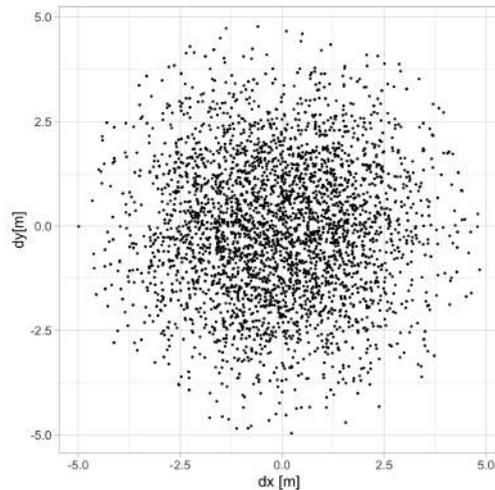


FIGURE 4.2: Positional accuracy expressed as the difference between detected tree location (tree tops) and reference locations (tree stems) in x and y direction.

4.2 Tree height accuracy

The histogram in Figure 4.3 shows that the ALS-derived height of the matched trees has a similar distribution and mean to the field inventory reference data. Nevertheless, the peak at 30 m is more pronounced for ALS estimates, while the inventory measurements are more often higher than that. The individually matched tree heights are shown as a scatterplot in Figure 4.4. While

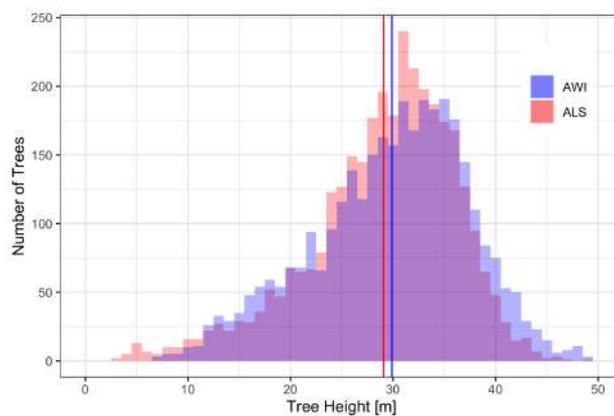


FIGURE 4.3: Height distribution of Aargau inventory trees (AWI) and matched ALS trees. The vertical lines indicate the respective mean values.

the main bulk of trees exhibits relatively small height differences, there are some outliers with a considerable effect on the linear regression fit, leading to an offset of 11.2 m, a slope of 0.64 and a R^2 of 0.38. The average RMSE lies at 6.5 m, which corresponds to about 22% of the average tree height in the study area. In most cases, height differences of more than 10 m are assumed to be caused by mismatches (i.e. not the same tree). The additional robust linear regression fit decreases the weight of outliers, and therefore serves as an approximation to a reference dataset without mismatches. Although the residuals have been artificially reduced in this approach (it is

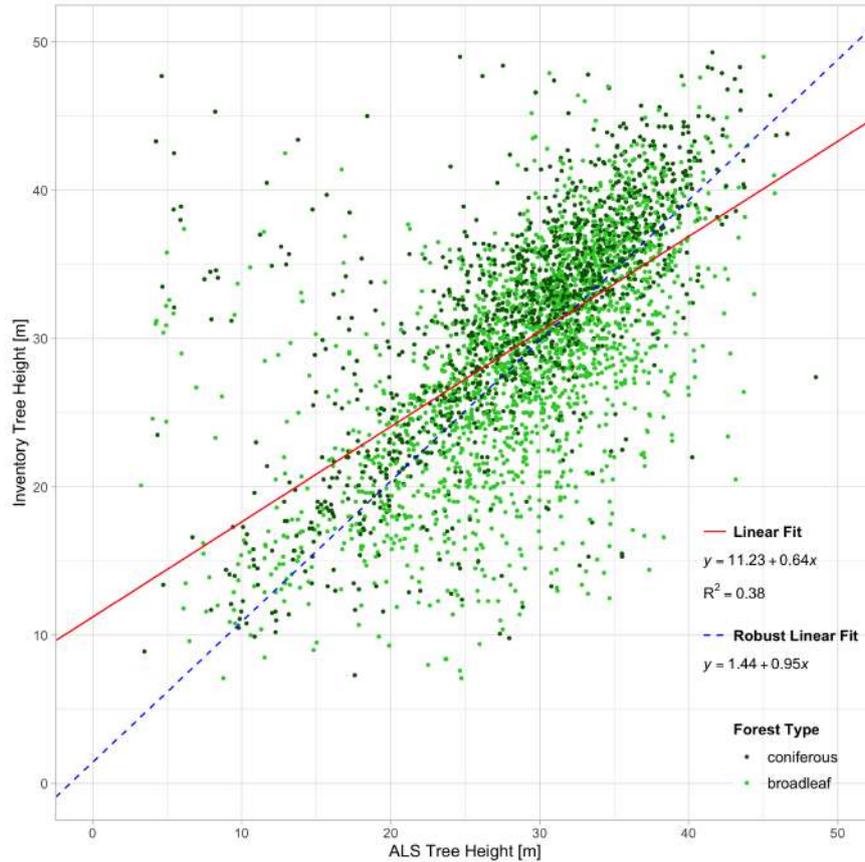


FIGURE 4.4: Scatterplot and linear regression of matched individual tree heights.

therefore not meaningful to calculate R^2), there is still an offset of 1.4 m. This systematic underestimation of tree height is typical for discrete ALS and has been observed in many cases (Gaveau and Hill, 2003; Morsdorf et al., 2004). Figure 4.5 illustrates that height underestimation primarily affects conifer trees, while there is no such systematic tendency for broadleaf species. The mean height differences for conifers lie between -3 and -2.4 m, whereas they are distributed around 0 m for broadleaf trees. This is due to the difference between the steeply-sloped crown shapes of conifers and the flatter, less horizontally variable broadleaf crowns (Disney et al., 2010). Figure 3.2 illustrates an example of this effect: the uppermost tip of the spruce (*Picea abies*) was detected in the leaf-on but not the leaf-off acquisition.

4.3 Classification results

4.3.1 Relevance of input datasets

Table 4.2 provides the results of each RF classification model listed in Table 4.1. The accuracy was assessed as described in Chapter 3.5.6. All of these models were built with 30 features (except RGBI only contains 27) to ensure their comparability. Regarding the ALS datasets, leaf-off features on their own lead to an overall accuracy (OA) of 53%, while leaf-on features only reach 49%. Using only the difference between these datasets (ALS-diff) as input features slightly increases

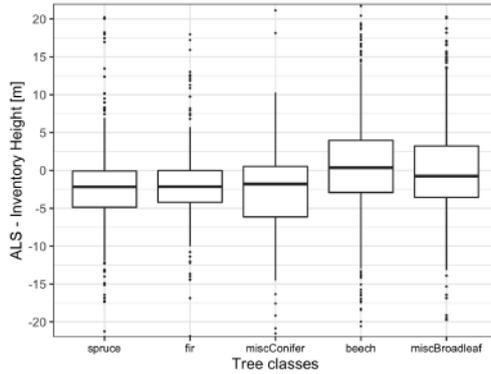


FIGURE 4.5: Difference between matched ALS-derived and inventory tree height, grouped by simplified classes.

TABLE 4.1: Classification datasets comprising different combinations of spectral and structural features. The number of features as selected by RFE and used for classification is given for each dataset.

Dataset name	ALS features	RGBI features	Description
ALS-leafOn	30		Features derived from ALS leaf-on acquisition
ALS-leafOff	30		Features derived from ALS leaf-off acquisition
ALS-diff	30		Differential (leaf-on – leaf-off) ALS features
ALS-all	30		30 most important features selected from all ALS features
RGBI		27	All features derived from RGBI-imagery
ALS-RGBI	27	3	30 most important features selected from all available features

($OA = 50\%$) compared to only leaf-on. A model built on the best 30 features of all ALS data (ALS-all) leads to an increase in OA of 57%. Spectral features (RGBI), though, perform poorly on their own ($OA = 40\%$). Selecting the 30 most important features from all available datasets (ALS-RGBI) leads to the best results ($OA = 58\%$), but the increase in OA compared to ALS-all is only 0.5%. The pairwise McNemar’s test (Table 4.3) shows accordingly that including spectral features makes no significant difference for this classification ($p = 0.6$). There is also no significant difference in correct class allocations between ALS-leafOn and ALS-diff ($p = 0.3$) as well as between ALS-leafOff and ALS-diff ($p = 0.2$). Meanwhile, the difference observed between ALS-leafOn and ALS-leafOff is still significant (though very low; $p < 0.05$). This shows that when used on their own, leaf-off and leaf-on may result in similar classification accuracy, but combining both can increase OA , in this case by about 6%.

When focusing on individual classes, beech and spruce possess the highest user’s and producer’s accuracy in (almost) all classifications. This results from a larger number of training samples available for these dominant species compared to other classes. If class imbalances were not accounted for by applying BRF, this effect would be even more pronounced. Moreover, beech consistently exhibits higher UA than PA , meaning that more beech trees were misclassified as other species than other trees were misclassified as beech.

The confusion matrix of the ALS-RGBI model (Table 4.4) reveals the pairs of classes which frequently confused. Ash is often misclassified as oak (22% out of 79 ash trees are predicted as oak) and beech (18%). Similarly, maple and oak are both most often misclassified as ash (25% and 18%) and beech (23% and 22%). Maple has both the lowest PA (27%) and UA (23%). Silver fir is primar-

TABLE 4.2: Comparison of classification accuracies among the different datasets. Class-specific accuracy is estimated by user’s accuracy (UA) and producer’s accuracy (PA), overall performance by Cohen’s kappa coefficient (κ) and overall accuracy (OA).

Class name	Accuracy	ALS-leafOn	ALS-leafOff	ALS-diff	ALS-all	RGBI	ALS-RGBI
spruce	UA	0.58	0.63	0.64	0.69	0.48	0.69
	PA	0.65	0.60	0.42	0.66	0.40	0.64
fir	UA	0.43	0.59	0.46	0.63	0.37	0.64
	PA	0.48	0.46	0.48	0.51	0.44	0.51
pine	UA	0.10	0.23	0.19	0.28	0.15	0.23
	PA	0.12	0.37	0.33	0.30	0.42	0.30
beech	UA	0.72	0.77	0.75	0.79	0.67	0.78
	PA	0.55	0.60	0.65	0.66	0.43	0.68
maple	UA	0.16	0.13	0.19	0.22	0.12	0.23
	PA	0.25	0.23	0.30	0.25	0.23	0.27
ash	UA	0.25	0.19	0.19	0.24	0.22	0.26
	PA	0.19	0.34	0.31	0.44	0.28	0.43
oak	UA	0.16	0.29	0.24	0.26	0.17	0.25
	PA	0.24	0.39	0.35	0.39	0.24	0.38
	κ	0.338	0.397	0.364	0.450	0.247	0.454
	OA	0.487	0.528	0.504	0.575	0.395	0.580

TABLE 4.3: P-values for the McNemar’s test. P-values < 0.05 indicate a significant difference in the number of correct and incorrect class allocations.

