

# **Combined use of airborne laser scanning and hyperspectral imaging for forest inventories**

**Thèse N° 9033**

Présentée le 11 janvier 2019

À la Faculté de l'environnement naturel, architectural et construit  
Laboratoire de systèmes d'information géographique  
Programme doctoral en génie civil et environnement

pour l'obtention du grade de Docteur ès Sciences

par

**MATTHEW JOSEF PARKAN**

Acceptée sur proposition du jury

Prof. A. Martinoli, président du jury  
Prof. F. Golay, Prof. D. Tuia, directeurs de thèse  
Prof. J. Hyppä, rapporteur  
Dr C. Mallet, rapporteur  
Prof. A. Buttler, rapporteur

2019



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

## Overview

---

### Project

Title: *Combined use of airborne laser scanning and hyperspectral imaging for forest inventories*

Candidate: Matthew Parkan (LASIG, EPFL)

Supervision: Prof. François Golay (LASIG, EPFL), director  
Prof. Devis Tuia (LGRS, Wageningen University, Netherlands), co-director

Enrollment date: Jan. 1, 2014

Host institute: École Polytechnique Fédérale de Lausanne (EPFL), Switzerland  
Doctoral Program in Civil and Environmental Engineering (EDCE)  
Geographic Information Systems Laboratory (LASIG)

Federal funding: Swiss Forestry and Wood Research Fund (WHFF, OFEV), project 2013.18

State funding: State of Neuchâtel  
State of Vaud  
State of Fribourg  
State of Jura

### Contact

Postal address: EPFL ENAC IIE LASIG  
GC D2 414 (Bâtiment GC)  
Station 18  
CH-1015 Lausanne  
Switzerland

Office: GCD2414

Phone: +41 79 789 81 51

Email: matthew.parkan@epfl.ch

## Abstract

---

The thesis presented hereby proposes as a general objective the estimation of forest inventory parameters (e.g. trunk location, height, basal area, crown area, species, etc..) from the combination of Airborne Laser Scanning (ALS) and Hyperspectral Imaging (HI). The research is centered around three main topics: the development of new individual tree segmentation algorithms, the assessment of direct and indirect dendrometry methods, tree species classification based on ALS and HI features. A common dependency of these topics is the availability of reliable reference datasets for the calibration and validation (error assessment) of algorithms. This requirement is addressed with the development of an interactive software application and procedures to facilitate the manual extraction of trees and visual identification of species from ALS points clouds. The results of this research can be useful to the operational domain in several ways: providing tools and procedures to characterize areas that are not covered by field inventories (e.g. private forests, low accessibility areas), act as a decision support (e.g. preparing plot maps, identifying priority intervention zones, etc.) when planning field surveys or logging, improving the integration of field and remote sensing measurements for forest inventories.

**Keywords:** Airborne Laser Scanning (ALS), LiDAR, forest inventory, tree modeling, point cloud segmentation, machine learning, Hyperspectral Imaging (HI)

## Résumé

---

La thèse présentée ici propose comme objectif général l'estimation des paramètres d'inventaire forestier (p. ex. position du tronc, hauteur, surface terrière, surface de la couronne, espèces, etc.) à partir de la combinaison de Relevés altimétriques par Laser Aéroporté (RLA) et d'Imagerie Hyperspectrale (IH). La recherche s'articule autour de trois thèmes principaux: le développement de nouveaux algorithmes de segmentation d'arbres individuels, l'évaluation de méthodes de dendrométrie directe et indirecte, la classification d'espèces basée sur les caractéristiques du RLA et de l'IH. Une dépendance commune de ces sujets est la disponibilité de données de référence fiables pour la calibration et la validation (évaluation des erreurs) des algorithmes. Cette exigence est abordée par le développement d'une application interactive et de procédures pour faciliter l'extraction manuelle des arbres et l'identification visuelle des espèces à partir des nuages de points issus de RLA. Les résultats de cette recherche peuvent être utiles au domaine opérationnel de plusieurs façons: fournir des outils et des procédures pour caractériser les zones qui ne sont pas couvertes par les inventaires de terrain (p. ex. forêts privées, zones à faible accessibilité), agir comme support décisionnel (p. ex. préparation de plan de situation, identification de zones d'intervention prioritaire, etc) lors de la planification des relevés de terrain ou des coupes, améliorer l'intégration des mesures de terrain et de télédétection dans les inventaires forestiers.

**Mots clés:** Relevés altimétriques par Laser Aéroporté (RLA), LiDAR, inventaires forestiers, modélisation des arbres, segmentation de nuages de points, apprentissage automatique, Imagerie Hyperspectrale (IH)

## Acknowledgements

---

This work would not have been possible without the exceptional kindness, enthusiasm and commitment of a few individuals:

My wife, Dai.

My Family, who gave me the opportunity to study.

Marc Riedo, Gilles Gachet and Pascal Junod, who showed persistent interest in my work, gave technical support and provided most of the data.

Marie Guillot, Rita Bütler Sauvain, Jean-François Métraux and Jean Laurent Pfund, whose advice was essential in setting up the research project and obtaining complementary funds.

My advisors, François Golay and Devis Tuia, who encouraged me to undertake this work, ensured scientific, administrative and diplomatic assistance.

My colleagues from the LASIG and TOPO labs whose friendship I have enjoyed daily: Kévin Barbeau, Romain Beauvais, Alain Clappier, Philipp Clausen, Emmanuel Clédat, Heydi Contreras, Solange Duruz, Stéphane Joost, Gabriel Laupré, Marie-Christine Lehmann, Jessie Madrazo, Estelle Rochat, Oliver Selmoni, Marc Soutter, Emanuele Strano and Elia Vajana.

The members of my thesis committee, Alexandre Buttler, Juha Hyypä, Clément Mallet and Alcherio Martinoli, who openly accepted to evaluate and discuss my work.

## List of abbreviations

---

The following abbreviations are used:

<b>AGL</b>	Above Ground Level
<b>AHI</b>	Airborne Hyperspectral Imaging
<b>ALS</b>	Airborne Laser Scanning
<b>APEX</b>	Airborne Prism EXperiment
<b>ASL</b>	Above Sea Level
<b>BRDF</b>	Bidirectional Reflectance Distribution Function
<b>CHM</b>	Canopy Height Model
<b>DBH</b>	Diameter at Breast Height
<b>DFT</b>	Digital Forestry Toolbox
<b>DHM</b>	Digital Height Model
<b>DSM</b>	Digital Surface Model
<b>DTM</b>	Digital Terrain Model
<b>FN</b>	False Negative
<b>FP</b>	False Positive
<b>FW</b>	Full Waveform
<b>GCP</b>	Ground Control Point
<b>GNSS</b>	Global Navigation Satellite System
<b>HI</b>	Hyperspectral Imaging
<b>IMU</b>	Inertial Measurement Unit
<b>SAR</b>	Synthetic Aperture Radar
<b>IoU</b>	Intersection over Union
<b>ITC</b>	Individual Tree Crown
<b>MLS</b>	Mobile Laser Scanning
<b>MNF</b>	Maximum Noise Fraction
<b>NFI</b>	National Forest Inventory
<b>NIR</b>	Near Infrared
<b>PCA</b>	Principal Component Analysis
<b>PRR</b>	Pulse Repetition Rate
<b>RANSAC</b>	RANdom SAmple Consensus
<b>RBF</b>	Radial Basis Function
<b>RF</b>	Random Forest
<b>RGB</b>	Red-Green-Blue
<b>SVM</b>	Support Vector Machine
<b>SVR</b>	Support Vector Regression
<b>SWIR</b>	Short Wave Infrared
<b>TIN</b>	Triangular Irregular Network
<b>TLS</b>	Terrestrial Laser Scanning
<b>TP</b>	True Positive
<b>UAV</b>	Unmanned Aerial Vehicle
<b>VIS</b>	Visible
<b>VLR</b>	Variable Length Records
<b>WSL</b>	Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft

## List of publications

---

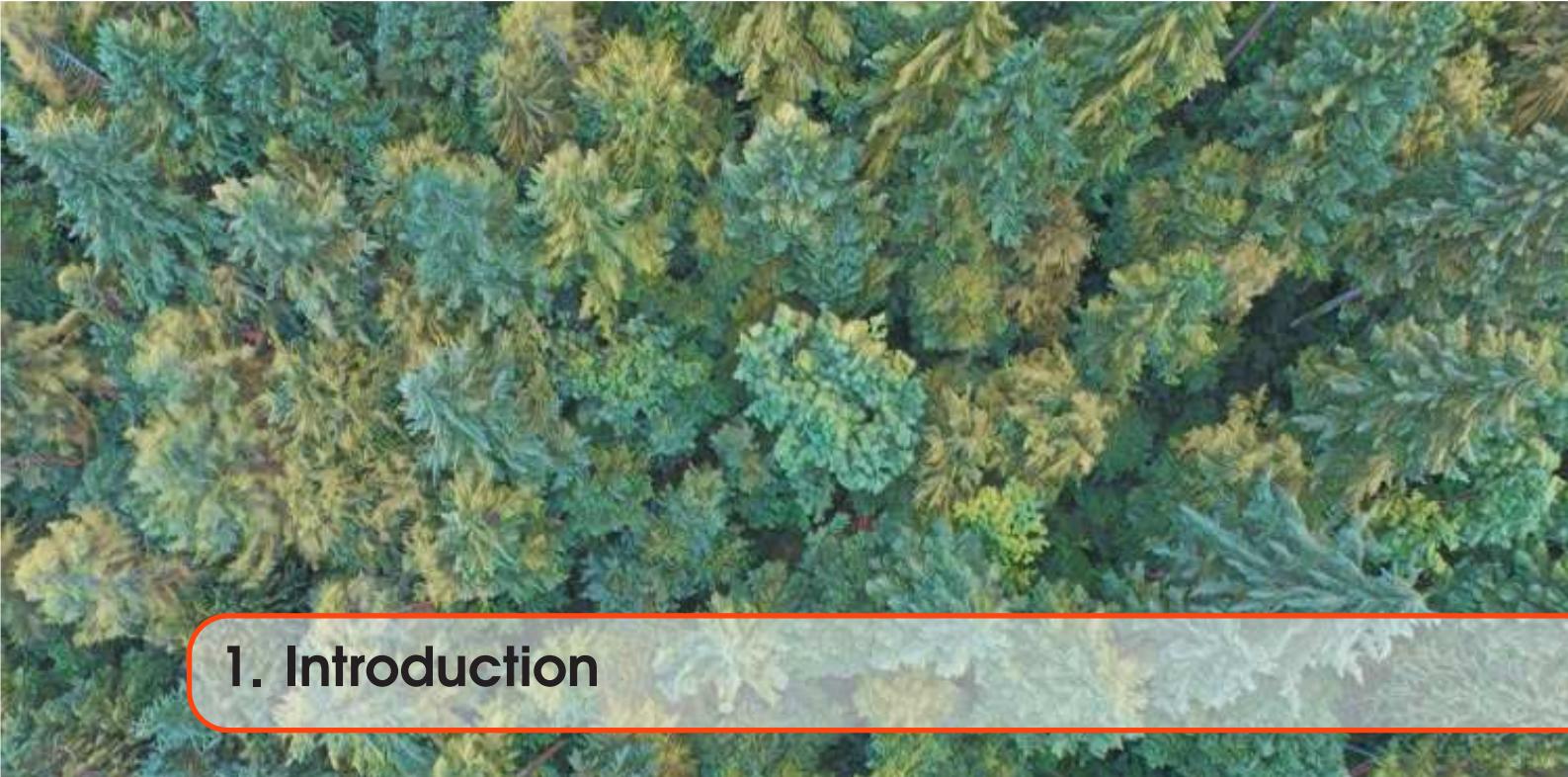
- Parkan, M., Tuia, D., 2015. Individual tree segmentation in deciduous forests using geodesic voting. In: Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, pp. 637–640. Recipient of the 2015 IEEE Geoscience and Remote Sensing Society - Symposium Interactive Prize Paper Award (ISPPA).
- Parkan, M., Aug. 2017a. Digital-Forestry-Toolbox: A collection of digital forestry tools for Matlab [<https://github.com/mparkan/Digital-Forestry-Toolbox>]
- Parkan, M., Mar. 2017b. Données LiDAR aériennes: Pertinence de l'interprétation visuelle pour la foresterie. Schweizerische Zeitschrift fur Forstwesen 168 (3), 127–133
- Parkan, M., Tuia, D., Feb. 2018. Estimating Uncertainty of Point-Cloud Based Single-Tree Segmentation with Ensemble Based Filtering. Remote Sensing 10 (2), 335
- Parkan, M., Junod, P., Lugrin, R., Ginzler, C., 2018. A reference airborne LiDAR dataset for forest research. Remote Sensing (Submitted). Submitted - pending corrections.



# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Rationale and organization	9
1.1.1	Context	9
1.1.2	Objectives	11
1.1.3	Outreach	11
1.2	Forest inventories	12
1.3	Airborne laser scanning	14
1.4	Airborne hyperspectral imaging	18
<b>2</b>	<b>Data</b>	<b>21</b>
2.1	Overview	21
2.2	Creation of a benchmark ALS dataset	23
2.2.1	ALS data preparation	24
2.2.2	Manual segmentation	26
2.2.3	Characterization	28
2.3	Sampling strategy	34
2.4	Synthesis	34
<b>3</b>	<b>Individual tree segmentation</b>	<b>35</b>
3.1	State of the art	35
3.2	Error assessment framework	41
3.3	Tree segmentation with geodesic voting	50
3.3.1	Description	50
3.3.2	Results	56
3.3.3	Discussion	64

<b>3.4</b>	<b>Evaluating and improving segmentation with ensemble filtering</b>	<b>65</b>
3.4.1	Description .....	65
3.4.2	Results .....	70
3.4.3	Discussion .....	75
<b>3.5</b>	<b>Stem detection with layered morphological analysis</b>	<b>76</b>
3.5.1	Description .....	76
3.5.2	Results .....	81
3.5.3	Discussion .....	85
<b>3.6</b>	<b>Synthesis</b>	<b>86</b>
<b>4</b>	<b>Diameter estimation</b>	<b>87</b>
<b>4.1</b>	<b>State of the art</b>	<b>87</b>
<b>4.2</b>	<b>Error assessment framework</b>	<b>93</b>
<b>4.3</b>	<b>Indirect diameter estimation</b>	<b>94</b>
4.3.1	Description .....	94
4.3.2	Results .....	96
4.3.3	Discussion .....	102
<b>4.4</b>	<b>Direct diameter measurement</b>	<b>103</b>
4.4.1	Description .....	103
4.4.2	Results .....	108
4.4.3	Discussion .....	112
<b>4.5</b>	<b>Synthesis</b>	<b>113</b>
<b>5</b>	<b>Tree species classification</b>	<b>115</b>
<b>5.1</b>	<b>State of the art</b>	<b>115</b>
<b>5.2</b>	<b>Error assessment framework</b>	<b>126</b>
<b>5.3</b>	<b>General tree characteristics</b>	<b>127</b>
5.3.1	Structural separability .....	127
5.3.2	Spectral separability .....	132
<b>5.4</b>	<b>ALS based classification</b>	<b>138</b>
5.4.1	Description .....	138
5.4.2	Results .....	142
5.4.3	Discussion .....	145
<b>5.5</b>	<b>AHI based classification</b>	<b>146</b>
5.5.1	Description .....	146
5.5.2	Results .....	150
5.5.3	Discussion .....	153
<b>5.6</b>	<b>Combined ALS and AHI classification</b>	<b>155</b>
5.6.1	Description .....	155
5.6.2	Results .....	157
5.6.3	Discussion .....	159
<b>5.7</b>	<b>Synthesis</b>	<b>161</b>
<b>6</b>	<b>Conclusion</b>	<b>163</b>
	<b>Bibliography</b>	<b>167</b>
	<b>Appendix</b>	<b>204</b>



# 1. Introduction

This chapter introduces the context and objectives of the thesis. It also provides the basic technical background required to understand the presented work. It is structured in the following way:

- 1.1 introduces the role of remote sensing in forest monitoring and management in the current global and Swiss contexts. It also summarizes the research questions, the objectives and the outreach activities/products of the thesis.
- 1.2 introduces general notions about forest inventories and their use in Switzerland.
- 1.3 presents the Airborne Laser Scanning technique (ALS).
- 1.4 presents the Airborne Hyperspectral Imaging technique (AHI).

## 1.1 Rationale and organization

### 1.1.1 Context

Forests are a fundamental component of terrestrial ecosystems. They constitute diverse and irreplaceable habitats for a large part of living organisms. Forests also provide invaluable ecosystem services including oxygen production, carbon sequestration, water filtering, climate regulation, protection against avalanches, erosion and desertification. The exploitation of forest timber and non-timber resources brings direct economic benefits (FAO, 2016; Radkau, 2012). Forests also have an important social, spiritual and therapeutic role in human societies. Yet, in many regions, they are unmonitored, unprotected or poorly managed. The unwillingness or inability to enforce protection and setup sustainable management policies are the main reasons for this situation. The problem is reinforced by the fact that many developed countries with rich forest resources are not able to successfully exploit them, instead relying on wood products from countries with lower exploitation costs and less stringent work or environmental regulations. Finally, because of Earth's rapidly changing climate, a better characterization of forest ecosystems is required to identify vulnerabilities and possibility mitigate irreversible damage.

Remote sensing can help improve this situation by providing timely and objective surveys over large forest extents at a low cost (Finer et al., 2018; Waser et al., 2017a; White et al., 2016; Eitel et al., 2016; FAO, 2016; Wulder et al., 2012; McRoberts and Tomppo, 2007). From global spaceborne missions to local drone flights, remote sensing technologies are now routinely used to map forests with unprecedented detail. Information derived from these maps can be used across all decision scales, from local forest managers to strategic policy makers, for research, management, conservation and commercial activities. For example, using satellite images, Hansen et al. (2013) created a global interactive map of forest change starting from year 2000 to nowadays, objectively quantifying massive degradation in tropical areas. Mascaro et al. (2014) suggested that with only 5% of the funding pledged to reduce carbon emissions in the tropics, all tropical forests could be mapped with Airborne Laser Scanning (ALS), potentially greatly improving biomass measurement and habitat modeling. In the United States, Snyder (2012) estimated the potential yearly business benefit derived from elevation data to be 62 million U.S. dollars for forest resources management alone and 159 million for planning and response to wildfires.

The possibility to accurately measure the 3D geometry of forest landscapes is particularly valuable, because it allows among other things to characterize terrain topography, canopy height, forest stratification, growth rates and wood volume (biomass). Aerial photogrammetry, Synthetic Aperture Radar (SAR) and laser scanning all provide this capacity, but the latter technique is currently the most suitable for forests (St-Onge et al., 2008; Sexton et al., 2009; White et al., 2013b; Vastaranta et al., 2013). Current ALS systems, however, do not capture detailed radiometric or texture characteristics which can help distinguish structurally similar species and diagnose leaf health. In this regard, airborne multispectral and especially hyperspectral imaging is complementary to ALS.

In North America and Europe, many national and regional mapping agencies now acquire high resolution imagery and elevation models on a regular basis. Commonly and sometimes openly, this data is provided as a basic geospatial service to other public and private users. However, data availability by itself is insufficient to fully leverage the use of remote sensing in forestry. Successful integration lies in the ability to analyze the raw data, extract useful information and communicate it in a format that can support and improve existing workflows and decisional procedures.

With its difficult to access mountain forests, large amount of private plots, high timber exploitation costs, varying environmental agency budgets, strong regulations and minimal intervention silviculture practices (selection cutting), Switzerland is a prime example of challenging management conditions. In Switzerland, field inventories still provide the bulk of forest measurements used for planning. However, these operations are work intensive, expensive, restricted to accessible areas and biased by subjectivity and observer skills. These conditions, coupled with the availability of high quality data, are strong incentives to investigate the use of remote sensing in support of traditional forestry. Moreover, because inventories are a nodal procedure in both forest science and management, their improvement can be considered a priority research topic. Overall, the new measurement and data processing techniques, the survey of previously unexplored forest environments, the accumulated observations, and the need for regular benchmarking and replication experiments provide a steady supply of research questions.

### 1.1.2 Objectives

In direct relation with the above rationale, using remote sensing techniques, this thesis aims to replicate three fundamental operations of traditional field inventories: locating trees, measuring their diameter and identifying their genus/species. Formally, this can be translated into several research questions:

- How and under which conditions can Airborne Laser Scanning (ALS) and Hyperspectral Imaging (AHI) be used to map forest characteristics at the individual tree scale (i.e. tree location, diameter and genus/species)?
- How can the reliability of such maps be quantified?
- What is the relative value of ALS and AHI to support current forest inventory needs?

These research questions are investigated by following a set of core scientific and technical objectives:

- The development of a novel **workflow and tool to visually interpret ALS point clouds and interactively extract tree models** to create a large and reliable reference dataset. This development is also a common dependency for the calibration and validation (error assessment) of automatic individual tree segmentation, species classification and stem diameter estimation algorithms.
- The development of a new **specification to store individual tree attributes in the ASPRS LAS format**.
- The development of new **individual tree segmentation** algorithms with a particular emphasis on deciduous broadleaf forests.
- The development of a **rigorous framework to assess the performance of individual tree segmentation** algorithms.
- The development of a new **method to directly estimate stem diameter and taper** from high density ALS point clouds.
- the development of **tree species classification** methods to classify the main tree genus/species encountered in Swiss forests from ALS and AHI.

### 1.1.3 Outreach

Finally, a strong emphasis was put on the transfer of technical know-how to the operational domain (i.e. forest and geomatic state services). This requirement was addressed in the following ways:

- The **development of open source software and tutorials** for ALS data analysis - the Digital Forestry Toolbox (Parkan, 2017a) - based on the Matlab/Octave programming language.
- The **open distribution of a benchmark dataset** composed of several thousand reliable 3D tree models extracted from high density ALS point clouds.
- The **coordination of an interstate work group** with Marc Riedo (Neuchâtel) and Dr. Gilles Gachet (Vaud) involving about 30 participants from the public and private forestry, remote sensing and geomatics sectors. Yearly group meetings were organized for technical knowledge transfer and end-user requirement analysis.
- **Direct data and information exchanges** with private and public organizations active in forestry, for example through presentations in continuing education workshops (Fortschreibung Wald und Landschaft - FoWaLa).
- **Creation and maintenance of a website** dedicated to explaining and promoting the use of ALS in forestry.

## 1.2 Forest inventories

Forest inventories are the primary monitoring operation to assess the state of forests. They aim at quantifying the economic, security, protective, sanitary and ecological characteristics of forests at a given time. A review of the main forest inventory parameters and measurement techniques currently in use is provided in table 1.1 and figure 1.1. In Switzerland, inventories are conducted by various public and private organizations at different spatial scales and time frequencies:

- The **National Forest Inventory (NFI)** is conducted by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) and the Federal Office for the Environment (FOEN). The first NFI was finished in 1985 and subsequent NFI have been updated approximatively every ten years (1985, 1995, 2006, 2017). The fifth NFI was started in 2018 and is currently underway. The first NFI was conducted over 12000 plots distributed on a 1x1 km grid (systematic sampling) covering the whole country, further NFI significantly reduced the amounts of plots and increased reliance on aerial photo interpretation. Field data collection and aerial image interpretation are conducted following rigorous measurement procedures published by WSL (Düggelin and Keller, 2017; Ginzler et al., 2005).
- The **Cantonal (state) forest inventories** are conducted by state forest services using different sampling approaches (systematic, stratified random).
- **Sanitary inventories** (Sanasilva) have been conducted by WSL yearly since 1985. These inventories cover a sample of about 1100 trees distributed across Switzerland and are used to characterize defoliation, foliage color, growth and various other sanitary attributes.
- **Private forest plot inventories** are conducted by state forest services or by private firms mandated by the owner.
- **Research plot inventories** are the most detailed and are generally conducted by WSL (Schaub et al., 2011). They are used for long term monitoring of forest ecosystems.
- **Urban tree inventories** are conducted in some cities (e.g. Geneva, Lausanne) by park services in public areas. In some cases, they are updated systematically at each intervention (cut or plantation).

The aforementioned inventories are completed by a systematic registration of silvicultural interventions (cutting or planting).

Currently, the integration of remote sensing in the NFI is mostly limited to stereo photointerpretation of high resolution visible (VIS) and near infrared (NIR) imagery. The main applications of this approach are the determination of coniferous/deciduous proportions, forest type (structure), height estimation and forest/stand edge delineation (Barrett et al., 2016). Recently, Waser et al. (2017b) also developed a country-wide coniferous/deciduous map using a combination of VIS/NIR imagery and ALS derived canopy height models.

At the sub-national level, the integration of remote sensing in inventories is very variable and mostly dependent on personal forester skills/interests. Foresters do not receive systematic training on how to integrate remote sensing products in their workflow and exposure to such techniques remains episodic.

With national ALS and high resolution RGB imagery acquisition campaigns currently underway, it can be expected that both national and sub-national forest inventories will further increase reliance on remote sensing in the near future. It can also be expected that the current systematic sampling scheme will evolve towards stratified sampling, to reduce field work.

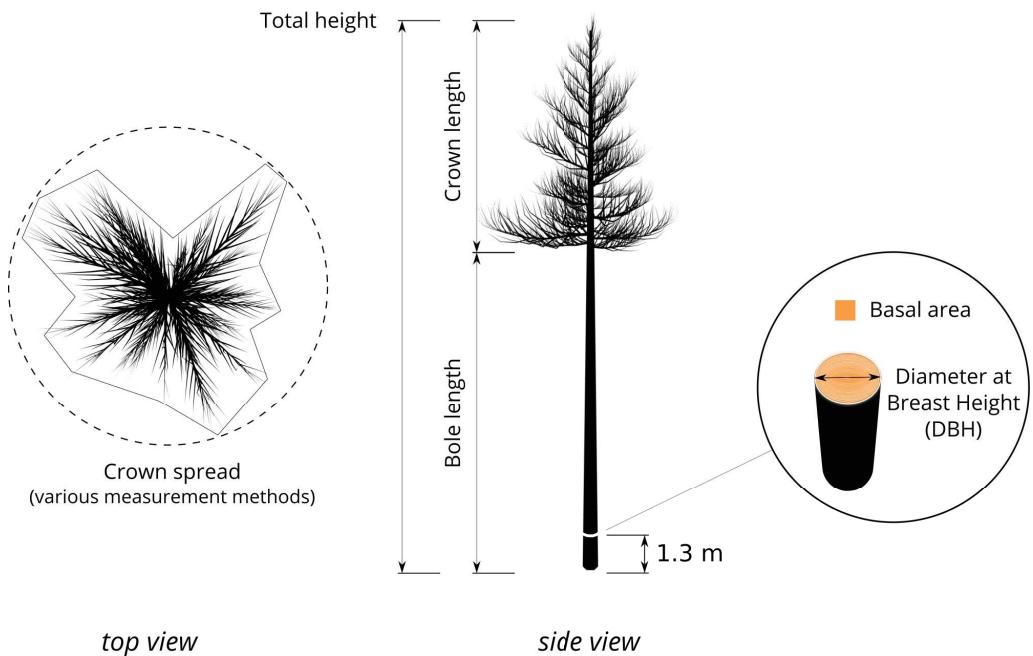


Figure 1.1: Main inventory forest parameters measured at the individual tree scale.

Table 1.1: Main forest inventory parameters and direct measurement methods.

Tree scale	Plot scale	Measurement methods
Stem position	Stem density [ $\text{ha}^{-1}$ ]	GNSS positioning, compass and metric tape, reference grid materialization, total station, terrestrial/aerial laser scanning, aerial photointerpretation
Total / bole height [m]	Height statistics (sd, max., min., mean)	Hypsometer, terrestrial/aerial laser scanning, aerial photointerpretation
Basal area [ $\text{m}^2$ ]	Total basal area [ $\text{m}^2/\text{ha}$ ]	Diameter tape, caliper, Biltmore stick, relascope, terrestrial/aerial laser scanning
Taper [cm/m]	-	Diameter tape, caliper, Biltmore stick, terrestrial/aerial laser scanning
Volume [ $\text{m}^3$ ]	Total volume [ $\text{m}^3/\text{ha}$ ]	terrestrial/aerial laser scanning, destructive sampling
Crown spread	Plot limits	Measurement tape, aerial photointerpretation, terrestrial/aerial laser scanning
Species	Species distribution	Field observation, aerial photointerpretation

### 1.3 Airborne laser scanning

Airborne Laser Scanning (ALS), also called airborne LiDAR, is a remote sensing technique which produces high resolution 3D models of the land surface. ALS systems generally include four components:

- an **aircraft** which can be manned or remotely operated. It precisely follows planned flight trajectories to fully cover the area of interest with overlapping scan swaths.
- a **laser scanner** used to measure the range ( $R$ ) between the sensor and the surface. Current laser ranging uses either direct (time of flight) or indirect (phase shift) measurements to determine the distance to the target. The former method is generally used for long range measurements and the latter for short to medium ranges. A mechanism continuously rotates the laser to generate a scan pattern (which may differ between instruments), allowing it to sample a swath of the land surface.
- a differential **Global Navigation Satellite System (GNSS) receiver** used to measure the exact position ( $X, Y, Z$ ) of the aircraft.
- an **Inertial Measurement Unit (IMU)** used to measure the orientation (attitude) angles (roll, pitch, yaw) of the aircraft and to improve GNSS derived positions (integrated sensor orientation).

ALS works by emitting high frequency (typically hundreds of kHz) light pulses (usually in the infrared or green domain) towards the land surface and measuring the intensity of the reflection as a function of time (waveform). Peaks (also called echoes) in the return signal of each pulse correspond to locations where the laser beam intersected an object (cf. figure 1.2). By simultaneously measuring the scan angle ( $\theta$ ) and the range ( $R$ ), it is possible to determine the 3D position of these intersections (reflection peaks) relative to the sensor. These peaks can be then characterized in terms of amplitude and width. Moreover, since the accurate location and orientation of the aircraft are continuously measured by GNSS/IMU navigation, the laser intersection point coordinates can also be determined in an absolute spatial reference frame.

The interaction of the laser with the atmosphere and surface can be modeled by the RaDAR/LiDAR range equation (Vain and Kaasalainen, 2011; Kashani et al., 2015):

$$P_r = \frac{D_r^2 \eta_{atm} \eta_{sys} \sigma P_t}{4\pi R^4 \beta_t^2} \quad (1.1)$$

$$\sigma = \frac{4\pi}{\Omega} \rho A_t \quad (1.2)$$

where:

- $P_r$  is the received power [W];  
 $P_t$  is the transmitted power [W];  
 $D_r$  is the aperture diameter [m];  
 $\eta_{atm}$  is the atmospheric transmittance;  
 $\eta_{sys}$  is the system transmittance;  
 $\sigma$  is the effective target cross-section [ $m^2$ ];  
 $R$  is the range from sensor to target [m];  
 $\beta_t$  is the width of the laser beam [m];  
 $\Omega$  is the scattering solid angle [sr];  
 $\rho$  is the reflectance of the target;  
 $A_t$  is the area of the target [ $m^2$ ].

Thus, ALS produces a collection of georeferenced waveforms (i.e. continuous representation) and dense 3D point clouds (i.e. discrete representation). Current long range ALS systems routinely attain sub-decimetric point position accuracy. Moreover, ALS is theoretically able to measure the reflectance of a target (cf. equation 1.1). However, in practice accurately estimating all of the parameters affecting reflectance is very difficult and the measured value can generally only be interpreted in a relative sense.

Although it preserves only part of the original measurement, the 3D point cloud representation is the most commonly used, because it requires much less storage space and is computationally less expensive to process. From a practical point of view, it can also be noted that software to process full waveform data is currently relatively scarce compared to what is available to handle discrete data. After acquisition, the points are usually classified into several standard land cover categories (e.g. ground, water, buildings, high vegetation) defined by the American Society of Photogrammetry and Remote Sensing (ASPRS), using a combination of automatic and manual classification procedures. The final dataset is generally stored and distributed using the binary ASPRS LAS format (American Society for Photogrammetry and Remote Sensing, 2013).

Most of the currently operational ALS systems are manufactured by RIEGL, Teledyne Optech, Leica Geosystems and Velodyne.

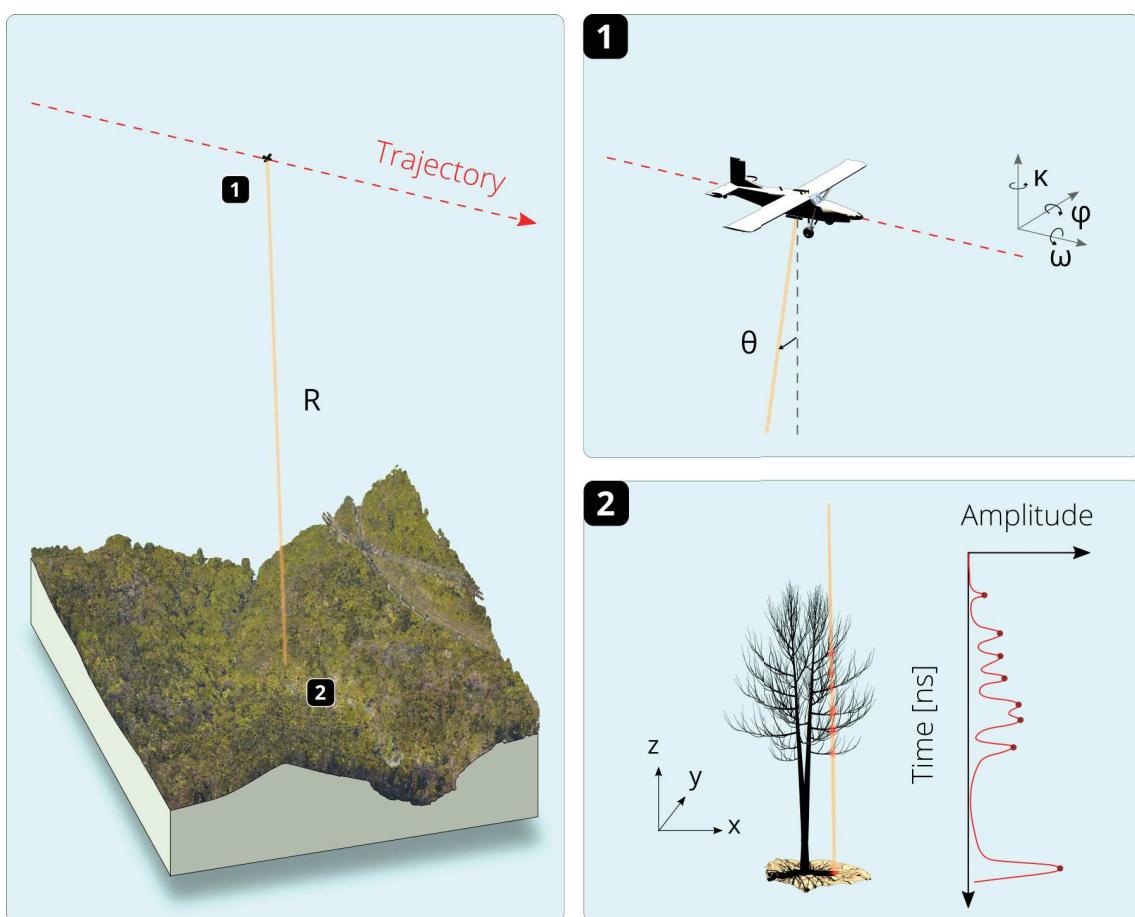


Figure 1.2: Airborne Laser Scanning (ALS) basics. Background ALS data is a courtesy of the state of Neuchâtel.

The following paragraphs provide a brief review of the main ALS acquisition parameters that are important for forest characterization. Several authors (e.g. Næsset (2005); Hyppä et al. (2005); Ørka et al. (2010); Disney et al. (2010); Jakubowski et al. (2013a); White et al. (2013a); Hovi et al. (2016)) have investigated the influence of acquisition conditions on products derived from ALS, we refer the reader to their work for further details.

The following parameters influence ALS data characteristics and should be considered when analyzing forest areas:

- the **phenological phase** of the vegetation;
- the **instrument and flight parameters** (e.g. pulse repetition rate, scan angle, flying height, aircraft speed);
- the **processing of the full waveform signal** (echo intensity in particular).

The phenological phase determines the penetration distance of the laser through the canopy and thus influences the vertical distribution of the echoes (cf. figure 1.3). The presence of leaves reduces the penetration distance resulting in poor sampling of the structures underlying the canopy (i.e. branches, stem, terrain). This can also induce significant errors in terrain modeling which then propagate to tree height estimates. In leaf-off acquisitions, differential canopy opacity can be exploited to map persistent and deciduous species, because the opacity difference affects the number of echoes per pulse and their intensity (Liang et al., 2007). The strong opacity of coniferous canopies results in fewer and more intense echoes, while the weak opacity of deciduous canopies results in numerous lower intensity echoes. This contrast is further increased by the strong reflectance of photosynthetic pigments in the near infrared wavelength used by most ALS systems.

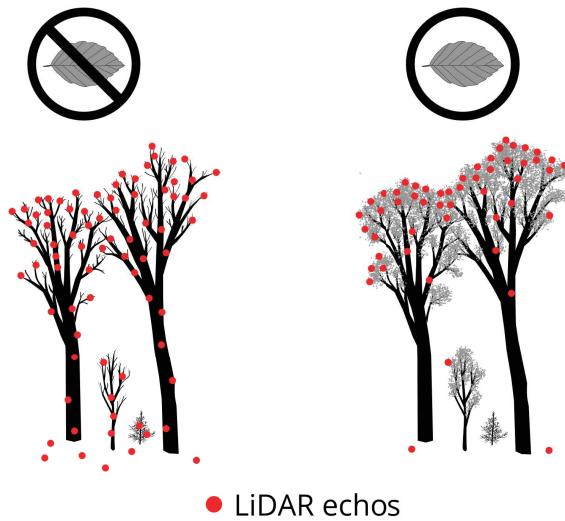


Figure 1.3: Effects of the phenological phase on ALS point distribution.

Combined with the phenological phase, the point density will determine how much structural detail is observable. This parameter is dependent on the flight configuration (height, speed, swath overlap), instrument settings (transmitted energy, pulse repetition rate, number of echoes per pulse, digitization of the waveform) and post-processing (e.g. filtering of the points). The lower the point density, the greater the reliance on interpolation, when calculating elevation models (terrain, surface and canopy height). Point density is often inhomogeneous, either because of voluntary changes in resolution (e.g. a differentiation between urban, rural and / or mountainous areas), or

because of the variability of scan swath overlap. It can also be pointed out that very high point densities ( $> 100$  points /  $m^2$ ) make it possible to directly measure stem diameter in the point cloud without resorting to allometric estimates (cf. chapter 4.4). Point density has also been shown to be positively correlated with tree species classification performances (Li et al., 2013).

The flying height of the aircraft and the scan angle of the laser beam determine the distance between the sensor and the observed surface. Associated with the divergence of the laser beam, this distance in turn determines the size of the projected footprint on the target surfaces. Decreasing the flying height reduces the footprint (concentrates the energy of the pulse) and generally allows better laser penetration through the canopy, but reduces the likelihood of having multiple echoes. Conversely, increasing the flying height widens the footprint (diffuses the energy of the pulse) which increases the probability of obtaining multiple echoes. Finally, the planimetric and altimetric accuracy of the 3D points is inversely proportional to the flying height.

The last important parameter is the processing of the full waveform signal and in particular the echo intensity. Proper calibration of echo intensity ensures comparability of measurements made under different conditions (e.g. flying height, topography, different instruments). It is particularly important for analysis related to foliage persistence and species classification. However, rigorous calibration / correction of intensity requires relatively complex modeling (Höfle and Pfeifer, 2007; Kashani et al., 2015) which is rarely implemented in practice. Only the correction related to the distance between the sensor and the observed surface is sometimes applied.

Considering the importance of these parameters, the implementation of an optimal acquisition protocol and the standardization of metadata, like the steps undertaken by the North American forest and geological services (Heidemann, 2014; White et al., 2013a; Gatziolis et al., 2008) would ensure better integration of ALS into Swiss forestry.

Current and near future developments in ALS include new low-cost high performance sensors (e.g. flash LiDAR) and aircrafts (in particular unmanned). New ALS systems are able to produce very dense 3D point clouds ( $> 100$  points /  $m^2$ ) which offer a lot of potential for work at the individual tree scale including direct stem diameter estimation (Jaakkola et al., 2017) and derivation of inner crown characteristics (Harikumar et al., 2017b). Multispectral ALS sensors such as the Optech Titan (triple wavelength) and Riegl VQ-1560i-DW (dual wavelength) also provide additional features that could help land cover and tree species characterization (Vauhkonen et al., 2013; Hopkinson et al., 2016; Yu et al., 2017; Axelsson et al., 2018).

## 1.4 Airborne hyperspectral imaging

Airborne Hyperspectral Imaging (AHI) is a remote sensing technique which is used to measure how materials interact with solar radiation (light). More specifically, AHI is used to estimate for a given surface the fraction of incoming solar radiation that is reflected as a function of wavelength; a property called reflectance and designated by the greek letter  $\rho$  (cf. figure 1.4).