Data source	ALS-RGBI	ALS-leafOn	ALS-leafOff	ALS-diff	ALS-all
ALS-leafOn	1.03×10^{-10}				
ALS-leafOff	3.59×10^{-5}	0.021			
ALS-diff	7.25×10^{-9}	0.348	0.237		
ALS-all	0.668	1.31×10^{-11}	9.33×10^{-5}	2.31×10^{-9}	
RGBI	2.55×10^{-24}	5.56×10^{-7}	1.29×10^{-12}	1.02×10^{-8}	1.03×10^{-21}

ily confused with the similarly-shaped spruce (19%), albeit more with beech (10%) and ash (8%) than with pine (5%). Generally, different broadleaf species are the most challenging to distinguish with this model, especially when there are few reference samples.

4.3.2 Individual feature importance

The ranking of RFE-derived feature importance scores enables the assessment of individual features and their category and data source.

Figure 4.6 shows the overall feature importance scores, which represent the average of all classes. Out of the 30 top-ranked features, 23 are intensity-based, 5 are derived from return number (echo category), and only 2 are height related. Simultaneously, 16 of them are calculated from the difference between leaf-on and leaf-off features, while 13 are based solely on the leaf-off dataset. The most important features, which are based on the seasonal difference in intensity, are the median and mean of first returns (*medif.diff* and *meaif.diff*), upper percentiles (e.g. *ip80.diff*) and the kurtosis (*kurti.diff*). Additionally, the fraction of last returns during leaf-off season (*fome.off*) has a

TABLE 4.4: Confusion matrix for ALS-RGBI classification (based on 30 features). Tree-segments where the model and the inventory reference are in agreement (diagonal) are bold-faced. The total number of trees is 1145, of which 58% were correctly classified.

Classified as	Reference							UA
	ash	beech	fir	maple	oak	pine	spruce	
ash	34	37	12	13	15	4	14	0.26
beech	15	320	16	12	18	7	22	0.78
fir	0	9	80	0	4	6	25	0.65
maple	6	21	3	14	6	1	9	0.23
oak	18	40	7	5	30	2	15	0.26
pine	2	15	8	3	2	13	12	0.24
spruce	4	24	30	5	4	10	173	0.69
PA	0.43	0.69	0.51	0.27	0.38	0.30	0.64	
OA								0.58
κ								0.45

certain relevance as well.

Figure 4.7 depicts the class-specific RFE feature importance scores for the classes beech, fir, maple

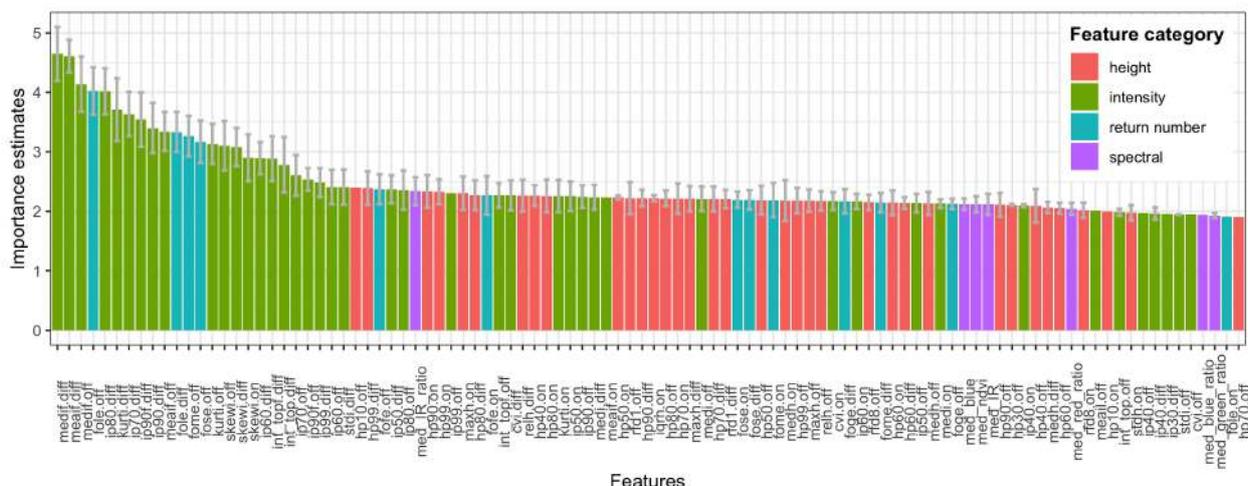


FIGURE 4.6: Overall RFE feature importance estimates of the 100 most important features, coloured by feature category (for description see Tables 3.1 and 3.2). The bar height represents the mean importance derived from 3-times repeated 10-fold cross-validation and the standard deviation is indicated by grey errorbars.

and pine (bar-charts for the remaining classes are provided in the appendix). Next to the generally important features based on seasonal difference in intensity, fraction of intermediate returns during leaf-off season (*fome.off*) seems especially crucial for the classification of beech trees. Tree height normalised by crown diameter (*relh*) also serves as a defining characteristic for beech. For the classification of fir trees, the fractions of last, intermediate and single echoes in leaf-off state (*fole.off*, *fome.off*, *fose.off*) are key descriptors. The feature *rfd1.off*, which is a measure for point density in the lowest height bin, also appears to be of certain relevance.

Compared to other classes, 15 of the 30 top-ranked features for the maple class consist of height-related features, especially seasonal difference in upper height percentiles (*hp90.diff*, *hp80.diff* etc.). The pine class contains the most spectral features among the 30 top-ranked features (6), of which the normalised infrared band (*med_IR_ratio*) ranked highest, followed by the similar NDVI value (*med_ndvi*).

In summary, the importance scores show that varying types of features are relevant to classify specific tree classes, while features based on echo intensity and return number make up the majority of top-ranked features.

4.3.3 Predictive mapping

Table 4.5 shows the confusion matrix of the simplified classification, which was trained on one half and validated on the other half of the reference data. An overall accuracy of 62% and a kappa coefficient of 0.49 were reached. Similarly to the observations in Chapter 4.3.1, the miscellaneous broadleaf class is most often confused with beech (24% of beech trees are predicted as misc. broadleaf) and vice versa (23%). Contrarily, spruce is also most often misclassified as miscellaneous broadleaf (13%) and beech (9%). Beech reaches the highest producer’s and user’s accuracy, while the misc. conifer class only reaches a *PA* of 24% and an *UA* of 29%. This means that reference trees of misc. conifer are classified almost randomly into one of the 5 available classes. For comparison, a binary forest type classification (conifer or broadleaf) was also performed with the same dataset and features. An overall accuracy of 83% and a kappa coefficient of 0.64 were achieved.

TABLE 4.5: Confusion Matrix for simplified classification (based on 48 features). Tree-segments where the models and the inventory reference are in agreement (diagonal) are bold-faced. The total number of trees is 1517, of which 62% were correctly classified. Classes: beech (*Fagus sylvatica*), spruce (*Picea abies*), fir (*Abies alba*), miscellaneous Broadleaf (remaining broadleaf species), miscellaneous Conifer (remaining conifer species).

Classified as	Reference					<i>UA</i>
	miscB.	beech	fir	miscC.	spruce	
miscB.	214	135	26	23	42	0.49
beech	80	392	11	16	30	0.74
fir	8	8	103	12	25	0.66
miscC.	15	14	10	24	20	0.29
spruce	15	23	39	23	209	0.68
<i>PA</i>	0.64	0.69	0.54	0.24	0.64	
<i>OA</i>						0.62
κ						0.49

In addition to model evaluation with reference data, the simplified classification was applied to approx. 80 ha of forest near the town of Bremgarten, to qualitatively evaluate the prediction of a larger area. Figure 4.8 shows the results for the full extent of this test area. Grey polygons

indicate trees that were excluded from classification, because they are smaller than 4 m and/or contain less than 50 ALS echoes. Pure spruce stands, appearing as blue areas, were consistently well detected, independently of height. Large broadleaf trees were mainly classified as beech, which is very plausible, as the old growth forest in this area is dominated by beech. The large misc. conifer polygons scattered among those beeches possibly indicate larch (*Larix decidua*). A majority of smaller trees were predicted as misc. broadleaf, which is plausible, as these young stands are often dominated by maple. However, an underestimation of the fraction of beech trees in this area is probable, due to the frequent misclassification between beech and misc. broadleaf (cf. Table 4.5).

Figure 4.9 shows a detail located in the northeastern corner of the map in Figure 4.8 including train tracks. This image exemplifies that the train's catenary masts can sometimes be mistaken as trees by the ITD method. In this case, they were not classified, as they presumably contain less than 50 echoes. The forest in this area is characterised by both spruce stands with several other conifers at the edges and stands of mostly young broadleaf trees.

Figure 4.10 shows a detail located in the southwestern corner of the map in Figure 4.8. A visual comparison with the orthoimage reveals that broadleaf and coniferous trees can be reliably discerned. As compared to other areas, there are several non-beech broadleaf trees with relatively large crowns.