The distinction between multispectral and hyperspectral sensors is not precisely defined. However, generally speaking, hyperspectral sensors have a number of bands at least an order of magnitude above multispectral sensors, typically in the 50-300 range. Another distinction is spectral resolution (bandwidth); hyperspectral sensors have narrow bands (a few nanometers wide), while multispectral sensors tend to have broader bands (tens of nanometers wide). AHI systems generally include the following components:

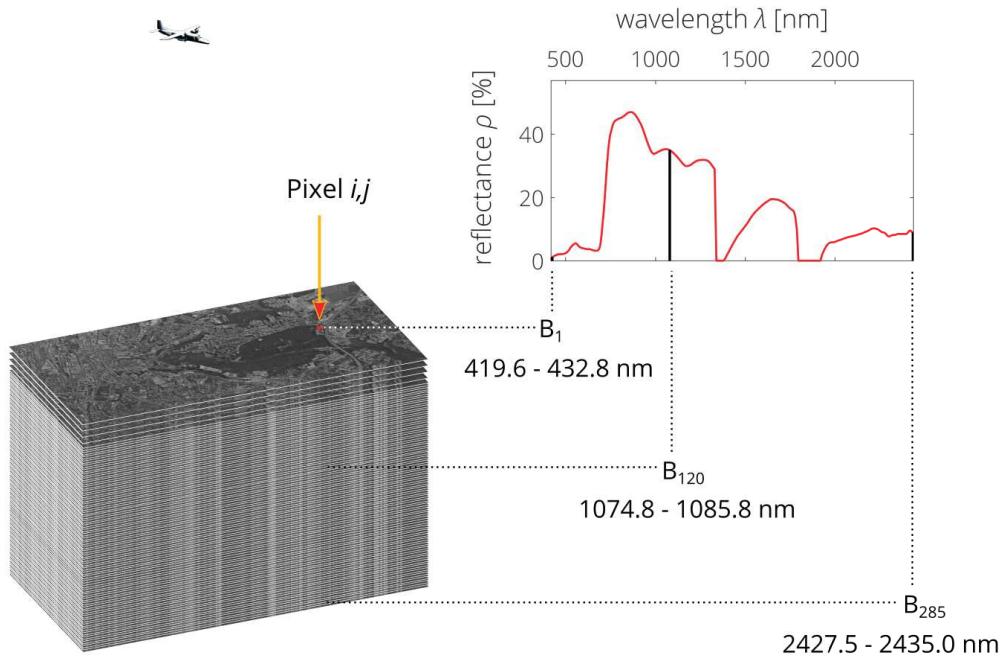
- an **aircraft** which can be manned or remotely operated.
- a **hyperspectral camera** used to measure solar radiation reflected off the observed surface. The reflected light passes through the camera lens and different regions of the spectrum are separated with a dispersive prism/grating or a series of optical bandpass filters. The individual filtered light bands are then redirected to an array of photosensitive elements. These elements, which can be either semiconductor Charge Coupled Devices (CCD) or Complementary Metal Oxide Semiconductors (CMOS), convert the separated light bands into a digital signal. The configuration of the photosensitive element array may vary, with the most common being along track (pushbroom) and across-track (whiskbroom) linear arrays, and less commonly snapshot (staring) arrays. Hyperspectral cameras require rigorous laboratory and/or in flight calibration with reference targets of known reflectance and regular measurement consistency checks.
- a **solar (irradiance) photometer** used to measure incoming solar radiation. It may be located either near the observed surface (on the ground) or on the aircraft.
- (Optionally) a differential **Global Navigation Satellite System (GNSS) receiver** used to measure the exact position (X, Y, Z) of the aircraft.
- (Optionally) an **Inertial Measurement Unit (IMU)** used to measure the orientation (attitude) angles (roll, pitch, yaw) of the aircraft and to improve GNSS derived positions (integrated sensor orientation).

Accurate measurement of reflectance is complicated, because the characteristics of light that arrive at the sensor depend on many factors, including the surface geometry, the illumination angle and the viewing angle (for non-Lambertian surfaces). For a given surface, the effect of these factors can be physically modeled and corrected with a Bidirectional Reflectance Distribution Function (BRDF) defined for each wavelength. In practice however, detailed physical BRDF modeling of the observed surface is often too complex to be carried out and simpler empirical BRDF correction procedures are applied instead (Schlerf and Atzberger, 2006a). In addition to BRDF corrections, an atmospheric correction is applied to compensate light scattering and absorption by particles and gases present in the atmosphere (Gao et al., 2009). Geometric correction (orthorectification) is also applied to remove lens distortion and perspective effects. This step involves removing radial/tangential distortions and projecting the images onto a 3D model of the observed surface (obtained by photogrammetry or by an auxiliary technique such as ALS). Finally, georeferencing is applied to position the images in an absolute spatial reference system. If GNSS/IMU navigation was available on the aircraft, direct georeferencing is possible. Otherwise, indirect georeferencing using Ground Control Points (GCP) or spatial co-registration with other imagery is used. The direct and indirect approaches can also be combined to improve image georeferencing accuracy (McGlone et al., 2013).

Several AHI acquisition parameters are critical for vegetation analysis:

- the **phenological phase** influences leaf geometry, pigmentation, nutrient content and water content. Most studies use hyperspectral data that was acquired during full leaf deployment periods. However, the effect of phenology at time of acquisition on spectral separability of species has not been extensively investigated, with only a limited number of studies (Key et al., 2001a; Dennison and Roberts, 2003; Voss and Sugumaran, 2008; Hill et al., 2010; Hesketh and Sánchez-Azofeifa, 2012; Somers and Asner, 2013, 2014; Nikopensius et al., 2015; Richter et al., 2016) and no consensual conclusions on a single optimal observation time. The use of multi-temporal acquisitions has been shown to be useful for tree species and health characterization (Key et al., 2001a; Liu et al., 2006; Hill et al., 2010; Richter et al., 2016; Tigges et al., 2013).
- the **illumination conditions**. Since AHI relies entirely on natural illumination, it is strongly affected by shadowing and atmospheric effects. To reduce shadows and illumination intensity variation, acquisitions are preferably conducted when the sun is at its highest elevation and the sky is clear or fully overcast. Sunlit observations have been shown to help spectral separability of species (Leckie et al., 2005; Clark et al., 2005; Puttonen et al., 2009).
- the **spatial resolution**, also called Ground Sampling Distance (GSD), which is essentially constrained by flying height, aircraft speed and instrument design. Measured reflectance has been shown to be dependent on observation scale for vegetation (Williams, 1991; Roberts et al., 2004; Clark et al., 2005). Spatial resolutions similar to or larger than tree crown sizes necessarily induces spectral mixing. Systematic investigations of the effects of spatial resolution for forest analysis are scarce (Rahman et al., 2003; Nijland et al., 2009; Roth et al., 2015). It has also been suggested that the resolution should not exceed half the size of the target object (tree crown) size (Hengl, 2006) (i.e. at least four pixels per crown). On the other hand, Nagendra and Rocchini (2008) have argued that using pixels smaller than individual tree crowns increases the variability of the spectral signatures of species when aggregated at the crown scale (thus making species identification more difficult). However, there is no clear consensus on this issue and it has also been shown that finer spatial resolutions could provide better tree species spectral separability (Clark et al., 2005; Baldeck et al., 2015). The optimal spatial resolution may depend on the application (Aplin, 2006; Stoy et al., 2009).
- the **spectral resolution** (bandwidth), the number and location of bands determine how well the reflectance spectra is sampled. Subtle discriminative spectral features might not be observable with broad bands and/or insufficient sampling of the different spectral regions.
- the **quality of georeferencing and geometric correction** (orthorectification) is particularly important because it ensures adequate co-registration with other datasets (e.g. field surveys, ALS). AHI images used for vegetation analysis should be orthorectified using a surface model (not a terrain model).
- the **quality of the BRDF correction** influences the variability of the reflectance maps and ability to conduct species classification (Korpela et al., 2011).

Current operational AHI systems are mostly developed and used in the realm of research. Such systems include the Carnegie Airborne Observatory (CAO), the Airborne Visible / Infrared Imaging Spectrometer (AVIRIS), the Compact Airborne Spectrographic Imager (CASI) and the Airborne Prism Experiment (APEX). Several commercial manufacturers, including HySpex, Headwall Photonics (e.g. Nano-Hyperspec VNIR, Micro-Hyperspec SWIR), Integrated Spectronics (e.g. HyMap), Specim (e.g. AISA Eagle, AisaFENIX) also provide AHI systems. Due to scarce commercial service availability, processing complexity and high costs, it is unclear if AHI will be adopted in forestry in the near future in the same way as ALS. Forestry applications might instead rely on less expensive and more available very high resolution RGB and/or multispectral imagery.



(a) A hyperspectral image consists in a stack of reflectance bands each covering a different part of the light spectrum. Looking at the reflectance of a single pixel across all bands provides the spectral signature of the observed surface.



(b) True color composite image (red: 650 nm, green: 540 nm, blue: 440 nm) over the Boudry area (Neuchâtel, Switzerland) obtained with the APEX hyperspectral sensor.

(c) False color composite image (red: 800 nm, green: 650 nm, blue: 540 nm) over the Boudry area (Neuchâtel, Switzerland) obtained with the APEX hyperspectral sensor.

Figure 1.4: Airborne Hyperspectral Imaging (AHI) basics.



## 2. Data

This chapter presents the study sites, remote sensing data characteristics and describes the creation of a novel reference dataset used across the thesis. It is structured in the following way:

- 2.1 provides an overview of forest characteristics and ALS/AHI remote sensing data availability in Switzerland. It also presents the location of the sites studied. Due to its voluminous format, the detailed metadata about study sites and remote sensing data is reported in the appendix.
- 2.2 presents the workflow used to create a novel reference dataset of 3D tree models extracted from ALS data. The content of this section is adapted from Parkan (2017b) and Parkan et al. (2018).
- 2.3 reviews important sampling considerations.

### 2.1 Overview

Forests cover about one third of Switzerland. Their composition and structure vary between regions, mostly influenced by environmental conditions and silvicultural practices. Switzerland can be divided into three distinct geographic regions:

- the **Jura** mountain chain in the North which culminates at 1720 m (Le Crêt de la Neige, France). It is characterized by mixed forests in lower stages, coniferous dominated forests and woodland pastures in the upper part.
- the **Plateau** located between the Jura and the Alps (altitude ranging from about 400 to 600 m). It is characterized by mixed and deciduous forests.
- the **Alps** in the South which culminate at 4810 m (Mont Blanc, France). In this region, the strong altitude gradient and many valleys provide a wide range of environmental conditions. The high average altitude means that most species in this region are coniferous.

Tree species encountered in Switzerland are typical of European temperate and alpine forests with beech (*Fagus sylvatica*), Norway spruce (*Picea abies*), silver fir (*Abies alba*), sessile oak (*Quercus petraea*), pedunculate oak (*Quercus robur*), larch (*Larix decidua*), Scots pine (*Pinus sylvestris*) and ash (*Fraxinus excelsior*) being the most frequently observed and also those that have the highest commercial value. Several exotic species have been introduced for the timber industry, the most noticeable being the Douglas fir (*Pseudotsuga menziesii*).

Since year 2000, ALS data has been acquired regularly over swiss cantons, with varying acquisition parameters (cf. appendix E.1). Most of the swiss ALS data has a point density greater than 10 points/m<sup>2</sup> and some of the most recent datasets have densities which can reach 80 points/m<sup>2</sup> in flight line overlap areas. Leaf-off acquisitions have generally been favored to obtain a better sampling of the terrain, but a significant fraction of acquisitions have also been acquired with leafs. ALS data is well integrated in state geomatic services and currently reacquired every 4 to 6 years.

In July 2014, medium resolution (2.7-3 m) AHI data was acquired with the Airborne Prism EXperiment (APEX) hyperspectral sensor over Lausanne and Boudry (cf. appendix D.1), for the purpose of this study. AHI is not currently used by any state service and renewal of AHI data is not planned in the near future.

The main criteria used to select the forest survey sites were ALS acquisition dates (leaf-off data was favored), point density (high was favored) and field survey availability. Secondary criteria were diversity of topographic (altitude, slope, aspect) and biological characteristics. Figure 2.1 illustrates the spatial distribution of the study sites, tables in appendix B.1, E.1 and D.1 present the field, AHI and ALS survey metadata. Pictures of selected sites are also presented in appendix C.1.

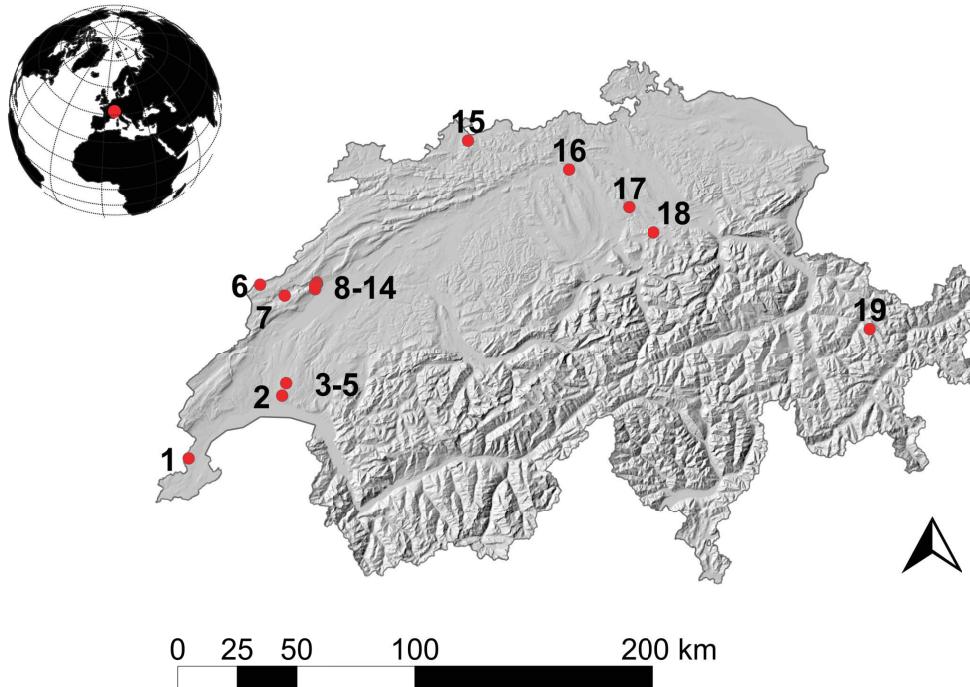


Figure 2.1: Geographic context and map of selected inventory sites in Switzerland. The asterisk next to the site name indicates if it is covered by hyperspectral data. **1:** Versoix, **2:** Sauvabelin\*, **3:** Benenté\*, **4:** Jorat south\*, **5:** Gésiaux\*, **6:** La Brévine, **7:** Couvet, **8:** Cortaillod D4\*, **9:** Boudry D20\*, **10:** Boudry D19\*, **11:** Boudry D1\*, **12:** Chambrelien\*, **13:** Rochefort\*, **14:** Bevaix\*, **15:** Grosszinggibrunn, **16:** Ottmarsingen, **17:** Sihlwald, **18:** Oberaegeri, **19:** Dischma. Background terrain elevation data courtesy of the Swiss Federal Office of Topography (swisstopo).

## 2.2 Creation of a benchmark ALS dataset

This section is adapted from Parkan (2017b) and Parkan et al. (2018). It presents a novel reference ALS dataset composed of several forest inventory sites across Switzerland. The dataset is representative of central European temperate broadleaf, mixed and coniferous forests. Individual trees were manually delineated from point clouds to produce a set of detailed 3D reference segments. These segments were then geometrically characterized (height, crown area and volume) and were matched with field surveys to assign Diameter at Breast Height (DBH) and species attributes.

The workflow used to create derivative products from ALS data typically involves a combination of area (stand) or object (tree) based segmentation, classification and regression (allometric modeling) algorithms. A critical aspect of this workflow is quantitative error assessment. This step is generally accomplished by comparing the output of an algorithm to reference values. These reference values may be either simulated or measured with a reliable independent method (typically visual interpretation and/or an intensive field survey).

Simulation could theoretically allow to explore the effects of different ALS acquisition parameters and provide perfect reference datasets. However, modeling the interactions between physical (e.g. topography, atmosphere, surface reflectance, illumination), biological (e.g. stem density, crown diameter distribution, clustering, layering, tree architecture) and sensor (e.g. flying height and speed, pulse repetition rate) related parameters is very complex. The Discrete Anisotropic Radiative Transfer (DART) physical model (Gastellu-Etchegorry et al., 2015), for example, attempts to fully integrate all these parameters. Simpler models have also been used successfully to validate segmentation algorithms. Wang et al. (2011), for example, modeled ALS data by sampling a set of simple 3D geometric primitives (ellipsoids) with different sampling densities, location accuracies, ranging accuracies, and volumetric backscatter probabilities. Similarly, Zhang et al. (2014) set up a library of 3D tree templates (extracted from an ALS point cloud) and then used point processes to model plots with different stem density and crown overlap. Calders et al. (2018) developed a reconfigurable virtual forest stand based on Terrestrial Laser Scanning (TLS) derived tree models.

When compared to simulated ALS datasets, the use of independent measurements as a reference is currently much more widespread. Benchmark studies have been periodically conducted to compare algorithm performance using a common dataset and error assessment framework (Kaartinen et al., 2012a; Vauhkonen et al., 2012; Eysn et al., 2015; Wang et al., 2016b). Recent benchmark datasets, however, lack high density leaf-off data, full-waveform data and 3D tree (crown) boundaries. Some organizations have deployed efforts to openly distribute ALS, field survey data and validation procedures (e.g. NEW technologies for a better mountain FOREst timber mobilization - NewFOR<sup>1</sup>, National Ecological Observatory Network - NEON<sup>2</sup>, Harvard Forest). Nonetheless, the number of available benchmark datasets customized for forest research remains relatively small when compared to what is found in other research communities (e.g. computer vision, machine learning).

Yin and Wang (2016) reviewed qualitative and quantitative error assessment procedures used in remote sensing for forest inventories. They divided the procedures into three categories: summary metrics (e.g. tree detection rate), position accuracy metrics (e.g. stem location accuracy) and attribute accuracy metrics (e.g. dendrometric attributes, species). For each category, they evaluated the strengths and weaknesses of different indices. They concluded that quality assessment could be improved by combining multiple assessment techniques and indices, using optimal sampling schemes when creating validation datasets, using multiscale evaluation (tree and plot scales) and conducting sensitivity analysis. The same authors also noted that 3D crown boundary delineation accuracy has generally been ignored in existing studies and underlined the need of high-quality reference data for benchmark studies. Overall, the inherent reference dataset differences and lack

<sup>1</sup>NewFOR, [<http://www.newfor.net/>], accessed January 23, 2018.

<sup>2</sup>NEON, [<http://www.neonscience.org/data-collection>], accessed January 23, 2018.

of standardization in error assessment procedures also means that the results of different studies are often difficult to compare (Zhen et al., 2016).

The novel dataset presented here addresses these error assessment problems by providing a reliable individual tree segmentation. The dataset also has potential uses beyond the validation of algorithms. It can for example be used to simulate different forest configurations, by removing or duplicating tree models and reconfiguring their spatial arrangement. Such simulated forests can in turn be used with high performance tree detection and classification algorithms (such as deep neural networks) that require very large training datasets. By adding variable amounts of noise to the individual tree models, the dataset can also be used to assess the effect of segmentation quality and to train tree species classification algorithms to handle poor segmentation (Ko et al., 2016). The dataset can also be used to create accurate allometric models. The procedures and tool developed to create the dataset can also be used to prepare detailed and accurate individual tree maps for research or trainings plots, without relying on surveyor expertise.

### 2.2.1 ALS data preparation

All ALS point clouds were converted to the LAS 1.4 format defined by the American Society for Photogrammetry and Remote Sensing (2013). Point data formats 6 and 7 (same as 6 with the addition of RGB colors) were used. The standard LAS 1.4 classification definition was applied to each dataset and non vegetation classes (except terrain) were removed (cf. appendix F.1). When present, the overlap points were preserved and the associated overlap bit was set in the classification record.

A coordinate transformation to the local Swiss reference frame LV95 was applied to the original 3D point clouds, when necessary. This system is composed of the CH1903+ (EPSG:2056) planimetric and LHN95 altimetric reference frames. The metadata for this coordinate reference system was stored in the Variable Length Records (VLR) of the LAS files. The 3D point clouds were then clipped to the extent of the survey sites.

In many cases, the state services that mandate ALS acquisitions are only interested in geometric information and do not require the full waveforms and/or derived attributes such as amplitude, pulse width or pulse deviation. Thus, it is frequent that delivered ALS data only contains a generic 16 bit "intensity" attribute with little or no metadata on how it was processed or how the value should be interpreted. Nonetheless, some information regarding intensity processing was obtained by directly contacting data providers and instrument manufacturers.

Many LiDAR instruments have a high dynamic range meaning the ratio of the minimum and maximum measurable intensity (also called amplitude) value is very large and may span several orders of magnitude (Riegl Laser Measurement Systems, 2017). Thus, it is common to use decibel units (i.e. a logarithmic scale) to represent the intensity:

$$I[dB] = 10 \cdot \log_{10} \left( \frac{P_t}{P_{ref}} \right) \quad (2.1)$$

where:

$P_t$  is the transmitted power [W];

$P_{ref}$  is the minimum detectable power [W].

A relatively standard procedure found across LiDAR software is to apply a scale and offset transformation to the original intensity values to fit the 16 bit (0-65535) storage range used in the LAS format. The scaled value can be unscaled, using the following formula:

$$v_u = \frac{(v_{max} - v_{min}) \cdot v_s}{65535} + v_{min} \quad (2.2)$$

where:

- $v_u$  is the unscaled value;
- $v_s$  is the scaled value;
- $v_{max}$  is the maximum value;
- $v_{min}$  is the minimum value.

Using this procedure, the data acquired with the RIEGL LMS-Q1560 was scaled using a minimum value of -20.0 dB and a maximum value of 15.0 dB. Thus, for example, a value of 24'000 in the LAS file can be unscaled using:

$$v_u = \frac{(15 + 20) \cdot 24000}{65535} - 15 = -2.182 \quad [dB] \quad (2.3)$$

Another relatively common procedure is to apply a correction for the sensor to target range by applying the formula described in Luzum et al. (2004):

$$I_n = I \cdot \left( \frac{r}{r_{ref}} \right)^2 \quad (2.4)$$

where:

- $I_n$  is the normalized intensity;
- $I$  is the raw intensity;
- $r$  is the range between the sensor and the point;
- $r_{ref}$  is an arbitrary constant reference range (1000 m was used here).

Riegl LiDAR software allows exporting either the amplitude (range dependent value) or a range corrected value called reflectance to the "intensity" LAS field. For the datasets acquired with the Riegl LMS-Q1560 (cf. appendix E.1), the intensity stored in the LAS files is the reflectance (i.e. range independent value) scaled to a 16 bit range (0-65535) (Riegl Laser Measurement Systems, 2017). For datasets acquired with the Optech ALTM Gemini (cf. appendix E.1), the data provider has indicated that the raw intensity was corrected for range dependence using the formula described in Luzum et al. (2004). No information was obtained about the intensity processing of data acquired with the Riegl LMS-Q680i, Riegl LMS-Q780 and Trimble AX60.

Thematic attributes were stored directly in the LAS 1.4 files. To identify individual tree segments, a 32 bit Locally Unique Identifier (LUID) extra field was added to the point record. The "user\_data" field was used to store phenology flags (0 = no leafs, 1 = partial leaf deployment, 2 = full leaf deployment). The LAS 1.4 specification provides a mechanism (called LAS Domain Profile) which allows the definition of custom records for domain specific applications (such as forestry). This mechanism was used to store the dendrometric attributes in the Extended Variable Length Records (cf. table G.1).

## 2.2.2 Manual segmentation

Using a custom Matlab based interactive application, each tree was manually delineated from the 3D point clouds by iteratively cropping out points until only the region (tree) of interest remained (cf. figure 2.2). As points were cropped out, the view was progressively zoomed in and rotated as necessary, to help identification of finer details. Horizontal cross-sections at different heights were also employed in some cases, to distinguish adjacent or interlocking structures. Multiple features were used to delineate and identify individual trees:

- **Spacing between crowns** was particularly useful in low density plots and in some older broadleaf forests where individual trees exhibit crown shyness;
- **Intensity difference between structures within the tree**; larger opaque structures (stem, large branches) and photosynthetically active parts typically result in higher echo intensity. This feature was also useful in mixed forest to discriminate deciduous and persistent foliage;
- When available, **color** helped to discriminate foliage persistence types;
- **Prior knowledge about the tree shape and color** (Oester, 2003; Johnson, 2006; Sayn-Wittgenstein, 1978) of different species;
- **High resolution drone aerial photography** was acquired on some of the sites to help species identification;
- **Field survey maps** which indicated the location, diameter and species of individual trees. The accuracy of stem positions was highly variable depending on the type of field survey and ranged from sub-metric for scientific research plots to sub-decametric for non-research plots. Field survey data originated from private inventories, long term monitoring and research plots (Schaub et al., 2011) and forest educational/training plots (Junod and Ammann, 2018).

The above indicators were insufficient to unambiguously delineate all tree shapes. Thus, some of the observations included in the dataset were flagged as ambiguous (e.g. suspicion of multiple trees, strongly interlocking branches, no match with field survey). These ambiguous observations were excluded from the counts provided in the introduction and in appendix E.1. In some areas, individual tree shapes could not be delineated at all, due to clutter and/or low point densities. When the tree species could not be determined with certainty, only a visual distinction between angiosperms and gymnosperms was conducted.

Manual segmentation of CHM has previously been used to validate segmentation algorithms (e.g. Heinzel et al. (2011)). However, to the best of the author's knowledge, manual segmentation of the point cloud over large extents has not been used previously for this purpose. This approach also has several potential practical applications in precision forestry; for example creating detailed reference maps to support field operations like inventories or timber marking (cruising). In this context, the use of direct visual interpretation is interesting, because it is much easier to implement than automatic segmentation algorithms and allows a quick and reliable identification of zonal (e.g. forest edge, deciduous/persistent ratio), structural (stratification) and point (e.g. tree/stem position and height) features. It could also be used as a complement and eventually replace current aerial stereo photointerpretation procedures used in the national forest inventory.

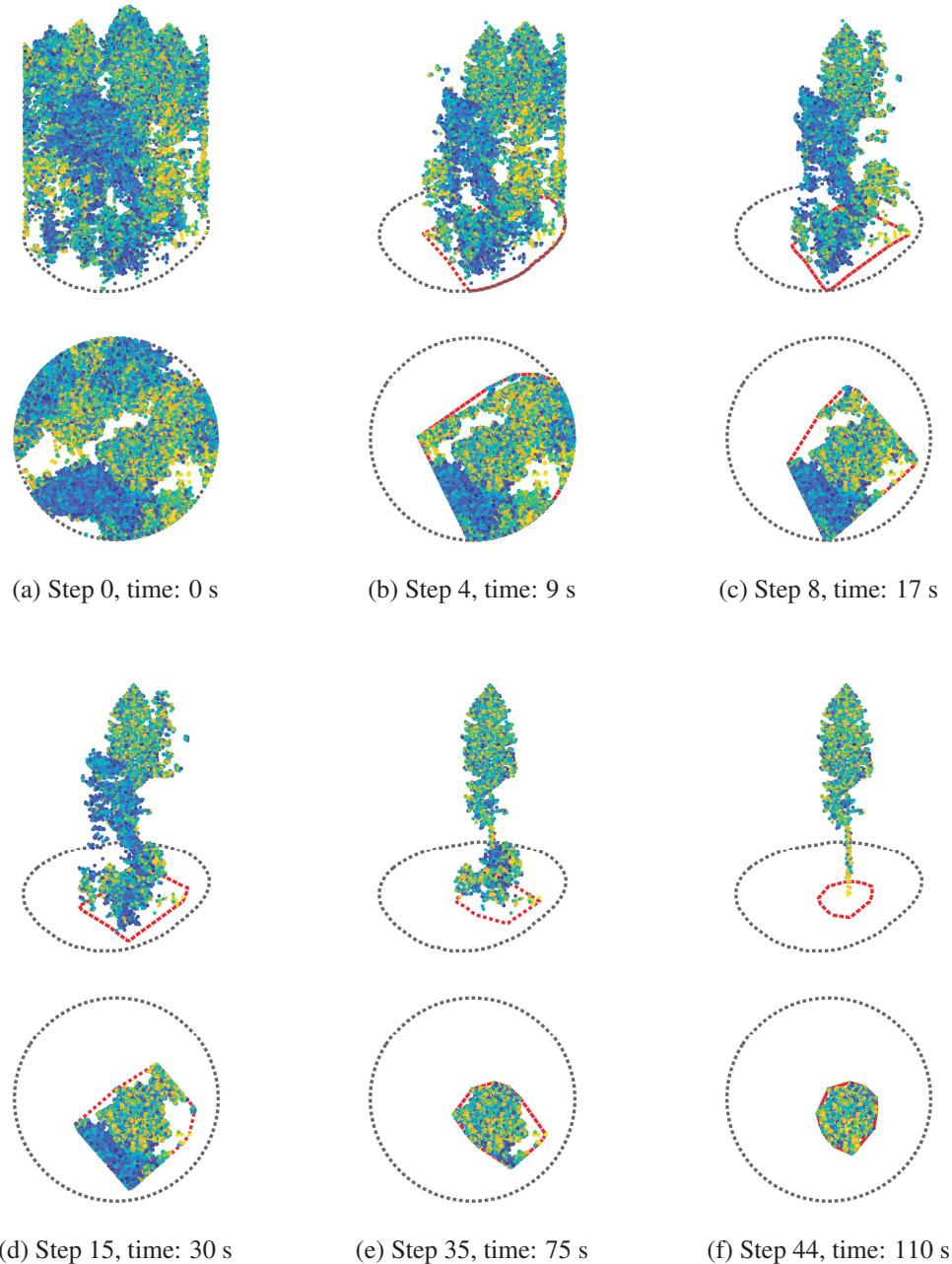


Figure 2.2: Side and top view of a manual tree segmentation sequence in 44 steps (duration 110 s). The black circle represents the outline of the original sample area. The red polygon represents the convex outline of the remaining subset of points after each cropping step. The points are progressively removed from the sample until only the region of interest remains. The points are colored by return intensity to simplify the distinction of deciduous/coniferous crown edge limits, large branches and stems.

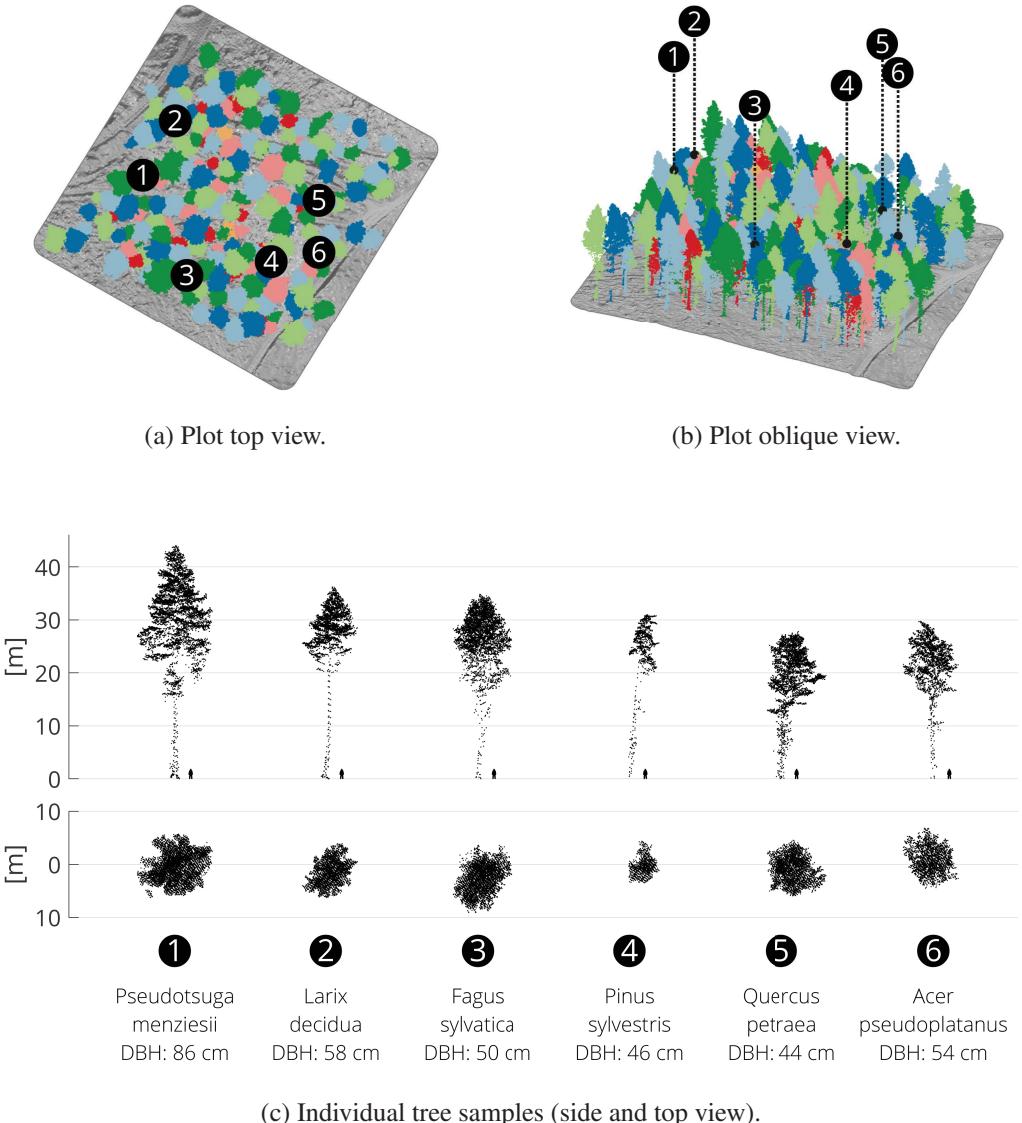


Figure 2.3: Central part of the ALS based virtual survey at the Boudry D20 site (Neuchâtel). Topological coloring (Welsh and Powell, 1967) of the point cloud is used to visually differentiate adjacent trees in sub-figures (a) and (b). A shaded terrain model helps to provide contrast and highlights topographic features (e.g. roads, ditches, skid trails). To avoid visual clutter, only surveyed trees are displayed.

### 2.2.3 Characterization

All the manually delineated segments were characterized by several geometric attributes (stem position, height, crown volume and area). Depending on the spatial point distribution within the segments, the stem position was determined using either the base, apex or centroid xy coordinate as a proxy and projecting it onto the terrain model (cf. figure 2.4). The tree height was computed as the distance between the stem position and the apex of the segment (cf. figure 2.5a). The crown area and volume were respectively derived from the 2D and 3D single region concave hulls ( $\alpha$  shapes) (Edelsbrunner and Mücke, 1994) of the segments (cf. figures 2.5b and 2.5c). The resulting stem positions and associated geometric attributes were then exported to GIS software (Quantum GIS) in order to match them with previously established field survey maps of stem position, diameter and species. The matching was based on spatial proximity, height-diameter coherence and species

traits (cf. figure 5.3). This step allowed to assign a diameter at breast height (1.3 m above ground) and species to most of the 3D segments.

No uncertainty estimate is available for the diameter. This value can be affected by observer bias and measurement technique (e.g. caliper, diameter tape, terrestrial laser scanner). Studies on diameter measurement differences resulting from observer bias have generally reported relatively small errors. For example, Omule (1980) reports a bias of 0.09%, Elzinga et al. (2005) reports that 1.9% of observations had a diameter difference  $\geq 10\%$ , Luoma et al. (2017) report standard deviations ranging from 1.3% to 1.9%. In this regard, the asynchronous collection of field and ALS data is a much more significant source of diameter uncertainty, with a time shift of up to 4 years in the dataset (cf. appendices B.1 and E.1).

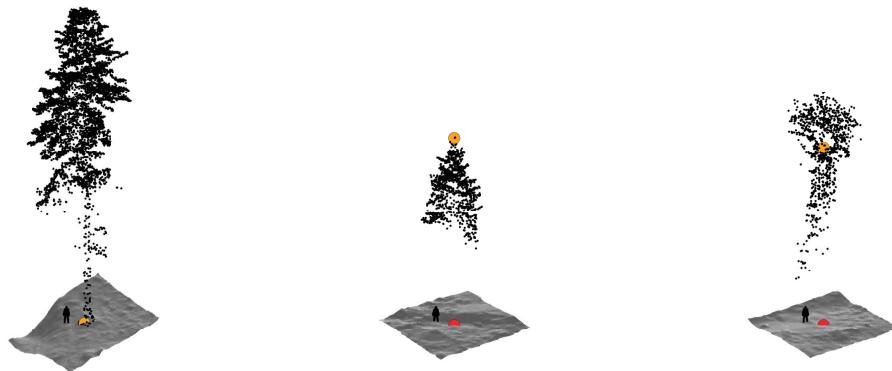
Observations from two of the sites (Benenté and Ottmarsingen) were used to compare field and ALS measured tree positions and heights. To match the two sets of observations (field and ALS), a simple algorithm was used. A match occurs, if an ALS observation is within a 2.5 m horizontal distance and 20% height difference from a field observation. In cases where multiple ALS observations fulfill these criteria, the closest (in terms of 3D euclidean distance between tree tops) is retained. The comparison results are reported in table 2.1 and maps of the matched trees are presented in figures 2.6 and 2.7.

Table 2.1: Differences between ALS and field measured positions and heights for the Benenté and Ottmarsingen sites.

Site	Matches	Position		Height		
		Bias $\pm 2\sigma$ [m]	RMSE [m]	Bias $\pm 2\sigma$ [m]	Rel. Bias $\pm 2\sigma$	
Benenté	144	0.81 $\pm 0.61$ m	2.68 m	-1.85 $\pm 3.87$ m	-0.08 $\pm 0.12$	
Ottmarsingen	89	1.14 $\pm 0.75$ m	3.06 m	0.39 $\pm 6.11$ m	0.02 $\pm 0.16$	

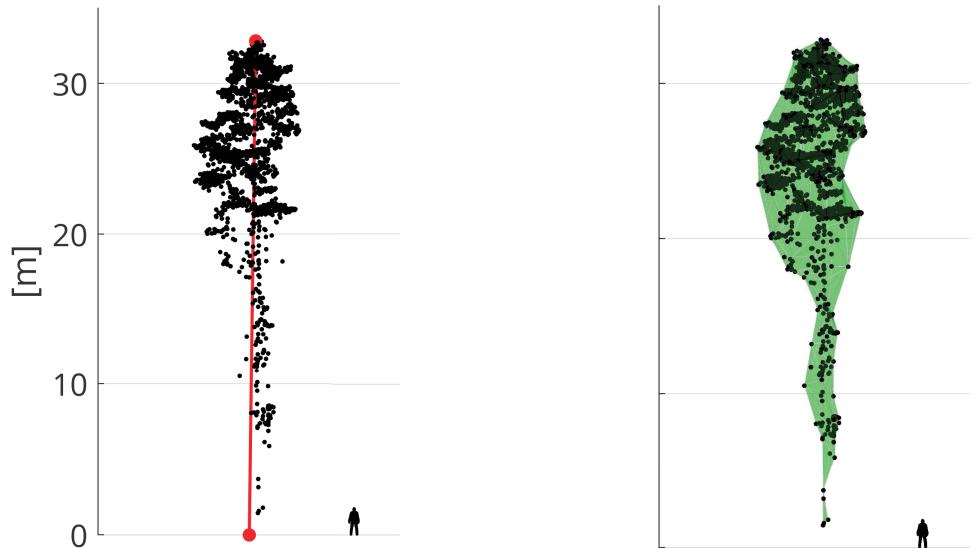
It has been reported that the difference between ALS and field measured tree heights is generally within  $\pm 1.5$  m (Andersen et al., 2006). Thus, the differences reported in table 2.1 are consistent with previous studies. When comparing ALS and field measured tree heights, it is important to keep in mind that accurate field measurement of height is difficult (in particular for tall trees) and may contain a significant amount of error. Moreover, as noted previously, the field and ALS measurement did not occur at the same time. Thus, results of the above height comparison should not be interpreted as an accuracy assessment but rather as a coarse validation of coherence.

The accuracy of crown edge delineation was not evaluated, as no reference field measurements were available for this attribute. Moreover, crown edge limits are notoriously difficult to measure accurately in the field for large trees (in particular for those with asymmetric crowns).



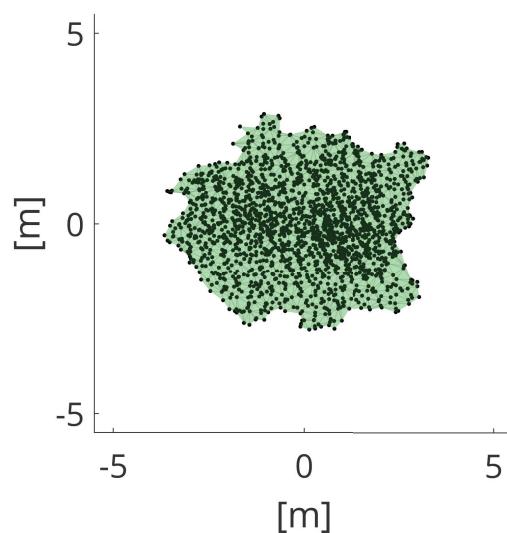
- (a) The mean XY position of points located in the lower 50 cm of the segment (root point) is used as a stem proxy.
- (b) The mean XY position of points located in the upper 50 cm of the segment (apex point) is used as a stem proxy.
- (c) Neither the root nor the apex is well defined, so the mean XY position of all points located in the segment (centroid point) is used as a stem proxy.

Figure 2.4: Estimating the stem position. Depending on the 3D segment geometry, three different stem position proxies can be used (i.e. root, apex, centroid). The proxy point (orange dot) is then projected onto the terrain model, to obtain the estimated stem position (red dot).



(a) The tree height is estimated from the distance (red line) between the root point (stem proxy) and the apex (highest) point.

(b) The tree volume is estimated from the single region 3D alpha shape (in green) of the XYZ point coordinates



(c) The tree area is estimated from the single region 2D alpha shape (in green) of the XY point coordinates.

Figure 2.5: Geometric attributes derived from the 3D segments.

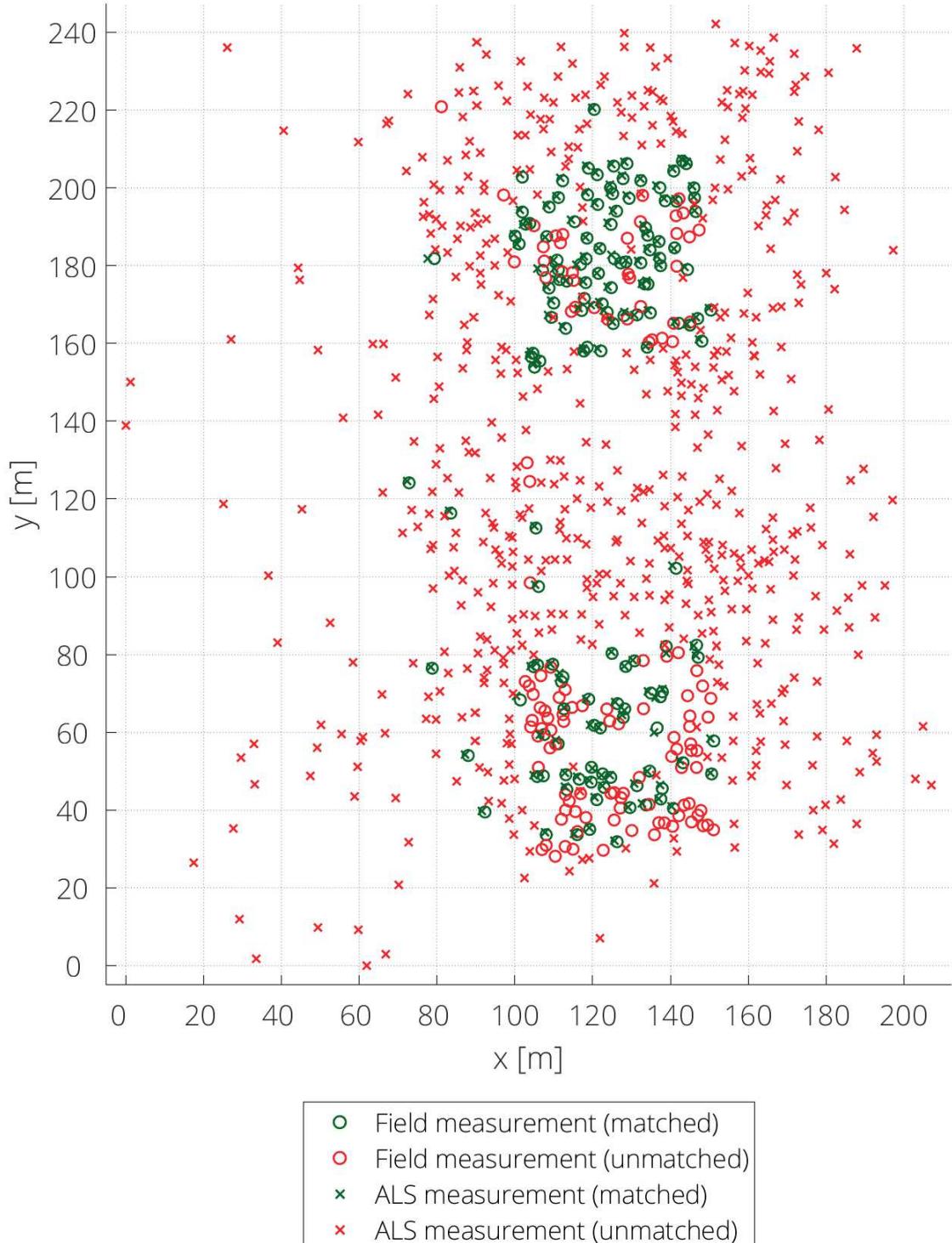


Figure 2.6: Matching of field and ALS tree positions for the Benenté site. Note that only field measurements where the tree height was available were used to assess height and position differences.

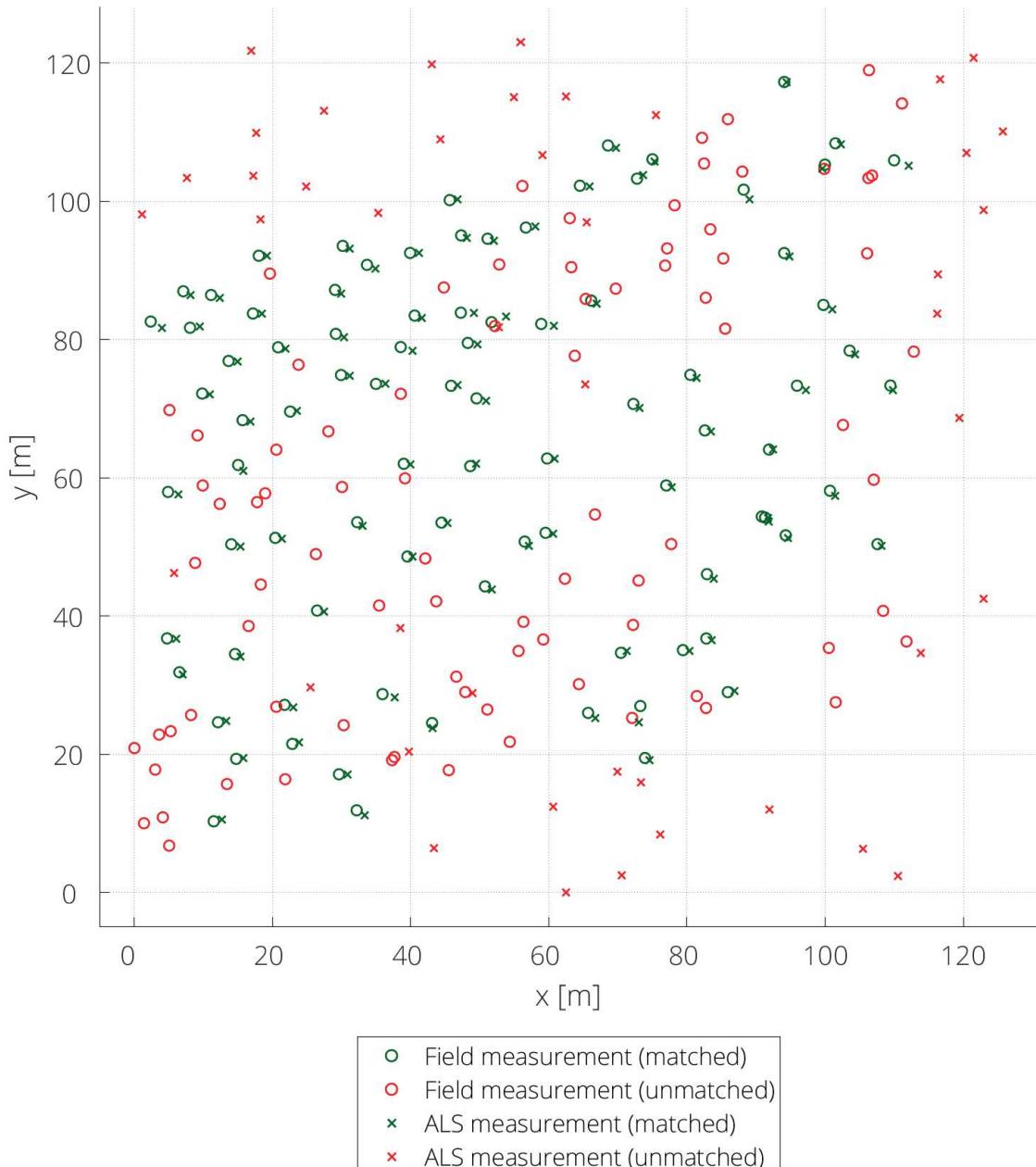


Figure 2.7: Matching of field and ALS tree positions for the Ottmarsingen site. Note that only field measurements where the tree height was available were used to assess height and position differences.

## 2.3 Sampling strategy

In many forest settings, tree species and height frequency distributions are imbalanced (i.e. presence of dominant species and non-uniform height/shape distributions). If this imbalance is not taken into account, models become overly adapted to the most frequent species/height classes and do not generalize well to other species/height classes. Consequently, this leads to an overly optimistic or pessimistic error assessment. Commonly used strategies to mitigate the effect of imbalanced species/height distributions include oversampling minority classes or undersampling majority classes, introducing classification/regression weights (costs) inversely proportional to class frequencies, working with variables aggregated by class (Duncanson et al., 2015; Jucker et al., 2016). For classification algorithms that optimize hyper-parameters based a performance metric, using a metric that takes into account the classification score on each class (e.g. average precision) rather than the overall accuracy can also help mitigate the effects of class imbalance.

Additionally, a standard procedure to improve model generalization is to train the classification/regression model on a subset of the observations and validate it on the remaining observations. This procedure can be repeated multiple times with different training and validation subsets (i.e. cross-validation) and the error metrics can be averaged to provide a more robust error assessment.

For classification and regression problems, during the setup of the training and validation sets, grouping of observations by species and height stratification is necessary. This ensures that the training and validation sets have approximately the same height distributions for each species and that the classification model is able to handle all height classes. Additional stratification may be applied. For example stratifying the samples by segment shape quality has been shown to reduce the classification error (Ko et al., 2016).

Scale and location dependent features can inadvertently be discriminative for species identification within a sample, even though these feature are not discriminative for the overall population. A typical case where such a problem would arise is if trees for a given species were all sampled in an even aged stand. Thus, in the presence of imbalanced height/shape distributions, particular care should be taken to use features that are independent of scale. However, this is not a requirement if the full range of height/shape variability is uniformly distributed in the training/validation sets for each species. Finally, when reporting the results, the error metrics may be stratified (e.g. by species, height, diameter, social class, region, etc) to provide a more detailed diagnosis of performance.

## 2.4 Synthesis

In this chapter, an overview of the data and forest environment considered in this work was presented. The following contributions were made:

- A novel ALS dataset containing over 5000 manually delineated 3D tree models including observations from multiple sensors and at different forest sites across Switzerland. The dataset has multiple research applications including the simulation of different forest configurations, the rigorous validation of segmentation algorithms, development of genus/species classification models and calibration allometric models.
- An extended variable length record specification (cf. Appendix G.1) for the ASRPS LAS point cloud format to store forest inventory attributes at the individual tree scale directly in the file. This specification is a practical solution to store and distribute forest inventory parameters directly with 3D tree models (point clouds) in a single self-contained unit.



## 3. Individual tree segmentation

This chapter covers the topic of individual tree segmentation from ALS data. It is structured in the following way:

- 3.1 introduces the topic and describes the state of the art.
- 3.2 proposes a rigorous error assessment framework to evaluate the performance of individual tree segmentation methods.
- 3.3 presents a novel graph-based tree segmentation method called geodesic voting. The content of this section is adapted from Parkan and Tuia (2015).
- 3.4 presents a novel ensemble method used to estimate segmentation error and improve tree shape delineation in coniferous forest. The content of this section is adapted from Parkan and Tuia (2018).
- 3.5 presents a novel stem detection method called layered morphological analysis.

### 3.1 State of the art

Individual tree segmentation is a bottleneck problem for many forestry related applications including dendrometry (e.g. height, diameter at breast height, basal area, crown spread), tree species classification and forest stand delineation. The main difficulties associated with this problem arise from the potentially complex spatial configuration of trees in forest environments. This includes tree crown adjacency or intersection, heterogeneity of shape within and across species, significant changes in shape with age and canopy layering (understory vegetation). Individual tree segmentation methods rely (often implicitly) on several assumptions:

- the **spacing** between individual trees and/or their spatial distribution;
- the **outer/inner shape/structure** of trees (e.g. structural opacity, vertical growth, pointy top, minimum/maximum crown area, minimum/maximum height, branch size distribution, etc);
- the **radiometric properties** of trees and their parts (e.g. echo intensity, reflectance, color).

By combining these assumptions, it is possible to define a set of clustering rules that can be used to merge similar parts and split dissimilar parts of the data. The strength of the assumptions regarding tree characteristics varies between algorithms. Strict assumptions may reduce commission errors and lead to good performances in forests with low shape variability. Conversely, algorithms that do not rely on strict assumptions tend to perform better in more structurally diverse forests, often at the cost of more commission errors. In practice, a balance between prior assumptions and adaptability is often desirable. Most segmentation methods also require some level of manual parameter tuning which generally affects the trade-off between detection/delineation recall and precision (cf. section 3.2). Overall, some desirable characteristics of a tree segmentation algorithm are:

- **simplicity** (e.g. easy implementation, limited number of parameters, minimal parameter tuning);
- **adaptability to different crown geometries;**
- ability to **segment vegetation in all forest strata;**
- **provision of a measure of segmentation uncertainty** (e.g. by allowing fuzzy membership for intersecting crowns).
- **small computational complexity** and running times on large areas

Unsurprisingly, many well established data clustering algorithms (e.g. K-means, DBSCAN, hierarchical clustering, mean-shift, graphs-cuts, etc) have been applied directly or in a modified form to the problem of individual tree segmentation (Koch et al., 2014; Zhen et al., 2016; Lindberg and Holmgren, 2017). Often, these algorithms require an initial solution which is then iteratively optimized according to the clustering rules. This initial solution can be random or it can be based on preliminary detection of tree proxies (i.e. stem, top, centroid). In the latter case, different approaches for deciduous and persistent tree species may be required. Many deciduous species form a relatively flat canopy with many local irregularities and indistinct tree tops. On the other hand, species with persistent foliage often form canopies with distinct tree tops. Canopy opacity is also an important factor to consider when designing a segmentation algorithm. Opacity can be related to the phenological phase and/or the type of foliage (deciduous/persistent). In leaf-on conditions, a representative sampling of the branching structure is usually not possible with long range ALS. This is particularly the case with young or sparse coniferous forests (e.g. in woodland pastures) where self-pruning is limited and dense branching occurs from the ground up. Conversely, in leaf-off conditions, ALS can sample most of the tree structure and thus provide a better characterization of branching configuration (e.g. stem position, branch size distribution).

Divide and conquer approaches have been used to reduce the complexity of forest structures and subsequently apply segmentation to each of the partitions separately. For example, partitioning point clouds according to horizontal layers (strata) and processing each of the layers separately has been shown to be an effective approach to deal with understory trees (Wang et al., 2008a; Rahman et al., 2009; Duncanson et al., 2014; Paris et al., 2016; Hamraz et al., 2017).

Most of the current individual tree segmentation algorithms belong to one of three categories, depending on the data representations they use:

- **Raster** algorithms which convert the raw point cloud to 2D (e.g. canopy height model) or 3D gridded models (e.g. voxel model). Methods that use only the CHM have the disadvantage of excluding all the sub-canopy information contained in the original dataset and are affected by canopy height errors (especially in sloped terrain). This category of algorithms has been the most widely investigated, in the past decade. A list of such algorithms is provided in table 3.1;
- **Vector** algorithms which use of all the raw information but tend to be more computationally

intensive. A list of such algorithms is provided in table 3.2;

- **Mixed representation** algorithms which generally attempt to combine the advantages of raster and vector representations, but do not necessarily exploit the full structural information available in the point cloud. A list of such algorithms is provided in table 3.3;

Special mention should also be made of a subcategory of segmentation algorithms which focuses exclusively on the detection of stems. Most of the stem detection methods found in recent scientific literature were developed for terrestrial laser scanning. However, with the increasing availability of high density ALS, some of these stem detection methods are now also applicable to ALS. Similarly to full tree segmentation algorithms, methods to detect stems are difficult to categorize precisely as they often combine different approaches including:

- Fitting 3D lines to the point cloud with RANSAC (Reitberger et al., 2009; Lamprecht et al., 2015);
- Fitting circles, ellipse, cylinders to horizontal cross-sections of the point cloud with RANSAC, simple/robust least squares or Hough transforms (Simonse et al., 2003; Aschoff and Spiecker, 2004; Bienert et al., 2007; Maas et al., 2008; Moskal and Zheng, 2011; Lindberg et al., 2012; McDaniel et al., 2012; Olofsson et al., 2014; Wang et al., 2016a; Wieser et al., 2017; Cabo et al., 2018; Calders et al., 2018);
- Using local spatial covariance features and/or surface normals. A frequently used approach is applying a local Principal Component Analysis (PCA) and computing eigenvalue based indices (e.g. ratio of first over sum of second and third eigenvalues) to identify local linearity or planarity (Lalonde et al., 2006; Liang et al., 2012; Xia et al., 2015; Lamprecht et al., 2015; Wang et al., 2016a; Amiri et al., 2017; Burt, 2017; Wang et al., 2018);
- Clustering based on spatial separation (Bienert et al., 2007; Maas et al., 2008; Brolly and Király, 2009; Yao et al., 2011; McDaniel et al., 2012; Fritz et al., 2013; Lu et al., 2014; Lamprecht et al., 2015; Shendryk et al., 2016; Amiri et al., 2017; Bock et al., 2017; Cabo et al., 2018);
- Morphological operations on voxels (Heinzel and Huber, 2016);
- Graph based analysis (Gorte and Winterhalder, 2004a; Côté et al., 2009; Bucksch et al., 2014; Parkan and Tuia, 2015);
- Projected point density analysis (Rahman and Gorte, 2009; Wang et al., 2016a)
- Laser pulse timing analysis (Bock et al., 2017)
- Radiometric properties analysis, e.g. reflectance, echo width (Yao et al., 2011; Lu et al., 2014; Shendryk et al., 2016; Wang et al., 2018)

Comparing the results of segmentation algorithms reported in separate studies is difficult, because of significant differences in data characteristics, types of forest and error assessment protocols. In particular, the vast majority of individual tree segmentation studies do not validate the 3D shape of segments and instead only consider tree position, height and sometimes crown extent. Several benchmarking studies have been conducted to compare a limited set of algorithms using common datasets and error assessment procedures (Larsen et al., 2011; Kaartinen et al., 2012b; Vauhkonen et al., 2012; Jakubowski et al., 2013b; Eysn et al., 2015; Dalponte et al., 2015; Wang et al., 2016b; Pirotti et al., 2017). However, no algorithm has been shown to perform well across all forest types (Zhen et al., 2016) and the development of a universally adaptable algorithm is an ongoing research area.

Near future developments in individual tree segmentation are likely to make more use of 3D structural and radiometric information (e.g. from multispectral LiDAR and/or simultaneous photo acquisitions). It can also be expected that tree segmentation and characterization using deep learning

algorithms trained on manually delineated point clouds and/or simulated tree/forest models will be investigated.

Table 3.1: Selected raster algorithms for individual tree segmentation from ALS data.

Algorithm	Multi-strata	Reference
Region growing	×	Hyyppa et al. (2001)
Parabolic surface fitting	×	Persson et al. (2002)
Parabolic surface fitting	×	Holmgren et al. (2003a)
Multiscale region merging	×	Brandtberg et al. (2003)
Variable window size tree top detection	×	Popescu and Wynne (2004)
Region growing	×	Solberg et al. (2006)
Watershed	×	Chen et al. (2006)
Spatial wavelet analysis	×	Falkowski et al. (2006)
Watershed and region merging	×	Koch et al. (2006)
Watershed	×	Kwak et al. (2007)
Multiscale template matching	×	Korpela et al. (2007)
Gaussian template matching	×	Pirotti (2010)
Watershed with morphological corrections	×	Heinzel et al. (2011)
Watershed with morphological corrections	×	Ene et al. (2012)
Correlation surface analysis and region merging	×	Holmgren and Lindberg (2013)
Elevation contour analysis	×	Tang et al. (2013)
Spoke wheel	×	Liu et al. (2013)
Watershed and shape correction	×	Zhang et al. (2014)
Fishing Net Dragging	×	Liu et al. (2015)
Graph based clustering	×	Strîmbu and Strîmbu (2015)
Elevation contour analysis	×	Wu et al. (2016)
Horizontal cross-section analysis	×	Zhao et al. (2017a)
Constrained region growing (using multispectral ALS)	×	Naveed and Hu (2017)
Gradient orientation clustering	×	Dong et al. (2018)

Table 3.2: Selected vector algorithms for individual tree segmentation from ALS data.

Algorithm	Multi-strata	Reference
Normalized graph cut	✓	Reitberger et al. (2007)
Delaunay triangulation	✗	Alexander (2009)
Skeletonization based on graph-reduction	✓	Bucksch et al. (2009)
Active contour and hill climbing	✗	Ke et al. (2010b)
Adaptive region growing and merging	✗	Lee et al. (2010)
Paraboloid surface fitting with RANSAC	✗	Tittmann et al. (2011)
Modified single linkage clustering	✗	Li et al. (2012)
Mean-shift	✓	Ferraz et al. (2012)
Normalized graph cut	✓	Yao et al. (2012)
Normalized graph cut and mean shift	✓	Yao et al. (2013)
Layered K-means clustering	✓	Kandare et al. (2014)
Modified single linkage clustering	✗	Lu et al. (2014)
Bayesian template fitting	✗	Lahivaara et al. (2014)
3D ellipsoid fitting	✓	Lindberg et al. (2014)
Multiscale point cloud analysis	✓	Vega et al. (2014)
Geodesic vote	✓	Parkan and Tuia (2015)
Adaptive mean shift	✓	Ferraz et al. (2016)
Layered vertical profile analysis	✓	Hamraz et al. (2016)
Adaptive mean shift	✓	Xiao et al. (2016)
Adaptive mean shifts	✓	Hu et al. (2017b)
Normalized graph cut	✓	Hu et al. (2017a)
Supervoxel clustering	✗	Xu et al. (2018)

Table 3.3: Selected mixed representation algorithms for individual tree segmentation from ALS data.

Algorithm	Multi-strata	Reference
K-means clustering	×	Morsdorf et al. (2004)
Region growing	×	Tiede et al. (2005)
Density of high points and watershed	×	Rahman and Gorte (2008)
Layered hierarchical morphological analysis	✓	Wang et al. (2008b)
Paraboloid surface fitting with RANSAC	×	Tittmann et al. (2011)
Layered marker controlled watershed	✓	Duncanson et al. (2014)
Voxel space morphological analysis	✓	Mongus and Žalik (2015)
Layered and compartmentalized clustering	✓	Paris et al. (2016)
Layer stacking	✓	Ayrey et al. (2017)

### 3.2 Error assessment framework

The segmentation error indicates how well a procedure is able to partition data into individual tree instances. Quantifying this error is particularly important when working at the tree scale, because it influences the quality of the derived features/predictors used in subsequent analysis (e.g. regression and classification). The reporting of error across publications on individual tree segmentation is very variable in form. Many authors validate their results based only on the horizontal position difference of tree proxies (e.g. tree tops) relative to reference positions. Others combine positional and tree height difference criteria. A smaller subset of authors validate the segmentation based on the 2D overlap of segmented and reference tree crown boundaries. However, the validation of the 3D shape of trees has generally been ignored (Zhen et al., 2016). As a consequence, the results obtained in different studies are often not comparable. To address this problem, a rigorous framework for 3D segment validation is proposed in this section.

The segmentation error can be characterized in terms of:

- The **stem position error** (cf. table 3.4) which quantifies the distance to the reference tree position.
- The **shape delineation error** (cf. table 3.4) which quantifies how similar the tested and reference tree shapes are in terms of relative height, area and volume difference.

Reporting stem position errors without associated shape delineation errors can be deceptive, as segments may still have strongly erroneous shapes which may preclude further analysis. Thus, to be counted as a correct detection (true positive), a segment should fulfill both stem position and shape delineation quality criteria (cf. figure 3.1). First, its stem proxy should be located within a maximum horizontal distance  $\varepsilon_{xy,max}$  from the reference position. The maximum tolerable stem position error  $\varepsilon_{xy,max}$  (cf. figure 3.2) can either be a fixed value (e.g. empirically determined based on stem density) or it can be a variable value. In the latter case, the following formula may be used:

$$\varepsilon_{xy,max} = \frac{f}{2}DBH + \varepsilon_{xy,ref} + \varepsilon_{xy,als} \quad (3.1)$$

where:

- $f$  is the stem cross section eccentricity factor;
- $DBH$  is the Diameter at Breast Height;
- $\varepsilon_{xy,ref}$  is the estimated planimetric error of the reference stem position;
- $\varepsilon_{xy,als}$  is the estimated planimetric error of the laser scanning points.

Equation 3.1 takes into account the possible cumulation of field and ALS measurement errors. It is also important to note that stems are often sampled irregularly on a single side with most ALS systems, although overlapping scan swaths or dual laser systems may sample stems on multiple sides (possibly allowing direct measurement of stem diameter).

Second and more importantly, the tested segment shape should spatially overlap with the reference shape. Formally, the correct detection rate  $d$  can be defined as:

$$d = \frac{\sum_{i=1}^N I_C(\varepsilon_{xy,i}, \varepsilon_{h,i}, \dots)}{N} \in [0, 1] \quad (3.2)$$

Where  $N$  is the total number of segments and  $I_C$  is the indicator function which determines if segment  $i$  is correct:

$$I_C(\varepsilon_{xy}, \varepsilon_h, \dots) = \begin{cases} 1 & \text{if } \varepsilon_{xy} \leq \varepsilon_{xy,max} \text{ AND } \varepsilon_h \leq \varepsilon_{h,max} \text{ AND } \dots \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

The  $I_C$  function may contain multiple shape quality criteria. In addition to the height error, area, volume and point pattern overlap metrics should be included in  $I_C$  (cf. table 3.4). Point pattern overlap metrics ( $J_P$ ,  $r$ ,  $p$ ,  $F$ ) are sensitive to point density. Thus, they can be misleading when points are distributed non-homogeneously within a segment (for example within flight line overlap bands). Conversely, area ( $J_A$ ) and volume ( $J_V$ ) overlap metrics are insensitive to inhomogeneous point density, but are affected by the location of each individual point (boundary defects). Thus, for a given segment, area/volume overlap errors will often be larger than point pattern overlap errors (cf. figure 3.3).

Based on the number of True Positives ( $n_{TP}$ ), False Negatives ( $n_{FN}$ ), False Positives ( $n_{FP}$ ), common detection metrics such as recall ( $r$ ), precision ( $p$ ) and F-score ( $F$ ) can be computed:

$$r = \frac{n_{TP}}{n_{TP} + n_{FN}} \in [0, 1] \quad (3.4)$$

$$p = \frac{n_{TP}}{n_{TP} + n_{FP}} \in [0, 1] \quad (3.5)$$

$$F = 2 \cdot \frac{p \cdot r}{p + r} \in [0, 1] \quad (3.6)$$

The interpretation of these detection metrics is provided in figure 3.6.

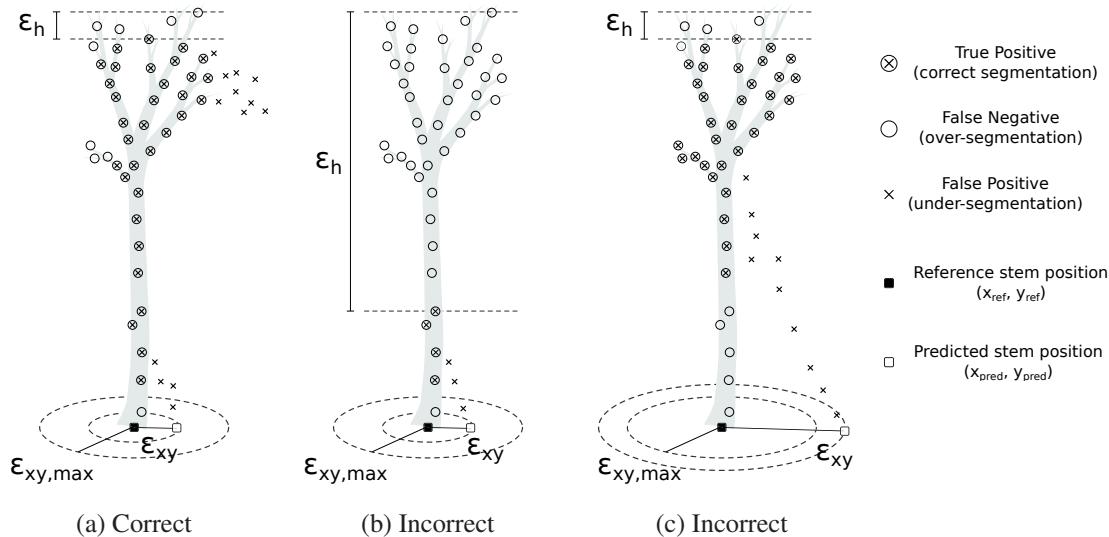


Figure 3.1: Correct and incorrect tree detection cases. Combined stem position and shape delineation errors determine if the detection is correct.

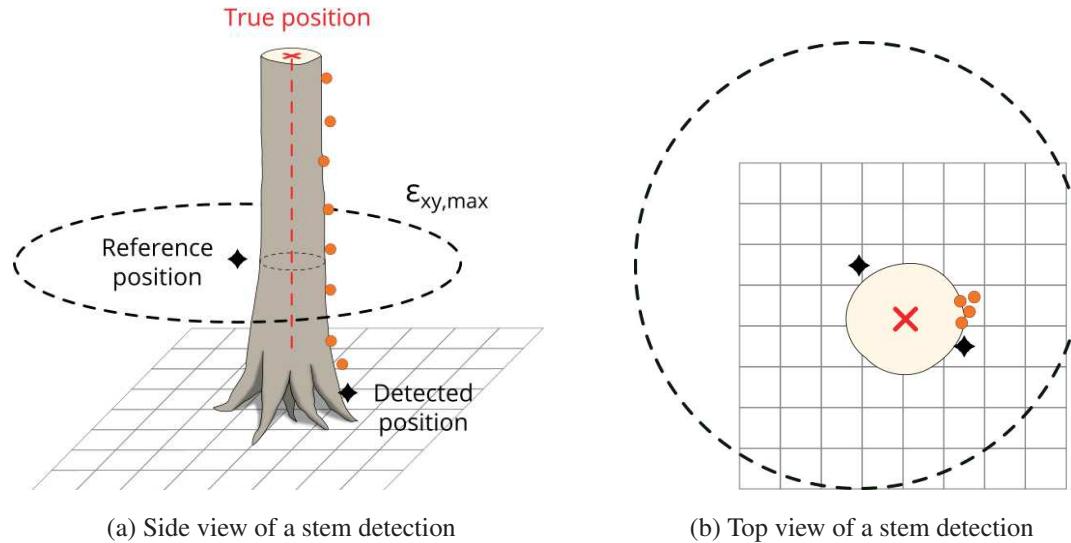
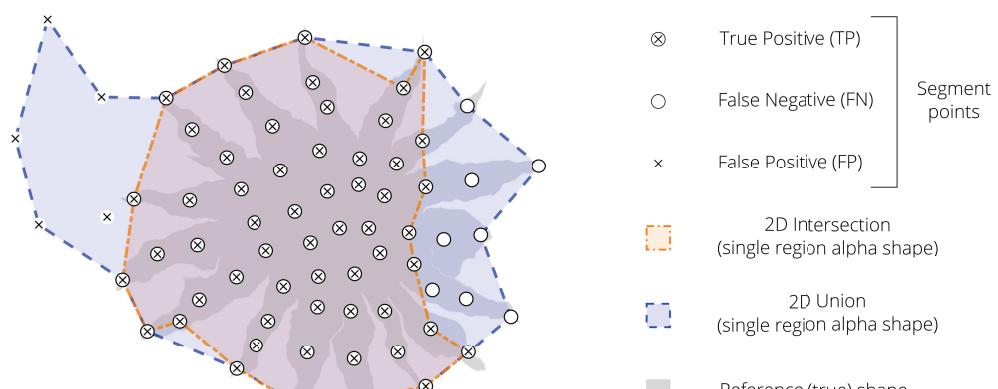


Figure 3.2: Stem position error modeling.

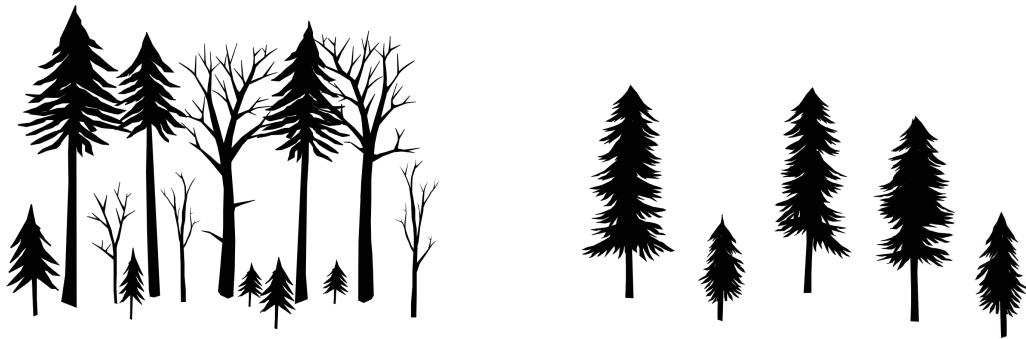


$$J_P = \frac{n_{TP}}{n_{TP} + n_{FN} + n_{FP}} = \frac{57}{70} = 0.81$$

$$J_A = \frac{A_i}{A_u} = \frac{27.97}{47.06} = 0.66$$

Figure 3.3: Difference between the point pattern overlap ( $J_P$ ) and the area overlap ( $J_A$ ) computed with the Jaccard coefficient (intersection over union). Area and volume overlap metrics are more sensitive to shape boundary defects.

A very important aspect of individual tree segmentation is characterizing the difficulty of the problem. Clearly, the difficulty is related to local structural complexity and is not the same for a sparse woodland pasture than for a dense tropical forest. Essentially, in dense settings, there is more potential for tree shape/edge confusion.



(a) **Complex setting:** dense, multi-layered, uneven-aged, mixed species (e.g. selection cutting).      (b) **Simpler setting:** sparse, single-layered, even-aged, pure species (e.g. woodland pasture)

Figure 3.4: Schematic examples of forest environments with different structural complexity.

One way of quantifying the segmentation difficulty is by examining the spatial adjacency (cluttering) of trees. More specifically, given a labeled point cloud, where each unique label defines a segment (tree), the following approach can be used:

1. Remove scan overlap swaths from the point cloud and/or resample the point cloud using a regular 3D grid, to obtain an approximately homogeneous point density (cf. figure 3.5a).
2. For each point  $i$  in the point cloud, find all points located within radius  $R_A$  (including self) and compute the fraction of these points that have a different label than the one of point  $i$ . This fraction is called the adjacency factor and is designated by the symbol  $f_{A,i}$  (cf. figure 3.5b). The search radius  $R_A$  is set empirically and should be sufficiently large to ensure that in a given forest setting, any segment with an aggregated adjacency factor close to zero can unambiguously be delineated.
3. Aggregate the point scale adjacency factor at the segment (tree) scale by consecutively computing their respective average or any other aggregation function (cf. figures 3.5c and 3.5d). At the plot scale, both the average and the standard deviation of the aggregation index can be used to characterize the overall difficulty of segmentation.

$$f_{A,i} = \frac{\sum_{j=1}^{N_{A,i}} I_A(L_i, L_j)}{N_{A,i}} \in [0, 1[ \quad (3.7)$$

Where  $N_{A,i}$  is the number of points located within radius  $R_A$  of point  $i$  and  $I_A$  is the indicator function which determines if the label  $L_i$  of point  $i$  is different than the label  $L_j$  of point  $j$ :

$$I_A(L_i, L_j) = \begin{cases} 1 & \text{if } L_i \neq L_j \\ 0 & \text{otherwise} \end{cases} \quad (3.8)$$

In a sparse forest settings, trees have a low adjacency factor and a simple distance-based clustering algorithm such as single linkage or density based clustering (DBSCAN) may perform well for segmentation. Conversely, the same algorithm may perform poorly in dense settings. Thus, stratifying the segmentation scores according to the adjacency factor helps to interpret algorithm performance.

Tree adjacency is probably the most determinant factor of segmentation difficulty, but it is not the only one. Tree shape diversity, can also add complexity to the problem. Thus, a complementary metric to assess segmentation difficulty could be a tree shape dissimilarity index. Such an index could quantify coarse shape dissimilarity (e.g. simply combining height and crown area attributes) or it could use more sophisticated shape comparison approaches (Veltkamp, 2001; Cardone et al., 2003). Moreover, the visibility of the stem and of the apex is an important factor contributing to segmentation difficulty.

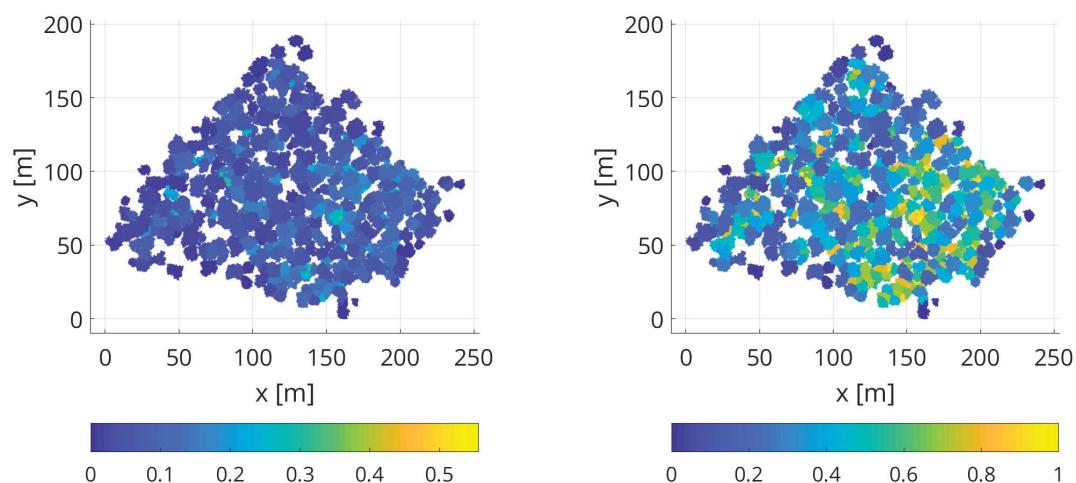
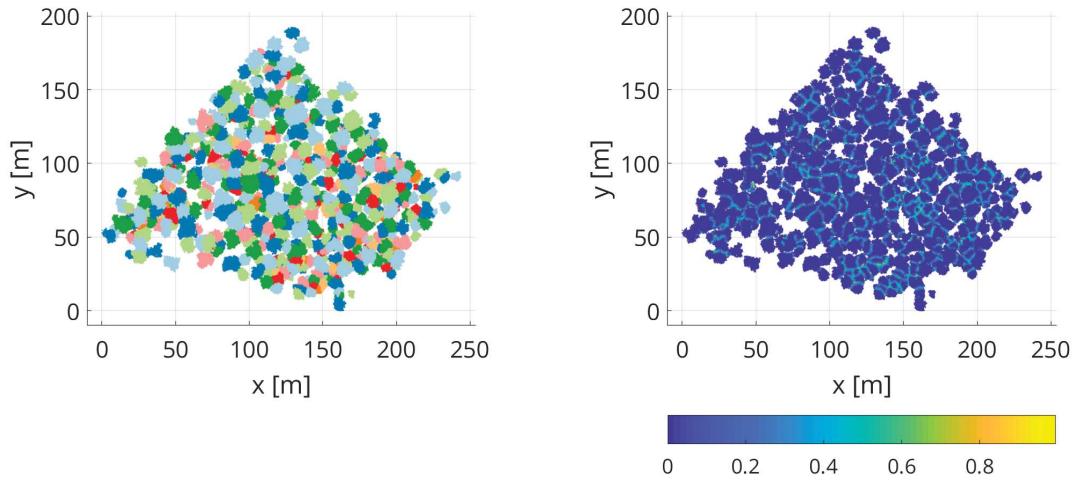


Figure 3.5: Adjacency factor for tree segments from the Boudry D1 reference site.

Table 3.4: Stem position and shape delineation error metrics.

Metric	Formula	Interpretation
Stem position error	$\epsilon_{xy} = \sqrt{(x_p - x_r)^2 + (y_p - y_r)^2}$	Horizontal distance between reference $(x_r, y_r)$ and predicted $(x_p, y_p)$ stems
Height error	$\epsilon_h =  h_p - h_r $	Difference between reference $(h_r)$ and predicted $(h_p)$ height
Point wise Jaccard index	$J_P = \frac{n_{TP}}{n_{TP} + n_{FN} + n_{FP}} \in [0, 1]$	Relative overlap between reference and predicted point pattern (cf. figure 3.3)
Area wise Jaccard index	$J_A = \frac{A_i}{A_u} \in [0, 1]$ $A_i$ : area of the intersection $A_u$ : area of the union	Relative area overlap between the reference and predicted single region 2D $\alpha$ shape (cf. figure 3.3)
Volume wise Jaccard index	$J_V = \frac{V_i}{V_u} \in [0, 1]$ $V_i$ : volume of the intersection $V_u$ : volume of the union	Relative volume overlap between the reference and predicted single region 3D $\alpha$ shape (cf. figure 3.3)
Recall (Producer's Accuracy)	$r = \frac{n_{TP}}{n_{TP} + n_{FN}} \in [0, 1]$	Tendency to completely include points (sensitivity) (cf. figure 3.7)
Precision (User's Accuracy)	$p = \frac{n_{TP}}{n_{TP} + n_{FP}} \in [0, 1]$	Tendency to correctly include points (cf. figure 3.7)
F score	$F = 2 \cdot \frac{p \cdot r}{p + r} \in [0, 1]$	Harmonic mean of $r$ and $p$ . The harmonic mean is used instead of the simple mean, because it penalizes extreme values more

In the context of tree or stem detection, the following interpretation of recall and precision can be made (cf. figure 3.6):

- **Low recall:** the algorithm is failing to detect a lot of existing trees.
  - **High recall:** the algorithm is detecting most or all of the existing trees but may also be detecting non-existing trees.
  - **Low precision:** the algorithm is detecting many non-existing trees.
  - **High precision:** the algorithm is detecting mostly or only existing trees.