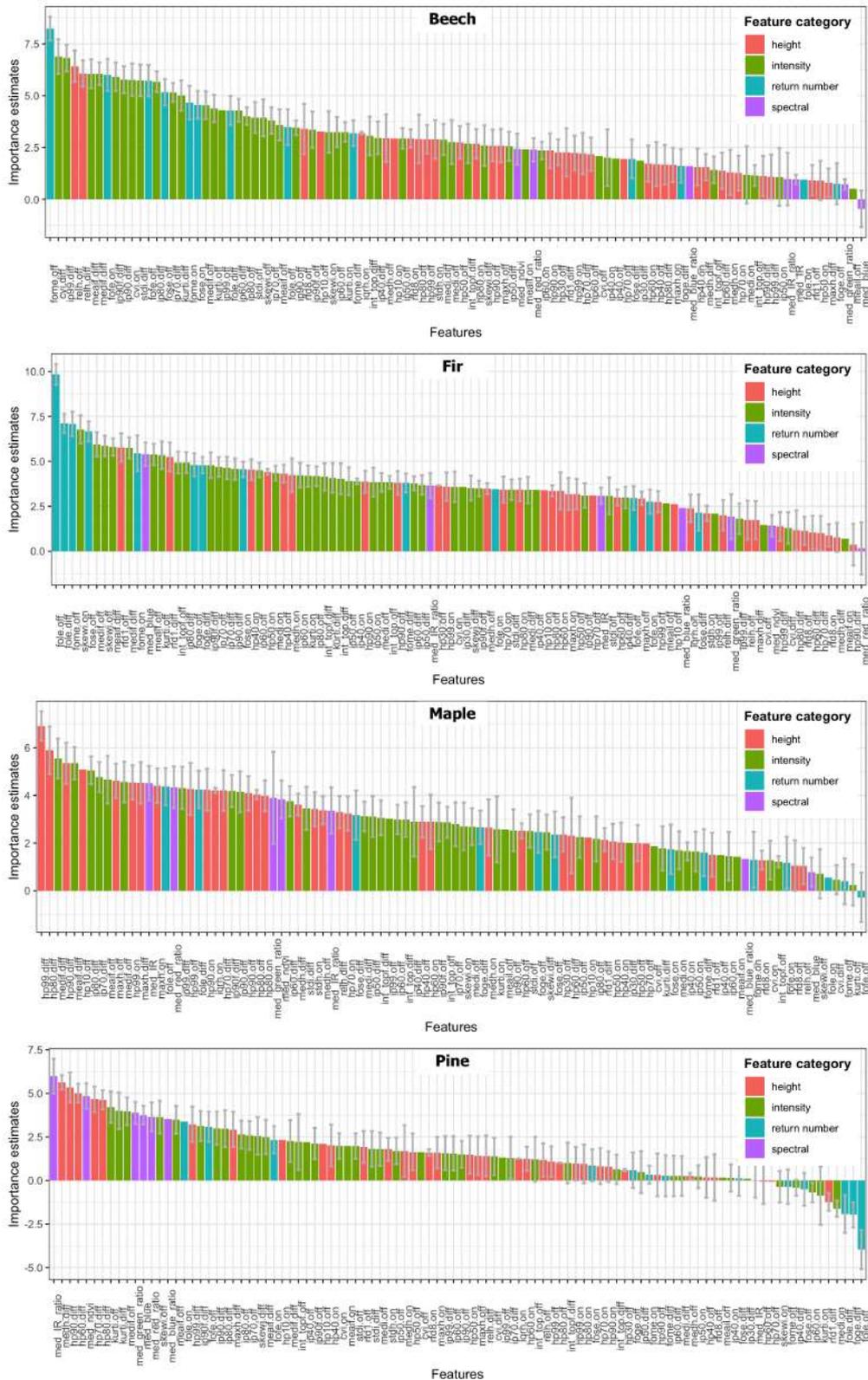


FIGURE 4.7: Class-specific RFE feature importance estimates of the 100 most important features for each of the classes beech, fir, maple and pine. The bars are coloured by feature category (for description see Tables 3.1 and 3.2). The bar height represents mean importance derived from 3-times repeated 10-fold cross-validation and the standard deviation is indicated by grey errorbars.

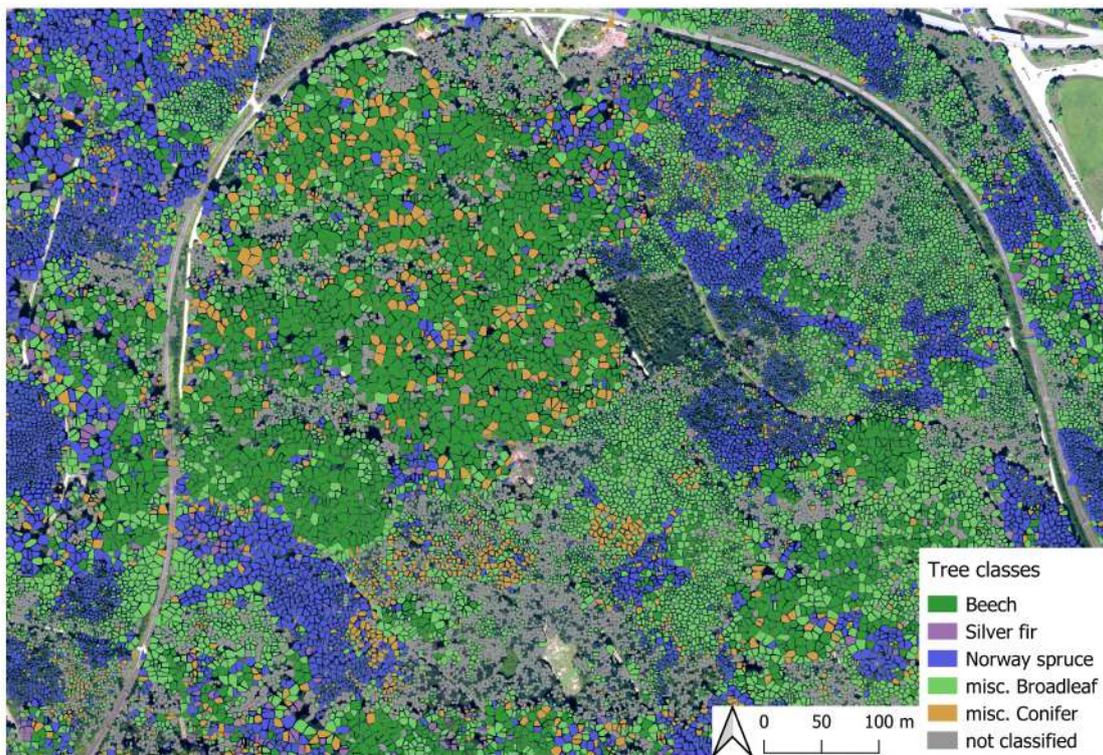


FIGURE 4.8: Simplified classification applied to a test area near Bremgarten. The approx. centre coordinates are X:666'500 and Y:245'110. The background is a histogram equalised RGB orthophoto.

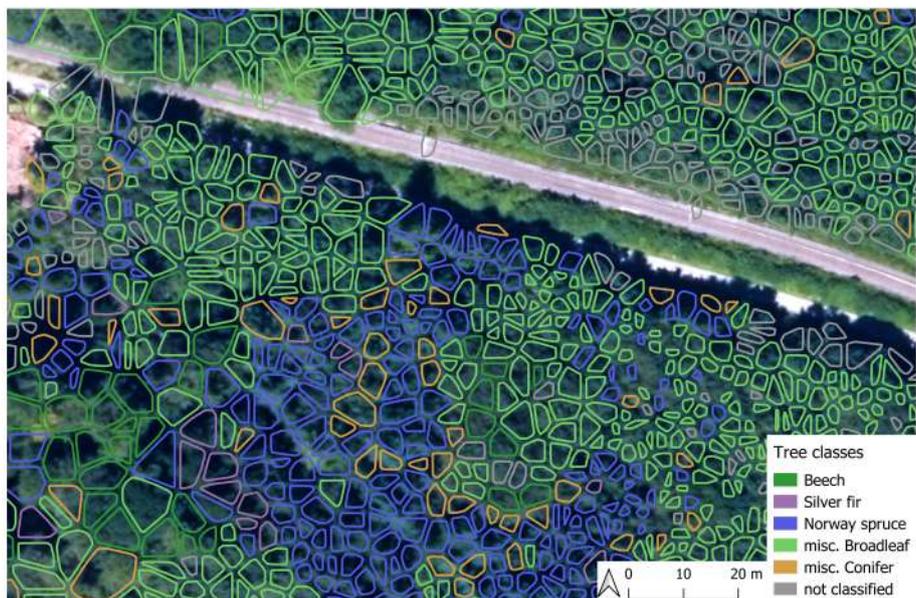


FIGURE 4.9: Detail of Figure 4.8 including train tracks, northeast. The approx. centre coordinates are X:666'620 and Y:245'330.

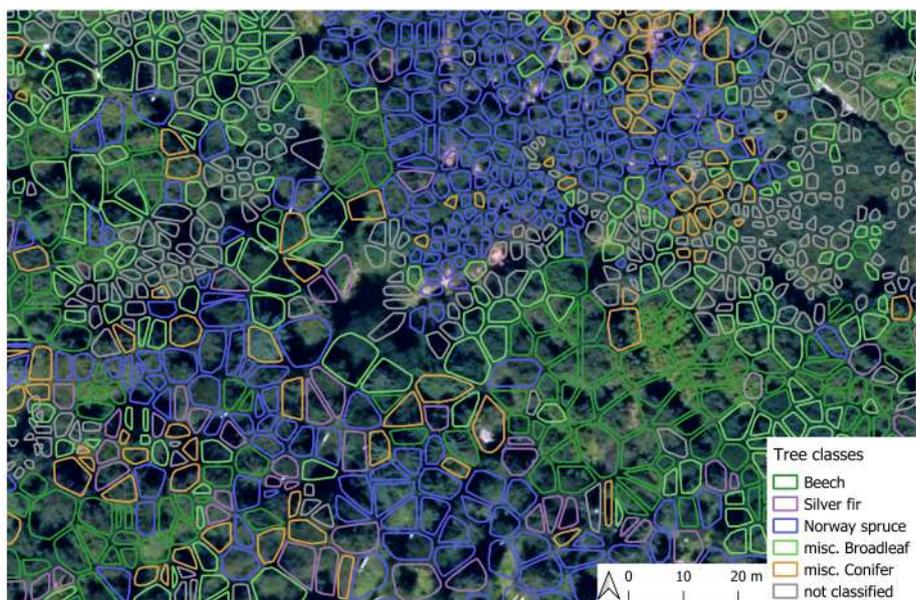


FIGURE 4.10: Detail of Figure 4.8, southwest. The approx. centre coordinates are X:666279 and Y:244766.

5 Discussion

This chapter follows the workflow in Figure 5.1, indicating uncertainties or limitations associated with this large-scale approach next to the corresponding data products or processing steps.

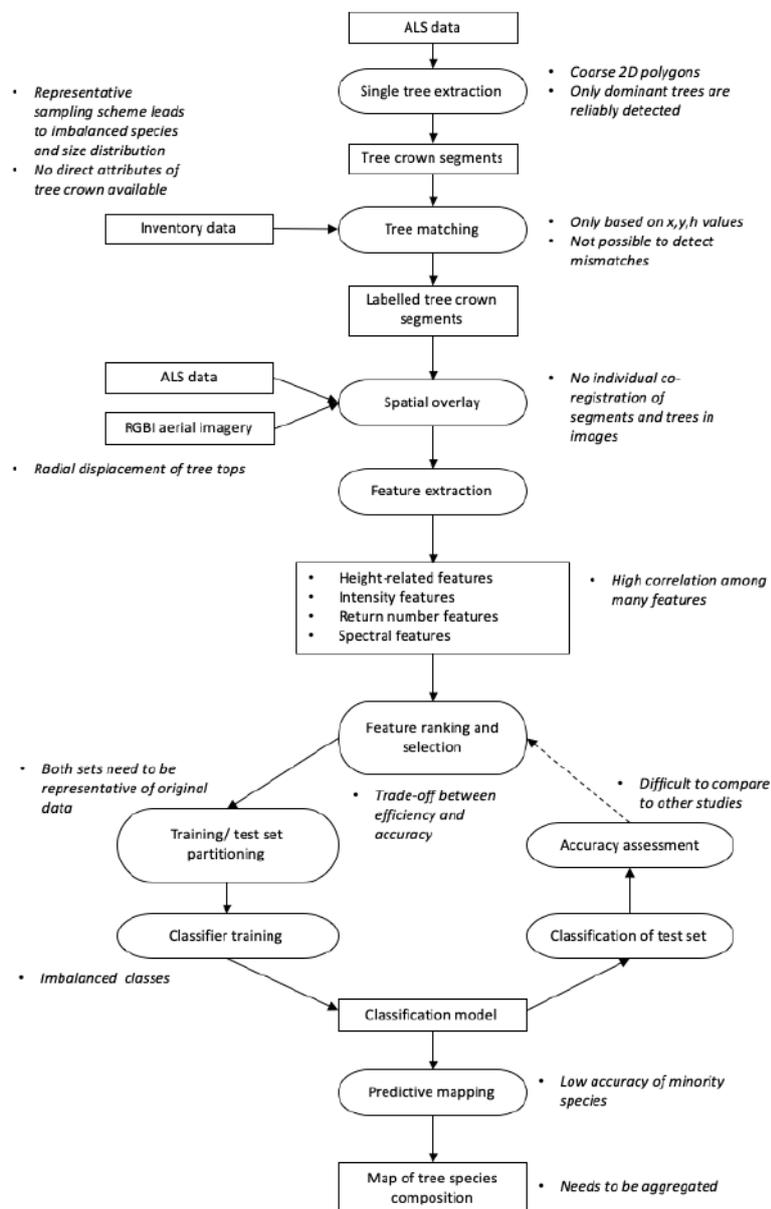


FIGURE 5.1: Processing and classification workflow. Uncertainties or limitations associated with each step are annotated in italics.

5.1 Compatibility of inventory data and ALS tree segments

Due to the large study area (1403 km²), individual trees were extracted as 2D crown polygons by applying a CHM-based watershed segmentation. 3D alpha-shapes, which would provide a more detailed characterisation of tree shape (Korpela et al., 2010; Yao et al., 2012), were therefore not available. Figure 4.10 illustrates that most polygons possess around 5 to 10 vertices, and thus do not represent exact crown outlines. However, when the polygon is intersected with the point cloud, the crown shape is often discernible (see Figure 3.2 for example). As an inherent property of CHM-based treetop detection techniques, subdominant tree crowns are often not detected and delineated, since they are (partially) occluded by dominant tree crowns and less prominent in the CHM. Thus, the available ALS segments essentially represent dominant and co-dominant trees (Harikumar et al., 2019). Although not every tree might be detected, an advantage of object-based classification over pixel-based is that no other land cover type is mistaken as one of the tree classes in the final product. The only exception for this are power supply and railroad masts, as visible in Figure 4.9, which were removed in the reference dataset.

In order to provide ALS segments with tree species labels, inventory records in the style of the Swiss national forest inventory were used. According to Fassnacht et al. (2016), reference or training and validation data has to fulfil certain criteria, the first being that data should be representative for the site under investigation and match the spatial scale. The use of operational inventory data is advantageous in this regard, since it is designed to be a representative sample of the forest in the respective region (in this case Canton of Aargau), and it contains measurements of individual trees from uniformly distributed sampling plots. As a consequence of representative sampling, tree species proportions in the reference data follow natural occurrence. We have shown that despite adapting the classifier to imbalanced classes, this training set does not allow to accurately classify tree species representing less than 10% of total trees.

Another criteria mentioned by Fassnacht et al. (2016) is that observation errors should be known and their impact on the results should be discussed. While errors in species label are not expected in the field inventory data, the matching algorithm represents the main source of uncertainty. For the creation of reference data, it is fundamental that the inventory label and ALS measurement of the exact same tree are combined. Although the algorithm found a match for 92% of reference trees, it was not possible to verify how many of those are actually the same tree, due to the spatial scale and number of reference trees. In small-scale experiments, mismatches are usually excluded by visually verifying each sample tree based on field measurements and very high resolution aerial photographs (Shii et al., 2018) or even by recording tree positions from the field directly onto remote sensing images (Dalponte et al., 2012).

An indication for the presence of mismatches in our reference data is that the overall accuracy of the binary conifer/broadleaf classification was 83%. This is comparably low, as 95-99% have been reached previously when using four-band aerial images (Waser et al., 2017). Also, there are around 100 trees where the height difference between ALS and inventory measurement amounts to more than 10 m (these were excluded in the classification). The fact that there was no directional bias in position indicates mismatches to be caused by the structure of individual trees rather than by systematic registration errors. Due to the different perspectives of airborne acquisitions and field

inventory, position and height are momentarily the only metrics on the basis of which individual trees can be matched. This is associated with several uncertainties such as the underestimation of conifer height by ALS. Also, in tree detection from ALS data, position is defined by the local maximum of the canopy, while in field inventory, the tree stem position is recorded. If a tree is leaning or shows an asymmetric growth pattern, tree top and stem do not have to be directly above each other. In addition, the accuracy of GPS measurements of inventory plots centres influences the positional accuracy of all trees inside the corresponding plot (Maltamo et al., 2009).

Despite these uncertainties, the automated matching of field inventory and ALS data on individual tree level is an important step in order to generate species labels and other (species-specific) parameters on large scales. If the discussed challenges are addressed, existing inventory sources can be utilised without requiring additional data collection or manual delineation and matching steps

5.2 Combining two ALS acquisitions and aerial imagery

In this study, we performed a classification based on features from only leaf-on ALS data, only leaf-off, the difference of those two, all ALS-derived features combined, only spectral features from orthophotos and the most important features selected from all of the datasets (cf. Table 4.1).

5.2.1 Leaf-off and leaf-on ALS datasets

The overall accuracy when using only leaf-off was just 3% higher than leaf-on (McNemar's $p = 0.02$). In other studies (with fewer classes), significantly higher accuracies were achieved with leaf-off than with leaf-on. For instance, Ørka et al. (2010) reported an *OA* of around 90% for leaf-on and 97% for leaf-off for three species (spruce, birch, aspen). Focusing on intensity features, Kim et al. (2009) also increased the separability of various broadleaved and coniferous species by around 10% when using leaf-off instead of leaf-on. However, Shii et al. (2018) observed no statistically significant difference in tree species mapping accuracy between the use of leaf-on and leaf-off features, whereas combining both datasets increased their overall accuracy from 58.2% (leaf-on) and 62.0% (leaf-off) to 66.5%. Our study using similar species classes produced a comparable effect, where *OA* was increased from 52% (leaf-off) to 57.5% (ALS-all), while keeping the number of features constant. Opposite to Shii et al. (2018), we also included the differences between leaf-on and leaf-off as features (Torabzadeh et al., 2019), which on their own performed similar to leaf-on. These results indicate a combination of LiDAR data acquired at different phenological stages to be particularly beneficial when aiming to classify mixed forests featuring 5-10 tree species.

5.2.2 RGBI orthophotos

Feature selection has shown that spectral features only marginally contribute to the classification models tested in this thesis. They are therefore not included in the simplified classification. Nevertheless, multispectral imagery has previously been successfully used in tree species classification, be it aerial photographs with three to four bands (e.g. Waser et al., 2011) or multispectral satellite imagery (e.g. Ke et al., 2010; Zhang and Liu, 2013). In the last few years, the best results in

tree species classification have been achieved by combining ALS with imaging spectroscopy data (Heinzel and Koch, 2012; Fassnacht et al., 2016; Shi et al., 2018; Trier et al., 2018; Parkan and Tuia, 2019). A high number of spectral bands allows to distinguish different broadleaved species, such as oak and maple, based on reflectance. However, spatial coverage of imaging spectroscopy data is still limited compared to multispectral or RGB aerial imagery, as the latter is often acquired wall-to-wall by national agencies (Waser et al., 2017). The high spatial resolution of aerial imagery (0.25 m in our case) can also be used as a source of textural information (Heinzel and Koch, 2012; Zhang and Liu, 2013; Schumacher and Nord-Larsen, 2014). This indicates a general expedience in combining ALS data with aerial imagery for tree species classification and it shows that the low relevance of spectral features in our case does not originate from the use of unsuitable data (Xu et al., 2015).

When integrating LiDAR and optical data, some studies perform individual tree segmentation based on layers from both data sources (Waser et al., 2011; Zhang and Liu, 2013). In comparison, our segments are derived solely from ALS data using a watershed method. Consequently, these tree polygons are not necessarily congruent with all tree crowns in orthoimages, when spatially overlaid. Although the images are true orthophotos, trees can still exhibit radial displacement, as is visible in Figures 4.1 and 4.10. Even if the tree crown in the image and the ALS-derived polygon partially overlap, it is still probable for this intersection to mainly cover shaded areas, as the ALS polygons are often smaller than the whole tree crown. These issues are particularly relevant for large scale applications, as the data could have been acquired under different geometric circumstances (angle, direction, height etc.). To further address these challenges, a tree top based registration of aerial images to ALS data could be applied in the future, for instance by following the approach presented by Lee et al. (2016).

5.3 Evaluation of ALS features

Figure 5.2 presents a selection of features of different categories, which achieved high importance scores during the RFE process. Features derived from return number were frequently found among most species classes' the top-ranking features. Proportions of single, last and intermediate echoes have been successfully used before. Ørka et al. (2012) combined them with density features and received an accuracy of 74–77% in the classification of pine, spruce and deciduous trees in Norway. Sasaki et al. (2012) also emphasised these features' use in discriminating trees with different canopy structures, because denser canopy reduces penetration depth and leads to fewer multiple returns per laser pulse. Parkan and Tuia (2019) call them opacity or echo rank distribution metrics and relate these features to branch size (diameter) distribution within a tree.