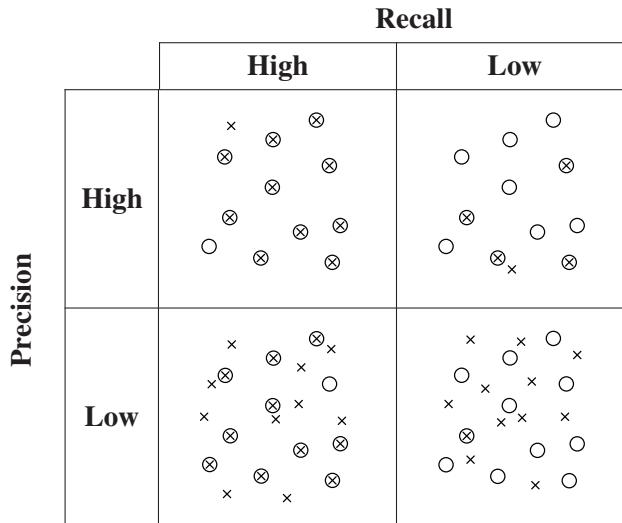


Figure 3.6: Interpretation of recall and precision metrics for tree detection. Symbols:  $\otimes$  = True Positive,  $\times$  = False Positive,  $\circ$  = False Negative

In the context of tree shape delineation, the following interpretation of recall and precision can be made (cf. figure 3.7):

- **Low recall:** the tested segment is failing to include a lot of points that belong to the reference segment.
- **High recall:** the tested segment includes most or all of the points that belong to the reference segment, but may also include points that do not belong to it (false positives).
- **Low precision:** the tested segment includes many points that do not belong to the reference segment (false positives).
- **High precision:** the tested segment includes mostly or only points that belong to the reference segment (true positives).

The combination of high precision and low recall is called over-segmentation. Inversely, the combination of low precision and high recall is called under-segmentation.

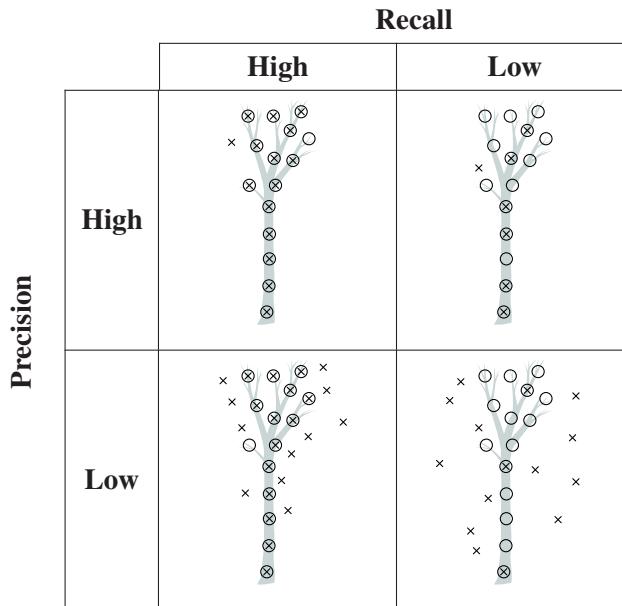


Figure 3.7: Interpretation of recall and precision metrics for tree delineation. Symbols:  $\otimes$  = True Positive,  $\times$  = False Positive,  $\circ$  = False Negative

### 3.3 Tree segmentation with geodesic voting

In this section, an algorithm to extract individual deciduous trees (in leaf-off conditions) from high-density discrete return ALS data is presented and its performance is evaluated under different forest configurations. The method is based on quantifying the topological hierarchy of different tree structures (i.e. from trunk to leaf) within a  $K$  nearest neighbour graph derived from the ALS point cloud. This section expands the work presented in Parkan and Tuia (2015) with new study sites, a more rigorous segmentation error assessment, modifications to the method, a parameter sensitivity analysis and a more detailed discussion.

#### 3.3.1 Description

The proposed method is based on the following assumptions:

1. forests and trees can be represented as (directed) acyclic graphs.
2. the set of shortest paths (geodesic graph) linking each tree node to the terrain within the graph representation derived from the point cloud can be used as approximations of real branching structures.
3. the nodes which are part of central structures (i.e. trunks, main branches) are more frequently traversed by terrain-linking geodesics than other nodes.

Hypotheses 1 and 3 are implicit. Hypothesis 2 is biologically justifiable, considering that natural selection has tended to maximize metabolic capacity and internal efficiency by respectively maximizing the exchange surface areas and minimizing transport distances (Leopold, 1971; West et al., 1999a,b). Such tree-like energy minimizing structures are observed across nature; for example in the food gathering patterns used by some ants (Deneubourg et al., 1989), slime molds path finding (Tero et al., 2007), vascular tissues, neural networks, hydrological drainage bassins, and lightning (Bejan, 2000). The analogy between graph representations and botanical trees is so apparent that mathematicians reuse botanical terms to describe some types of graphs. Thus, a connected acyclic graph is called a tree and a graph whose connected component are all trees is called a forest. In addition to being inherently well adapted to botanical tree representation, graphs have been studied extensively and many descriptors have been developed to characterize them. Interestingly, some of these descriptors such as the branching order (Horton, 1945; Strahler, 1952), the branching angles (Honda and Fisher, 1978; Bayer et al., 2013), the branch lengths (Honda and Fisher, 1979; Bayer et al., 2013), the bifurcation ratios (Oohata and Shidei, 1971; Whitney, 1976; Borchert and Slade, 1981) and the fractal (self similarity) dimension (Zeide, 1991; Lorimer et al., 1994; West et al., 1999a; Godin and Ferraro, 2010) have a meaningful botanical interpretation and may be used to differentiate species (Ferraro and Godin, 2000). Thus, there exists both functional and structural justifications to employ geodesic graph based methods in tree segmentation.

Unsurprisingly, algorithms based on geodesic graphs have found applications for both artistic rendering and accurate structural reconstruction of trees from 3D point clouds. In this regard, much attention has been directed to reconstruct single tree structures from dense 3D point clouds often collected with a terrestrial laser scanner (Gorte and Winterhalder, 2004b; Gorte and Pfeifer, 2004; Xu et al., 2007; Bucksch et al., 2009, 2010; Yan et al., 2009; Côté et al., 2009; Preuksakarn et al., 2010; Livny et al., 2010; Gatziolis et al., 2010; Xu and Mould, 2012; Delagrange et al., 2014; Hu et al., 2017a). However, there has been little investigation on using geodesic graphs for simultaneous segmentation of multiple trees in forest environments (Parkan and Tuia, 2015; Tao et al., 2015; Shendryk et al., 2016; Méndez et al., 2016).

The method we propose quantifies the topological importance of the different branching structures which compose individual trees. More specifically, for each graph node it assigns a centrality value which can be used to differentiate high hierarchy structures (i.e. trunk, main branches) from lower hierarchy structures (i.e. small branches, leaves). This indexing can then be

used to discriminate the location of individual trunks and assign lower hierarchy structures to each of these trunks through connectivity (connected component analysis).

The intuition to use the frequency at which a network node is located on the shortest path linking pairs of other nodes as an indicator of topological centrality was popularized by Freeman (1977) who developed a set of node centrality measures. One of the measures described in Freeman (1977) is the partial betweenness - an index commonly employed to identify important nodes within communication networks. This index is defined in the following way. Given a node  $p_k$  in a graph and an unordered pair of nodes (source and destination)  $p_i, p_j$  where  $i \neq j \neq k$ , the partial betweenness  $b_{ij}(p_k)$  of  $(p_k)$  with respect to  $(p_i, p_j)$  is:

$$b_{ij}(p_k) = \begin{cases} \frac{g_{ij}(p_k)}{g_{ij}} & \text{if } p_i \text{ and } p_j \text{ are connected} \\ 0, & \text{if } p_i \text{ and } p_j \text{ are unconnected} \end{cases} \quad (3.9)$$

where:

- $g_{ij}(p_k)$  is the number of geodesics linking  $p_i$  and  $p_j$  that contain  $p_k$ ;
- $g_{ij}$  is the number of geodesics linking  $p_i$  and  $p_j$ .

Thus,  $b_{ij}(p_k)$  can be considered as the probability that  $p_k$  is located on a randomly selected geodesic linking  $p_i$  and  $p_j$ . The terminology used in Freeman (1977) was expanded by Rouchdy and Cohen (2013) who coined the terms "geodesic voting score" or "geodesic density" to designate  $g_{ij}(p_k)$ . This terminology is used in the remainder of this section.

The algorithm proposed here is specifically designed for ALS and differs from the biomedical image segmentation method described by Rouchdy and Cohen (2013) in several ways. First, the geodesic density at each node ( $g_{ij}(p_k)$ ) is based on the  $K$ -NN graph and not on the optimization of a flow potential function. Secondly, the method described here does not require the manual input of an initialization point by the user. Third, the introduction of multiple (randomly scattered) end points is not required, we use a single source point (tie node). Moreover, the characteristics of the 3D data we are considering differ largely from 3D biomedical imagery. ALS point clouds have a relatively low and non-uniform resolution (point density diminishes when nearing the terrain). Additionally, the forest environment is composed of multiple unconnected tree structures.

### Algorithm

The key steps of the method are illustrated in figure 3.8. Below follows a detailed description of each step:

- (a) **Setup the vegetation, terrain and tie nodes** (cf. figure 3.8a).
  - 1 Create a Digital Terrain Model (DTM) with horizontal resolution  $d_{xy}$  (e.g.  $d_{xy} = 0.5$  m).
  - 2 Merge the DTM and vegetation points to form  $p_i$ .
  - 3 (Optional) Apply the vertical scaling factor  $f_v$  to the z coordinate of the points, to penalize the horizontal displacement of geodesics.
  - 4 Add a global source node  $p_j$  (tie node).
- (b) **Build the K-NN graph** (cf. figure 3.8b).
  - 1 Construct the vegetation to vegetation and terrain to vegetation  $K$  Nearest Neighbor ( $K$ -NN) graphs separately and merge them.
  - 2 (Optional) Apply a weighing function to the graph edges. Such a function can for example be used to penalize paths that pass through long edges and can improve structural reconstruction. A function of the form  $w(L) = (L + 1)^a$  can be used, where  $L$  is the length of the edge in meters and  $a$  is an exponent larger than 1. Note that graph edge weighing has also been used in Shendryk et al. (2016).
  - 3 Extract the largest connected component in the graph (i.e. remove disconnected parts of the graph).
  - 4 Within the graph adjacency matrix, add equally weighted edges from all terrain nodes to the tie node.
- (c) **Compute geodesics linking each node ( $p_i$ ) to the tie node ( $p_j$ ) using Dijkstra's shortest path algorithm (Dijkstra, 1959)** (cf. figure 3.8c).
- (d) **Compute the geodesic density  $g_{ij}(p_k)$  at each node  $k$**  (cf. figure 3.8d).
- (e) **Identify individual trees** (cf. figure 3.8e).
  - 1 Identify all the root nodes (i.e. those on the DTM) within geodesics, excluding the tie node.
  - 2 Remove roots which are linked with too few nodes, by thresholding the geodesic density (voting), i.e.  $g_{ij}(p_k) \geq g_{min}$ .
  - 3 Assign geodesics to their respective root nodes.
  - 4 Merge adjacent trees based on a separation distance criterion  $d_{adj}$ .
  - 5 (Optional) Filter erroneous points by applying additional geometric criteria.
- (f) **Compute the branching order** (cf. figure 3.8f) and other graph descriptors.

### Parameters

- $d_{xy}$**  Raster DTM resolution (nominal value: 0.25 m);
- $K$**  Number of nearest neighbors used to build graph (nominal value: 15);
- $f_v$**  Vertical scaling factor (nominal value: 1);
- $a$**  Edge weight exponent (nominal value: 2);
- $g_{min}$**  Minimum geodesic density (nominal value: 30);
- $d_{adj}$**  Minimum stem separation distance (nominal value: 0.75 m).

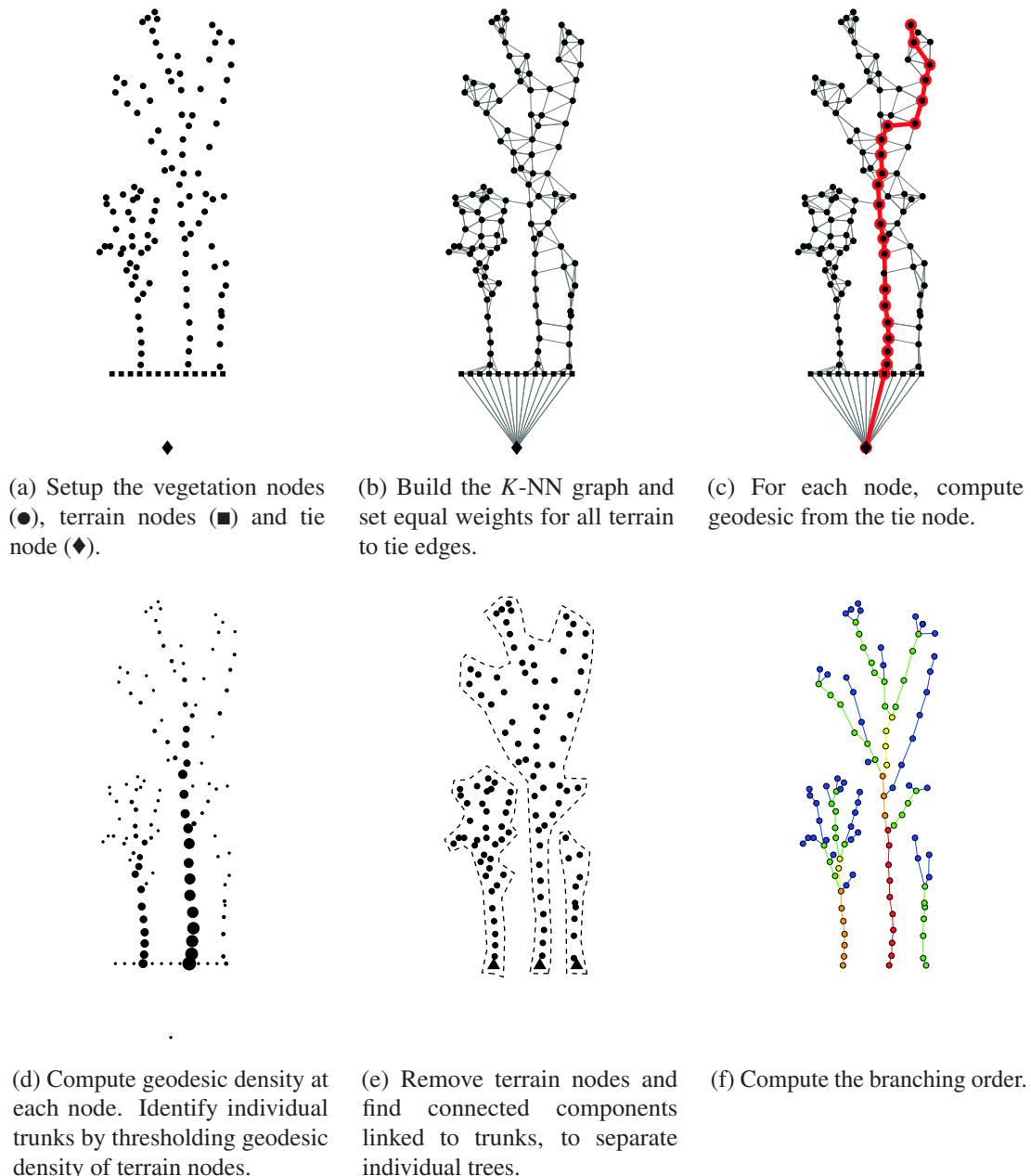
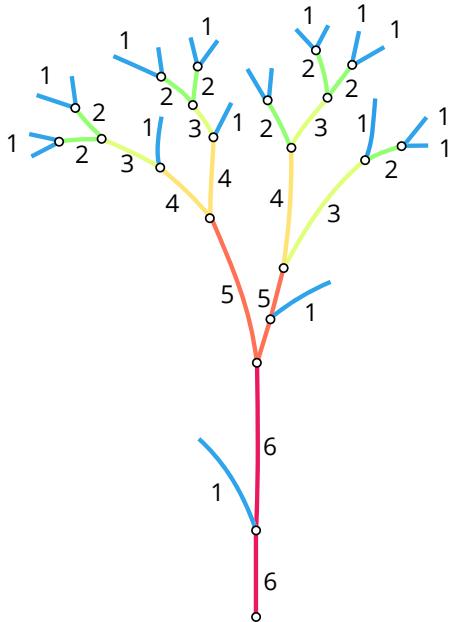


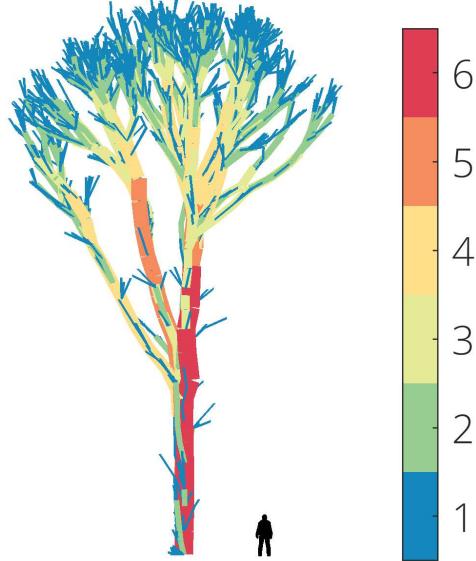
Figure 3.8: Main steps in the individual tree segmentation. Adapted from Parkan and Tuia (2015).

### Characterization of the branching hierarchy

The method produces a graph for each tree. Each point has an associated relative geodesic density ( $b$ ) which can be used to reveal different levels of branching hierarchy (cf. figure 3.14). The graph representation can also be used to characterize the branching order with the Horton-Strahler number (cf. figure 3.9).



(a) Conceptual example of the Horton-Strahler number.



(b) The Horton-Strahler number computed for an individual tree segment.

Figure 3.9: The Horton-Strahler number can be used to characterize the branching structure of trees.

### Correcting under-segmentation

Under-segmentation can be detected and reduced by using the ratio of the geodesic and linear distances as a measure of shape uncertainty (cf. figure 3.10). Each point within a segment can be filtered with the following equation:

$$R_i = \frac{D_{G,i}}{D_{L,i}} \quad (3.10)$$

$$I_F(R_i) = \begin{cases} 1 & \text{if } R_i \leq \rho \\ 0 & \text{otherwise} \end{cases} \quad (3.11)$$

where:

$D_{G,i}$  is the geodesic distance between point  $i$  and the root of its assigned segment;

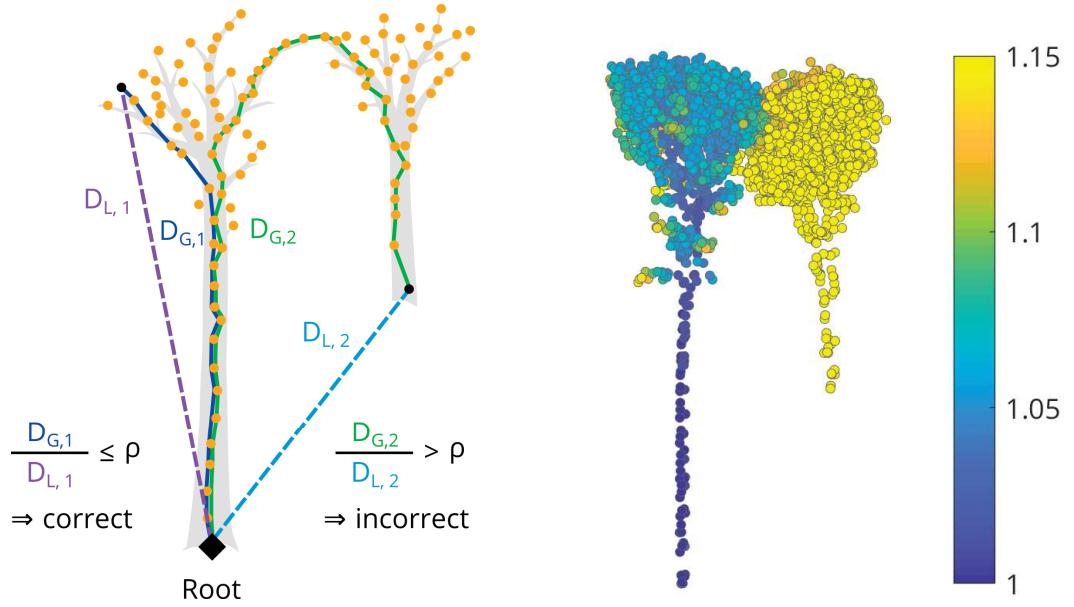
$D_{L,i}$  is the linear distance between point  $i$  and the root of its assigned segment;

$I_F$  is the indicator function which determine if a point  $i$  is valid ( $I_F(R_i) = 1$ ) or not ( $I_F(R_i) = 0$ );

$\rho$  is the maximum allowed value of  $R$  for a point to be included in a segment.

The value of  $\rho$  can be estimated by examining the probability distribution of  $R$  for the reference (manually delineated) deciduous trees (cf. figure 3.11). In the considered study sites, most of the

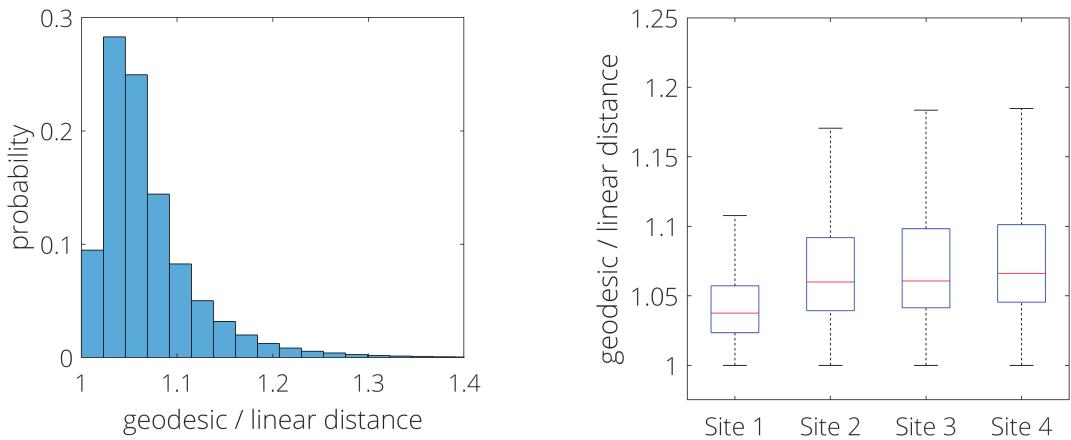
$R$  values range from 1.03 to 1.1. This range of values is valid in dense forests, however it can be expected that  $R$  can be larger in lower density forests or when considering isolated trees with more prominent lateral branching.



(a) Conceptual example of the geodesic to linear distance ratio.

(b) Under-segmentation example in the Beneté site. Points with  $R \geq 1.15$  are colored in yellow and indicate potential outliers in the segment.

Figure 3.10: The ratio ( $R$ ) between the geodesic and linear distance from a point to its root can be used to identify erroneous points (in case of under-segmentation) and detect segments that require a posteriori shape correction.



(a) Probability distribution of the distance ratio for all sites.

(b) Boxplots of the distance ratio for each study site. Outliers are not shown.

Figure 3.11: Distribution of the geodesic and linear distances ratio ( $R$ ) computed on the reference (manually delineated) trees (deciduous only).

### 3.3.2 Results

The method was applied to the Versoix, Sauvabelin, Benenté and Boudry D1 sites. These sites were selected because they are broadleaf deciduous dominated forests covered by high density leaf-off ALS data. The Versoix, Benenté and Boudry sites also have a coniferous understory layer. Since a fully labeled point cloud is required for 3D shape validation, simulated forest plots were created by considering only labeled trees from these sites. In other words, the simulated plots were identical to the real plots, but they contained only labeled high vegetation (manually delineated trees), unlabeled low vegetation ( $< 1$  m) and terrain points. Moreover, to assess the influence of understory coniferous trees on segmentation performance, two simulation subsets were created for each of the four study sites: one including all species, the other with only deciduous species.

To evaluate the effect of vertical scaling and edge weight exponent, a full factorial sensitivity analysis (cf. figure 3.13) was conducted on the simulated pure deciduous forest sites. The effect of the adjacency factor on the correct detection rate (recall) was also evaluated (cf. figure 3.12). Finally, the effect of correction for under-segmentation was examined (cf. table 3.9).

The proposed method was validated using the nominal parameters suggested in section 3.3.1 and compared to local maxima detection and marker controlled watershed segmentation (Meyer and Beucher, 1990; Meyer, 1994; Soille, 2013). The local maxima detection (which also provided markers for the watershed segmentation) was applied to 0.5 m resolution raster Canopy Height Models (CHM) derived from the 3D point clouds and smoothed using a Gaussian 3x3 lowpass filter. It used a height ( $h$ ) dependent search radius ( $r$ ) defined by:

$$r(h) = 1 + 0.25 \cdot \log(\max_h(h, 1)) \quad (3.12)$$

The comparison of the two methods is reported in tables 3.5-3.6 for segmentation and in tables 3.7-3.8 for stem detection. Qualitative examples of segmentation with geodesic vote for the Boudry and Versoix study site are presented in figures 3.14 and 3.15.

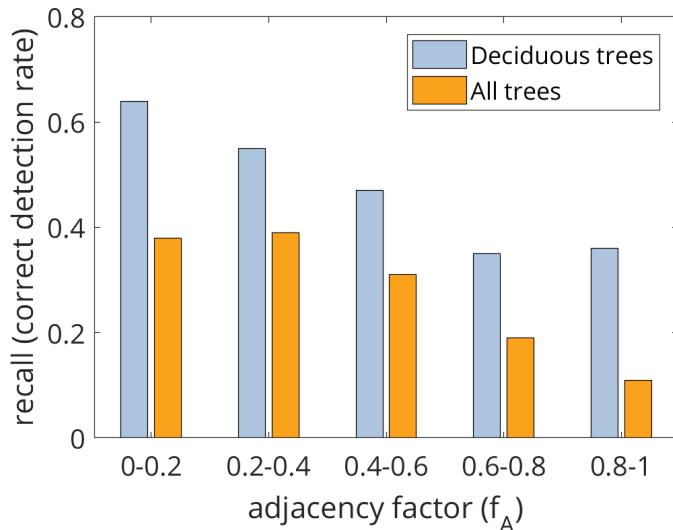


Figure 3.12: Correct detection rate as a function of adjacency factor (all sites combined). Correct detection criteria:  $\epsilon_{xy} \leqslant 2$  m,  $\epsilon_h \leqslant 2$  m,  $J_P > 0.5$ .

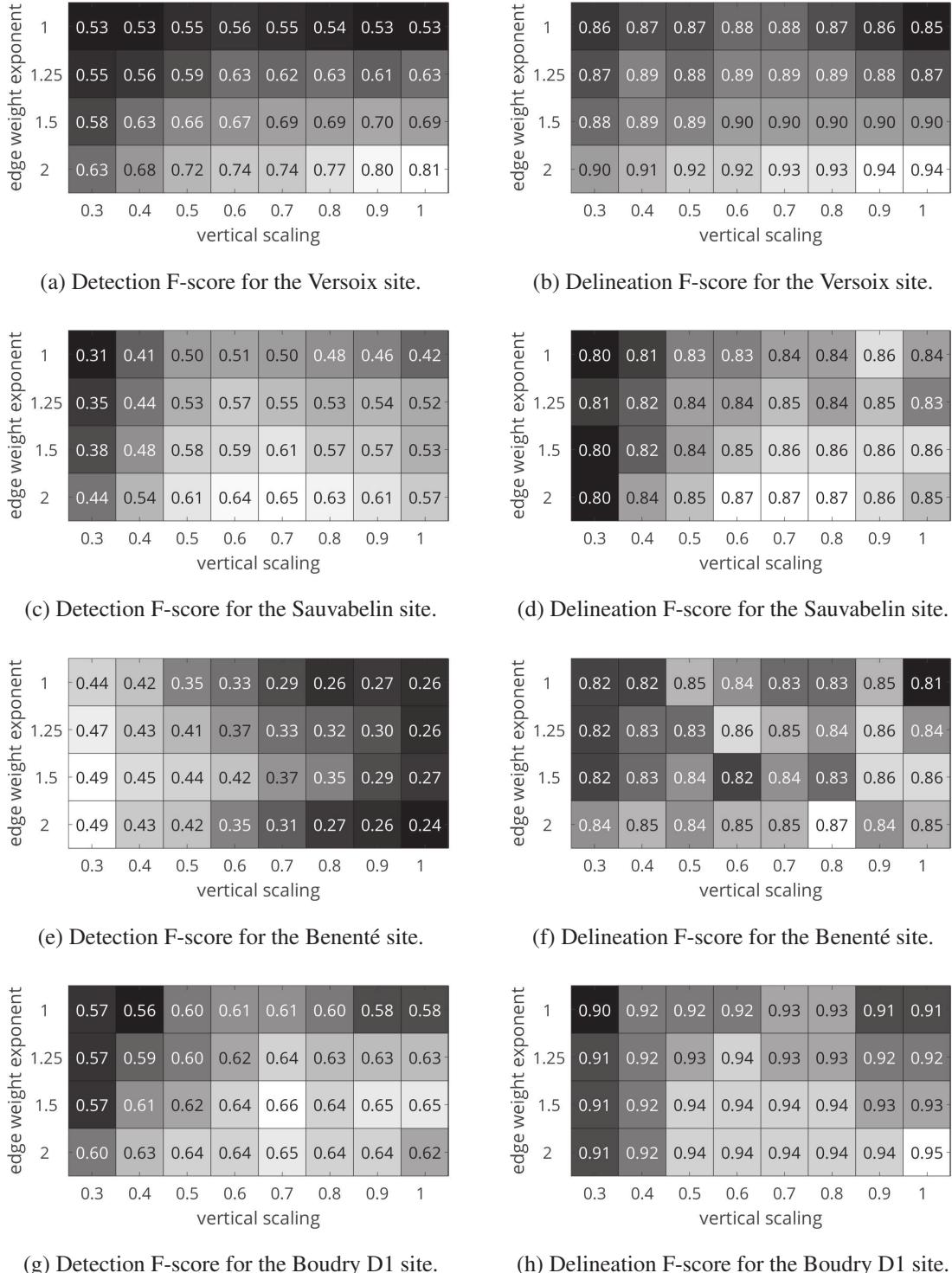


Figure 3.13: Detection and delineation F-score as a function of vertical scaling and graph edge weight exponent. Correct detection criteria:  $\epsilon_{xy} \leq 2$  m,  $\epsilon_h \leq 2$  m,  $J_P > 0.5$ . All the other segmentation parameters are fixed, i.e.  $d_{xy} = 0.25$  m and  $kNN = 15$ .

Table 3.5: Detection and delineation performance of geodesic voting versus marker controlled watershed segmentation for deciduous and coniferous trees. For delineation scores the mean and standard deviation are reported. Correct detection criteria:  $\varepsilon_{xy} \leq 2$  m,  $\varepsilon_h \leq 2$  m,  $J_P > 0.5$ . Note that the delineation scores are computed on the correctly detected tree only, which explains why they are systematically high. The best performing method is indicated with an asterisk.

Site	Obs.	$f_A$	Method	Detection			Delineation					
				$p$	$r$	$F$	$p$	$r$	$F$	$J_P$	$J_A$	$J_V$
Versoix	372	$0.5 \pm 0.22$	Geodesic vote*	0.57	0.62	0.59	$0.87 \pm 0.14$	$0.92 \pm 0.1$	$0.88 \pm 0.9$	$0.8 \pm 0.14$	$0.75 \pm 0.16$	$0.72 \pm 0.19$
			Watershed	0.23	0.26	0.24	$0.86 \pm 0.13$	$0.86 \pm 0.14$	$0.84 \pm 0.1$	$0.74 \pm 0.15$	$0.68 \pm 0.16$	$0.61 \pm 0.21$
Sauvabelin	171	$0.35 \pm 0.2$	Geodesic vote*	0.62	0.53	0.57	$0.85 \pm 0.14$	$0.89 \pm 0.11$	$0.85 \pm 0.09$	$0.75 \pm 0.13$	$0.75 \pm 0.14$	$0.7 \pm 0.17$
			Watershed	0.07	0.09	0.08	$0.87 \pm 0.14$	$0.81 \pm 0.11$	$0.83 \pm 0.08$	$0.72 \pm 0.13$	$0.67 \pm 0.12$	$0.64 \pm 0.17$
Benenté	699	$0.38 \pm 0.26$	Geodesic vote	0.19	0.09	0.12	$0.86 \pm 0.16$	$0.9 \pm 0.13$	$0.86 \pm 0.11$	$0.78 \pm 0.18$	$0.73 \pm 0.21$	$0.69 \pm 0.24$
			Watershed*	0.32	0.25	0.28	$0.89 \pm 0.14$	$0.91 \pm 0.12$	$0.89 \pm 0.1$	$0.81 \pm 0.16$	$0.8 \pm 0.17$	$0.76 \pm 0.21$
Boudry D1	535	$0.47 \pm 0.25$	Geodesic vote*	0.31	0.3	0.3	$0.81 \pm 0.16$	$0.93 \pm 0.1$	$0.85 \pm 0.1$	$0.76 \pm 0.16$	$0.77 \pm 0.15$	$0.72 \pm 0.18$
			Watershed	0.18	0.19	0.18	$0.8 \pm 0.15$	$0.84 \pm 0.14$	$0.8 \pm 0.1$	$0.68 \pm 0.15$	$0.68 \pm 0.16$	$0.61 \pm 0.18$
<b>Overall</b>	1777	$0.44 \pm 0.25$	Geodesic vote*	0.38	0.31	0.34	$0.85 \pm 0.15$	$0.91 \pm 0.11$	$0.87 \pm 0.1$	$0.78 \pm 0.15$	$0.75 \pm 0.16$	$0.71 \pm 0.19$
			Watershed	0.22	0.22	0.22	$0.86 \pm 0.15$	$0.87 \pm 0.13$	$0.85 \pm 0.11$	$0.76 \pm 0.16$	$0.74 \pm 0.17$	$0.68 \pm 0.21$

Table 3.6: Detection and delineation performances of geodesic voting versus marker controlled watershed segmentation for deciduous trees only. For delineation scores the mean and standard deviation are reported. Correct detection criteria:  $\epsilon_{xy} \leq 2$  m,  $\epsilon_h \leq 2$  m,  $J_P > 0.5$ . Note that the delineation scores are computed on the correctly detected tree only, which explains why they are systematically high. The best performing method is indicated with an asterisk.

Site	Obs.	$f_A$	Method	Detection			Delineation					
				$p$	$r$	$F$	$p$	$r$	$F$	$J_P$	$J_A$	$J_V$
Versoix	301	$0.46 \pm 0.2$	Geodesic vote*	0.81	0.8	0.8	$0.9 \pm 0.12$	$0.93 \pm 0.09$	$0.91 \pm 0.08$	$0.84 \pm 0.13$	$0.77 \pm 0.16$	$0.75 \pm 0.19$
			Watershed	0.26	0.37	0.31	$0.9 \pm 0.1$	$0.83 \pm 0.15$	$0.85 \pm 0.1$	$0.76 \pm 0.15$	$0.68 \pm 0.16$	$0.62 \pm 0.2$
Sauvabelin	171	$0.35 \pm 0.2$	Geodesic vote*	0.6	0.52	0.56	$0.85 \pm 0.14$	$0.89 \pm 0.1$	$0.86 \pm 0.09$	$0.76 \pm 0.13$	$0.75 \pm 0.14$	$0.71 \pm 0.17$
			Watershed	0.07	0.09	0.08	$0.87 \pm 0.14$	$0.81 \pm 0.11$	$0.83 \pm 0.08$	$0.72 \pm 0.13$	$0.67 \pm 0.12$	$0.64 \pm 0.17$
Benenté	271	$0.38 \pm 0.21$	Geodesic vote*	0.36	0.18	0.24	$0.82 \pm 0.17$	$0.94 \pm 0.1$	$0.86 \pm 0.12$	$0.78 \pm 0.18$	$0.74 \pm 0.21$	$0.72 \pm 0.23$
			Watershed	0.19	0.32	0.24	$0.91 \pm 0.1$	$0.81 \pm 0.15$	$0.85 \pm 0.1$	$0.75 \pm 0.15$	$0.7 \pm 0.17$	$0.65 \pm 0.19$
Boudry D1	268	$0.26 \pm 0.19$	Geodesic vote*	0.7	0.57	0.63	$0.88 \pm 0.14$	$0.95 \pm 0.08$	$0.9 \pm 0.1$	$0.84 \pm 0.15$	$0.8 \pm 0.16$	$0.79 \pm 0.18$
			Watershed	0.19	0.38	0.25	$0.92 \pm 0.1$	$0.8 \pm 0.16$	$0.84 \pm 0.1$	$0.74 \pm 0.15$	$0.69 \pm 0.17$	$0.67 \pm 0.17$
<b>Overall</b>	1011	$0.37 \pm 0.21$	Geodesic vote*	0.66	0.52	0.58	$0.88 \pm 0.14$	$0.93 \pm 0.09$	$0.89 \pm 0.09$	$0.82 \pm 0.14$	$0.77 \pm 0.16$	$0.75 \pm 0.19$
			Watershed	0.19	0.31	0.24	$0.91 \pm 0.11$	$0.82 \pm 0.15$	$0.85 \pm 0.1$	$0.75 \pm 0.15$	$0.69 \pm 0.16$	$0.64 \pm 0.19$

Table 3.7: Comparison of stem detection performance with geodesic voting and local maxima detection for deciduous and coniferous trees. Correct detection criteria:  $\varepsilon_{xy} \leq 2$  m. *obs*: number of observations, *p*: precision, *r*: recall, *F*: F-score,  $\varepsilon_{xy}$ : mean and standard deviation of stem position error. The best performing method is indicated with an asterisk.

<b>Site</b>	<b>Obs.</b>	<b>Method</b>	<i>p</i>	<i>r</i>	<i>F</i>	$\varepsilon_{xy}[m]$
Versoix	372	Geodesic vote*	0.84	0.91	0.87	$0.26 \pm 0.29$
		Local maxima	0.52	0.59	0.55	$1.17 \pm 1.01$
Sauvabelin	171	Geodesic vote*	0.97	0.83	0.89	$0.32 \pm 0.22$
		Local maxima	0.17	0.23	0.2	$1.4 \pm 0.86$
Benenté	699	Geodesic vote	0.7	0.35	0.47	$0.66 \pm 0.58$
		Local maxima*	0.56	0.45	0.5	$0.98 \pm 1.22$
Boudry D1	535	Geodesic vote*	0.65	0.64	0.64	$0.53 \pm 0.5$
		Local maxima	0.46	0.49	0.47	$1.03 \pm 1.08$
<b>Overall</b>	1777	Geodesic vote*	0.75	0.6	0.67	$0.45 \pm 0.46$
		Local maxima	0.47	0.47	0.47	$1.07 \pm 0.56$

Table 3.8: Comparison of stem detection performance with geodesic voting and local maxima detection for deciduous trees only. The best performing method is indicated with an asterisk.

<b>Site</b>	<b>Obs.</b>	<b>Method</b>	<i>p</i>	<i>r</i>	<i>F</i>	$\varepsilon_{xy}[m]$
Versoix	301	Geodesic vote*	0.97	0.94	0.95	$0.2 \pm 0.14$
		Local maxima	0.46	0.65	0.54	$1.19 \pm 1$
Sauvabelin	171	Geodesic vote*	0.96	0.83	0.89	$0.32 \pm 0.22$
		Local maxima	0.17	0.23	0.2	$1.4 \pm 0.86$
Benenté	271	Geodesic vote*	1	0.5	0.67	$0.27 \pm 0.19$
		Local maxima	0.36	0.59	0.45	$1.31 \pm 0.86$
Boudry D1	268	Geodesic vote*	0.91	0.74	0.82	$0.3 \pm 0.23$
		Local maxima	0.33	0.65	0.44	$1.17 \pm 0.92$
<b>Overall</b>	1011	Geodesic vote*	0.95	0.75	0.84	$0.26 \pm 0.19$
		Local maxima	0.34	0.56	0.42	$1.23 \pm 0.47$

Table 3.9: Detection and delineation performances before and after correcting under-segmentation (for deciduous trees only).

<b>Site</b>	<b>Obs.</b>	<b>f<sub>A</sub></b>	<b>Detection</b>			<b>Delineation</b>					
			<i>p</i>	<i>r</i>	<i>F</i>	<i>p</i>	<i>r</i>	<i>F</i>	<i>J<sub>P</sub></i>	<i>J<sub>A</sub></i>	<i>J<sub>V</sub></i>
Versoix	301	0.46				0.9	0.93	0.91	0.84	0.77	0.75
		±	0.81	0.8	0.8	±	±	±	±	±	±
		0.2				0.12	0.09	0.08	0.13	0.16	0.19
Versoix (corrected)	"	"	0.79	0.77	0.78	0.94	0.84	0.88	0.79	0.7	0.68
						±	±	±	±	±	±
						0.09	0.14	0.08	0.13	0.16	0.18
Sauvabelin	171	0.35				0.85	0.89	0.86	0.76	0.75	0.71
		±	0.6	0.52	0.56	±	±	±	±	±	±
		0.2				0.14	0.1	0.09	0.13	0.14	0.17
Sauvabelin (corrected)	"	"	0.41	0.36	0.38	0.93	0.69	0.78	0.64	0.64	0.56
						±	±	±	±	±	±
						0.09	0.12	0.07	0.1	0.1	0.12
Benenté	271	0.38				0.82	0.94	0.86	0.78	0.74	0.72
		±	0.36	0.18	0.24	±	±	±	±	±	±
		0.21				0.17	0.1	0.12	0.18	0.21	0.23
Benenté (corrected)	"	"	0.53	0.27	0.36	0.85	0.81	0.81	0.7	0.66	0.6
						±	±	±	±	±	±
						0.14	0.12	0.09	0.13	0.14	0.15
Boudry D1	268	0.26				0.88	0.95	0.9	0.84	0.8	0.79
		±	0.7	0.57	0.63	±	±	±	±	±	±
		0.19				0.14	0.08	0.1	0.15	0.16	0.18
Boudry D1 (corrected)	"	"	0.39	0.31	0.35	0.97	0.67	0.79	0.66	0.63	0.61
						±	±	±	±	±	±
						0.05	0.1	0.07	0.1	0.11	0.11



(a) All the labeled points.