The boxplots of features *fome.diff* and *fome.off* (Figure 5.2) indicate beech trees to be characterised by a high fraction of intermediate returns in leaf-off state. This high number of pulses with multiple returns (i.e. an increased proportion of partial hits) could be explained by beech having many narrow and pointy twigs, as compared to e.g. sparser but thicker branches of ash (for backscattering mechanisms see Korpela et al., 2013). Comparatively, silver fir exhibits lower values for *fome.off* than Norway spruce. Silver firs possess flat needles and horizontal twigs while spruce needles are arranged spirally around the often hanging twig. Such morphological differences also

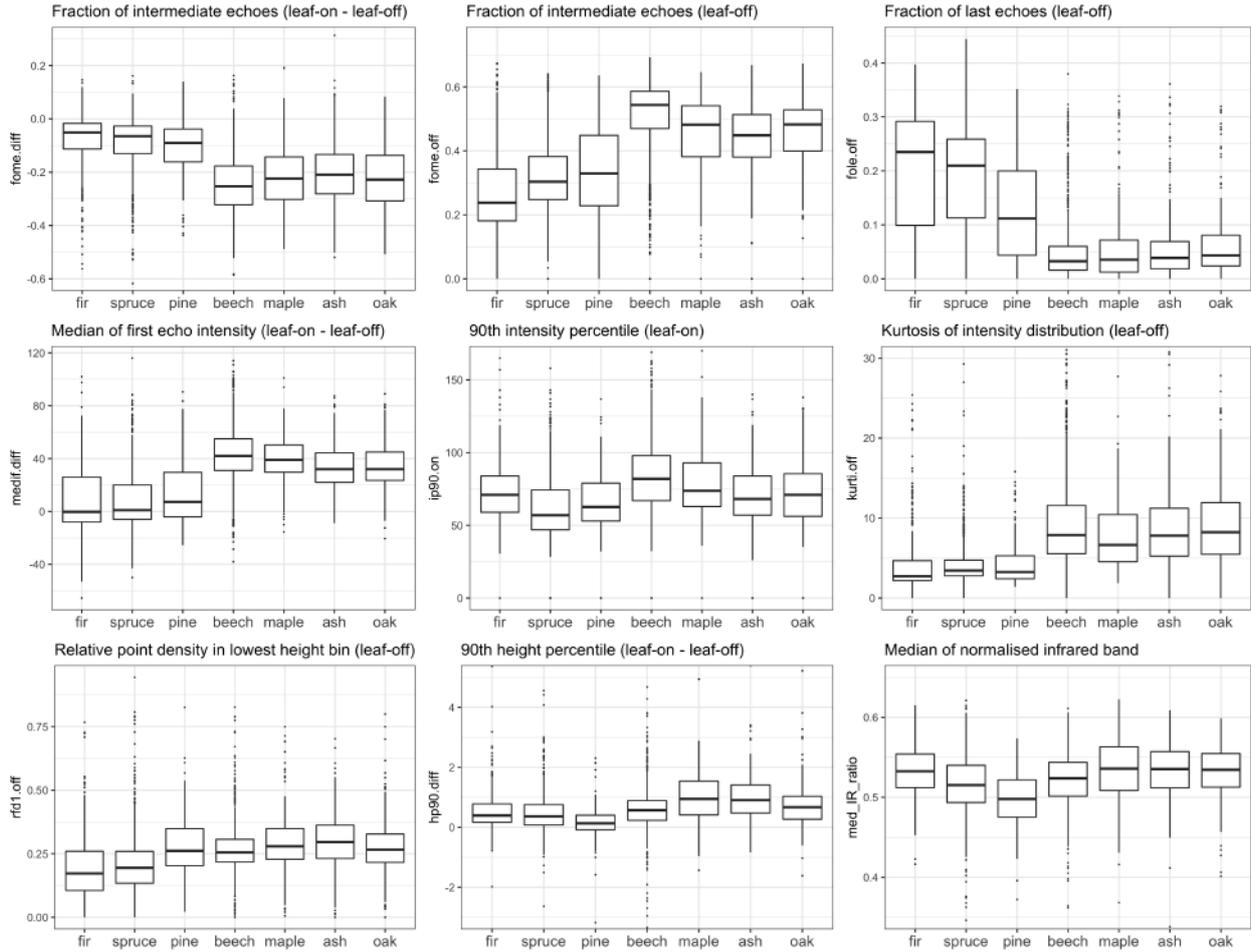


FIGURE 5.2: Boxplots of exemplary features per class.

seem to be reflected in the observed patterns of intensity features (e.g. *medif.diff* and *ip90.on*).

The RFE scores (e.g. Figure 4.6) indicate features derived from echo intensity to play an important role in the classification of all present species. This is in accordance with findings previously reported by Ørka et al. (2009), Vauhkonen et al. (2010) and Parkan and Tuia (2019). Intensity describes some properties related to the reflectivity in the wavelength of the emitted pulse (1550 nm in this case) but has also been found to be affected by structure, such as the sizes, shapes and arrangements of leaves and branches (e.g. Korpela et al., 2010; Vauhkonen et al., 2014).

Most broadleaved trees produce higher average intensities than conifers during foliated stage and lower intensities during leaf-off season (Ørka et al., 2009). Consequently, median or mean of first and single echo intensity has repeatedly been reported as one of the most important features (e.g. Kim et al., 2009; Vauhkonen et al., 2010; Shii et al., 2018). Achieving highest overall RFE scores for *medif.diff* and *medif.off*, our results corroborate these findings.

Beech exhibits the highest values in *medif.diff*, presumably due to its dense summer foliage and frequent multiple scattering during leaf-off. As first returns are usually triggered by the outer crown layers, these features effectively characterise the structural differences influencing the opacity or reflectivity of top crown layers (Shii et al., 2018). The kurtosis of the intensity distribution inside each tree (*kurti.off*) indicates that conifers produce similar intensity values throughout the

crown, while broadleaves exhibit a higher variation during leaf-off.

While features derived from a combination of return number and intensity also contain geometric information, features describing the height distribution of echoes proved to be less important. Normalising height-related features by tree height probably increases their relevance (Shii et al., 2018), but Parkan and Tuia (2019) observed no difference in performance when using only scale invariant features. Tree height is not the only cause of intra-species variability, as shape variability, especially due to crown plasticity, is strongly influenced by e.g. age, competition or environmental conditions (Jucker et al., 2015). Consequently, the most important feature derived from height distribution, point density on ground level (*rfd1.off*), is not related to tree shape but rather to crown density.

In summary, this suggests that small scale structures such as foliage arrangement represent the most consistent species-specific traits characterisable by ALS data. Nevertheless, structures influencing intensity and scattering mechanisms can be altered by e.g. tree health, vines and epiphytes (such as ivy or mistletoe) and differences in environmental niche including temperature or soil condition.

5.4 Predictive mapping

For less common species like maple or ash, prediction accuracy ranges mostly between between 22 and 30%. This leads to unreliable results when applying the classification to a larger area. Therefore, we devised a simplified classification, using 48 ALS features, with the aim to separate the three most common species (European beech, Norway spruce, silver fir) from the remaining broadleaved and coniferous species. The model produced an *OA* of 62%, which is a certain improvement compared to the classification with seven main species classes (*OA* = 58%). In general, the accuracy increases with a decreasing number of classes, as was for instance observed by Heinzel and Koch (2011). Their *OA* increased by ca. 20% when limiting the classification depth to cover four instead of six species. The smaller effect in our case can be explained by the inclusion of "miscellaneous classes". These are difficult to classify, particularly the class misc. conifer, because they contain comparably few training and validation samples while exhibiting high heterogeneity. For example, misc. conifer contains both Scots pine and European larch, which differ strongly in morphology and phenology. However, if these species were simply excluded from the training set, the results for the main species would be distorted by these minority species, when applying the model to a continuous area of mixed forest.

When the classification is used to predict a larger area (as illustrated in Figure 4.8), the resulting map cannot provide reliable information concerning individual trees. However, the prediction may be used to assess the spatial distribution and proportions of dominant versus sub-dominant tree species in an area. This can be applied as pre-stratification when aiming to produce species-specific parameter estimations of e.g. forest basal area, volume, or biomass (Korpela and Tokola, 2006; Vauhkonen et al., 2014). Interpretation of these predictive mapping results needs to take into account that they only provide information about dominant trees, i.e. small and/or understorey trees are systematically omitted, as they are not fully detectable by ALS.

5.5 Outlook

An important step in improving the classification results is to increase the reliability of the connection between ALS segments and field inventory records, thus reducing uncertainty in the reference data. As ALS data mainly delineates upper canopy layers, future inventory measurements of crown attributes, such as diameter, shape, volume or position of tree crown, could be beneficial. Close-range laser scanning, such as TLS (terrestrial laser scanning) and UAVLS (ultra-light aerial vehicle laser scanning), may have the potential to bridge this gap between airborne and ground measurements (Morsdorf et al., 2018).

More reliable reference data would also lead to an increase in the amount of usable reference samples for less common species. A reduction of intra-class variance can then lead to more accurate classification results for these classes. Also, better co-registration of tree segments and multispectral imagery, or even the use of imaging spectroscopy, would particularly benefit the distinction of broadleaved species. With these improvements, it should be possible to detect, for instance, European ash (*Fraxinus excelsior*) with reasonable accuracy, whereas less common or pioneer species, such as wych elm (*Ulmus glabra*) or black alder (*Alnus glutinosa*) would need a more specialised reference dataset. The assessment of classification accuracy would also profit from a validation on a completely independent reference site, ideally with manually labelled segments.

As this thesis' main focus was to explore the potential of different datasets and features, Random Forest was used as classifier due to its transparency and ease of implementation. However, for such large amounts of training data, deep learning approaches such as convolutional neural networks (CNN) should be considered in the future as well (Zhu et al., 2017). Deep learning may improve flexibility of species classification across different biogeographical regions and has frequently been applied to hyperspectral imagery (Trier et al., 2018; Fricker et al., 2019).

Despite limited classification accuracies, main tree species composition, as estimated by the approach presented in this thesis, can be aggregated from tree level to any larger unit. This versatility allows the results to be used as input for any analysis where species-specific pre-stratification might be beneficial. Also, summarising the results obtained from our study to stand level would probably improve the accuracy of species composition estimates (Ørka et al., 2013).

On the condition that leaf-off and leaf-on ALS acquisitions are available, this classification approach could theoretically be transferred to other temperate forest regions, as it does not require any site-specific parameters. The main restraint though is the availability of suitable reference data. For other areas in Switzerland, NFI data presents an option, as it is acquired with the same methods as the inventory data used in this study. The low density of NFI sampling plots, however, might cause difficulties in gathering sufficient reference data, especially if the classification is to be trained on a smaller area than our study site. To achieve a sufficient number of reference trees, a more reliable matching technique or a manual verification of reference data would have to be performed in that case. Also, the cantonal inventory used here was conducted during one year, while the NFI is designed as a continuous inventory system, resulting in plots being sampled in different years. Lastly, the large number of labelled individual tree point clouds that we gained for our study area could also be used as training set for estimating other parameters such as biomass on an individual tree level.

6 Conclusion

We have shown that large-scale object-based tree species classification is possible with an overall accuracy of around 60% when solely based on ALS features. In comparison to pixel-based approaches, individual tree based analysis has the advantage of being able to directly match the tree segments to trees recorded in the field. The availability of representative field measurements (such as the NFI) is crucial in order to train tree species or genera classification models over large areas. However, we found the automated creation of reliable reference data to be one of the main challenges, as the matching results of over 3000 inventory trees could not be validated directly. This uncertainty in training and validation data therefore needs to be taken into account when interpreting the classification accuracy. In addition, the classification results indicate that an automated approach with representative reference data allows to accurately classify dominant species, if a sufficient number of homogeneous training samples is available.

Regarding the relevance of input datasets, combining features derived from leaf-off and leaf-on ALS acquisitions significantly improves classification accuracy. When used separately, the datasets produce comparable results, while leaf-off performs slightly better in discriminating the most common species. Despite the coarse shape of our tree segments, many features found to be important in previous small scale studies were proven to be transferable to a representative sampling over a larger scale. Tree shape is subject to high intra-species variability and is generally not well captured by a CHM-based ITD approach. Features related to tree shape are therefore less important than features related to radiometric properties and scattering mechanisms caused by the arrangements and shapes of foliage and small branches.

Characterising mixed temperate forests is a challenging task when limited to datasets available over large spatial scales. However, applying a 5-class classification to a test area lead to promising results in terms of using aggregated species composition to estimate species-specific forest attributes.

References

- Barrett, F., McRoberts, R. E., Tomppo, E., Cienciala, E. & Waser, L. T. (2016). A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. *Remote Sensing of Environment*, 174, 279–289. <https://doi.org/10.1016/j.rse.2015.08.029>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, C., Liaw, A. & Breiman, L. (2004). *Using Random Forest to Learn Imbalanced Data* (tech. rep.). Dept. Statistics, Univ. California Berkeley. <https://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf>
- Chen, Q., Baldocchi, D., Gong, P. & Kelly, M. (2006). Isolating individual trees in a savanna woodland using small footprint lidar data. <https://doi.org/10.14358/PERS.72.8.923>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783–2792. <https://doi.org/10.1890/07-0539.1>
- Dalponte, M., Bruzzone, L. & Gianelle, D. (2012). Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment*, 123, 258–270. <https://doi.org/10.1016/j.rse.2012.03.013>
- Dalponte, M., Ørka, H. O., Gobakken, T., Gianelle, D. & Næsset, E. (2013). Tree species classification in boreal forests with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 2632–2645. <https://doi.org/10.1109/TGRS.2012.2216272>
- Departement Bau Verkehr und Umwelt. (2018). Zustand und Entwicklung des Aargauer Waldes. Ergebnisse der 2. Aargauer Waldinventur 2016. *Aarau: Departement Bau, Verkehr und Umwelt, Abteilung Wald*.
- Disney, M. I., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S. & Pfeifer, M. (2010). Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. *Remote Sensing of Environment*, 114(7), 1546–1560. <https://doi.org/10.1016/j.rse.2010.02.009>
- Donoghue, D. N., Watt, P. J., Cox, N. J. & Wilson, J. (2007). Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment*, 110(4), 509–522. <https://doi.org/10.1016/j.rse.2007.02.032>
- Duncanson, L. I., Cook, B. D., Hurtt, G. C. & Dubayah, R. O. (2014). An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosys-

- tems. *Remote Sensing of Environment*, 154, 378–386. <https://doi.org/10.1016/j.rse.2013.07.044>
- Duro, D. C., Franklin, S. E. & Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, 118, 259–272. <https://doi.org/10.1016/j.rse.2011.11.020>
- Fassnacht, F. E., Neumann, C., Forster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P. K. & Koch, B. (2014). Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central european test sites. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2547–2561. <https://doi.org/10.1109/JSTARS.2014.2329390>
- 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 Sensing of Environment*, 186, 64–87. <https://doi.org/10.1016/j.rse.2016.08.013>
- Fischer, C. & Traub, B. (2019). *Swiss National Forest Inventory – Methods and Models of the Fourth Assessment* (C. Fischer & B. Traub, Eds.). Springer. <https://doi.org/10.1007/978-3-030-19293-8>
- Foody, G. M. (2004). Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 70(5), 627–633. <https://doi.org/10.14358/PERS.70.5.627>
- Franklin, S. E. & Ahmed, O. S. (2018). Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *International Journal of Remote Sensing*, 39(15-16), 5236–5245. <https://doi.org/10.1080/01431161.2017.1363442>
- Fricker, G. A., Ventura, J. D., Wolf, J. A., North, M. P., Davis, F. W. & Franklin, J. (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>
- Gaveau, D. L. A. & Hill, R. A. (2003). Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, 29(5), 650–657. <https://doi.org/10.5589/m03-023>
- Ghosh, A., Fassnacht, F. E., Joshi, P. K. & Koch, B. (2014). A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation*, 26(1), 49–63. <https://doi.org/10.1016/j.jag.2013.05.017>
- Harikumar, A., Bovolo, F. & Bruzzone, L. (2019). A local projection-based approach to individual tree detection and 3-d crown delineation in multistoried coniferous forests using high-density airborne LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2), 1168–1182. <https://doi.org/10.1109/TGRS.2018.2865014>
- Hastings, J. H., Ollinger, S. V., Ouimette, A. P., Sanders-DeMott, R., Palace, M. W., Ducey, M. J., Sullivan, F. B., Basler, D. & Orwig, D. A. (2020). Tree species traits determine the success of LiDAR-based crown mapping in a mixed temperate forest. *Remote Sensing*, 12(2), 309. <https://doi.org/10.3390/rs12020309>

- Heinzel, J. & Koch, B. (2011). Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, 13(1), 152–160. <https://doi.org/10.1016/j.jag.2010.09.010>
- Heinzel, J. & Koch, B. (2012). Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *International Journal of Applied Earth Observation and Geoinformation*, 18(1), 101–110. <https://doi.org/10.1016/j.jag.2012.01.025>
- Hill, R. A. & Thomson, A. G. (2005). Mapping woodland species composition and structure using airborne spectral and LiDAR data. *International Journal of Remote Sensing*. <https://doi.org/10.1080/01431160500114706>
- Huber, P. J. (2011). Robust Statistics (M. Lovric, Ed.). In M. Lovric (Ed.), *International encyclopedia of statistical science*. Berlin, Heidelberg, Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_\}594
- Hyypä, J., Hyypä, H., Leckie, D., Gougeon, F., Yu, X. & Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29(5), 1339–1366. <https://doi.org/10.1080/01431160701736489>
- Jucker, T., Bouriaud, O. & Coomes, D. A. (2015). Crown plasticity enables trees to optimize canopy packing in mixed-species forests (J. Baltzer, Ed.). *Functional Ecology*, 29(8), 1078–1086. <https://doi.org/10.1111/1365-2435.12428>
- Kaartinen, H., Hyypä, J., Yu, X., Vastaranta, M., Hyypä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B. M. & Wu, J. C. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4(4), 950–974. <https://doi.org/10.3390/rs4040950>
- Kandare, K., Dalponte, M., Ørka, H. O., Frizzera, L. & Næsset, E. (2017). Prediction of species-specific volume using different inventory approaches by fusing airborne laser scanning and hyperspectral data. *Remote Sensing*, 9(5), 1–19. <https://doi.org/10.3390/rs9050400>
- 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. *Scandinavian Journal of Forest Research*, 33(4), 397–412. <https://doi.org/10.1080/02827581.2017.1416666>
- Kankare, V., Holopainen, M., Vastaranta, M., Liang, X., Yu, X., Kaartinen, H., Kukko, A. & Hyypä, J. (2017). Outlook for the Single-Tree-Level Forest Inventory in Nordic Countries (I. Ivan, A. Singleton, J. Horák & T. Inspektor, Eds.). In I. Ivan, A. Singleton, J. Horák & T. Inspektor (Eds.), *The rise of big spatial data*, Cham, Springer International Publishing. https://link.springer.com/chapter/10.1007%2F978-3-319-45123-7_14#citeas
- Kärvemo, S., Van Boeckel, T. P., Gilbert, M., Grégoire, J. C. & Schroeder, M. (2014). Large-scale risk mapping of an eruptive bark beetle - Importance of forest susceptibility and beetle pressure. *Forest Ecology and Management*, 318, 158–166. <https://doi.org/10.1016/j.foreco.2014.01.025>

- Ke, Y., Quackenbush, L. J. & Im, J. (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment*, 114(6), 1141–1154. <https://doi.org/10.1016/j.rse.2010.01.002>
- Kim, S., McGaughey, R. J., Andersen, H. E. & Schreuder, G. (2009). Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment*, 113(8), 1575–1586. <https://doi.org/10.1016/j.rse.2009.03.017>
- Ko, C., Sohn, G., Rimmel, T. K. & Miller, J. R. (2016). Maximizing the diversity of ensemble random forests for tree genera classification using high density LiDAR data. *Remote Sensing*, 8(8), 646. <https://doi.org/10.3390/rs8080646>
- Korpela, I., Hovi, A. & Korhonen, L. (2013). Backscattering of individual LiDAR pulses from forest canopies explained by photogrammetrically derived vegetation structure. *ISPRS Journal of Photogrammetry and Remote Sensing*, 83, 81–93. <https://doi.org/10.1016/j.isprsjprs.2013.06.002>
- Korpela, I., Ørka, H. O., Maltamo, M., Tokola, T. & Hyypä, J. (2010). Tree species classification using airborne LiDAR - effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, 44(2), 319–339. <https://doi.org/10.14214/sf.156>
- Korpela, I. & Tokola, T. (2006). Potential of aerial image-based monoscopic and multiview single-tree forest inventory: A simulation approach. Oxford Academic. <https://doi.org/10.1093/forestscience/52.2.136>
- Krzystek, P., Serebryanyk, A., Schnörr, C., Červenka, J. & Heurich, M. (2020). Large-scale mapping of tree species and dead trees in Sumava National Park and Bavarian Forest National Park using lidar and multispectral imagery. *Remote Sensing*, 12(4), 661. <https://doi.org/10.3390/rs12040661>
- Kuhn, M. & Johnson, K. (2019). *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press. <https://books.google.ch/books?id=q5alDwAAQBAJ>
- 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>
- Latifi, H., Fassnacht, F. & Koch, B. (2012). Forest structure modeling with combined airborne hyperspectral and LiDAR data. *Remote Sensing of Environment*, 121, 10–25. <https://doi.org/10.1016/j.rse.2012.01.015>
- Leckie, D. G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F. A., Cloney, E. & Paradine, D. (2005). Issues in species classification of trees in old growth conifer stands. *Canadian Journal of Remote Sensing*, 31(2), 175–190. <https://doi.org/10.5589/m05-004>
- Lee, J. H., Biging, G. S. & Fisher, J. B. (2016). An individual tree-based automated registration of aerial images to lidar data in a forested area. *Photogrammetric Engineering and Remote Sensing*, 82(9), 699–710. <https://doi.org/10.14358/PERS.82.9.699>
- Leiterer, R., Torabzadeh, H., Furrer, R., Schaepman, M. E. & Morsdorf, F. (2015). Towards automated characterization of canopy layering in mixed temperate forests using airborne laser scanning. *Forests*, 6(11), 4146–4167. <https://doi.org/10.3390/f6114146>
- Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18–22. <http://www.stat.berkeley.edu/>