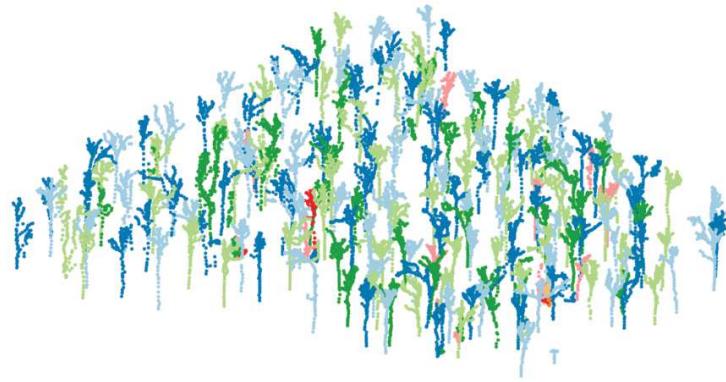
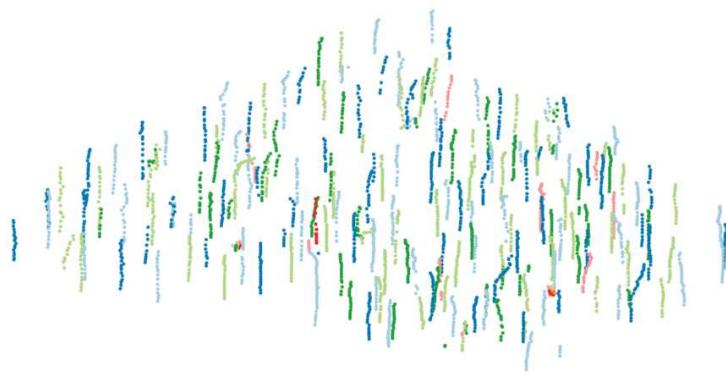
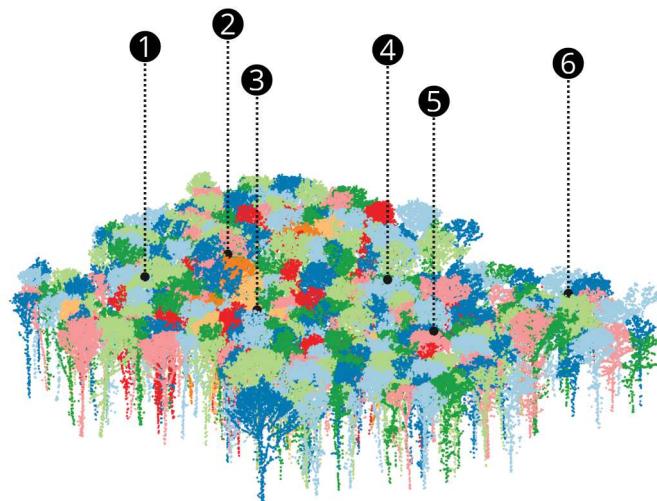
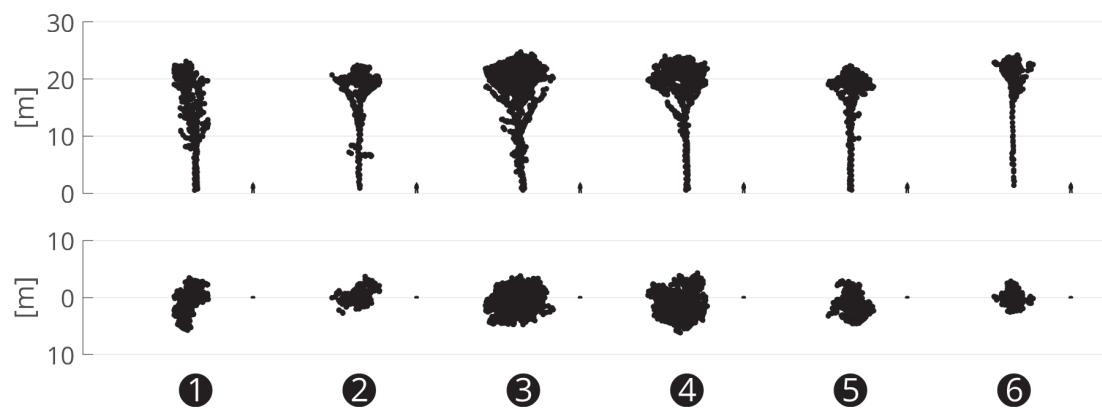
(b) Labeled points with  $b \geq 0.03$ .(c) Labeled points with  $b \geq 0.5$ .

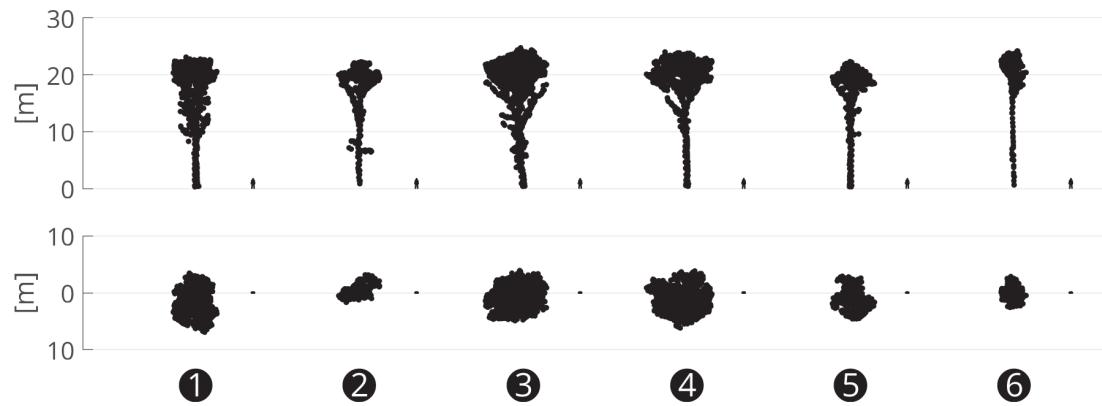
Figure 3.14: Individual tree segments obtained with geodesic voting on the Boudry D1 site (deciduous trees only). Thresholding the relative geodesic density ( $b$ ) reveals different branching hierarchies.



(a) All the segments (topological coloring)



(b) Example segments obtained with the algorithm. Side (first row) and top (second row) view.



(c) The corresponding manually delineated reference segments. Side (first row) and top (second row) view.

Figure 3.15: Individual tree segments obtained with geodesic voting on the Versoix site

### 3.3.3 Discussion

The results systematically show that the geodesic vote algorithm performs better (cf. tables 3.5, 3.6, 3.7, 3.8) in pure deciduous forests with a moderate adjacency factor (cf. figure 3.12). Thus, the presence of coniferous trees in the understory significantly reduces segmentation performance. This is not a surprising result, as the algorithm was not designed to handle coniferous trees. Removing the coniferous trees from the simulated mixed forest plots has a double effect: a reduction of the spatial adjacency factor (i.e. lower density) and a reduction of trees which do not have a well sampled stem and branching structures.

When considering only stem detection (cf. tables 3.7 and 3.8), the performances are generally high for the pure deciduous forest plots (overall F-score = 0.84). With only a few cases of false positive detections, the algorithm can be considered reliable for stem detection. The large performance differences between the stem detection (cf. tables 3.7, 3.8) and full shape delineation performances (cf. tables 3.5, 3.6) also illustrate the importance of conducting an error assessment based on the 3D shape delineation and not just on the tree position/height/surface when characterizing segmentation performance. The comparison with marker controlled watershed segmentation also shows that geodesic vote has a much higher rate of correct detections both for full tree detection (cf. tables 3.5, 3.6) and stem position detection (cf. tables 3.7, 3.8).

The parameter sensitivity analysis shows that the edge weight exponent  $a = 2$  systemically produces better segmentation performances, although the improvement is only marginal on some sites. The effect of vertical scaling is less clear and is more site dependent. Generally speaking, a moderate amount of vertical scaling ( $> 0.75$ ) seems to be beneficial. As it can be seen on the Benenté site, reducing the vertical scale below 0.5 can help improve the segmentation when lower parts of the stem have not been well sampled.

The correction for under-segmentation is detrimental to the segmentation performance on all sites except Benenté. This can be explained by the fact there are many trees with missing stems on this site. Thus, applying the correction is only recommendable when frequent under-segmentation is observed.

The main advantages of the geodesic vote algorithm are its direct applicability to the ALS point cloud, its robustness to noisy points, its relatively good resilience to density down sampling, the fact that no normalization of the point cloud with respect to the terrain is required and its ability to directly provide information on branching hierarchy. Its main drawbacks are the necessity to use high density point clouds acquired in leaf-off conditions, its tendency to split large trunks into separate trees (i.e. false positives) and its inapplicability to evergreen (e.g. coniferous) trees. It can also be noted that applying the algorithm to large areas requires splitting the point cloud into overlapping tiles and processing each tile separately. The resulting segmented tiles can then be merged to reconstitute the original point cloud. This method also highlights the potential of leaf-off acquisitions to characterize deciduous forests at the individual tree scale. In the past, leaf-on conditions have often been favored because most forestry applications were based on the creation of raster CHM for which information about the tree branching structure is not required.

There are several possible modifications and additional processing steps that could be applied to the algorithm to potentially improve its performance. The preliminary filtering of deciduous trees (i.e. removing all coniferous trees including those located in the understory) with a classification/clustering (cf. chapter 5) algorithm is very likely to improve the results. The point elevation could be normalized with respect to the terrain and the algorithm could be iteratively reapplied using a series of multilevel terrain planes (considering at each iteration only the points located above the plane). This modification could possibly solve the problem of poor sampling in lower parts of the stems. More sophisticated graph edge weight functions could be used (for example integrating intensity, echo width, color). Finally, more advanced a posteriori filtering approaches could be used, for example based on shape probability indexes.

### 3.4 Evaluating and improving segmentation with ensemble filtering

This section presents a method to estimate and improve individual tree segmentation error based on the ensemble theory framework. It is adapted from Parkan and Tuia (2018).

Errors in individual tree shape delineation propagate in further processing steps (e.g. timber volume, biomass and species prediction), so it is important to quantify them. However, the majority of segmentation algorithms do not directly provide any information about shape delineation uncertainty. Instead, the evaluation of shape delineation accuracy is usually done by comparing the segmentation with an independently produced and reliable reference (e.g. manual segmentation and/or field surveys). However, this approach is limited by the low availability of individual tree shape reference data over large areas. For this reason, 3D shape delineation accuracy is very often not evaluated in individual tree segmentation studies (Yin and Wang, 2016). Even though independent tree shape validation data may not be available, quantifying segmentation uncertainty is still necessary. One possible solution is to use algorithms which compare observed values with model based expectations (geostatistics for example model spatial autocorrelation as function of range). In addition to providing a prediction uncertainty, these methods also typically produce better predictions, because they incorporate prior knowledge about the investigated phenomenon. In the context of segmentation, such approaches involve modeling the spatial distribution and/or the shape of trees based on prior botanical and ecological knowledge. Many tree species exhibit an increase in crown geometry variability (heteroscedasticity) as a function of age (height) and environmental conditions. However, some coniferous species (such as Spruce and Fir) exhibit less geometric variability and are generally easier to model. For this reason, model based segmentation algorithms are generally better suited for coniferous forests.

In this section, a method which models tree shape probability directly from the ALS data (i.e. without the need for a predefined model) is described. The method uses the ensemble learning (model averaging) framework (Schapire, 1990; Breiman, 1996; Kuncheva, 2004). An ensemble is a group of segments which share similar (geometric and radiometric) features. We make the hypothesis that segments in an ensemble can be considered as noisy instances of the same tree shape template. By comparing all shape instances within an ensemble, inconsistencies between the shapes can be detected and an estimate of a probable underlying tree shape is obtained. The proposed method depends on several assumptions:

- Tree top geometric features can be used as proxies of overall tree shape.
- Tree crowns exhibit approximate radial symmetry.
- Tree growth is approximately vertical.

#### 3.4.1 Description

The proposed methods starts with a set of individual tree segments (obtained with any generic segmentation algorithm, such as marker controlled watershed (Meyer and Beucher, 1990; Meyer, 1994; Soille, 2013)). Each segment is then characterized by a set of descriptive (geometric and radiometric) features and matched with similar segments to form ensembles (groups). Within each ensemble, the 3D alpha shapes (concave hulls) (Edelsbrunner and Mücke, 1994) derived from the point cloud segments are mutually overlaid to detect common regions and determine shape probability. A threshold is then applied to the probability, to filter out erroneous points from the initial segmentation.

The five main steps of the method are summarized in figure 3.16 and each step is explained in the following subsections. The method was implemented in Matlab r2016b using custom functions part of which were included in the *Digital Forestry Toolbox* (Parkan, 2017a).

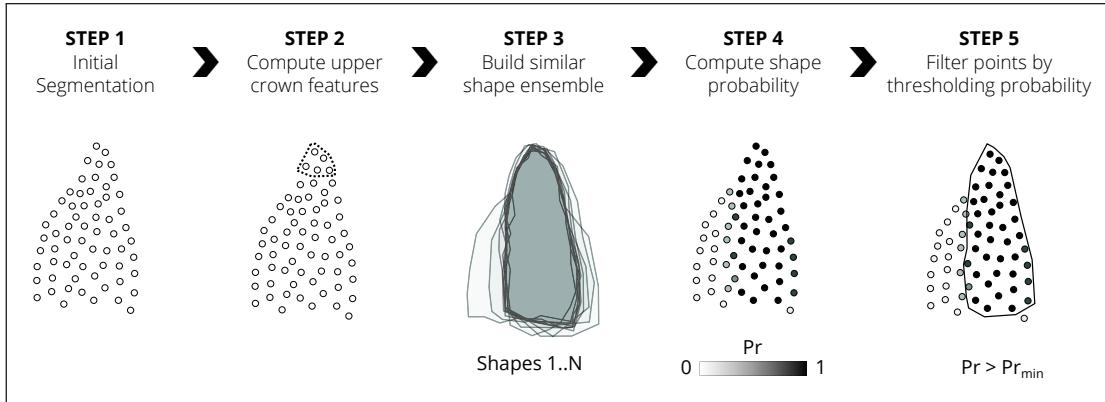


Figure 3.16: Main steps used to compute shape probability and subsequently filter the initial segment shape.

### Step 1 - Initial segmentation

A 0.4 m resolution raster Canopy Height Model (CHM) is first derived from the classified 3D point clouds (for the three sample sites). The CHM is smoothed using a Gaussian 6x6 lowpass filter. Tree top (local maxima) detection is then performed using a variable radius ( $r$ ) convolution window defined by a function of the pixel metric height ( $h$ ):

$$r(h) = 0.5 + 0.25 \cdot \log(\max_h(h, 1)) \quad (3.13)$$

The local maxima are merged and the highest point is retained, if separated by less than the 3D adjacency distance defined by function  $d_{adj}(h)$ :

$$d_{adj}(h) = \min_h(0.5 + 0.5 \cdot \log(\max_h(h, 1)), 4) \quad (3.14)$$

The choice of a logarithmic variable radius in eq. 3.13 and 3.14 is based on the observed relationship between upper crown radius and tree height in the region of interest. However, this relationship may vary significantly between forest types (Duncanson et al., 2015) and other variable radius functions such as those proposed in Pitkänen et al. (2004); Popescu and Wynne (2004); Chen et al. (2006) may be used in place of eq. 3.13 and 3.14.

The detected local maxima (cf. figure 3.17a) are subsequently used as markers (i.e. seed points) in watershed segmentation (Meyer and Beucher, 1990; Meyer, 1994; Soille, 2013) to label individual tree crowns (cf. figure 3.17b). The CHM labels are then assigned to their nearest 3D points, to obtain a 3D labeled point cloud. The presence of partial tree crowns would bias the shape probability estimates. Thus, segments located within 10 m of the edge of the point cloud are excluded. For the same reason, the segmentation parameter values (regardless of the segmentation algorithm) should be set to avoid over-segmentation.

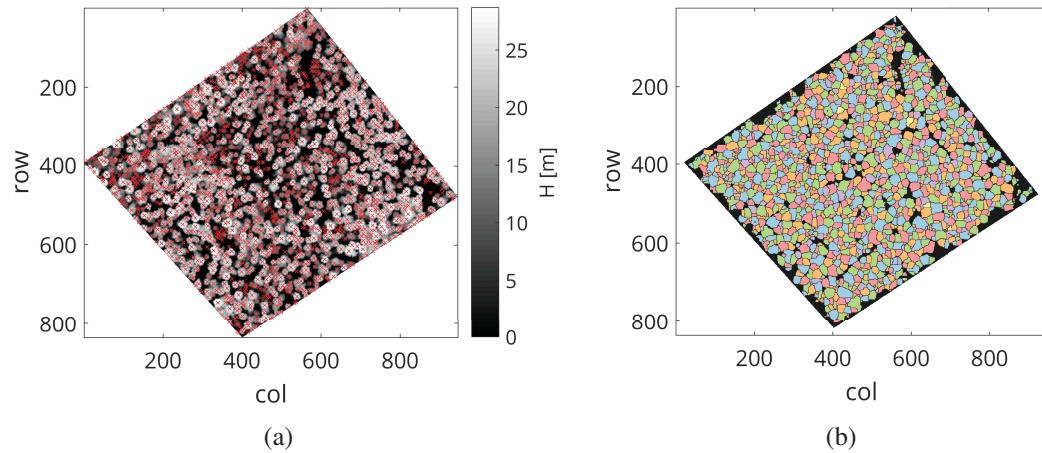


Figure 3.17: (a) Tree top detection results with variable size convolution window. (b) Raster CHM segmentation obtained with the marker controlled watershed algorithm.

### Step 2 - Computing upper crown features

In order to compare and group tree shapes in step 3, a set of descriptive features is required. Thus, the total height  $h$ , upper crown (i.e. points located in the upper 15% of the crown) convex volume  $v$  and median return intensity  $i$  (normalized by the [0.05, 0.95] quantile range) are computed for each segment. These upper crown features were chosen because for trees with a conical shape, they are less affected by poor segmentation than features that describe the lower parts of segments (cf. figure 3.18a).

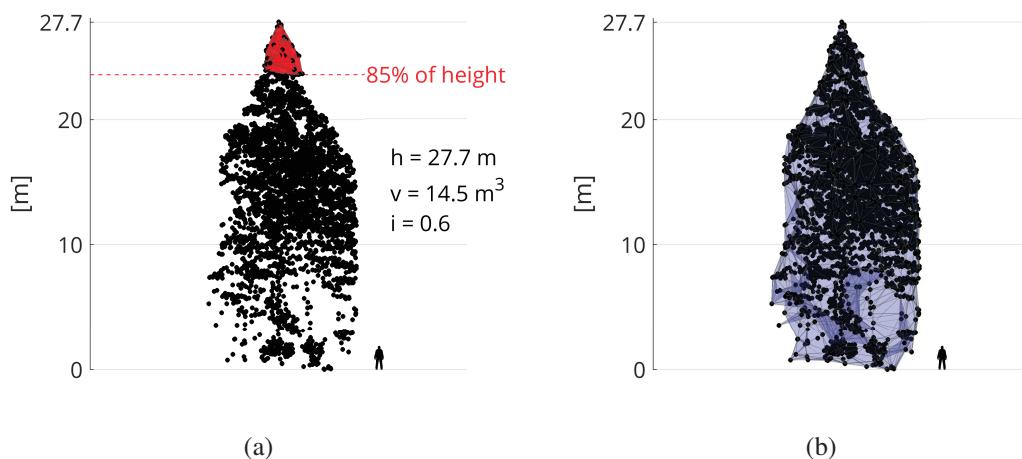


Figure 3.18: (a) The total height ( $h$ ), the 3D convex alpha shape (in red) volume ( $v$ ) and the median intensity ( $i$ ) of points located in the top 15 % of the tree crown are used as features because they are less affected by poor segmentation. (b) The single region 3D alpha shape (outlined in blue) derived from the point cloud segment.

### Step 3 - Building shape ensembles (grouping similar segments)

First, the XYZ point coordinates of each segment  $i$  are normalized so that the segment origin is vertically aligned with the tree top:

$$\mathbf{XYZ}_{norm}^{(i)} = \mathbf{XYZ}^{(i)} - \mathbf{J} \cdot \mathbf{XYZ}_{root}^{(i)} \quad (3.15)$$

where:

- N is the number of points in segment  $i$ ;
- $\mathbf{XYZ}_{norm}^{(i)}$  is a Nx3 matrix containing the normalized 3D point coordinates of segment  $i$ ;
- $\mathbf{XYZ}^{(i)}$  is a Nx3 matrix containing the original 3D point coordinates of segment  $i$ ;
- $\mathbf{XYZ}_{root}^{(i)}$  is a 1x3 matrix containing the root coordinate of the segment  $i$  (i.e. the projection of the tree top on the terrain model);
- $\mathbf{J}$  is a Nx1 vector of ones.

This coordinate normalization is required to overlay (stack) all shapes within an ensemble. Then, the single region 3D alpha shape (Edelsbrunner and Mücke, 1994) (cf. figure 3.18b) of each segment is computed and any holes in the shape are filled. Subsequently, ensembles (cf. figure 3.19) are constructed by grouping segments which share similar geometric and radiometric features (computed at step 2). Formally, given a segment  $i$  with total height  $h_i$ , upper crown convex volume  $v_i$  and upper crown median intensity  $i_i$ , all segments  $j$  with  $j \in [1 - N_{segments}]$  which fulfill the criteria listed in table 3.10 form the ensemble  $i$ .

Table 3.10: Criteria used to create ensembles (groups) of similar segments.

Feature	Criteria
Total height	$h_i - 0.5 \leq h_j \leq h_i + 1.15$
Upper crown convex hull volume	$v_i \leq v_j \leq 1.2 \cdot v_i$
Upper crown median intensity	$i_i - 0.2 \cdot i_{range} \leq i_j \leq i_i + 0.2 \cdot i_{range}$

The tolerances in terms of height, upper crown volume and median intensity differences used when matching segments are set empirically. There is a trade-off between these margins and the ensemble sizes. Tighter tolerances result in smaller ensembles and thus larger datasets are needed to reach the minimum required ensemble sizes.

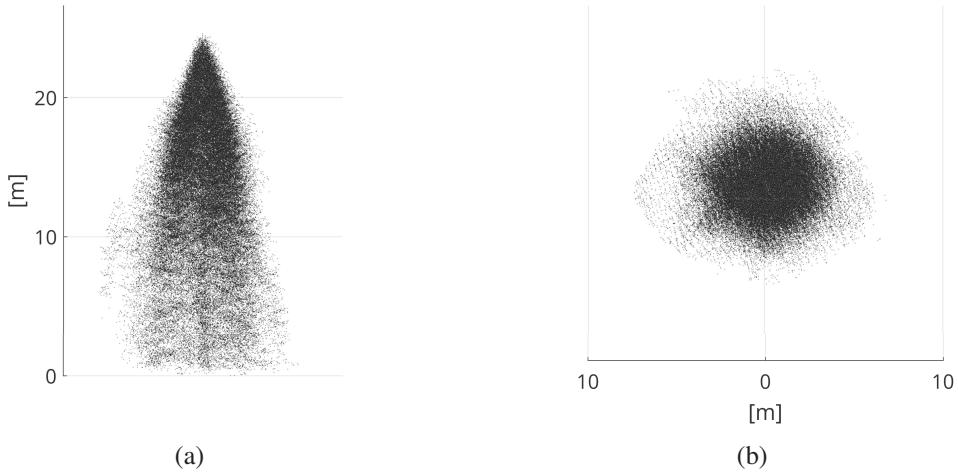


Figure 3.19: Example of an ensemble containing 69 overlaid segments with similar features. Dense point areas indicate high shape probability. **(a)** Side view **(b)** Top view

#### Step 4 - Computing shape probability

The shape probability is defined as the number of times a point was included in the alpha shapes of the ensemble divided by the ensemble size  $N$  (i.e. number of matching segments), as illustrated in figure 3.20. For each set of points  $P_0$  which form a segment, the shape probability  $Pr(P_0 \subseteq S_{0..N})$  is given by:

$$Pr(P_0 \subseteq S_{0..N}) = \begin{cases} \frac{\sum_{i=0}^N P_0 \subseteq S_i}{N} & \text{if } N \geq N_{min} \\ 0 & \text{otherwise} \end{cases} \quad (3.16)$$

where:

- $S_i$  is the alpha shape of segment  $i$ ;
- $S_{0..N}$  is the set of  $N$  alpha shapes with features similar to  $S_0$ ;
- $N$  is the number of segments in the ensemble  $i$ ;
- $N_{min}$  is the minimum number of segments per ensemble required to compute a reliable shape probability (10 was used here).

Thus, regions which are common to many alpha shapes in the ensemble obtain higher probability scores than regions that are only visible in few segments.

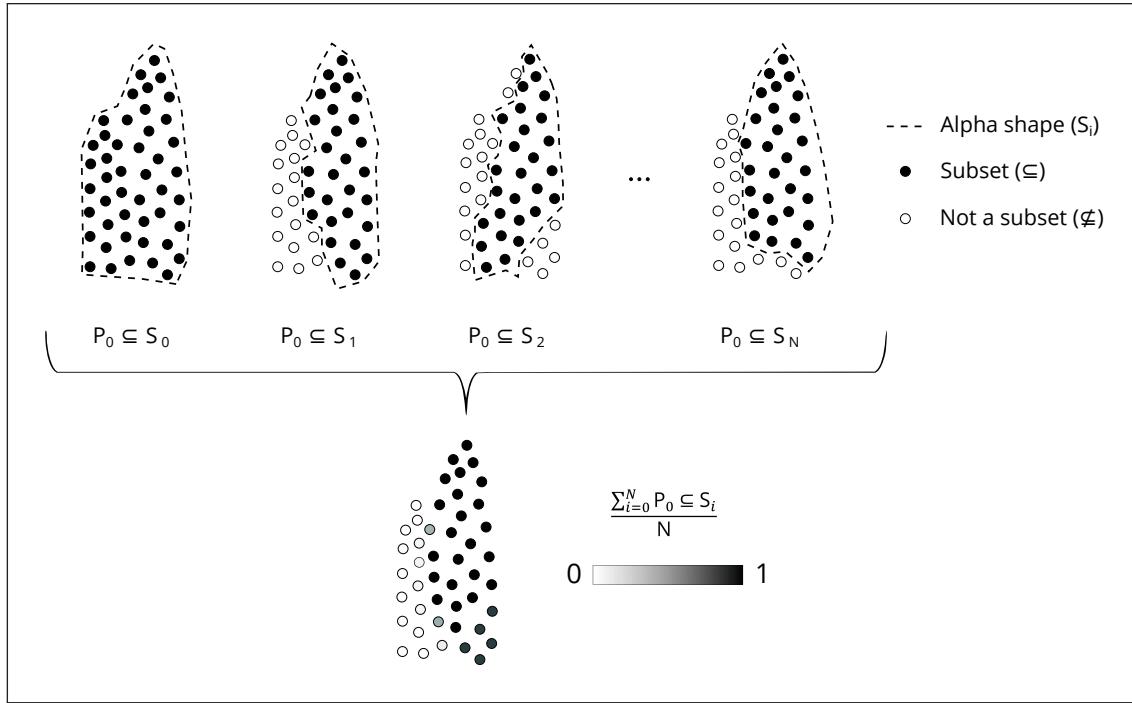


Figure 3.20: Each point cloud segment  $P_0$  is overlaid with the  $S_{0..N}$  alpha shapes of similar segments (including itself). Regions of the point cloud segment which occur more frequently inside  $S_{0..N}$  obtain a higher shape probability. Thus, inconsistencies between the shapes in the ensemble can be detected.

### Step 5 - Filtering

The points from the initial segments can be filtered by applying a threshold ( $Pr_{min}$ ) to the shape probability. The filtered point subset is defined by:

$$I_F(Pr_i) = \begin{cases} 1 & \text{if } Pr_i \geq Pr_{min} \\ 0 & \text{otherwise} \end{cases} \quad (3.17)$$

where:

$I_F$  is the indicator function which produces the filtered point subset;

$Pr_i$  is the shape probability associated with each point in segment  $i$ ;

$Pr_{min}$  is the minimum probability required to retain a point in the segment.

An optimal value of  $Pr_{min}$  can be set by visually examining the effect of applying different threshold values to a (height stratified) sample of the segments.

### 3.4.2 Results

In this section, the ensemble filtering method is applied to the Brévine study site and its performance is evaluated according to the error assessment framework presented in chapter 3.2.

The unfiltered (initial) and filtered shapes were compared to the manually delineated reference shapes. The delineation performance (cf. figure 3.1) was evaluated (for detected trees only) in terms of recall  $r$ , precision  $p$ , F-score  $F$  and Jaccard index  $J$  (also called Intersection over Union  $IoU$ ). The correct detection rate ( $d$ ), which is equivalent to the detection recall ( $r$ ), was computed as the proportion of segments with a delineation  $J_P > 0.5$  (i.e. a segment is considered to be detected if more than half of its points overlap with the reference points).

The detection metric and the median of each delineation metric (except F score) for different values of  $Pr_{min}$  are presented in figure 3.21. The same metrics for  $Pr_{min} = 0.25$  are also reported in table 3.11 stratified by height category. Figure 3.22 provides boxplots of the delineation metrics. Figure 3.24a illustrates the resulting probability map, figure 3.24b provides examples of individual tree shape probability and figure 3.24c shows the resulting filtered segments.

A one-sided Wilcoxon signed-rank test was used to compare the delineation before and after filtering with  $Pr_{min} = 0.25$ . This test was chosen because the before/after delineation scores are dependent and not normally distributed. In this test, the alternate hypothesis is that the score values after filtering minus those before filtering come from a distribution with a median greater than 0. Using a 0.5% (i.e.  $\alpha = 0.005$ ) significance level, the alternate hypothesis was accepted for all the delineation scores except recall. In other words, these delineation scores were significantly higher after filtering. The associated p-values of the comparison tests were  $p : 1.2176 \cdot 10^{-68}$ ,  $r : 1$ ,  $F : 1.4491 \cdot 10^{-6}$ ,  $J_P : 2.4291 \cdot 10^{-6}$ ,  $J_V : 9.6308 \cdot 10^{-7}$ ,  $J_A : 4.7761 \cdot 10^{-16}$ .

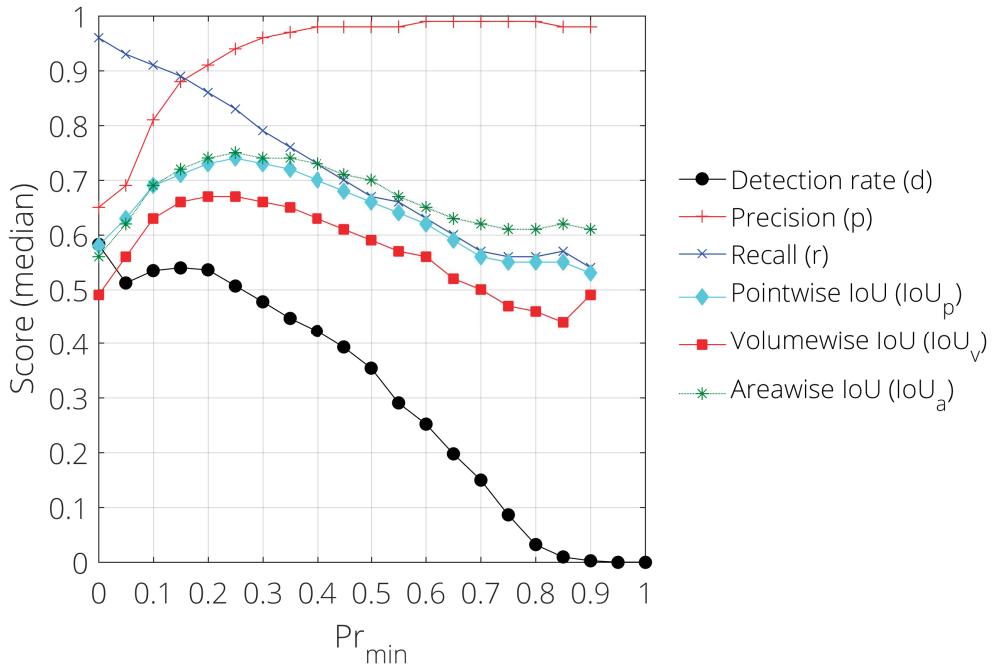


Figure 3.21: Sensitivity of the median validation scores to  $Pr_{min}$ . Notice that the delineation scores are undefined when the detection rate reaches 0.

Table 3.11: Comparison of detection and median delineation performance before and after filtering the segments with  $Pr_{min} = 0.25$ . Scores were rounded to the nearest second decimal.  $obs$ : number of observations,  $d$ : detection rate,  $p$ : precision,  $r$ : recall,  $F$ : F-score,  $J_P$ : pointwise intersection over union,  $J_V$ : volumewise intersection over union,  $J_A$ : areawise intersection over union.

height [m]	obs.	Detection			Delineation				
			$d$	$p$	$r$	$F$	$J_P$	$J_V$	$J_A$
$0 \leq h < 10$	118	no filter	0.58	0.65	0.96	0.74	0.58	0.47	0.58
		filter	0.49	0.92	0.83	0.84	0.73	0.65	0.73
$10 \leq h < 20$	182	no filter	0.57	0.60	0.94	0.71	0.55	0.48	0.53
		filter	0.51	0.93	0.81	0.81	0.69	0.64	0.71
$20 \leq h$	454	no filter	0.59	0.68	0.96	0.75	0.60	0.51	0.56
		filter	0.51	0.95	0.84	0.87	0.76	0.69	0.77
<b>Overall</b>	754	no filter	0.58	0.65	0.96	0.74	0.58	0.49	0.56
		filter	0.51	0.94	0.83	0.85	0.74	0.67	0.75

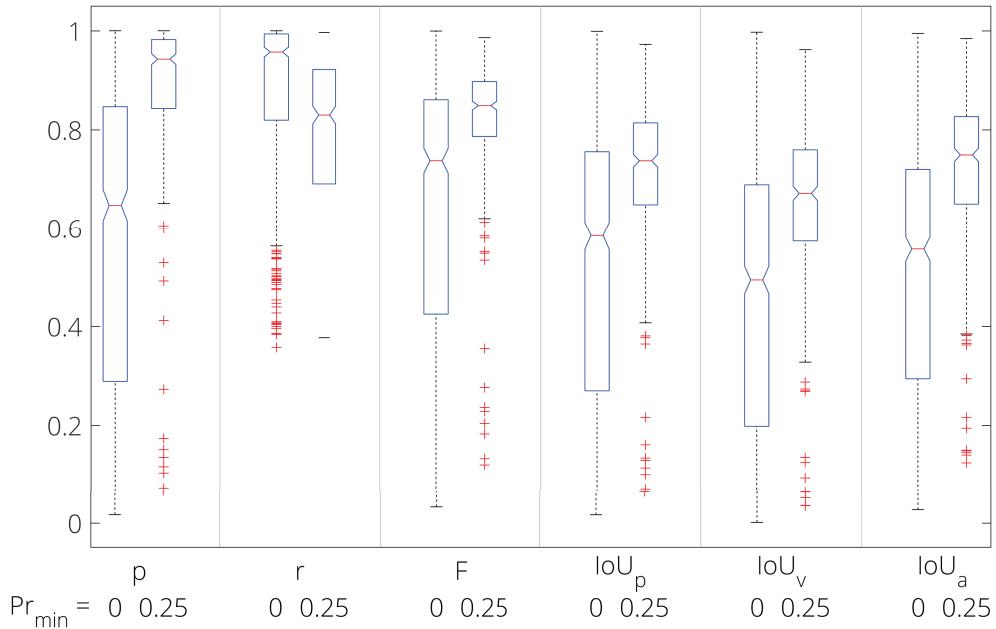


Figure 3.22: Boxplots of delineation scores before and after filtering the segments with  $Pr_{min} = 0.25$ . All the delineation scores except recall are significantly higher after filtering. It can also be noted that the filtering reduces the score spread.

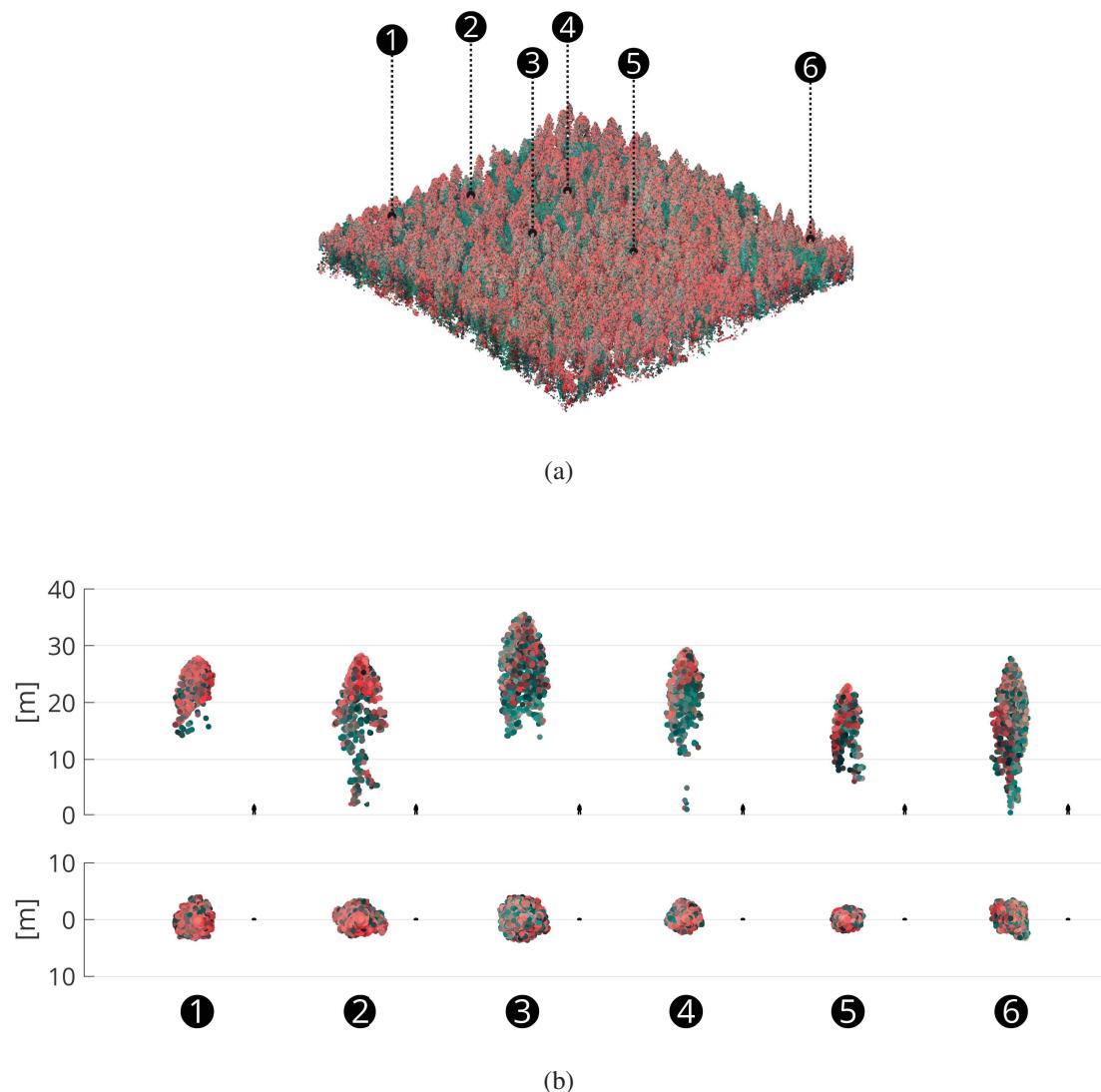


Figure 3.23: (a) False color composite (Red channel = ALS intensity rescaled to 0-1 range, Green channel = aerial image Red, Blue channel = aerial image Green) oblique view of the ALS point cloud (high vegetation only). (b) Side (first row) and top (second row) view of manually delineated tree examples.

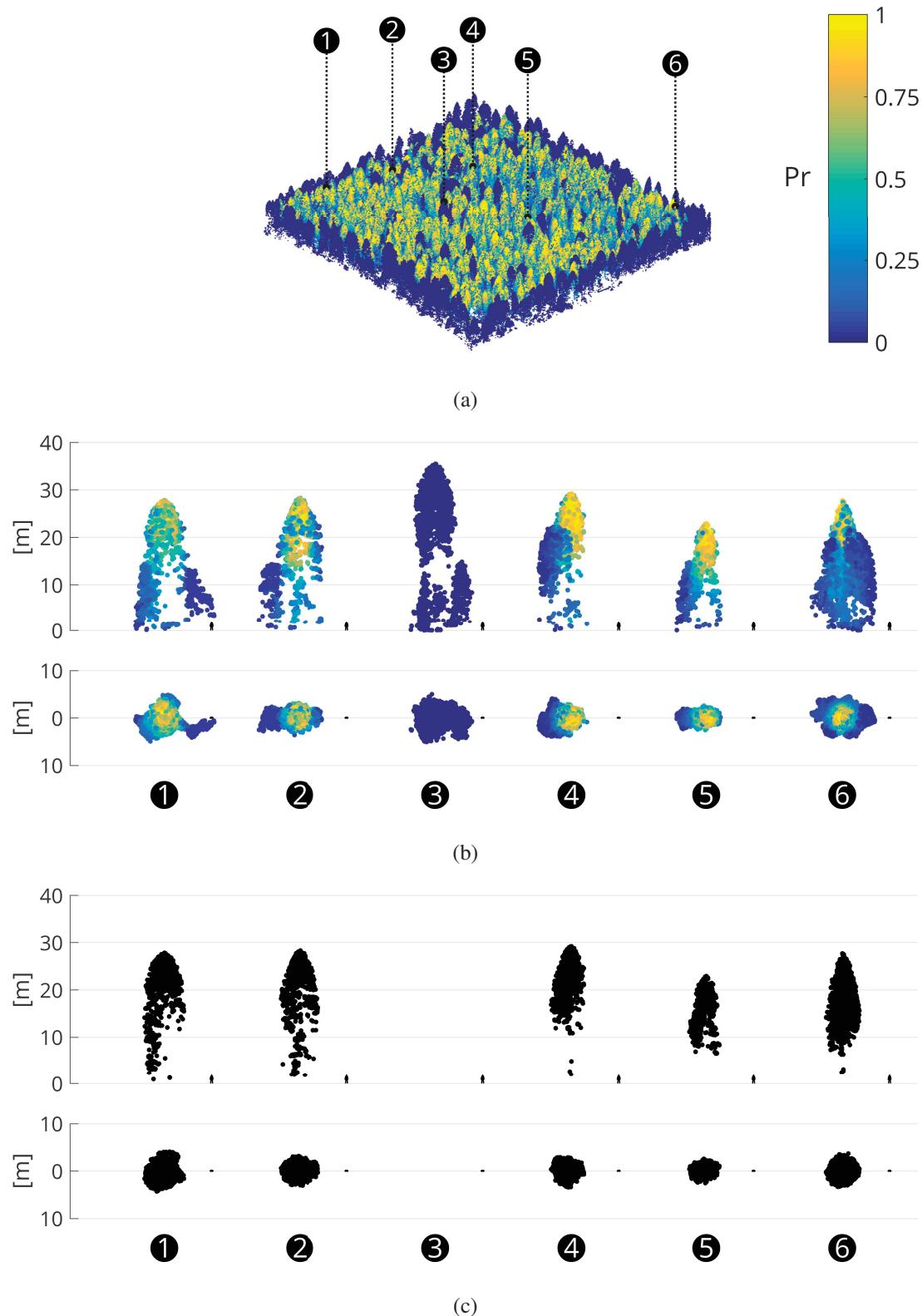


Figure 3.24: (a) Shape probability map (high vegetation only). (b) Side (first row) and top (second row) view of shape probability for six example segments. Notice that segment n°3 has null probability. This is explained by the fact it is a particularly high tree and there was an insufficient number of similar trees to form a reliable ensemble. (c) Filtered segments using  $Pr_{min} = 0.25$ .

### 3.4.3 Discussion

The results indicate that the proposed method can produce an estimate of tree shape delineation uncertainty. Shape probability at the point scale can be aggregated at the segment scale to produce a mean or median shape probability providing information on individual segment shape uncertainty.

In the presented study case, when using  $Pr_{min}$  values ranging from 0.1 to 0.6, the Jaccard index scores are improved with a peak at  $Pr_{min} = 0.25$ . The precision score is increased and the recall score is reduced for all values of  $Pr_{min}$ . The detection rate is also reduced for all values of  $Pr_{min}$ . This reduction in detection rate is due to the definition which requires  $Iou_p > 0.5$  for a segment to be counted as a correct detection. The better delineation scores obtained after filtering are due to the combined effect of removing erroneous points and discarding segments with  $Iou_p \leq 0.5$ .

For practical applications over large areas, it is sufficient to compute ensembles based on a subset of the area which includes most of the tree shape variability and a sufficient amount of redundant shape examples. By thresholding the resulting shape probability, a set of tree shape templates are produced. New segments (i.e. outside the sample area) can then be matched (i.e. using upper crown features) and compared with their most similar shape template to produce an estimate of their segmentation uncertainty.

The proposed ensemble based filtering method has several advantages. The segmentation/shape uncertainty estimate can be improved by adding additional observations to the ensembles. The method is adaptive because it does not rely on predefined allometric rules or 3D model templates. Moreover, although marker controlled watershed was used to produce the segmentation, any other automatic or manually delineated segments could be used instead in the first step. Finally, the method does not require high ALS point densities.

The main drawbacks of the method are its dependency on specific coniferous tree shapes, the need to use datasets with multiple examples of similar trees and the computational cost. Most of the computation time is used to compute the single region alpha shapes (~23%) and the shape probability inclusion tests (~74%). The total time to apply the method for the three sites used in this study was ~30 minutes. This computation time may be reduced by sub-sampling the point cloud (i.e. lowering density) and using fixed values of  $\alpha$  when computing the alpha shapes.

Further improvements could involve classification and separation of deciduous and coniferous trees, before running the algorithm. This separation step could be accomplished using ALS data alone using intensity (leaf-off), opacity (leaf-off) and/or shape features, for example with the approach described in Liang et al. (2007). The method is conditioned by the segmentation algorithm employed in step 1. In the current implementation, marker controlled watershed segmentation is used, thus points located in crown intersection regions cannot be allocated to a tree with certainty. This limitation could be improved by using a more sophisticated segmentation algorithm working at the inner crown level. Additional features could be included to improve the segment grouping step. These could include RGB or multispectral indices (e.g. from multiple wavelength LiDAR), geometric features (e.g. crown base height, convexity, surface area, projected area, etc). In particular, the addition of crown base height (which can be estimated for example with the approach used in Duncanson et al. (2014)) to the list of grouping features used at step 3 could possibly improve the shape uncertainty estimate beneath the crown base. Also, since it is assumed that coniferous trees exhibit approximate vertical radial symmetry, additional shape instances could be generated artificially by simply rotating the segments around the vertical (Z) axis. Finally, the alignment of segments could be improved by using a more elaborated co-registration algorithm. The segmentation and filtering procedure could theoretically be repeated and detected trees removed at each iteration until no more detectable trees were left in the point cloud. Finally, the method may also be employed to automatically create 3D tree shape templates which can be used in other processing routines.

### 3.5 Stem detection with layered morphological analysis

#### 3.5.1 Description

Tree crown extent is generally not measured in operational forest inventories and for many forestry applications the measurement of stem characteristics (diameter, height, taper) is sufficient. Moreover, tree crown shapes may evolve rapidly and their accurate delineation from ALS data can be difficult. Finally, high density ALS acquisitions which sample the stems with enough points to allow direct diameter measurement are becoming commonplace. For these reasons, it can be interesting to ignore tree crown delineation and investigate stem detection as a standalone problem.

This section describes a simple and efficient algorithm which is able to detect individual tree stems from ALS point clouds acquired in leaf-off conditions. The method relies on several assumptions:

- most trees have an approximately vertical growth;
- most stems are approximately linear;
- due to their large diameter relative to the laser footprint, stems have a higher probability of generating last returns (echoes) in leaf-off acquisitions. Although this observation is valid for most long range ALS acquisitions, it is dependent on flying height and laser beam divergence.

The proposed method is based on the morphological analysis of horizontal cross-sections (layers) sampled at regular height intervals. Splitting the point cloud into a set of horizontal cross-sections and clustering the points in each layer separately is a form of divide and conquer approach. It has been used for individual tree segmentation and has delivered promising results (Wang et al., 2008a; Moskal and Zheng, 2011; Tang et al., 2013; Kandare et al., 2016; Zhao et al., 2017b; Ayrey et al., 2017). The main advantages of layered clustering approaches is that they can generally handle multiple forest strata (under-story trees) and irregular crown geometries.

Segmentation algorithms that employ this approach rely on the same basic idea: the separability of individual trees is much better above and/or below the crown intersection/adjacency height. Due to the large inter-tree spacing in these upper/lower layers, simple algorithms such as hierarchical clustering, kmeans, connected component analysis and DBSCAN generally work well to identify individual tree locations. These initial (reliable) clusters can subsequently guide the segmentation process across the more cluttered layers. The resulting set of 2D clusters can then be combined to produce 3D clusters.

Each cross-section can either be processed directly in its raw vector representation or it can be converted into a raster. The latter representation generally allows faster and less computationally intensive processing.

The method exploits the fact that point densities are higher near the tree stems, an observation first made by Rahman et al. (2009). This phenomenon is further increased when considering only last returns, as larger branches and stems have a higher probability of completely intercepting the laser beam. The core principle of the proposed method is to identify clusters (connected components) in horizontal cross-sections, compute their centroids, and create a centroid density image by accumulating the centroids computed in each layer. This centroid density image can then be thresholded to identify the locations of individual stems.

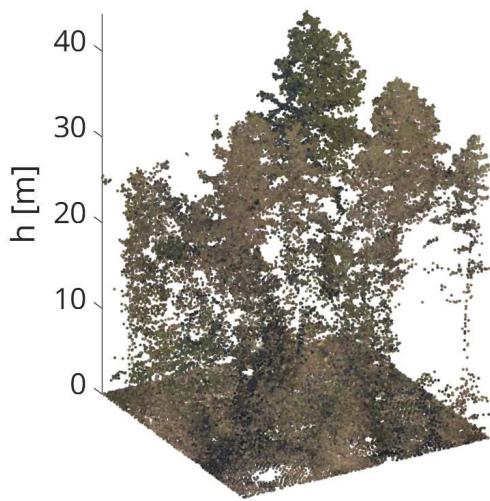
### Algorithm

The key steps of the method are illustrated in figure 3.25. Below is a detailed description of each step:

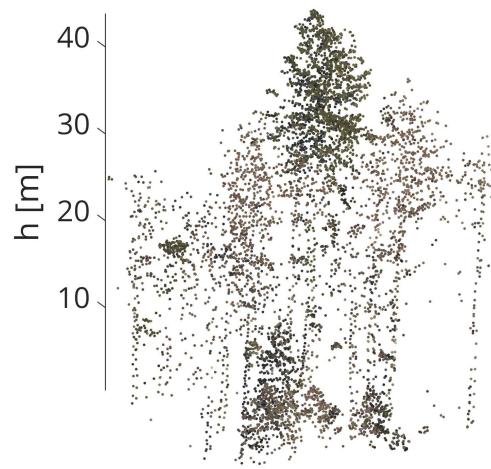
- (a) **Normalize the point elevation relative to the terrain elevation** (cf. figure 3.25a).
- (b) **Filter the points** (cf. figure 3.25b).
  - 1 Apply a vegetation class filter
  - 2 Apply a last return filter
  - 3 Apply an acquisition date filter (leaf-off)
  - 4 Apply a minimum height filter
  - 5 (Optional) Apply other filters (e.g. reflectance, color, etc...)
- (c) **Initialize the centroid density image to zero**
- (d) **Split the point cloud into a series of overlapping horizontal cross-sections with thickness w** (cf. figure 3.25c).
- (e) **For each layer, apply the following steps:** (cf. figures 3.25d and 3.25e).
  - 1 Rasterize the layer points to create a 2D boolean image, using raster resolution  $d_{xy}$ .
  - 2 Close small gaps with morphological closing (using a circular structuring element with radius  $r_1$ )
  - 3 Fill holes (flood filling)
  - 4 Identify connected components
  - 5 Compute the area ( $A_i$ ), perimeter ( $P_i$ ) and circularity ( $C_i = \frac{4\pi A_i}{P_i^2}$ ) of each connected component
  - 6 Filter components by imposing a maximum area ( $A_{max}$ ) and minimum circularity ( $C_{min}$ )
  - 7 Compute the centroid pixel of each filtered component and create a boolean image containing only the centroid pixels
  - 8 Buffer the centroid boolean image with morphological dilation (using a circular structuring element with radius  $r_2$ )
  - 9 Add the centroid boolean image to the density image (cf. figure 3.25f)
- (f) **Apply a threshold ( $\rho$ ) to the centroid density image** to create a boolean mask (cf. figure 3.25g) and buffer it with morphological dilation (using a circular structuring element with radius  $r_3$ ).
- (g) **Label the connected components in the boolean mask** (cf. figure 3.25h).
- (h) **Transfer the labels from the connected components to the unfiltered point cloud** (cf. figure 3.25i).
- (i) **Determine stem positions** by computing the median of XY coordinates in lower parts of the stem (cf. figure 3.25i).

### Parameters

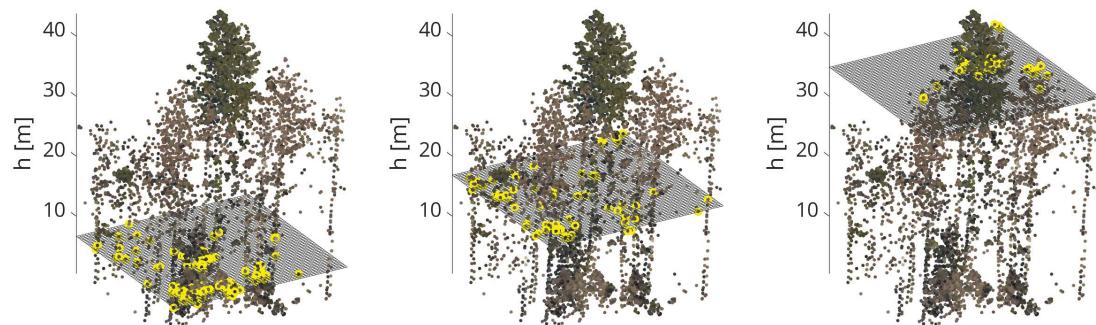
$d_{xy}$	Raster resolution (nominal value: 0.4 m);
$dz$	Vertical step (nominal value: 0.5 m);
$w$	Layer thickness (nominal value: 1 m);
$\rho$	Density threshold (nominal value: 5 m);
$A_{max}$	Maximum object area (nominal value: $\pi \cdot 4 \text{ m}^2$ );
$C_{min}$	Minimum object circularity (nominal value: 0.7);
$r_1, r_2, r_3$	Circular structuring element radius (nominal values: 1 m, 0.5 m, 0.5 m).



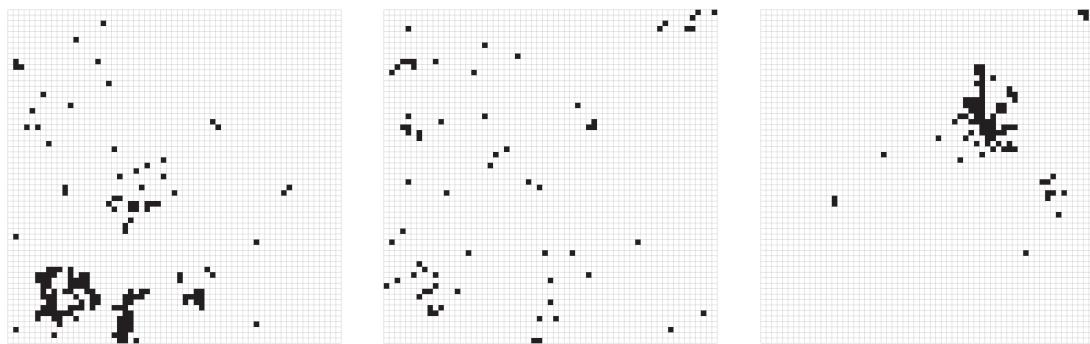
(a) Normalize the elevation of unfiltered point cloud.



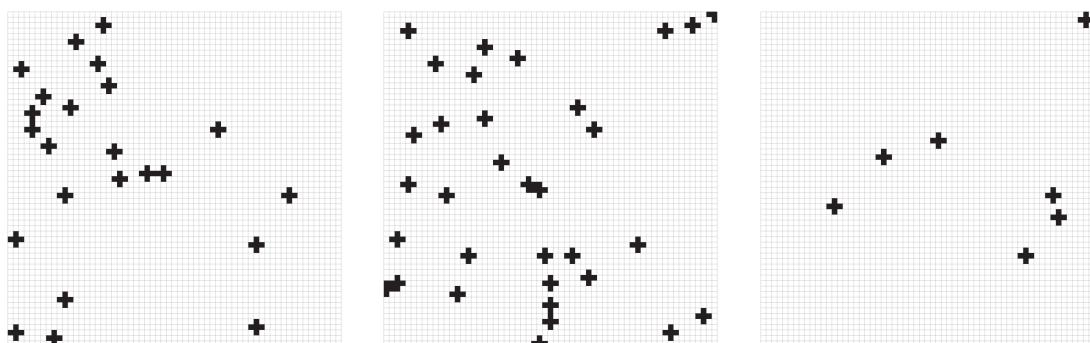
(b) Filter the point cloud (only last returns, leaf-off, vegetation class, min/max height)



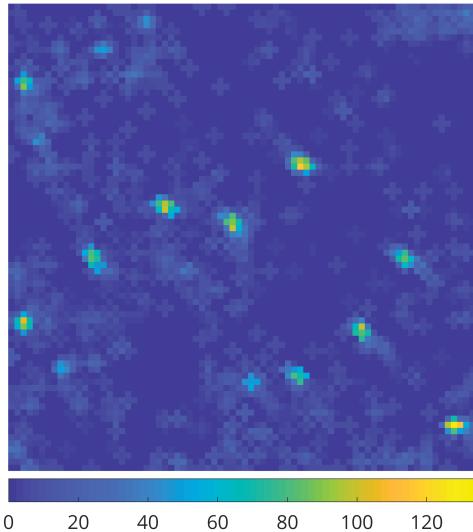
(c) Sample points that are located within vertical distance  $w/2$  above or below the cross-section plane.



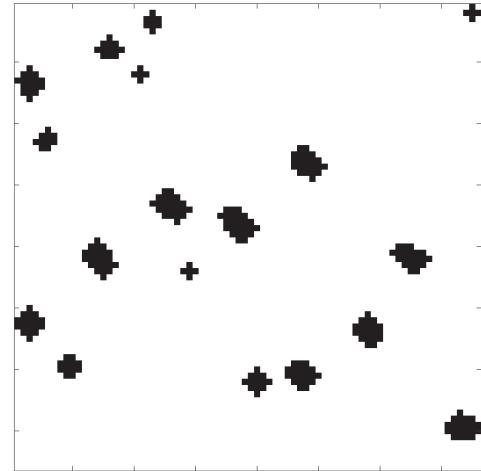
(d) Rasterize the layer points to create a boolean image.



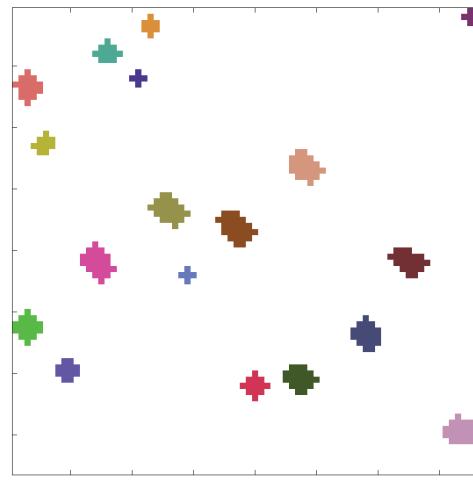
(e) Apply morphological operations to detect and filter region centroids.



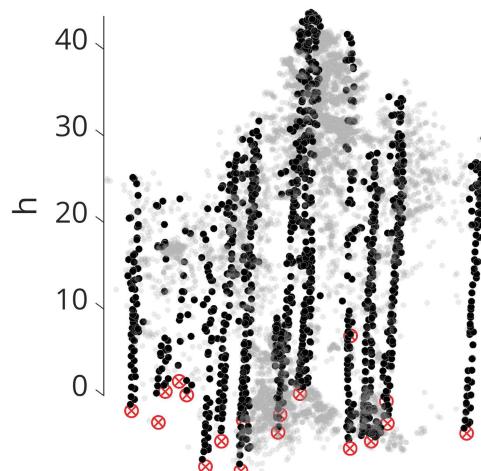
(f) Centroid density image.



(g) Boolean mask obtained by thresholding the centroid density image.



(h) Connected components detection in the boolean mask.



(i) Detected stem positions in the unfiltered 3D point cloud

Figure 3.25: Main steps in individual tree stem detection with layered morphological analysis.

### 3.5.2 Results

Since a fully labeled point cloud is required for validation, simulated forest plots were created by selecting only labeled trees in several sites of the reference (manually delineated) dataset presented in chapter 2.2. In other words, the simulated plots were identical to the real plots, but they contained only labeled high vegetation (manually delineated trees), unlabeled low vegetation (< 1 m) and terrain points.

A full factorial parameter sensitivity analysis was conducted on the Sauvabelin, Versoix, Boudry D20, Boudry D1, Rochefort and Sihlwald sites, to evaluate the effect of raster resolution ( $d_{xy}$ ), vertical step ( $dz$ ), cross-section thickness ( $w$ ) and centroid density threshold ( $\rho$ ) on the detection score. These sites were selected because they have different topographic and/or forest characteristics. The effects of varying maximum object area, circularity threshold and structuring element radius were not evaluated. Four values (levels) were tested for each parameter, thus 256 (=  $4^4$ ) runs of the algorithm were necessary for each site to evaluate all possible combinations (cf. table 3.12). Parameter values that consistently produced the best detection scores (i.e.  $\geq 90\%$  F-score quantile) were subsequently used as a nominal baseline in a one-at-time sensitivity analysis (cf. figure 3.26). The latter analysis was conducted on a smaller subset on the study sites (Versoix, Rochefort, Sihlwald).

The proposed method was validated using the nominal parameters suggested in section 3.5.1 and compared to local maxima detection. The local maxima detection was applied to 0.5 m resolution raster Canopy Height Models (CHM) derived from the 3D point clouds and smoothed using a Gaussian 3x3 lowpass filter. It used a height ( $h$ ) dependent search radius ( $r$ ) defined by:

$$r(h) = 1 + 0.25 \cdot \log(\max_h(h, 1)) \quad (3.18)$$

The comparison of the two methods is reported in table 3.13 and a qualitative example of results obtained with the proposed method is shown in figure 3.27.

Table 3.12: Optimal parameter values and corresponding detection scores obtained using a full factorial setup. The mean and standard deviation of parameter values that produced a detection F-score above the 90% quantile of all F-scores are reported. Correct detection criteria:  $\epsilon_{xy} \leq 2$  m.

Site	Obs.	Parameters				Scores		
		$d_{xy}$ [m]	$w$ [m]	$dz$ [m]	$\rho$ [m]	$p$	$r$	$F$
Versoix	372	0.2 ± 0	1.2 ± 0.6	0.6 ± 0.3	5.5 ± 0.6	0.89 ± 0.06	0.74 ± 0.02	0.81 ± 0.03
Sauvabelin	171	0.5 ± 0.3	1.1 ± 0.5	0.6 ± 0.3	5.8 ± 0.5	0.91 ± 0.05	0.64 ± 0.02	0.75 ± 0.01
Boudry D20	312	0.6 ± 0.2	1.2 ± 0.5	0.5 ± 0.3	5.3 ± 0.9	0.98 ± 0.02	0.86 ± 0.03	0.92 ± 0.01
Boudry D1	535	0.4 ± 0.2	0.8 ± 0.3	0.6 ± 0.3	4.9 ± 1	0.84 ± 0.08	0.58 ± 0.05	0.68 ± 0.01
Rochefort	377	0.7 ± 0.1	1.4 ± 0.5	0.6 ± 0.3	5.4 ± 0.7	0.91 ± 0.04	0.74 ± 0.02	0.81 ± 0.01
Sihlwald	1416	0.7 ± 0.1	1.1 ± 0.5	0.5 ± 0.3	5.6 ± 0.3	0.91 ± 0.04	0.81 ± 0.02	0.86 ± 0.01

Table 3.13: Comparison of stem detection results using the proposed morphological analysis method and local maxima with variable search window size. The following nominal parameter values were used for the proposed method: ( $d_{xy} = 0.4$  m,  $w = 1$  m,  $d_z = 0.5$  m,  $\rho = 5$  m). Correct detection criteria:  $\varepsilon_{xy} \leq 2$  m. Scores were rounded to the nearest second decimal. *Obs*: number of observations, *p*: precision, *r*: recall, *F*: F-score,  $\varepsilon_{xy}$ : mean and standard deviation of stem position error. The best performing method is indicated with an asterisk.

Site	Obs.	$f_A$	Method	<i>p</i>	<i>r</i>	<i>F</i>	$\varepsilon_{xy}[m]$
Versoix	372	$0.54 \pm 0.22$	Morpho. Analysis*	0.97	0.59	0.73	$0.32 \pm 0.4$
			Local maxima	0.52	0.55	0.55	$1.17 \pm 1$
Sauvabelin	171	$0.35 \pm 0.2$	Morpho. Analysis*	0.97	0.68	0.8	$0.42 \pm 0.28$
			Local maxima	0.17	0.23	0.2	$1.4 \pm 0.86$
Benenté	699	$0.38 \pm 0.26$	Morpho. Analysis*	0.9	0.36	0.51	$0.49 \pm 0.48$
			Local maxima	0.56	0.45	0.5	$0.98 \pm 1.22$
La Brévine	874	$0.4 \pm 0.27$	Morpho. Analysis	0.74	0.37	0.49	$0.82 \pm 0.52$
			Local maxima*	0.98	0.66	0.79	$0.36 \pm 0.65$
Couvet	256	$0.38 \pm 0.25$	Morpho. Analysis	0.71	0.5	0.59	$0.7 \pm 0.53$
			Local maxima*	0.96	0.65	0.78	$0.51 \pm 0.87$
Boudry D20	312	$0.25 \pm 0.2$	Morpho. Analysis*	0.98	0.9	0.94	$0.29 \pm 0.26$
			Local maxima	0.62	0.54	0.58	$1.09 \pm 0.96$
Boudry D19	320	$0.27 \pm 0.26$	Morpho. Analysis*	0.96	0.76	0.85	$0.34 \pm 0.3$
			Local maxima	0.54	0.49	0.51	$1.06 \pm 1.07$
Boudry D1	535	$0.47 \pm 0.25$	Morpho. Analysis*	0.92	0.55	0.69	$0.49 \pm 0.44$
			Local maxima	0.46	0.49	0.47	$1.03 \pm 1.08$
Chambrelien	224	$0.14 \pm 0.17$	Morpho. Analysis*	0.98	0.93	0.95	$0.23 \pm 0.23$
			Local maxima	0.51	0.58	0.54	$1.04 \pm 1.06$
Rochefort	287	$0.25 \pm 0.22$	Morpho. Analysis*	0.93	0.76	0.84	$0.42 \pm 0.34$
			Local maxima	0.38	0.44	0.41	$1.1 \pm 1.08$
Ottmarsingen	127	$0.25 \pm 0.17$	Morpho. Analysis*	0.97	0.87	0.92	$0.35 \pm 0.35$
			Local maxima	0.11	0.24	0.15	$1.32 \pm 1.04$
Sihlwald	1416	$0.39 \pm 0.22$	Morpho. Analysis*	0.96	0.73	0.83	$0.2 \pm 0.28$
			Local maxima	0.37	0.68	0.48	$1.14 \pm 1.02$
Oberaegeri	308	$0.48 \pm 0.24$	Morpho. Analysis	0.68	0.4	0.5	$0.61 \pm 0.61$
			Local maxima*	0.97	0.76	0.85	$0.59 \pm 0.8$

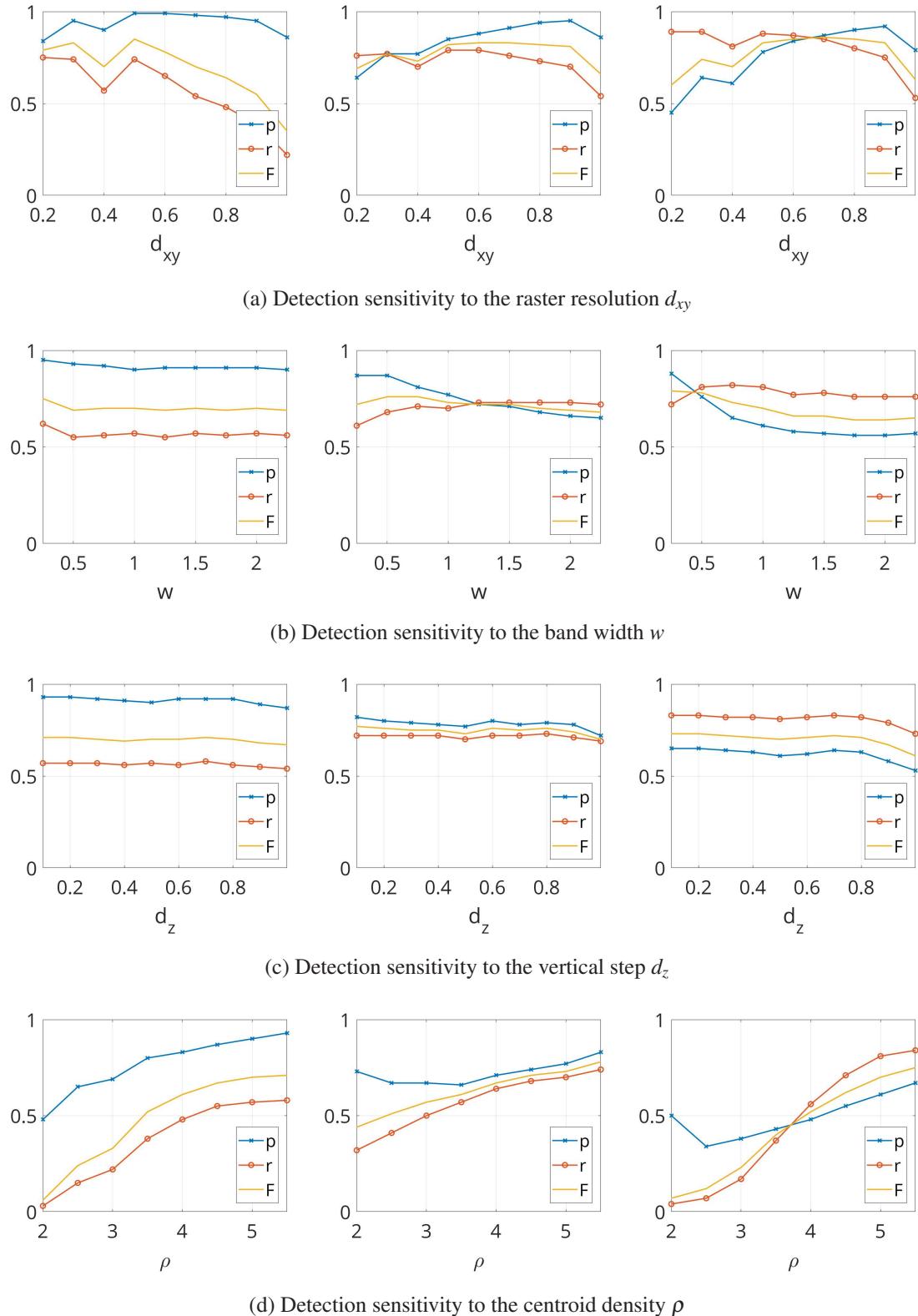
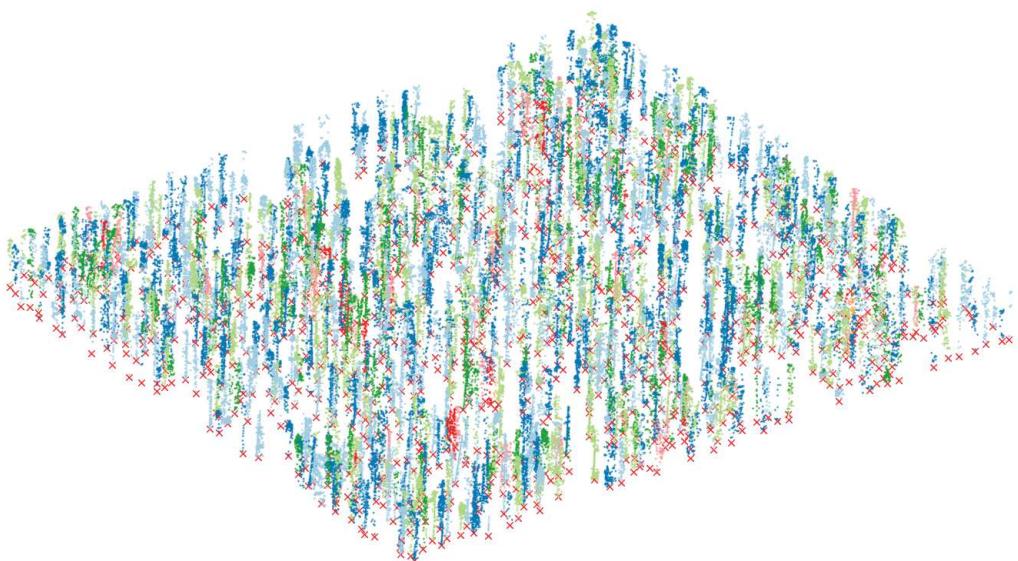


Figure 3.26: One-at-a-time parameter sensitivity for the Versoix (left), Rochefort (center) and Sihlwald (right) sites.



(a) The point cloud with elevation normalized relative to the terrain.



(b) Labeled stems.

Figure 3.27: Individual stems detected with the layered morphological analysis on the Boudry D20 site (raw data). Topological coloring is used to distinguish adjacent stems.

### 3.5.3 Discussion

The results show that the method is generally precise (low false positive detection rate), but may fail to detect a significant fraction of the stems. The best performances are observed on sites with many mature trees (high forests) and few understory trees (i.e. Sauvabelin, Boudry D20, Boudry D19, Chambrelien, Rochefort, Ottmarsingen, Sihlwald). The method also outperforms local maxima detection on all sites except La Brévine, Couvet and Oberaegeri. These three sites are dominated by coniferous trees with stems not apparent in the point cloud and thus particularly well suited for local maxima detection and poorly suited for stem detection.

The full factorial analysis shows that tuning the raster resolution  $d_{xy}$  may improve the detection results locally. This could be explained by the different stem densities and slopes on the tested sites. More specifically, finer raster resolutions may improve results when stem density is high and coarser resolution may improve the results when the slope is high (e.g. Rochefort). The one-at-time sensitivity analysis shows that the most critical parameters are the raster resolution  $d_{xy}$  and the centroid density  $\rho$ . Since the bandwidth  $w$  and vertical step  $d_z$  parameters do not influence the results very much, it is advantageous to set them to high values to decrease the processing time.

Since the size range of the detected objects in the cross-sections can be modulated with the  $A_{max}$  parameter, the algorithm is able to detect not only stems, but also the medial axis of trees that do not have a visible stem (e.g. young coniferous trees). This also means that the method can also be used to filter elements of a target size within a point cloud. Such filtering could be useful as a preprocessing step in other segmentation methods (e.g. geodesic vote).

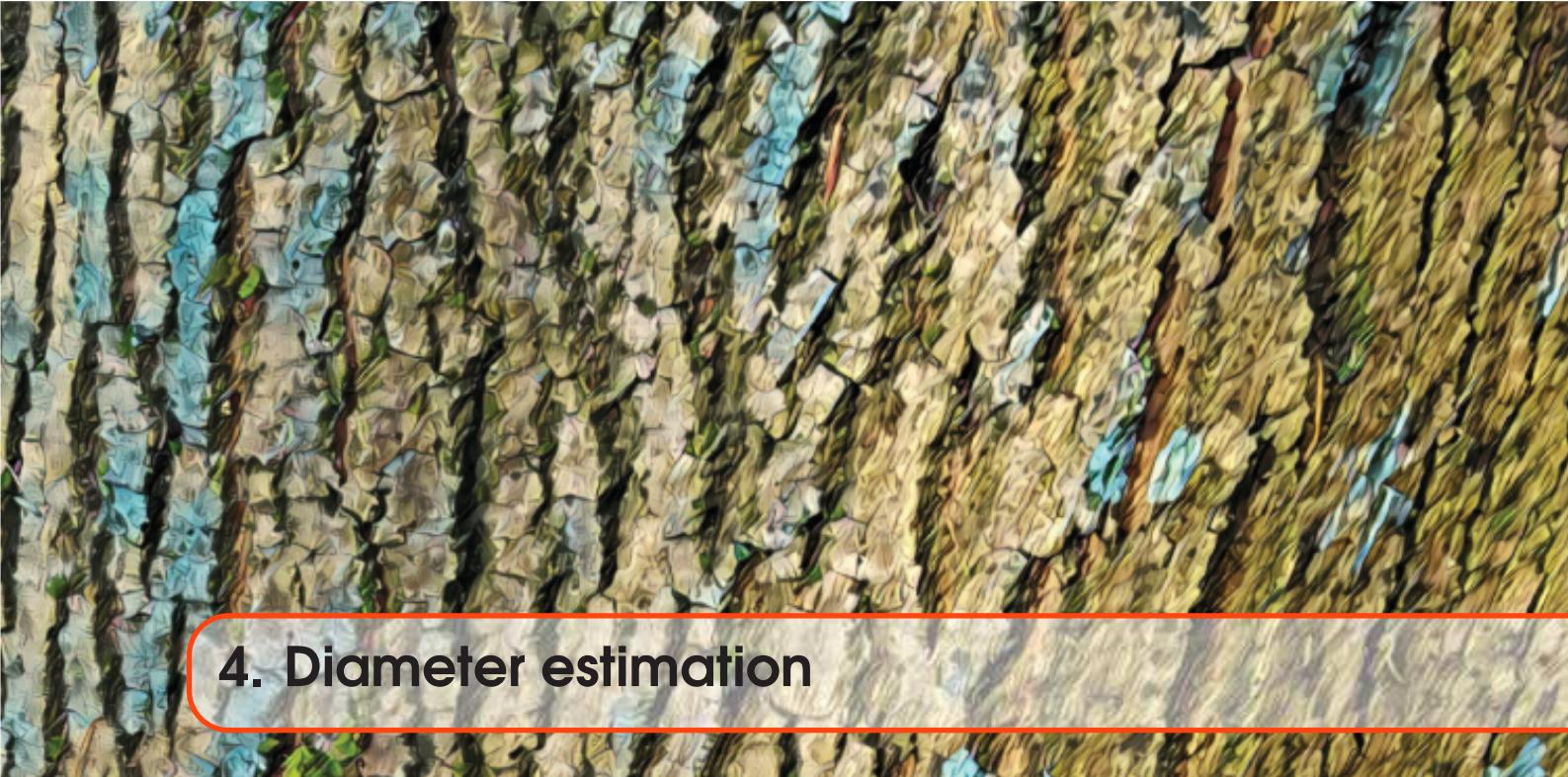
The main advantages of this method are its simplicity, its independence of point density (since the point cloud is rasterized), fast running times, low commission error rates and small positioning error. Its main drawback are its dependency on leaf-off ALS acquisitions, its inability to detect strongly inclined and/or poorly sampled stems.

Some possible improvements include: adding preliminary filtering steps (e.g. threshold on the echo width), adding a post-processing step to identify faulty segments (e.g. fitting a linear/quadratic model with robust least squares or RanSAC and checking the fit error).

### 3.6 Synthesis

In this chapter, a state of the art on individual tree and stem segmentation was first conducted. It revealed several shortage areas which require improvements or which have not been investigated extensively: the segmentation of trees/stems in general and in multi-layered deciduous and mixed forests in particular, the rigorous validation of the 3D shape of tree segments (absent from all the studies identified by the author), the standardization of problem difficulty and error quantification/reporting. Based on these premises, the following contributions were made:

- A rigorous error assessment framework to validate the 3D shape of tree segments.
- A simple procedure to objectively quantify the difficulty of tree segmentation problems in point clouds.
- A novel individual tree segmentation algorithm called geodesic voting based on a graph representation of the ALS point cloud. The 3D shape of the individual tree segments produced by the algorithm was validated by comparing it to manual delineation. This validation in simulated pure deciduous forest showed that the algorithm had a generally good stem detection rate (F-score between 67% and 95%) and a reliable delineation accuracy for 36% to 81% depending on test sites. The algorithm also systematically outperformed the commonly used marker controlled watershed segmentation algorithm. In the presence of coniferous trees in the dataset, the detection rate decreased significantly leading to the conclusion that a preliminary separation of deciduous and coniferous trees could be beneficial for segmentation when using this algorithm.
- A novel method to estimate segmentation uncertainty, produce tree shape templates and improve individual tree segmentation based on the ensemble learning framework. Essentially, the method combines many error-containing tree segments to estimate the true underlying tree shape. The method was validated on the Brévine study site and it was shown that it can be used to compensate for undersegmentation, in particular for tree with radial symmetry.
- A novel method to detect individual tree stems. This simple method based on the morphological analysis of point cloud horizontal cross-sections was tested on different study sites. The validation showed that it was able to achieve a detection F-score between 49% and 94% and systematically outperformed local maxima detection in complex forest settings.



## 4. Diameter estimation

This chapter covers the topic of stem diameter estimation from ALS data. It is structured in the following way:

- 4.1 introduces the topic and describes the state of the art.
- 4.2 describes the error assessment framework used to evaluate the performance of stem diameter estimation methods.
- 4.3 presents a replication study of the method proposed in Jucker et al. (2016) on the reference dataset described in chapter 2.
- 4.4 presents and evaluates a novel method to directly measure stem taper in ALS point clouds.

### 4.1 State of the art

Stem Diameter at Breast Height (DBH) is one of the most important and commonly used metrics in forestry. When combined with other characteristics (e.g. tree species, height, taper, crown diameter), it can be used to infer important variables such as wood volume and biomass. Moreover, the proven link between the increase in atmospheric carbon dioxide and global climate change has been a strong incentive to investigate the carbon cycle and to accurately quantify the carbon storage capacity of forests (Patenaude et al., 2005; Pan et al., 2011; Saatchi et al., 2011; Reich, 2011; Baccini et al., 2012; Calders et al., 2015).