- Lin, Y. & Hyyppä, J. (2016). A comprehensive but efficient framework of proposing and validating feature parameters from airborne LiDAR data for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, 46, 45–55. <https://doi.org/10.1016/j.jag.2015.11.010>
- Maltamo, M., Packalén, P., Suvanto, A., Korhonen, K. T., Mehtätalo, L. & Hyvönen, P. (2009). Combining ALS and NFI training data for forest management planning: a case study in Kuortane, Western Finland. *European Journal of Forest Research*, 128(3), 305–317. <https://doi.org/10.1007/s10342-009-0266-6>
- Marrs, J. & Ni-Meister, W. (2019). Machine Learning Techniques for Tree Species Classification Using Co-Registered LiDAR and Hyperspectral Data. *Remote Sensing*, 11(7), 819. <https://doi.org/10.3390/rs11070819>
- MATLAB. (2014). MATLAB 9.5. Natick, MA, The MathWorks Inc.
- McRoberts, R. E. & Tomppo, E. O. (2007). Remote sensing support for national forest inventories. *Remote Sensing of Environment*, 110(4), 412–419. <https://doi.org/10.1016/j.rse.2006.09.034>
- Morsdorf, F., Kükenbrink, D., Schneider, F. D., Abegg, M. & Schaepman, M. E. (2018). Close-range laser scanning in forests: Towards physically based semantics across scales. *Interface Focus*, 8(2), 20170046. <https://doi.org/10.1098/rsfs.2017.0046>
- Morsdorf, F., Meier, E., Kötz, B., Itten, K. I., Dobbertin, M. & Allgöwer, B. (2004). LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. *Remote Sensing of Environment*, 92(3), 353–362. <https://doi.org/10.1016/j.rse.2004.05.013>
- Naidoo, L., Cho, M. A., Mathieu, R. & Asner, G. (2012). Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69, 167–179. <https://doi.org/10.1016/j.isprsjprs.2012.03.005>
- Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, 56(1), 1–7. [https://doi.org/10.1016/0034-4257\(95\)00224-3](https://doi.org/10.1016/0034-4257(95)00224-3)
- Ørka, H. O., Dalponte, M., Gobakken, T., Næsset, E. & Ene, L. T. (2013). Characterizing forest species composition using multiple remote sensing data sources and inventory approaches. *Scandinavian Journal of Forest Research*, 28(7), 677–688. <https://doi.org/10.1080/02827581.2013.793386>
- Ørka, H. O., Gobakken, T., Næsset, E., Ene, L. & Lien, V. (2012). Simultaneously acquired airborne laser scanning and multispectral imagery for individual tree species identification. *Canadian Journal of Remote Sensing*, 38(2), 125–138. <https://doi.org/10.5589/m12-021>
- Ørka, H. O., Næsset, E. & Bollandsås, O. M. (2009). Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment*, 113(6), 1163–1174. <https://doi.org/10.1016/j.rse.2009.02.002>
- Ørka, H. O., Næsset, E. & Bollandsås, O. M. (2010). Effects of different sensors and leaf-on and leaf-off canopy conditions on echo distributions and individual tree properties derived from airborne laser scanning. *Remote Sensing of Environment*, 114(7), 1445–1461. <https://doi.org/10.1016/j.rse.2010.01.024>

- Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217–222. <https://doi.org/10.1080/01431160412331269698>
- Parkan, M. (2018). Digital Forestry Toolbox for Matlab/Octave. <https://doi.org/10.5281/zenodo.1213013>
- Parkan, M. & Tuia, D. (2019). *Combined use of airborne laser scanning and hyperspectral imaging for forest inventories* (Doctoral dissertation). EPFL. https://infoscience.epfl.ch/record/262809/files/EPFL_TH9033.pdf
- Perez-Riverol, Y., Kuhn, M., Vizcaíno, J. A., Hitz, M.-P. & Audain, E. (2017). Accurate and fast feature selection workflow for high-dimensional omics data. *PLoS one*, 12(12), e0189875. <https://doi.org/10.1371/journal.pone.0189875>
- Persson, M., Lindberg, E. & Reese, H. (2018). Tree species classification with multi-temporal Sentinel-2 data. *Remote Sensing*, 10(11), 1794. <https://doi.org/10.3390/rs10111794>
- Pullanagari, R. R., Kereszturi, G. & Yule, I. (2018). Integrating airborne hyperspectral, topographic, and soil data for estimating pasture quality using recursive feature elimination with random forest regression. *Remote Sensing*, 10(7), 1117. <https://doi.org/10.3390/rs10071117>
- Puumalainen, J., Kennedy, P. & Folving, S. (2003). Monitoring forest biodiversity: A European perspective with reference to temperate and boreal forest zone. *Journal of Environmental Management*, 67(1), 5–14. [https://doi.org/10.1016/S0301-4797\(02\)00183-4](https://doi.org/10.1016/S0301-4797(02)00183-4)
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Sasaki, T., Imanishi, J., Ioki, K., Morimoto, Y. & Kitada, K. (2012). Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data. *Landscape and Ecological Engineering*, 8(2), 157–171. <https://doi.org/10.1007/s11355-011-0158-z>
- Schumacher, J. & Nord-Larsen, T. (2014). Wall-to-wall tree type classification using airborne lidar data and CIR images. *International Journal of Remote Sensing*, 35(9), 3057–3073. <https://doi.org/10.1080/01431161.2014.894670>
- Shi, Y., Skidmore, A. K., Wang, T., Holzwarth, S., Heiden, U., Pinnel, N., Zhu, X. & Heurich, M. (2018). Tree species classification using plant functional traits from LiDAR and hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, 73, 207–219. <https://doi.org/10.1016/j.jag.2018.06.018>
- Shii, Y., Wang, T., Skidmore, A. K. & Heurich, M. (2018). Important LiDAR metrics for discriminating forest tree species in Central Europe. *ISPRS Journal of Photogrammetry and Remote Sensing*, 137, 163–174. <https://doi.org/10.1016/j.isprsjprs.2018.02.002>
- Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O. & Katila, M. (2008). Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment*, 112(5), 1982–1999. <https://doi.org/10.1016/j.rse.2007.03.032>
- Torabzadeh, H., Leiterer, R., Hueni, A., Schaepman, M. E. & Morsdorf, F. (2019). Tree species classification in a temperate mixed forest using a combination of imaging spectroscopy and airborne laser scanning. *Agricultural and Forest Meteorology*, 279, 107744. <https://doi.org/10.1016/j.agrformet.2019.107744>

- Trier, Ø. D., Salberg, A. B., Kermit, M., Rudjord, Ø., Gobakken, T., Næsset, E. & Aarsten, D. (2018). Tree species classification in Norway from airborne hyperspectral and airborne laser scanning data. *European Journal of Remote Sensing*, 51(1), 336–351. <https://doi.org/10.1080/22797254.2018.1434424>
- Vauhkonen, J., Korpela, I., Maltamo, M. & Tokola, T. (2010). Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sensing of Environment*, 114(6), 1263–1276. <https://doi.org/10.1016/j.rse.2010.01.016>
- Vauhkonen, J., Ørka, H. O., Holmgren, J., Dalponte, M., Heinzl, J. & Koch, B. (2014). Tree Species Recognition Based on Airborne Laser Scanning and Complementary Data Sources. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-8663-8{_}7
- Waser, L. T., Ginzler, C., Kuechler, M., Baltsavias, E. & Hurni, L. (2011). Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sensing of Environment*, 115(1), 76–85. <https://doi.org/10.1016/j.rse.2010.08.006>
- Waser, L. T. (2012). *Airborne remote sensing data for semi-automated extraction of tree area and classification of tree species* (Doctoral dissertation). <https://doi.org/10.3929/ethz-a-010782581>
- Waser, L. T., Ginzler, C. & Rehush, N. (2017). Wall-to-Wall tree type mapping from countrywide airborne remote sensing surveys. *Remote Sensing*, 9(8). <https://doi.org/10.3390/rs9080766>
- 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. *Canadian Journal of Remote Sensing*, 42(5), 619–641. <https://doi.org/10.1080/07038992.2016.1207484>
- Wulder, M. A., Coops, N. C., Hudak, A. T., Morsdorf, F., Nelson, R., Newnham, G. & Vastaranta, M. (2013). Status and prospects for LiDAR remote sensing of forested ecosystems. *Canadian Journal of Remote Sensing*, 39(SUPPL.1), S1–S5. <https://doi.org/10.5589/m13-051>
- Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W. & Gobakken, T. (2012). Lidar sampling for large-area forest characterization: A review. Elsevier. <https://doi.org/10.1016/j.rse.2012.02.001>
- Xu, C., Morgenroth, J. & Manley, B. (2015). Integrating data from discrete return airborne lidar and optical sensors to enhance the accuracy of forest description: A review. *Current Forestry Reports*, 1(3), 206–219. <https://doi.org/10.1007/s40725-015-0019-3>
- Yao, W., Krzystek, P. & Heurich, M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, 123, 368–380. <https://doi.org/10.1016/j.rse.2012.03.027>
- Yin, D. & Wang, L. (2016). How to assess the accuracy of the individual tree-based forest inventory derived from remotely sensed data: a review. *International Journal of Remote Sensing*, 37(19), 4521–4553. <https://doi.org/10.1080/01431161.2016.1214302>
- Zhang, Z. & Liu, X. (2013). WorldView-2 Satellite Imagery and Airborne LiDAR Data for Object-Based Forest Species Classification in a Cool Temperate Rainforest Environment (A. Abdul Rahman, P. Boguslawski, C. Gold & M. N. Said, Eds.). In A. Abdul Rahman, P. Boguslawski, C. Gold & M. N. Said (Eds.), *Developments in multidimensional spatial data models*. Berlin, Heidelberg, Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36379-5{_}7

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F. & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4), 8–36. <https://doi.org/10.1109/MGRS.2017.2762307>

A Appendix - Additional Figures

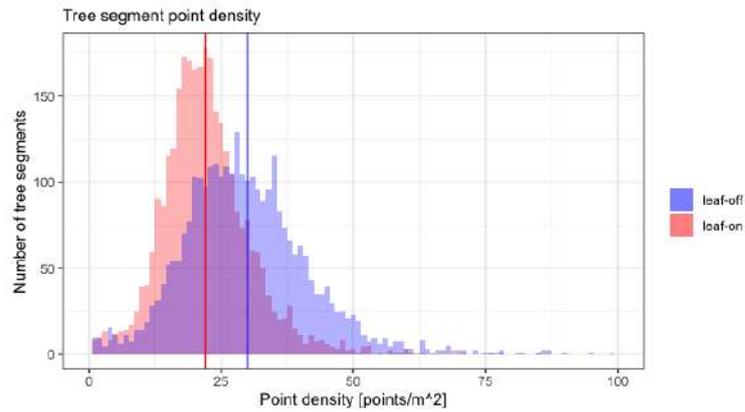


FIGURE A.1: Point density inside tree segments for leaf-on and leaf-off datasets.