Many approaches have been proposed to estimate DBH with remote sensing. For this purpose, the ability to accurately measure 3D forest geometry is critical and essentially limits the applicable techniques to photogrammetry, Synthetic Aperture Radar (SAR) and LiDAR. The latter option is currently the most suitable (St-Onge et al., 2008; Sexton et al., 2009; White et al., 2013b; Vastaranta et al., 2013). Approaches to estimate tree diameter from structural remote sensing data can be divided in two general categories:

- The **indirect estimation** of diameter from auxiliary geometrical variables at the tree scale (e.g.

height, crown diameter, crown volume) or plot scale (e.g. crown coverage, height percentiles) and environmental variables (e.g. altitude, resource availability). This is currently by far the most frequently used approach. Its disadvantage is that it requires field surveys to collect DBH measurements for calibration.

- The **direct measurement** of diameter based on a partial or full geometric modeling of the tree branching structure from a point cloud. This approach does not require field surveys (except for validation).

For the indirect approach, it is necessary to determine how the stem diameter changes in proportion to other dimensions (e.g. height, crown width) with growth and environmental factors. The concept of relative growth of different tree parts/dimensions was probably known for a very long time by foresters and its use in silviculture in the form of yield tables has been documented from the 18<sup>th</sup> century (Pretzsch, 2009). Through empirically investigation on the productivity of experimental forest plots, foresters gradually developed predictive growth and yield models that could be integrated in planning. However, the underlying physiological mechanisms and properties of these scaling models did not become a scientific paradigm before the beginning of the 20<sup>th</sup> century, when the topic was formalized under the name **allometry**. At that time, it was proposed (Dubois, 1897; Lapicque, 1907; Thompson, 1917; Huxley et al., 1932; Kleiber, 1932) that many biological scaling relations can be modeled mathematically with a power law of the form:

$$y = b \cdot x^k \quad (4.1)$$

Or its equivalent logarithmic form:

$$\log(y) = k \cdot \log(x) + \log(b) \quad (4.2)$$

where:

- $y$  is the dependant variable;
- $x$  is the independant variable;
- $b$  is a proportionality coefficient;
- $k$  is an allometric exponent.

Subsequently, allometry became the subject of reinterpretation and debate among biologists (Gould, 1966; Niklas, 1994; Gayon, 2000). In this context, the study of plant allometry gave rise to several theories attempting to explain structural and functional growth mechanisms. Shinozaki et al. (1964) suggested that a tree could be modeled as a bundle of pipes (linking the stem base to the leafs) acting as both mechanical support and conductive vessels. Their theory provided some insights into possible causes of tree shape and linked leaf size to sap wood area. It has been extended, reinterpreted and contested since its proposal (Lehnebach et al., 2018). Halle et al. (1978) developed 23 models describing the growth patterns and shape (which they called architecture) of all known tree species. More recently, in an attempt to explain the evolutionary origins of allometric scaling, West et al. (1997) argued that it is a consequence of the structural fractal self-similarity and space filling characteristics of vascular organisms. The authors suggested that there exists a universal allometric exponent  $k = 3/4$  which governs processes across functional scales from individual cells to ecosystems (West et al., 1999b; Enquist et al., 1999; Enquist and Niklas, 2001; Enquist, 2002). One implication of this theory is that allometric scaling is not related to environmental conditions. However, this theory has repeatedly been shown to be inconsistent with empirically evidence and not generalizable as originally claimed (Agutter and Wheatley, 2004; Kozłowski and Konarzewski, 2004; Li et al., 2005; Kozłowski and Konarzewski, 2005; Muller-Landau et al., 2006; Coomes, 2006; Pretzsch, 2006; Russo et al., 2007; Coomes et al., 2011; Lines et al., 2012; Rüger and Condit, 2012; Bentley et al., 2013; Muller-Landau et al., 2016).

Despite the large corpus of research on this topic, there is currently no consensus theory that is able to universally and accurately explain the structural development and allometric scaling of trees. For practical purposes, the most commonly used approaches in forestry remain simple allometric functions obtained by regression analysis of field and/or remote sensing measurements. A large number of such models have been proposed for different tree species and/or regional environmental conditions (Picard et al., 2012). In 2013, the Food and Agriculture Organization (FAO) and the French Agricultural Research Center for International Development (CIRAD) created GlobAllomeTree (Henry et al., 2013); an online repository of global tree allometric models.

Even though they are a practical solution, allometric models should be used with caution as they may be subject to large uncertainties due to sampling bias and error in the field or remote sensing data used to calibrate them. In this regard, Duncanson et al. (2015) showed that allometric models based on small sample sizes were biased in temperate forests. Based on this finding, Jucker et al. (2016) compiled a global database of field and remote sensing tree observations which was subsequently used to create a generic allometric model (using height and crown diameter as predictors). The parameters of their model can be adapted to different forest types and biogeographic regions to produce unbiased diameter and biomass estimations. The results obtained by Jucker et al. (2016) were later confirmed in a replication study by Dalponte et al. (2018b).

Indirect diameter estimation from ALS data has been investigated extensively, using both individual tree based (e.g. Chen et al. (2007); Peuhkurinen et al. (2007); Vauhkonen et al. (2010); Yu et al. (2011); Dalponte et al. (2011); Allouis et al. (2013); Lo and Lin (2013); Bucksch et al. (2014); Gonzalez-Benecke et al. (2014); Duncanson et al. (2015); Paris and Bruzzone (2016)) and area (plot) based (e.g. Means et al. (2000); Næsset (2002); Holmgren et al. (2003b); Gobakken and Næsset (2004); Popescu and Wynne (2004); Næsset (2004); Næsset et al. (2005); Gobakken and Næsset (2005); Thomas et al. (2006); Maltamo et al. (2006); Jensen et al. (2006); Hollaus et al. (2007); Mehtätalo et al. (2007); Anderson et al. (2008); Thomas et al. (2008); Straub et al. (2009); Maltamo et al. (2009); Hollaus et al. (2009a); Lindberg et al. (2010); Ioki et al. (2010); Monnet et al. (2011); Magnussen et al. (2012); Luther et al. (2014); Bouvier et al. (2015); Kankare et al. (2015); Wu et al. (2015)). Plot based approaches are an alternative to individual tree allometry. With this type of approach, the diameter distribution or total basal area of the forest plot is estimated by multiple regression on area descriptors (e.g. canopy coverage, ALS point height quantiles, proportion of deciduous and coniferous trees). Yu et al. (2010); Peuhkurinen et al. (2011) conducted a comparison of area based and tree based approaches to estimate forest plot attributes and found that both produced similar results. Lindberg and Hollaus (2012) on the other hand, reported that better results were obtained with area based approaches. Finally, Yu et al. (2010) suggested that increased ALS point density would likely improve estimations obtained with tree based approaches.

The influence of different regression methods used to fit diameter prediction models to tree/plot scale observations has also been evaluated. Dalponte et al. (2011) compared the performance of simple linear regression, linear and non linear Support Vector Regression (SVR) and found that all three methods produced similar results for area based approaches. Monnet et al. (2011) reached a similar conclusion, when comparing the performances of linear regression and SVR for predicting forest characteristics at the plot scale. They suggested that more complex regression methods like SVR could be more robust when dealing with a small number of training observations. Comparing the results of different studies is difficult, due to diversity of environments covered, prediction methods and error assessment procedures. The fact that many authors only report absolute errors also hinders comparison. Most studies report diameter prediction errors in the 5-40% range. However, it remains unclear how generalizable these results are and under which conditions (forest type, total basal area, stem count, topography, ALS point density, etc) they can be considered sufficiently reliable for operational forestry use.

The direct measurement of stem diameter from 3D point clouds is a much more recent development than allometric modelling. It is based on a partial or full 3D geometric reconstruction of the tree branching structure. This structural modeling typically involves fitting geometric primitives (e.g. circles, cylinders, ellipses) to subcomponents (stem, branches) of the tree or meshing the boundary points (e.g. with alpha shapes or splines). The most commonly used approach is circle fitting with either (robust) least squares regression, Random Sample Consensus (RanSaC) or Hough transform (cf. tables 4.1 and 4.2). The effect of different fitting methods has been investigated and most of them produce similar results (Pueschel et al., 2013; Koreň et al., 2017; Liu et al., 2018; Liang et al., 2018). Because most of the tree shape has to be sampled with points, direct stem measurement is only applicable to high density point clouds ( $> 100 \text{ points} / \text{m}^2$ ) preferably acquired in leaf-off conditions. Thus, it has almost exclusively been investigated over small areas using data from static or mobile Terrestrial Laser Scanning (TLS) which is typically orders of magnitude denser than ALS data. Most of the studies conducted with TLS report that the stem diameter measured in the point cloud is very close or equivalent to the reference field measurement. Thus, TLS can be considered sufficiently reliable for operational stem geometry measurement (Liang et al., 2018). As discussed in chapter 3.2, the environment in which methods are tested plays a leading role in determining their performance and general applicability. In this regard, environments with large spacing between stems (e.g. urban, even-aged high forests) and trees with linear stems are much easier to process; methods tested in these environments may not be transferable to more complex forest settings.

Recently, the combined commercial availability of Unmanned Aerial Vehicles (UAV) and small lightweight LiDAR sensors (e.g. Riegl VUX-1, Velodyne HDL-32E and VLP-16) has allowed the development of new intermediate range ALS systems. Data produced by these systems has point densities similar to TLS over much larger areas (several tens of hectares) and has been successfully used for direct measurement of stem diameter (Chisholm et al., 2013; Jaakkola et al., 2017; Brede et al., 2017; Wieser et al., 2017). Very little attention has been given to direct diameter measurement from the much more commonly available long range ALS data. One reason for this is the low availability of ALS data with sufficiently high point density below the canopy. Another reason is the prevalence of leaf-on ALS acquisitions which often hinders analysis of the stem geometry. Only two published studies could be found on this topic. The first study by Bucksch et al. (2014) uses as skeletonization algorithm to determine the medial axis (center) of the stem and subsequently measures its radius by computing the distance between the ALS points and the medial axis. They evaluate their method on 18 simulated trees ( $75 \text{ points} / \text{m}^2$ ) and 34 real trees manually extracted from point clouds produced with the FLI-MAP 400 instrument. They report a RMSE = 5 cm on the simulated dataset and RMSE = 11.6 cm on the real one. They also indicate that the mean stem diameter of the real trees is 40 cm, which means they obtain a relative error around 25%. More recently, Harikumar et al. (2017a) proposed a method to find the anchor points (i.e. where branches meet the stem) of coniferous branches and subsequently estimate stem diameter based on a non-linear least squares fitting of a 3D cone to the void central region delimited by the branch anchors. They tested their approach on 100 trees extracted from a high density ALS point cloud ( $50\text{-}200 \text{ points} / \text{m}^2$ ). They reported a MSE = 32.8 cm and a MAE = 4.98 cm, which based on the interpretation of their graphs, translates to a relative error around 10% (the authors do not explicitly present the diameter distribution of the studied trees, nor the relative error metrics). Interestingly, the method proposed by Harikumar et al. (2017a) does not rely on ALS points along the stem beneath the crown height (points below 50% of the tree height are removed). However, it is only applicable to coniferous trees with fully linear stems.

With the steady increase in point densities observed in ALS surveys, it can be expected that direct diameter measurement will be more generally applicable and could complement or replace allometric estimations in the near future.

Table 4.1: Selected publications on tree geometry modeling and direct stem diameter measurement from ALS data.

Algorithm	Parts	Reference
Single circle fitting (least squares)	Partial stem	Chisholm et al. (2013)
Distance to structural skeleton	Partial stem	Bucksch et al. (2014)
Multiple circle fitting (least squares)	Partial stem	Jaakkola et al. (2017)
Single circle fitting (least squares)	Partial stem	Brede et al. (2017)
Single cylinder fitting (least squares)	Partial stem	Wieser et al. (2017)
Single cone fitting (non-linear least squares)	Full stem (coniferous only)	Harikumar et al. (2017a)

Table 4.2: Selected publications on tree geometry modeling and direct stem diameter measurement from TLS data.

Algorithm	Parts	Reference
Circle fitting (Hough transform)	Partial stem	Simonse et al. (2003)
Single cylinder fitting (least squares)	Partial stem	Hopkinson et al. (2004)
Multiple cylinder fitting (least squares), B-Spline fitting	Full stem, main branches	Pfeifer et al. (2004)
Multiple cylinder fitting (least squares)	Full stem	Thies et al. (2004)
Multiple circle fitting (least squares)	Full stem	Henning and Radtke (2006)
Multiple circle fitting (least squares)	Full stem	Bienert et al. (2007)
Single circle/cylinder fitting (least squares)	Partial stem	Brolly and Király (2009)
Single cylinder fitting (least squares)	Partial stem	Moskal and Zheng (2011)
Multiple cylinder fitting (least squares)	Full tree	Dassot et al. (2012)
Voxel-based cross-section morphological analysis	Full stem, main branches	Vonderach et al. (2012)
Multiple cylinder fitting (robust regression, Tukey)	Full stem	Liang et al. (2012)
Single cylinder fitting (RanSaC)	Partial stem	Fritz et al. (2013)
Multiple cylinder fitting (least squares)	Full tree	Raumonen et al. (2013)
Single circle fitting (RanSaC)	Partial stem	Olofsson et al. (2014)
Multiple Cylinder fitting (non-linear least squares)	Full tree	Hackenberg et al. (2014)
Multiple cylinder fitting (RanSaC)	Full stem	Wang et al. (2016a)
Multiple cylinder fitting (least squares)	Full tree	Chen et al. (2018)

## 4.2 Error assessment framework

The regression error indicates how well a model is able to predict a continuous variable (e.g. diameter, biomass) from descriptive variables. Some of the most commonly used regression performance metrics (Congalton and Green, 2008) are provided in table 4.3.

No single metric can be used to unambiguously characterize model performance. Average model performance metrics such as the widely used RMSE should be accompanied by complementary dimensioned metrics such as the Mean absolute Error (MAE) and Bias and dimensionless metrics such as the relative bias and correlation coefficient. The reporting of relative (scaled, dimensionless) error metrics is particularly important because it allows comparison of results from different studies and datasets. These metrics should also be accompanied by correlation and bias graphs to help interpret performances

Table 4.3: Regression performance metrics.  $x_{p,i}$  is the ith predicted value,  $x_{r,i}$  is the ith reference value,  $\sigma_p$  is the standard deviation of the predicted values,  $\sigma_r$  is the standard deviation of the reference values

Metric	Formula	Interpretation
Root-Mean-Square Error	$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{p,i} - x_{r,i})^2}$	Unsigned quadratically weighted error (dimensioned)
Mean Absolute Error	$\text{MAE} = \frac{1}{N} \sum_{i=1}^N  x_{p,i} - x_{r,i} $	Unsigned error (dimensioned)
Bias	$\text{Bias} = \frac{1}{N} \sum_{i=1}^N x_{p,i} - x_{r,i}$	Signed error (dimensioned)
Relative bias	$\text{Rel. Bias} = \frac{1}{N} \sum_{i=1}^N \frac{x_{p,i} - x_{r,i}}{x_{r,i}}$	Signed error (dimensionless)
Correlation coefficient	$r = \frac{\text{cov}(x_r, x_p)}{\sigma_r \sigma_p} \in [-1, 1]$	Signed strength of linear relationship between the reference and predicted values
Regression line slope	$a = r \frac{\sigma_p}{\sigma_r} \in [0, 1]$	A slope close to 1 means a low bias

## 4.3 Indirect diameter estimation

In this section, observations from all the study sites are used to estimate the diameter at breast height with a general allometric model, using the approach proposed by Duncanson et al. (2015) and Jucker et al. (2016). This approach is chosen because it has been shown to work well across a wide range of forest types.

### 4.3.1 Description

The approach models the relation between the DBH and the area of the 2D longitudinal bounding box (i.e. height multiplied by crown diameter) of a tree. To compensate the non-uniform diameter distribution in the sample, observations are grouped into 50 logarithmic categories (diameter bins) and variables are averaged within each category. A log transformation is then applied to the averaged variables, so that a linear model can be adjusted (cf. equation 4.3):

$$\log(\overline{DBH}) = \alpha + \beta \cdot \log(\overline{H \cdot CD}) + \varepsilon \quad (4.3)$$

The diameter is then predicted with equation 4.4:

$$DBH = e^{\alpha + \beta \cdot \log(H \cdot CD)} \cdot e^{\frac{\sigma^2}{2}} \quad (4.4)$$

where:

- $\alpha$  is the intercept parameter of the linear model;
- $\beta$  is the slope parameter of the linear model;
- $\varepsilon$  is the regression error;
- $DBH$  is the diameter at breast height [cm];
- $H$  is the tree height [m];
- $CD$  is the crown diameter [m] estimated from the area  $A_c$  of the crown's 2D single region concave hull with ( $CD = 2\sqrt{\frac{A_c}{\pi}}$ );
- $\sigma^2$  is the estimate of the regression error variance (regression mean square error). This term replaces the  $\varepsilon$  found in the log form.

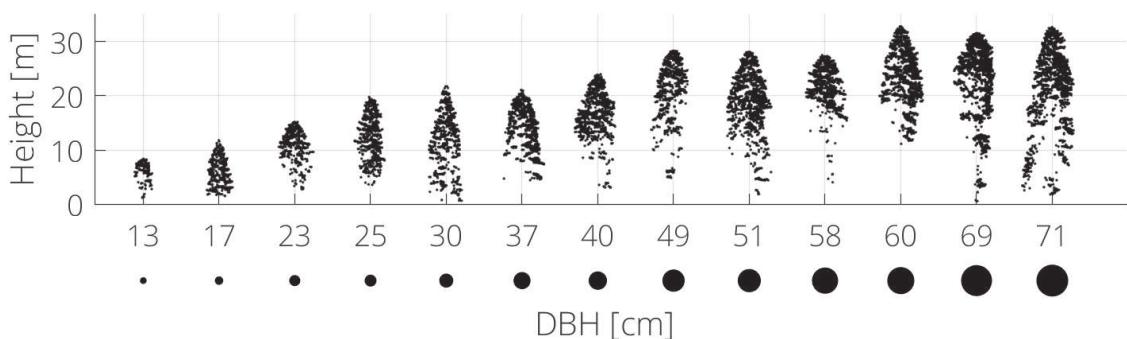


Figure 4.1: Sequence of Silver fir (*Abies alba*) tree profiles illustrating the non-linear relation between DBH and height. ALS data is from the Brévine study site.

When considering different sub-groupings (e.g. biome, taxonomy, diameter, etc), it is often the case that frequency distributions of observations are imbalanced. In other words, some groups may be over or under represented in different diameter categories. Figure 4.2 illustrates this problem, using observations from the reference dataset presented in chapter 2. If the group size imbalance is not taken into account when partitioning observations into training and test sets, the model parameters and the error metrics risk being biased in favor of the most represented group.

To compensate for this undesired effect when calibrating composite (i.e. non-group specific) models, the variables of equation 4.3 are computed separately for each group and are subsequently averaged. Thus, equation 4.3 can be reformulated into:

$$\frac{\sum_{i=1}^N \log(\overline{DBH}_i)}{N} = \alpha + \beta \cdot \frac{\sum_{i=1}^N \log(\overline{H_i \cdot CD_i})}{N} + \varepsilon \quad (4.5)$$

where:

$DBH_i$  is the diameter at breast height [cm] of observations in group  $i$ ;

$H_i$  is the tree height [m] of observations in group  $i$ ;

$CD_i$  is the crown diameter [m] of observations in group  $i$ ;

$N$  is the number of groups considered in the composite model;

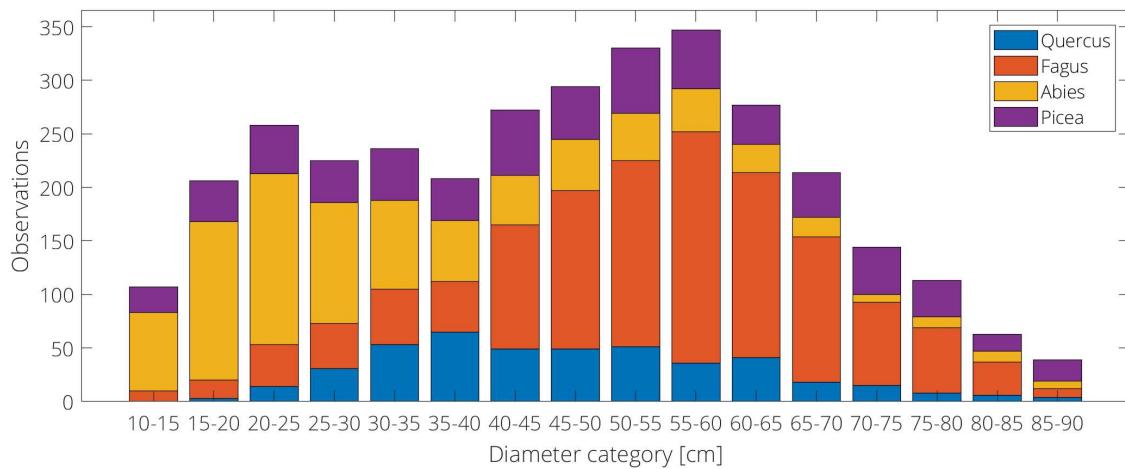


Figure 4.2: Diameter distributions for the four most frequent genus in the benchmark dataset. Note the imbalanced frequency distributions in terms of diameter categories as a function of genus.

### 4.3.2 Results

Only oak (*Quercus* sp.), beech (*Fagus* sp.), fir (*Abies* sp.) and spruce (*Picea* sp.) trees with a DBH ranging from 10 to 90 cm were considered in the analysis, because they are the most frequent in the reference dataset (cf. chapter 2).

Monte-Carlo cross-validation with 500 runs was used to determine the parameter ( $\alpha$ ,  $\beta$ ,  $\sigma$ ) values of equation 4.5 and to evaluate the prediction errors. In this procedure, at each run, a diameter stratified fraction (2/3) of the observations is randomly selected to train the model and the rest of the observations (1/3) are used for validation (error assessment). Since the partition is random, the same observations can occur multiple times in different runs. To compensate the effect of imbalanced diameter group sizes (cf. figure 4.2), average error metrics were computed for each diameter group, at each run. Finally, the parameter values and error metrics of the 500 runs were subsequently summarized in terms of mean and spread. This analysis was conducted for three different taxonomic ranks:

- **Rank 1 (R1) - composite model considering all genus**, cf. table 4.4 and figure 4.3
- **Rank 2 (R2) - composite models for each division (angiosperms, gymnosperms)**, cf. table 4.5 and figure 4.4
- **Rank 3 (R3) - specific models for each genus (*Quercus*, *Fagus*, *Abies*, *Picea*)**, cf. tables 4.6-4.7 and figures 4.5-4.6

The performances of these allometric models were subsequently compared, to evaluate the effect of including taxonomic information on the DBH prediction (cf. table 4.8). Using the mean parameter values obtained with Monte-Carlo cross-validation, diameter predictions were obtained for all observations. Then, a two-sided Wilcoxon signed-rank test was used to determine if the difference between the R1 or R2 and R3 diameter predictions has a zero median (i.e. the null hypothesis). Using a 0.1% (i.e.  $\alpha = 0.001$ ) significance level, the null hypothesis was rejected for all the predictions (with p-values systematically smaller than  $10^{-6}$ ). This means that the inclusion of taxonomic information when calibrating the model has a significant effect on the prediction.

A two-sided Wilcoxon signed-rank test was also used to determine if the difference between the predicted and reference diameters has a zero median (i.e. the null hypothesis). Using a 0.1% (i.e.  $\alpha = 0.001$ ) significance level, the null hypothesis was rejected for all the predictions (with p-values systematically smaller than  $10^{-2}$ ), meaning that a bias exists in the diameter predictions. Looking at the error metrics, indeed a small relative bias ranging from -1% to 7% can be observed for all the predictions (cf. tables 4.4-4.7).

Table 4.4: Composite model considering all genus. Average regression parameter values and error metrics based on 500 runs of Monte-Carlo cross validation. The dispersion of the scores is indicated by  $\pm 2\sigma$ .

Group	N <sub>train</sub>	N <sub>test</sub>	Parameters			Scores			
			$\alpha$	$\beta$	$\sigma$	RMSE	Bias	Rel. bias	$r^2$
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$	$\pm$ $2\sigma$
All	2224	1109	-0.24 $\pm$ 0.24	0.75 $\pm$ 0.04	0.143 $\pm$ 0.04	8.59 $\pm$ 0.7 cm	0.2 $\pm$ 0.74 cm	0 $\pm$ 0.04	0.98 $\pm$ 0

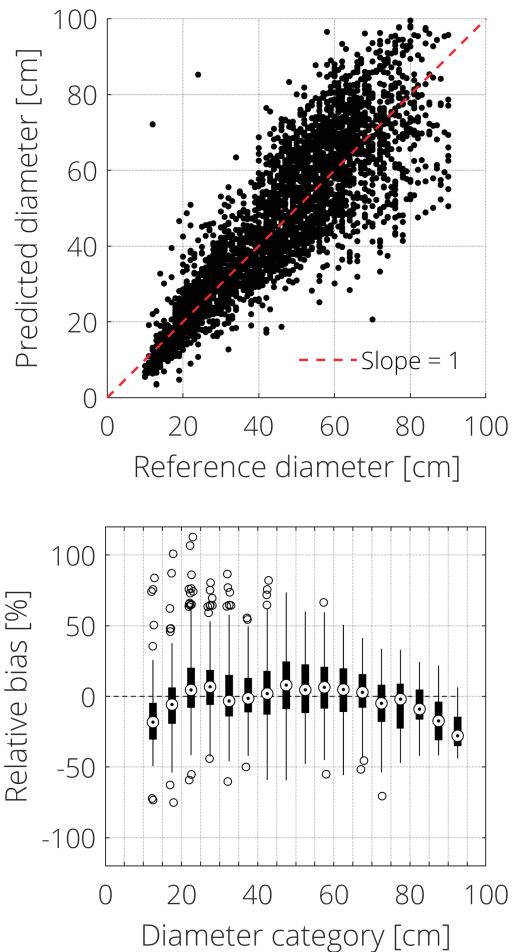


Figure 4.3: DBH regression error considering all available observations, using the average model parameter values from table 4.4. First row: correlation between predicted and reference diameter. Second row: boxplots of relative bias for 5 cm diameter categories.

Table 4.5: Composite models for each division (angiosperms, gymnosperms). Average regression parameter values and error metrics based on 500 runs of Monte-Carlo cross validation. The dispersion of the scores is indicated by  $\pm 2\sigma$ .

Group	N <sub>train</sub>	N <sub>test</sub>	Parameters			Scores			
			$\alpha$	$\beta$	$\sigma$	RMSE	Bias	Rel. bias	$r^2$
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$	$\pm$ $2\sigma$
Angiosperms	1195	596	-0.74 $\pm$ 0.24	0.82 $\pm$ 0.04	0.187 $\pm$ 0.114	9.19 $\pm$ 2.18 cm	1.53 $\pm$ 2.84 cm	0.06 $\pm$ 0.18	0.94 $\pm$ 0.16
Gymnosperms	1029	513	-0.19 $\pm$ 0.2	0.76 $\pm$ 0.04	0.127 $\pm$ 0.03	7.53 $\pm$ 0.52 cm	-0.88 $\pm$ 0.86 cm	0 $\pm$ 0.04	0.97 $\pm$ 0.02

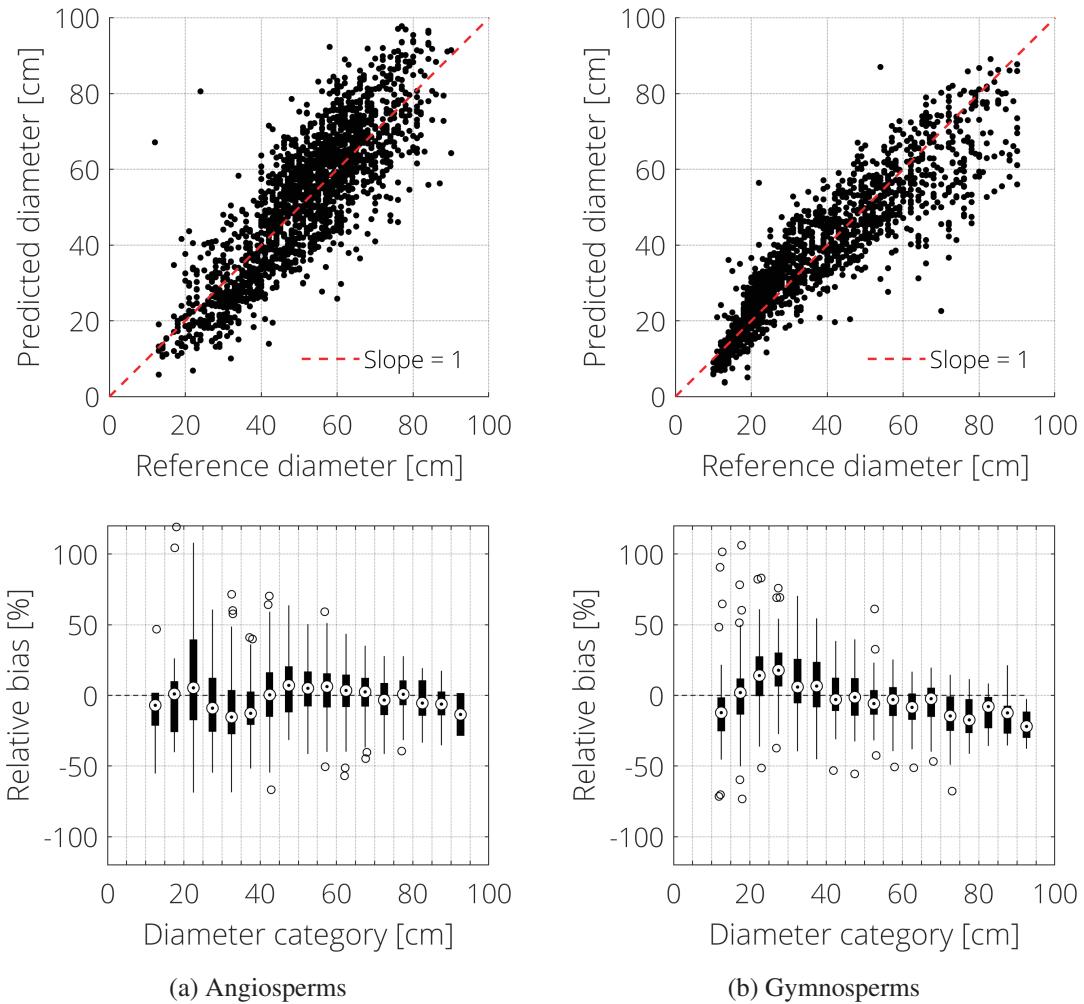


Figure 4.4: DBH regression error considering all available observations, using the average model parameter values from table 4.5. First row: correlation between predicted and reference diameter. Second row: boxplots of relative bias for 5 cm diameter categories.

Table 4.6: Specific models for oak (*Quercus* sp.) and beech (*Fagus* sp.). Average regression parameter values and error metrics based on 500 runs of Monte-Carlo cross validation. The dispersion of the scores is indicated by  $\pm 2\sigma$ .

Group	N <sub>train</sub>	N <sub>test</sub>	Parameters			Scores				
			$\alpha$	$\beta$	$\sigma$	RMSE	Bias	Rel. bias	$r^2$	
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$	$\pm$ $2\sigma$	
Quercus	296	147	0.04	0.72	0.115	6.21	0.37	0.01	0.94	
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$	$\pm$ $2\sigma$	
Fagus	899	449	-0.79	0.81	0.168	7.18	-0.36	0.02	0.94	
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$	$\pm$ $2\sigma$	
			0.28	0.04	0.076	2.26 cm	2.86 cm	0.1	0.12	
			0.26	0.04	0.09	1.62 cm	2.52 cm	0.16	0.16	

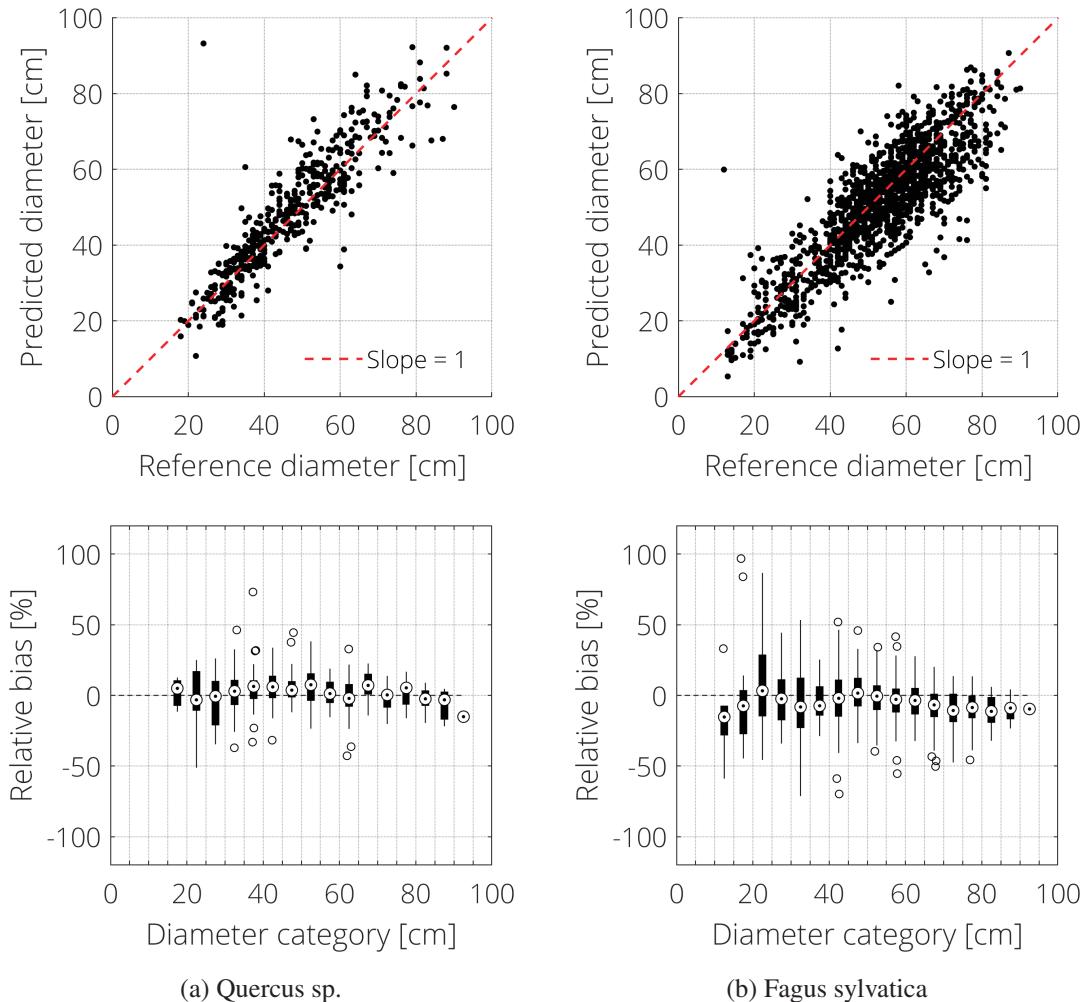


Figure 4.5: DBH regression error considering all available observations, using the average model parameter values from table 4.6. First row: correlation between predicted and reference diameter. Second row: boxplots of relative bias for 5 cm diameter categories.

Table 4.7: Specific models for fir (*Abies* sp.) and spruce (*Picea* sp.). Average regression parameter values and error metrics based on 500 runs of Monte-Carlo cross validation. The dispersion of the scores is indicated by  $\pm 2\sigma$ .

Group	N <sub>train</sub>	N <sub>test</sub>	Parameters			Scores				
			$\alpha$	$\beta$	$\sigma$	RMSE	Bias	Rel. bias	$r^2$	
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$ [cm]	$\pm$ $2\sigma$	$\pm$ $2\sigma$	
Abies	594	296	0.06	0.71	0.136	6.6	-0.7	0.01	0.96	
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$	$\pm$ $2\sigma$	
			0.14	0.02	0.018	0.56 cm	0.92 cm	0.04	0.02	
Picea	435	217	-0.57	0.84	0.14	7.96	-0.59	-0.01	0.97	
			$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$ cm	$\pm$ $2\sigma$	$\pm$ $2\sigma$	
			0.34	0.06	0.05	0.92 cm	1.76 cm	0.1	0.02	

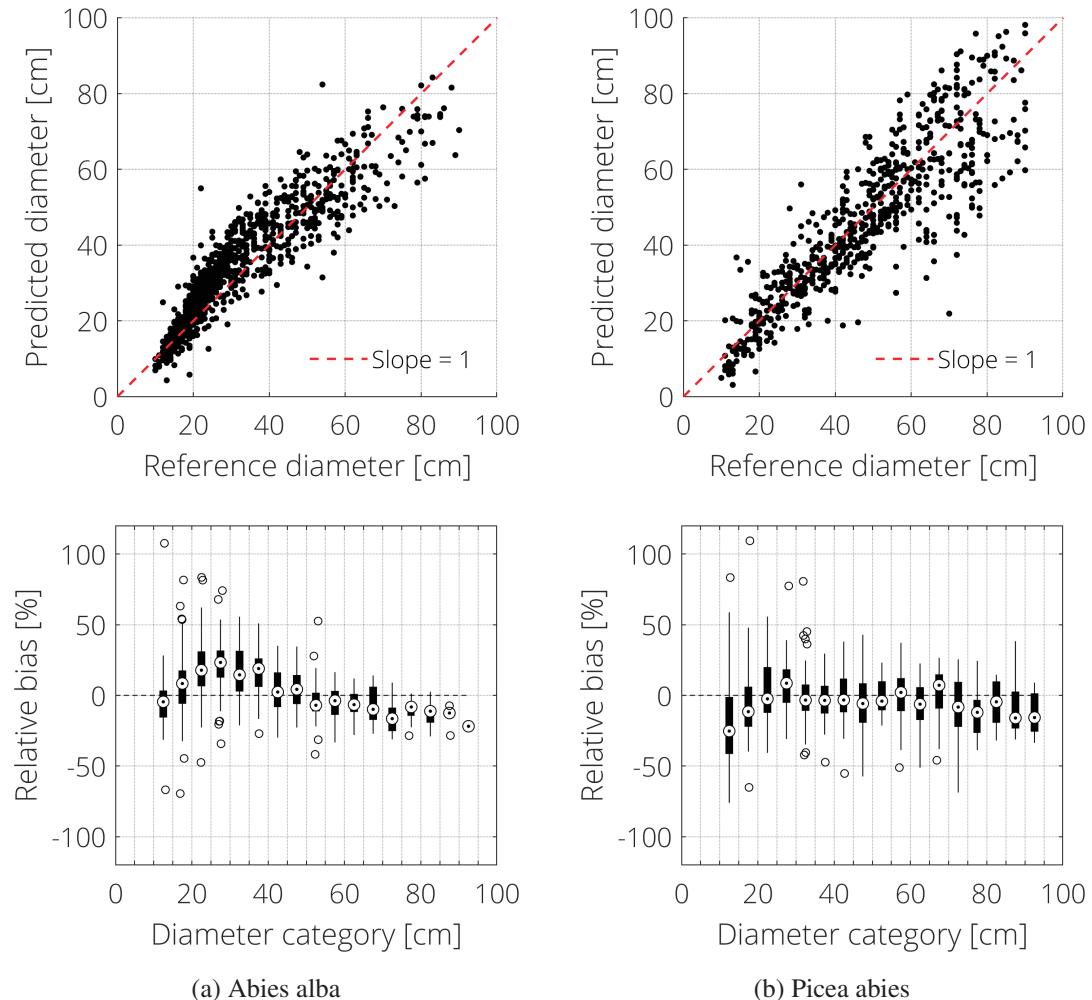


Figure 4.6: DBH regression error considering all available observations, using the average model parameter values from table 4.7. First row: correlation between predicted and reference diameter. Second row: boxplots of relative bias for 5 cm diameter categories.

Table 4.8: Error metrics for models calibrated at different taxonomic ranks. **R1**: composite (all), **R2**: division specific (Angiosperm or Gymnosperm), **R3**: genus specific. The parameter values used are the mean values obtained in cross-validation. The error metrics are based on all the observations. The best scores are indicated by the bold font and an asterisk (\*).

Genus	RMSE [cm]			Bias [cm]			Rel. Bias			$r^2$		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
Quercus	<b>7.89*</b>	10.88	8.11	-3.91	-8.25	<b>3.31*</b>	<b>-0.09*</b>	-0.2	<b>0.09*</b>	0.97	<b>0.98*</b>	0.97
Fagus	11.19	8.71	<b>7.74*</b>	7.52	3.13	<b>-1.19*</b>	0.26	0.12	<b>0*</b>	0.98	0.98	0.98
Abies	6.97	<b>6.73*</b>	6.74	-3.18	0.3	<b>-0.01*</b>	-0.07	<b>0.02*</b>	0.03	<b>0.97*</b>	<b>0.97*</b>	0.96
Picea	8.98	<b>8.09*</b>	8.28	-5.31	-2.09	<b>-0.97*</b>	-0.11	<b>-0.02*</b>	<b>-0.02*</b>	0.98	0.98	0.98

### 4.3.3 Discussion

The diameter prediction results are comparable (relative bias close to zero and  $0.96 \leq r^2 \leq 0.98$ ) to those reported in Duncanson et al. (2015), Jucker et al. (2016) and Dalponte et al. (2018b). Interestingly, the high geometric quality of the manually delineated trees does not seem to improve stem diameter estimates, when compared to the results obtained in Dalponte et al. (2018b). This might be a consequence of the increase in tree crown area variability as a function of height (Heteroscedasticity).