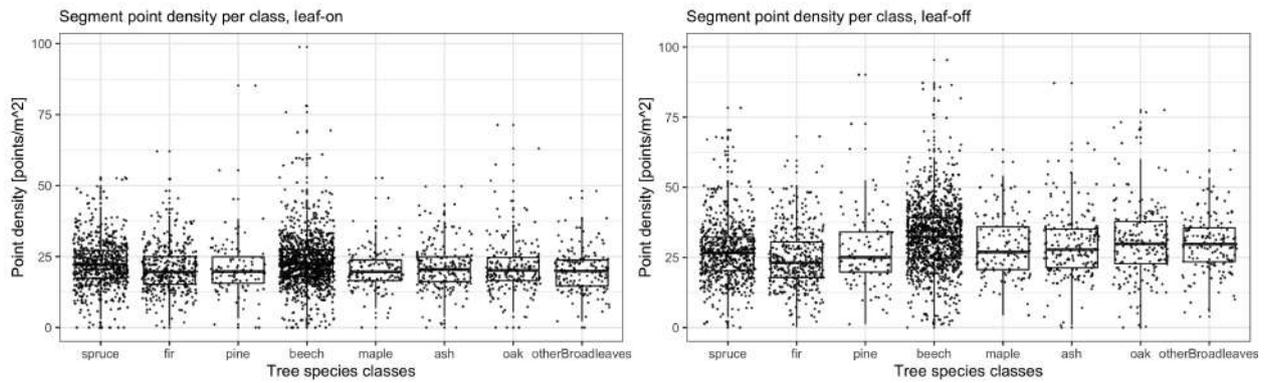


FIGURE A.2: Boxplots of tree segment point density per species class for leaf-on and leaf-off datasets.

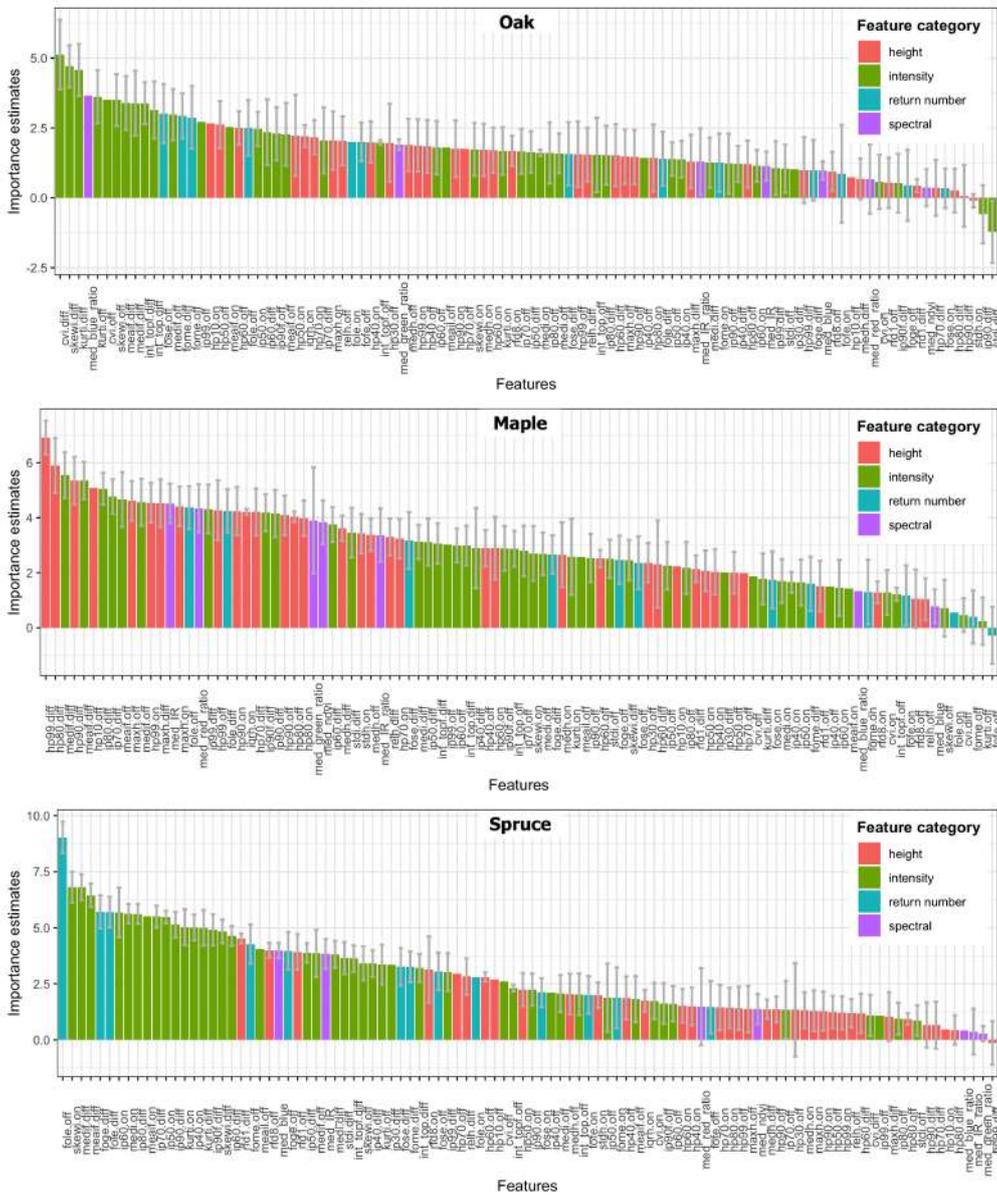


FIGURE A.3: Class-specific RFE feature importance estimates of the 100 most important features for each of the classes oak, maple and spruce. The bars are coloured by feature category (for description see Tables 3.1 and 3.2). The bar height represents mean importance derived from 3-times repeated 10-fold cross-validation and the standard deviation is indicated by grey errorbars.

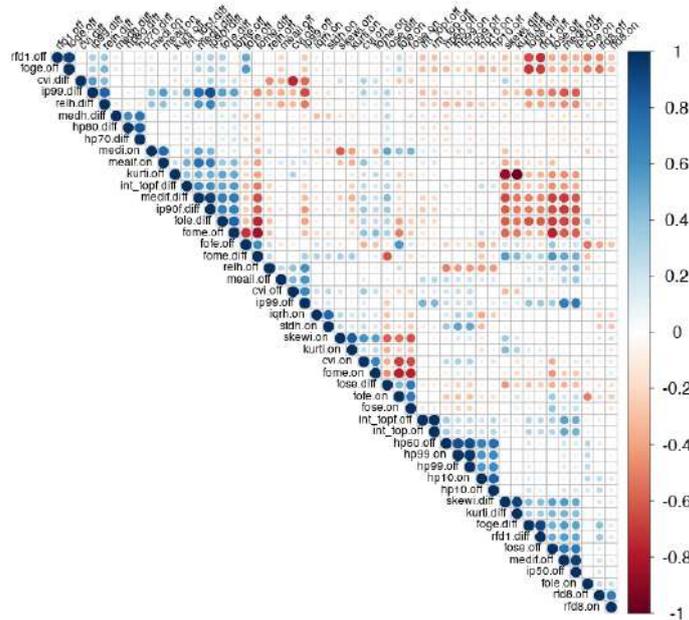


FIGURE A.5: Pearson correlation matrix of the 48 remaining features after removing highly correlated features. These were used in the simplified classification model. 1 = perfect positive correlation, -1 = perfect negative correlation.

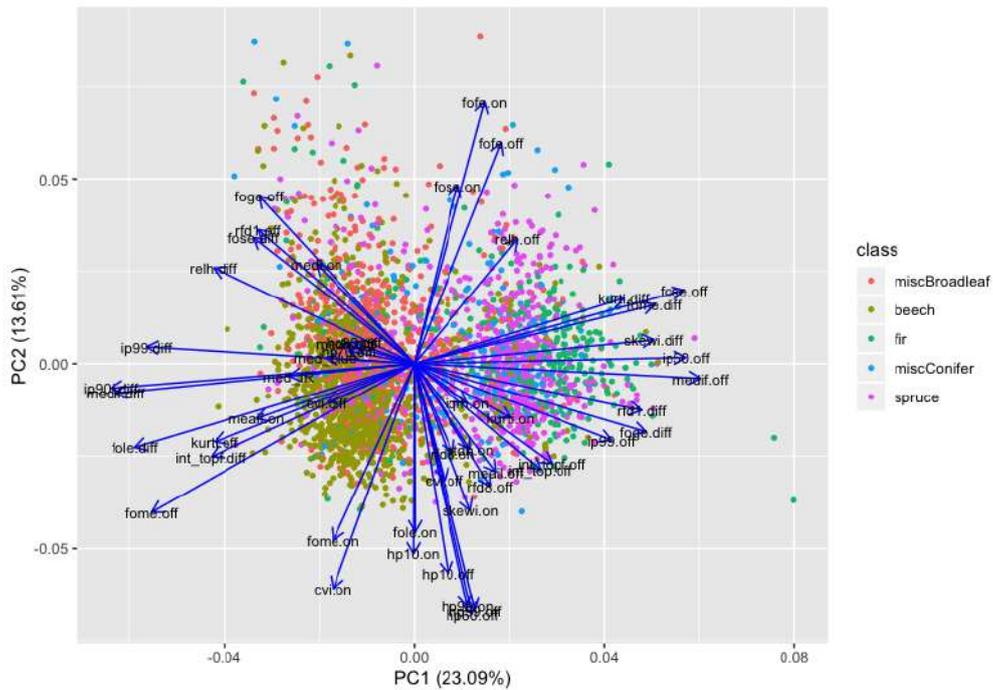


FIGURE A.6: Principal component plot of reference dataset used in simplified classification model. Blue arrows indicate the loadings of the features (cf. Table 3.1 for description).

Acknowledgements

I would like to thank my supervisor Felix for great advice, guidance and support during the completion of this Master thesis, in person as well as remotely, and my co-supervisor Meinrad for valuable discussions and input from a forest inventory perspective.

I would also like to thank the staff at the forestry department of Kanton Aargau for their interest, the uncomplicated provision of data and the possibility to present my thesis to them.

Furthermore, I would like to thank my fellow students and friends, for their help, support, patience and company. This includes everyone who kept me company in the old Master's "cave" as well as those who appeared in our new working space.

A big thank you goes to everyone carefully proofreading this thesis, Gwendolyn in particular but also Livia and Taja. I am grateful for everyone who supported me during this time, my friends and especially my family, for giving me the opportunity to study.

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Location, Date

Winterthur, July 31, 2020

Signature

Aline Bornand