There are several significant differences in the analytical approach and data used here, when compared to the work of Jucker et al. (2016) or Dalponte et al. (2018b). In the article by Jucker et al. (2016), observations are grouped (stratified) by biome and division (angiosperm/gymnosperm) when applying the regression, but the effect of possibly imbalanced species group sizes in each biome or division is not explicitly taken into account. Similarly, this effect was not taken into account by Dalponte et al. (2018b) when calibrating their composite (non group-specific) models. The dataset sizes and stem diameter ranges are also different: Jucker et al. (2016) use a global dataset (108'753 observations) with DBH ranging from 1 cm to more than 200 cm, most of the DBH observations (17'438 observations) used in Dalponte et al. (2018b) range from 5 to 50 cm, we use a much smaller dataset (3'333 observations) with DBH ranging from 10 to 90 cm. It can also be noted that Jucker et al. (2016) and Dalponte et al. (2018b) both use 90% of the data for training and only 10% for validation, while 2/3 and 1/3 are used here.

Most of the models show a negative bias for small ( $< 20$  cm) and large ( $> 80$  cm) diameters. It is unclear if this bias is a limitation of the allometric model itself or if it is due to the small sample sizes in these diameter categories. However, looking closely at the correlation plots in Jucker et al. (2016) and Dalponte et al. (2018b), a similar negative bias can be observed for small diameters, but not for large ones which seems to indicate that the bias observed here could be due to the small number of observations in the extreme diameter categories.

Using data from boreal and temperate forests, Dalponte et al. (2018b) found that locally calibrated species-specific allometric models did not perform significantly better than composite and regionally calibrated allometric models. The results obtained here are not in complete agreement with the finding of Dalponte et al. (2018b). In particular, it was found that the bias changes significantly between composite (R1), division specific (R2) and genus specific (R3) models and in this regard the latter models tended to produce the best results. This difference was not apparent when comparing the performances of the models using RMSE and  $r^2$ , with a difference in RMSE less than 2 cm and almost identical  $r^2$ . Here again, it remains unclear if this difference in bias could be due to the relatively small sample size used for calibration and validation or if it is an intrinsic limitation of the model. A definitive conclusion on the relative performance of R1, R2, R3 level models could possibly be made by repeating the analysis with additional observations.

## 4.4 Direct diameter measurement

This section presents a novel stem taper fitting method adapted to leaf-off high density ALS data. The method relies on preliminary segmentation of individual trees or stems (cf. chapter 3).

### 4.4.1 Description

Reliably measuring stem diameter and taper from long range ALS data is challenging for several reasons:

- The stem is sampled with a **low number of points**. Compared to TLS which may sample a single stem with thousands of measurements, long range ALS typically only samples a stem with several tens of measurements. This implies there is a low number of redundant observations.
- The measured points on the stem often have an **non-uniform spatial distribution**. Ideally, to identify the outer boundary of the stem, the points should be distributed uniformly around the stem circumference and along its lengths. However, due to a multitude of acquisition factors (e.g. occlusion, scan pattern, scan swath overlap), points are generally not uniformly distributed.
- The **point position error** may be high relative to the diameter of the stem. Although many ALS systems are able to routinely achieve sub-decimeter position errors, this value might vary considerably depending on acquisition conditions (Habib et al., 2009; Goulden and Hopkinson, 2010).
- The presence of a significant amount of **non-stem points (outliers)** in the input segment. Individual tree or stem segmentation is a prerequisite for stem diameter and taper measurement. However, segmentation methods often do not produce perfect results and a significant amount of non-stem elements (e.g. branches, low vegetation, parts of other trees) may still be present in the result. In other words, the input segment may have a low signal (stem points) to noise (non-stem points) ratio.

Despite these limiting factors, it is interesting to investigate the possibility of using long range ALS data for direct diameter measurement, because most of the near future ALS collected over large forest areas can be expected to have similar characteristics. Moreover, some long range ALS datasets already have point densities above 100 points per m<sup>2</sup> suitable for direct diameter measurement.

Directly measuring stem diameter and taper from a point cloud requires fitting geometric primitives (e.g. circle or ellipse) to subcomponents of the tree (e.g. stem, branches) or meshing (e.g. with alpha shapes or splines) the boundary points. Meshing approaches are able to reconstruct the tree shape with high accuracy, but they require very high point densities and homogeneous sampling of the tree structure. On the other hand, methods that fit geometric primitives can deal with a considerable amount of sampling inhomogeneity and lower point densities. For this type of approach, the use of robust methods to fit geometric primitives is generally necessary, when dealing with a low signal to noise ratio.

Tree stem geometry can be complex and may vary considerably depending on multiple factors such as species, age and environment. However, for many practical purposes and for the proposed taper fitting method, the following geometric simplifications are assumed:

- Approximate stem **linearity**. Many stems and in particular several coniferous species have an approximately linear stem. Heliophyte (i.e. shade intolerant) species (e.g. *Pinus sylvestris*) will often have distorted stems, but may still have a significant fraction of their stem that is linear.

- Approximate **vertical growth**. The growth direction of plants is governed by gravity (gravitropism) and incoming light (phototropism), both of which promote vertical growth for structural and light access reasons respectively. Heliophyte (i.e. shade intolerant) species (e.g. *Pinus sylvestris*) are more prone to have tilted stems.
- Approximate **circularity of stem cross-sections**. This assumption is generally true for coniferous species, but less adequate for deciduous (Matérn, 1956; West, 2013).
- Approximate **linear taper**. This assumption is not valid at the base of the stem and above the crown base height, but generally applicable in between (Larsen, 2017).

Mathematically, these assumptions imply that the taper and medial axis curve of the stem can be approximated with a polynomial of degree  $n$ :

$$p(x) = c_0 + c_1x + \dots + c_nx^n \quad (4.6)$$

where:

- $c_i$  is the coefficient for degree  $i$ ;  
 $x$  is the independent variable (e.g. diameter).

It can be assumed that  $n = 1$  for the taper model and  $1 \geq n \geq 3$  for the medial axis model. Moreover, the tolerable amount of stem non-linearity and tilt can be controlled by setting constraints on the coefficients  $c_i$  of the medial axis polynomial. Similarly, the DBH and taper ranges can also be constrained based on a priori knowledge about probable stem geometry.

Following this premise, a novel taper fitting method based on the Random Sample Consensus (RanSaC) algorithm is proposed. RanSaC is a robust model fitting method first described by Fischler and Bolles (1981). It has been successfully applied to tree modeling in the past; for example Reitberger et al. (2007) used it to detect linear structures in individual tree segments and Olofsson et al. (2014) used it to detect circles in TLS cross-sections. Generally speaking, RanSaC is used when the ratio of inlier observations is too low to obtain reliable results with less robust methods such as least-squares regression. It works by iteratively sampling the minimum subset of observations required to calibrate a model (e.g. 3 points for a circle) and computing the number of observations that conform to this candidate model (i.e. inliers) according to a predefined tolerance. The sampling is repeated until all combinations have been tested or a maximum number of iterations is reached. The number of iterations required to find suitable model parameters can be estimated with equation 4.7, if the approximate proportion of outliers is known:

$$N_{iter} = \frac{\log(1-p)}{\log(1-(1-e)^s)} \quad (4.7)$$

where:

- $N_{iter}$  is the number of required iterations (i.e. number of random draws);  
 $p$  is the desired probability of drawing a sample with only inliers;  
 $s$  is the minimum number of observations required to calibrate the model (e.g. 3 for a circle);  
 $e$  is the proportion of outliers.

After the maximum number of iterations is reached, the model parameters that result in the most inliers are retained. The following subsection explains how RanSaC is applied to address the problem of fitting the stem medial axis, diameter and taper simultaneously.

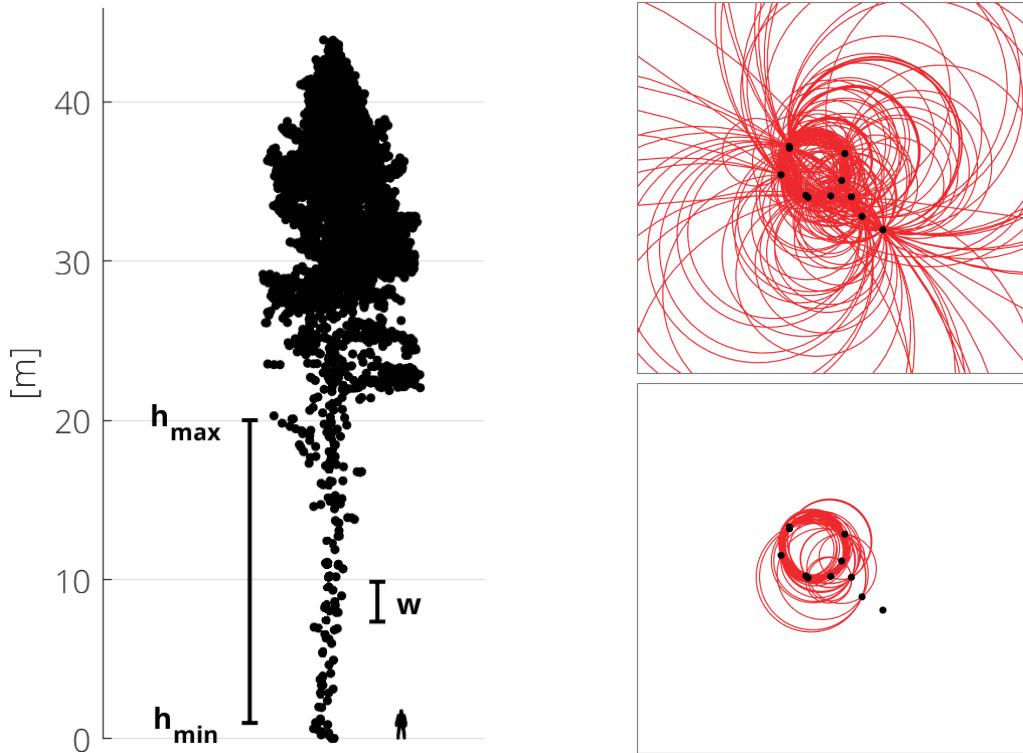
### Algorithm

The key steps of the method are illustrated in figures 4.7 to 4.9 and explained below:

- (a) **Isolate an individual tree or stem from a 3D point cloud** (cf. chapter 3).
- (b) **Normalize the height of the tree/stem points relative to the elevation of the root point (i.e. intersection of stem base and terrain).**
- (c) **Filter the points located in the sampling height range  $[h_{min}, h_{max}]$** , cf. figure 4.7a.
- (d) **Find all unique combinations of three points (triplets) separated by a vertical distance less than or equal to  $w$ . Each of these point triplets is projected on a horizontal plane and defines a circle which is potentially part of the stem boundary** (cf. figure 4.9a).
- (e) **Compute the geometric quality index of each circle** (cf. figure 4.8). This index provides an indication of reliability on the circle fit and is conceptually similar to the (covariance) error ellipse used to model the expected error of a least squares adjustment.
- (f) **Filter the circles based on the quality index, emptiness, a priori knowledge about stem diameter and taper ranges in the region** (cf. figures 4.9b and 4.7b).
  - 1 Exclude circles that do not fulfill a minimum geometric quality constraint  $Q_{min}$ .
  - 2 Exclude circles that do not fulfill the diameter range constraint  $[d_{min}, d_{max}]$ .
  - 3 Exclude circles that contain points further than  $\varepsilon_E$  from the boundary (within the vertical domain defined by the three points). This takes into account possible non-circularity of the stem and error in point positions.
- (g) **Simultaneously fit the medial axis and taper models with RanSaC** (cf. figure 4.9c):
  - 1 Randomly select (without repetition) a pair of circles that fulfill the taper  $[t_{min}, t_{max}]$  and DBH range constraints  $[d_{min}, d_{max}]$ .
  - 2 Find all circles with a diameter that fit these constraints within relative tolerance  $\varepsilon_D$ .
  - 3 Randomly select  $n_A + 1$  circles among the taper inliers and fit the medial axis model (polynomial of degree  $n_A$ ) to them (x and y coordinates are fitted independently).
  - 4 Count the number of circle centers that are located within a horizontal distance  $\varepsilon_A$  from the fitted medial axis model (i.e. number of inliers).
  - 5 Repeat steps 1-4 until all possible combinations of  $n_A + 1$  circles have been tested or the maximum number of iterations ( $N_{iter}$ ) is reached.
  - 6 Retain the set of circles that comply with the best fit.
- (h) **Check if the inlier circles cover at least length  $L$  of the stem.** This is done by dividing the stem into equidistant (0.5 m) bins and counting how many bins contain an inlier circle.
- (i) **Compute the refined taper and DBH** by using a least squares adjustment of diameter as a function of height on the set of inlier circles.

### Parameters

$n_A$	Degree of the polynomial used to model the stem medial axis (nominal value: 2);
$w$	Maximum vertical separation of point triplets (nominal value: 2 m);
$Q_{min}$	Minimum circle geometric quality (nominal value: 0.5);
$[h_{min}, h_{max}]$	Sampling height range (nominal values: [0, 15] m);
$[d_{min}, d_{max}]$	Diameter range (nominal values: [0.3, 1] m);
$[t_{min}, t_{max}]$	Taper range (nominal values: [0, 0.015]);
$\varepsilon_A$	Absolute tolerance on medial axis fit (nominal value: 0.05 m);
$\varepsilon_D$	Relative tolerance on circle diameter fit (nominal value: 0.15 m);
$\varepsilon_E$	Relative tolerance on circle emptiness (nominal value: 0.85 m);
$N_{iter}$	Maximum number of RanSaC iterations (nominal value: 3000);
$L$	Length of the stem covered by inlier circles (nominal value: 5 m);



(a) For each point  $i$ , circles are fitted to the points in the interval  $[h_i - w/2, h_i + w/2]$ . This process is repeated for all points between  $h_{\min}$  and  $h_{\max}$  and all unique circles are retained.

(b) **Above:** circles obtained with all point triplet combinations within  $[h_i - w/2, h_i + w/2]$ . **Below:** circles after applying the geometric quality, diameter and emptiness constraints.

Figure 4.7: Circle fitting example.

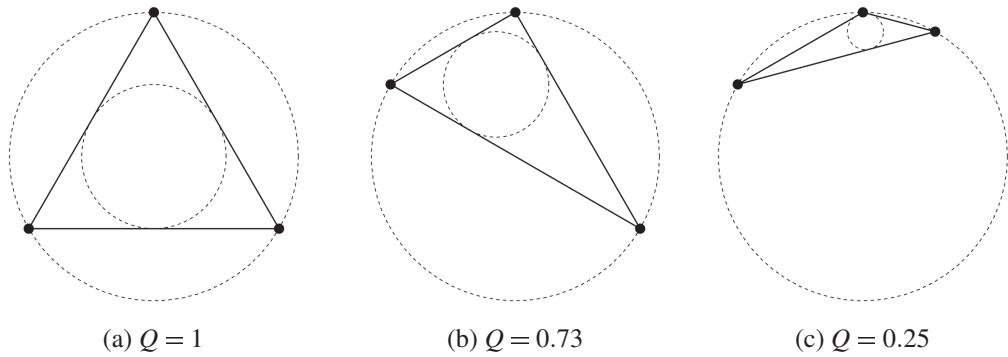
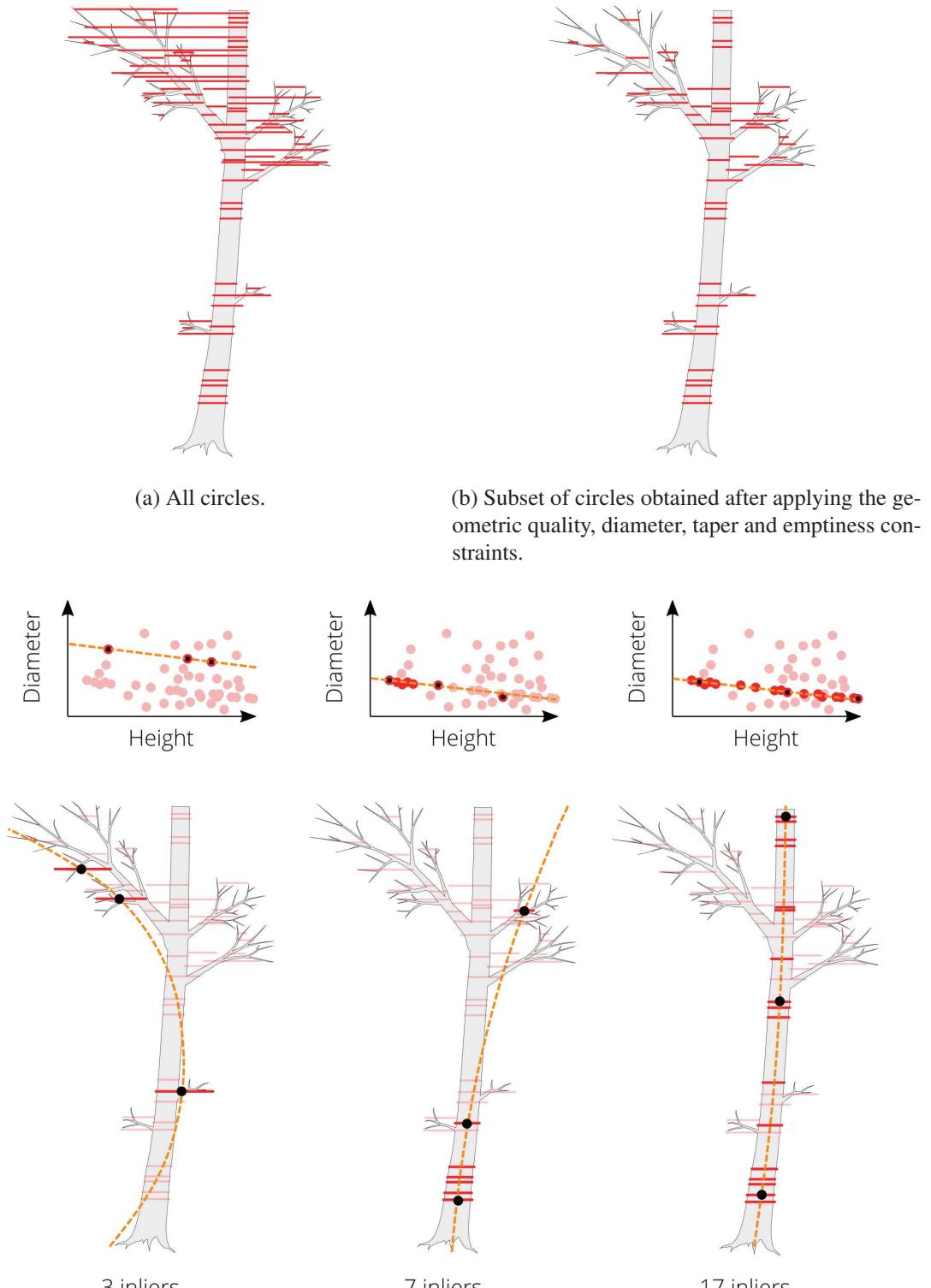


Figure 4.8: Examples of point triplet geometric quality  $Q$ . The index  $Q$  is equal to twice the ratio of the inscribed circle radius ( $r_{ic}$ ) to the circumscribed circle radius ( $r_{cc}$ ), i.e.  $Q = 2 \frac{r_{ic}}{r_{cc}}$ . Values of  $Q$  close to 1 (equilateral triangle) indicate that the circle is well defined. The geometric quality index  $Q$  is conceptually similar to the (covariance) error ellipse used to model the expected error of a least squares adjustment.



(c) At each iteration, RanSaC randomly selects  $n_A + 1$  circles (black dots) that define a taper (first row) and medial axis curve (second row). The candidate stem geometry model that results in the most inlier circles (bright red) according to chosen tolerances on  $\varepsilon_A$ ,  $\varepsilon_D$ , and  $\varepsilon_E$  is retained.

Figure 4.9: Simultaneous stem medial axis and taper fitting.

#### 4.4.2 Results

The proposed method was tested on a simulated stem and on real ALS data from the three study sites (Boudry D20, Boudry D19, Chambrelien) with the highest point density ( $\geq 70 \text{ m}^{-2}$ ). For the ALS simulation, a simple model (cf. figure 4.10) is used in which points on and around the stem boundary are defined by:

$$\begin{aligned} x(h) &= x_A(h) + r(h) \cdot \cos(\theta) + \varepsilon \\ y(h) &= y_A(h) + r(h) \cdot \sin(\theta) + \varepsilon \end{aligned} \quad (4.8)$$

where:

$h$  is the height above ground level,  $h \in [0, 15] \text{ m}$ ;

$\theta$  is an angle,  $\theta \in [0, 2\pi]$ .

The medial axis of the stem is modeled as a second degree polynomial:

$$\begin{aligned} x_A(h) &= 0.012 \cdot h^2 + 0.043 \cdot h \\ y_A(h) &= 0 \end{aligned} \quad (4.9)$$

The taper is modeled with a linear function and additional points are randomly added around the stem surface to simulate non-stem elements:

$$r(h) = \begin{cases} -0.008 \cdot h + 0.5 & \text{for stem points (inliers)} \\ -0.008 \cdot h + 0.5 + r_N & \text{for non-stem points (outliers)} \end{cases} \quad (4.10)$$

with  $r_N$  randomly sampled from the continuous uniform distribution on the interval  $[0, 1.5]$ :

$$r_N \sim U(a = 0, b = 1.5) \quad (4.11)$$

The ALS point position error  $\varepsilon$  is modeled as a random variable with a Gaussian distribution:

$$\varepsilon \sim \mathcal{N}(\mu = 0, \sigma^2 = 0.0025) \quad (4.12)$$

Finally, equations 4.8 are evaluated with  $N$  random values of  $h$  and  $\theta$ , considering a fraction  $f$  of stem points (inliers) and  $1 - f$  of non-stem points (outliers) in equation 4.10.

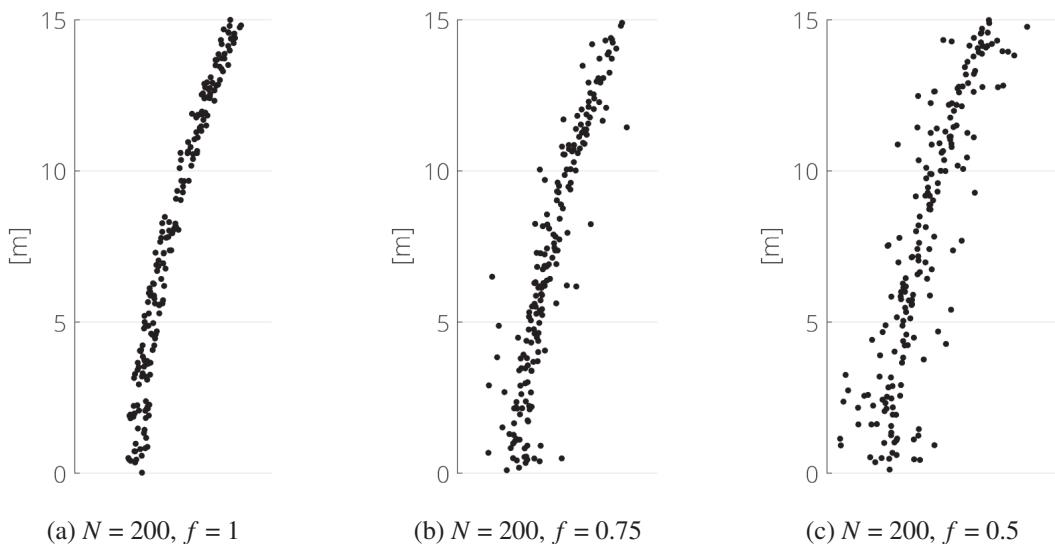


Figure 4.10: Side view of simulated stems with different inlier point fractions.

A full factorial performance sensitivity analysis was used to evaluate the combined effect of:

- point density,  $\forall N \in \{75, 150, 300\}$ ;
- fraction of stem (inlier) points,  $\forall f \in \{0.5, 0.75\}$ ;
- radial sampling homogeneity (only points with  $\theta \in [0, \theta_{max}]$  are considered),  $\forall \theta_{max} \in \{\pi, 2\pi\}$ .

The stem measurement algorithm was repeated 100 times for each combination of factor levels (with a new simulation at each iteration), using the nominal parameter values suggested in section 4.4.1. The mean results on the simulated data are reported in figure 4.11.

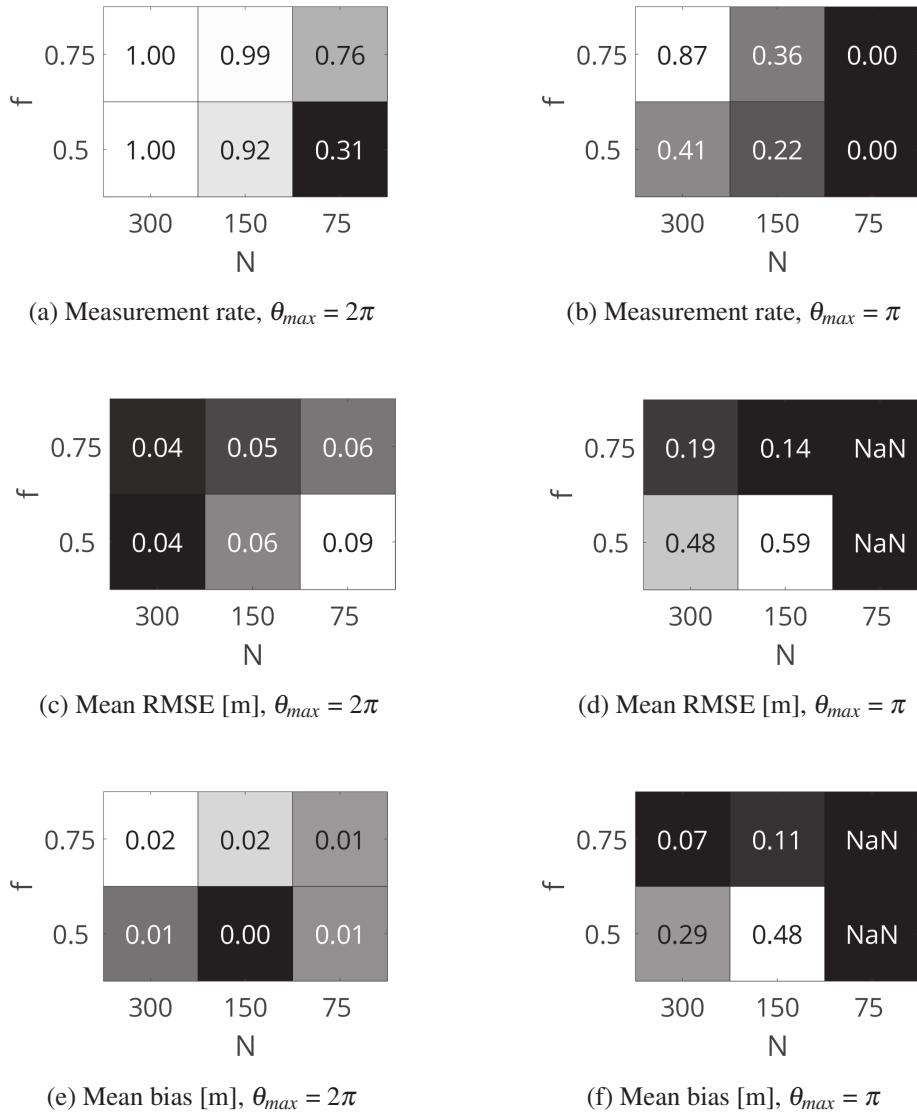


Figure 4.11: Diameter measurement sensitivity for simulated stems.

The performance of the algorithm was also evaluated on individual trees from the reference ALS dataset presented in chapter 2. The same nominal parameters as previously were used, except for the maximum number of iterations which was set to  $N_{\text{iter}} = 5000$ . The corresponding results are reported in table 4.9 and figure 4.12. Qualitative examples are illustrated in figure 4.13.

Table 4.9: Direct stem DBH measurement results.  $N_{\text{obs}}$  is the total number of observations,  $N_m$  is the number of stems that could be measured by the algorithm,  $N_v$  is the number of stems used for validation (i.e. where both field and ALS based measurements were available).

Site	$N_{\text{obs}}$	$N_m$ (%)	$N_v$ (%)	RMSE [cm]	MAE [cm]	Rel. bias	$r^2$
Boudry D20	312	116 (37%)	46 (15%)	8.7 cm	5.6 cm	0	0.7
Boudry D19	320	50 (16%)	24 (8%)	6.4 cm	4.4 cm	0.1	0.81
Chambrelien	224	65 (29%)	-	-	-	-	-
<b>Overall</b>	<b>856</b>	<b>231 (27%)</b>	<b>70 (8%)</b>	<b>8 cm</b>	<b>5.2 cm</b>	<b>0.03</b>	<b>0.75</b>

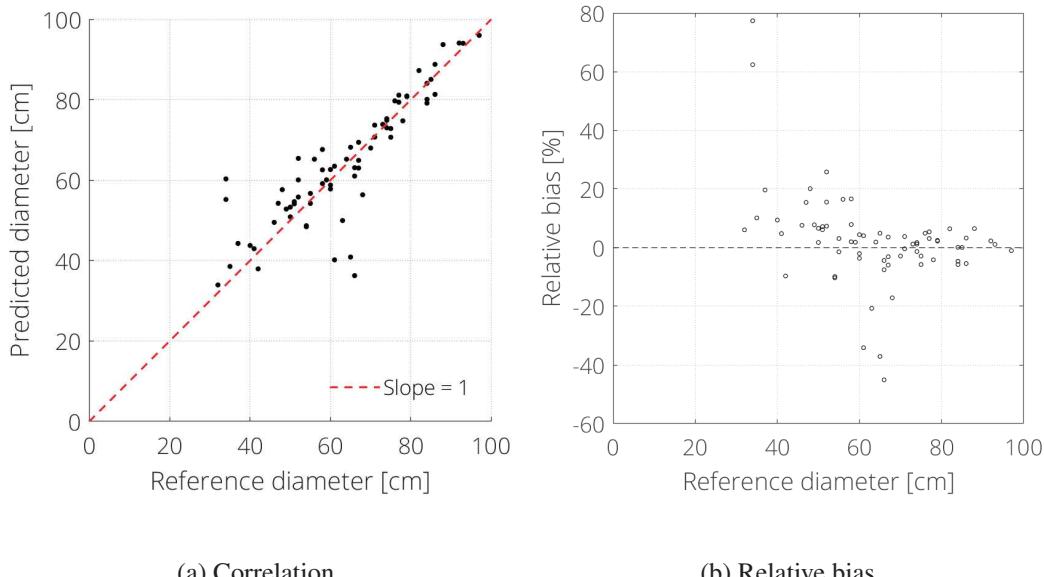


Figure 4.12: Diameter prediction errors for all available observations.

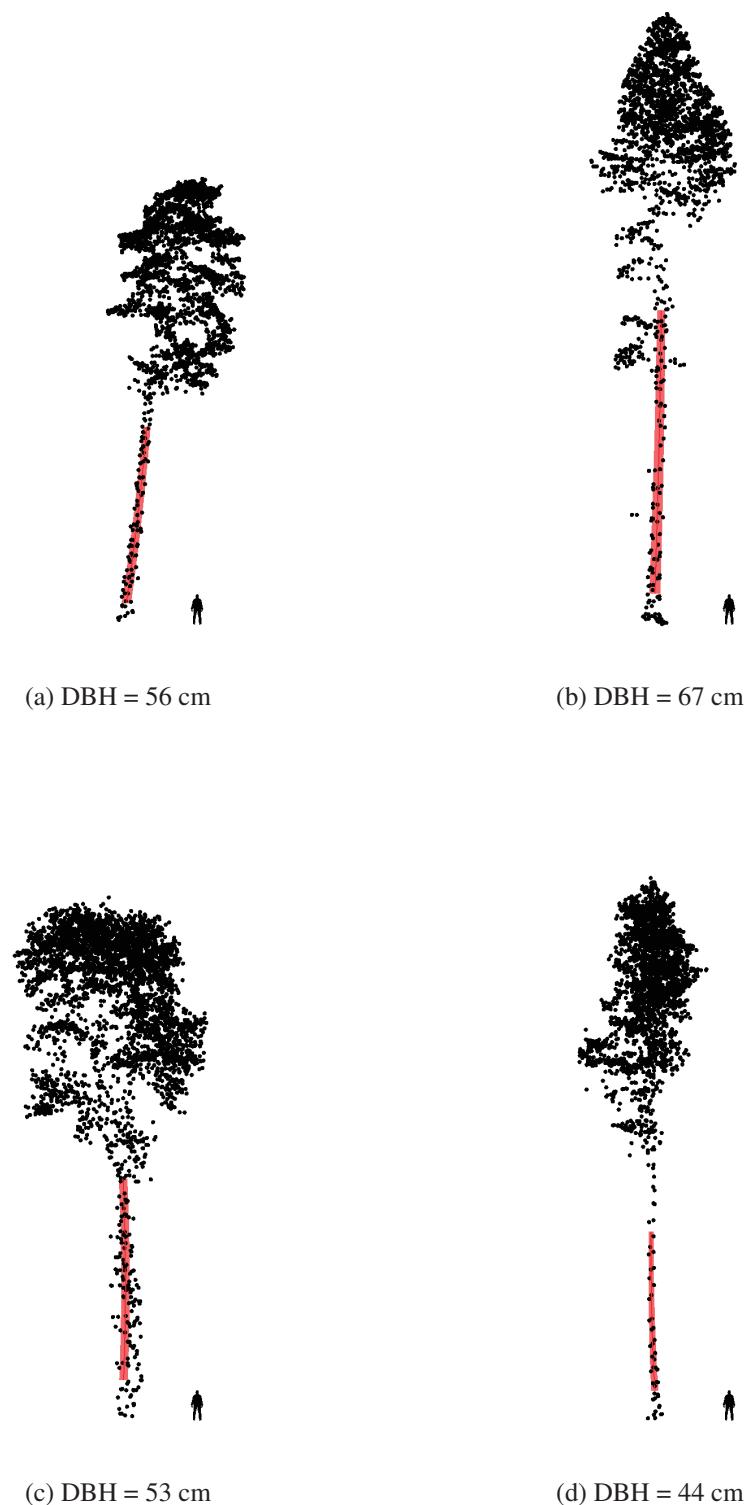


Figure 4.13: Qualitative examples of stem models fitted to ALS points. The fraction of the stem where circles could be adjusted is indicated in red.

#### 4.4.3 Discussion

The results on the simulated stems indicate that the algorithm can handle a considerable amount of outliers, as long as the point density is high ( $N > 150$  / stem) and the radial sampling is homogeneous ( $\theta_{max} = 2\pi$ ). Under these conditions the method consistently measured (> 92% of the time) the stem diameter with a 4-6 cm mean RMSE and a 0-2 cm mean bias. Decreasing the quality of the data, by simultaneously increasing the outlier ratio ( $f \geq 0.5$ ) and diminishing the number of points ( $N = 75$ ), resulted in lower measurement rates (31%) and an increase in RMSE (0.09 m). Reducing the radial sampling  $\theta_{max}$  to  $\pi$  significantly affected the capacity of the algorithm to accurately measure the stem. Under these conditions, only half of the points are considered ( $N/2$ ) and they are all distributed on the same side of the stem, which means there is a low probability that the algorithm's constraints are met.

On the real ALS data, the results indicate that the algorithm is able to measure about 25% of the stems. The algorithm was able to measures trees down to 30 cm DBH, but typically has higher measurement rates for larger diameters. The average error on DBH (MAE = 5.2 cm, RMSE = 8 cm, Rel. Bias = 0.03) is generally within the the uncertainties related to ALS and field surveys. These performances are slightly better than those obtained with allometric models 4.3. Note, that the evaluation of the method is based on relatively small validation set (70 observations).

This approach is advantageous because, unlike allometric models, it does not rely on field surveys (except for validation). The implementation is flexible and may be modified in a number of ways. First, the initial circle detection can be substituted with any other appropriate circle, ellipse, or other cross-section shape fitting methods (e.g. circular Hough transform, linear or non-linear least squares adjustment). The area of the cross-section can be used as the dependent variable in the taper equation, instead of the circle diameter. Both the medial axis and taper functions can also be changed to use different models. The main disadvantage of the method is its computation time ( 1-2 sec per tree) and reliance on random sampling of points. The latter is related to the use of RanSaC which may produce different results at each run and which does not use all the data when adjusting the model.

Steps to improve the results and increase the processing speed include: preliminary filtering of points that are likely to be on the stem surface (e.g. by using last returns only and/or applying a stem detection algorithm), using the curvilinear distance to the base of the stem instead of the 2D vertical distance (height) could improve results for trees that are heavily tilted and indirect diameter estimates with generic allometric models (cf. section 4.3) could be used to constrain the taper more tightly.

## 4.5 Synthesis

In this chapter, a state of the art on stem diameter estimation from ALS data was conducted. It was found that most methods to estimate stem diameter from ALS rely on allometry (i.e. indirect estimation from auxiliary geometric variables such as height or crown diameter) and that direct measurements of diameter/taper in long range ALS point clouds has not been extensively investigated. Based on this observation, two topics were investigated:

- The allometric diameter estimation method proposed by Jucker et al. (2016) was modified to take into account imbalanced species frequencies and its performance was evaluated on the benchmark dataset presented in chapter 2.2. It was shown that species-specific and non-specific models performed similarly in terms of RMSE and  $r^2$ , but produced different biases when evaluated on single species. RMSE values smaller than 10 cm were consistently obtained and no large systematic bias was observed. It can be concluded that the approach is applicable in an operational context, given a reliable individual tree crown segmentation. The advantage of the allometric approach is its simplicity and its applicability to low density point clouds. Conversely, the reliance on individual tree shape delineation is the main limitation of the approach because errors in the crown diameter or height of segments will propagate to the DBH estimation. Upscaling of the DBH at the plot scale also depends on the quality of the tree detection (precision and recall). Finally, with this approach identification of the species may be required to improve DBH predictions (in terms of bias at least) in some cases.
- A novel direct diameter and taper measurement method from high density ( $> 70$  points /  $m^2$ ) ALS point clouds was presented. The method is based on simultaneously fitting a stem curve and taper model to the point cloud using RanSAC. Validation on simulated and real trees extracted from ALS point clouds showed that the method is able to handle a significant amount of noise and could predict the diameter of about 25% of trees with a diameter larger than 30 cm. An overall RMSE of 8 cm, MAE of 5.2 cm and  $r^2$  were achieved. This type of direct measurement approach will be increasingly relevant, given the improving availability of high density ALS data. The main advantage of this method is that it only depends on the detection of stems (not the full tree shape) and does not require the species to be identified. Its main disadvantage is its reliance on sufficient sampling of the stem structure (which requires high density point clouds preferably acquired in leaf-off conditions) and computational complexity.





## 5. Tree species classification

This chapter covers the topic of tree species classification based on ALS and AHI derived features. It is structured in the following way:

- 5.1** introduces the topic and describes the state of the art.
- 5.2** describes the error assessment framework used to evaluate the performance of classification methods.
- 5.3** describes individual genus/species characteristics in terms of structural and spectral separability.
- 5.4** presents a tree species classification workflow based on ALS derived features only.
- 5.5** presents a tree species classification workflow based on AHI derived features only.
- 5.6** presents a tree species classification workflow based on combined ALS and AHI derived features. It also summarizes the overall performance of different ALS/AHI feature combinations and classifiers.

### 5.1 State of the art

The mapping of forest canopy characteristics over large areas has a history which coincides with the development of remote sensing technology. Thematic mapping of forest canopies has important applications including the conservation of areas with high species richness and/or rare species, the identification of biological habitat suitability, the quantification of biomass and timber volume by species, the detection of invasive neophytes, the evaluation of tree health (e.g. disease, water stress, nutrient deficiencies) and the localization of areas most vulnerable to natural disasters (e.g. forest fires, landslides). Of course, the quality of these thematic maps directly depends on the ability to reliably measure and analyze the structural and biological characteristics of forests. In particular, the capacity to differentiate tree functional traits and ultimately individual species is central. Early approaches to map canopy characteristics relied entirely on photo-interpretation and the mapping quality depended on the qualifications and experience of the interpreter. Photo-interpretation

consists in the analysis of texture, shape, color and spatial context of single or diachronic images (Boutin et al., 1953; Heller et al., 1964; Avery, 1969; Sayn-Wittgenstein, 1978; Hershey and Befort, 1995; Oester, 2003). If pairs of overlapping images are available, the relative height of trees can also be estimated with a stereoscope. Photo-interpretation was the main operational approach for large area forest mapping until the early 2000's and it is still actively used in some national forest inventories (in Switzerland for example). Following the pervasive adoption of digital sensors and computers, photo-interpretation is being progressively replaced by more efficient mapping techniques which can be automated and are not affected by individual bias. Most of these forest mapping techniques rely heavily on new active remote sensing technologies (such as airborne laser scanning or synthetic aperture radar) and digital photogrammetry which can accurately measure the forest's 3D structure. In this context, the combination of different remote sensing techniques and computational tools to improve forest mapping has become a popular research topic (cf. tables 5.2 to 5.7). Complementary information on tree structure and reflectance may be obtained with multi-sensor (e.g. laser scanner and multi/hyperspectral imager) or single sensors (multi/hyperspectral laser scanner, multi/hyperspectral photogrammetry) surveys. Currently, the combination of airborne laser scanning and multi/hyperspectral imaging is the most widely investigated solution and it has been shown to be effective for tree species and health mapping (Torabzadeh et al., 2014; Xu et al., 2015; Fassnacht et al., 2016; Kukkonen et al., 2018). Thanks to active research communities, ALS/AHI equipment and funding availability, the forests in developed countries have received significant attention (in particular Scandinavian taiga, European broadleaf forests and Puget lowland forests). Inter-tropical regions have received much less attention, even-though they are much richer in terms of biodiversity and more threatened by degradation.

Airborne and spaceborne multispectral imaging has been used extensively to map tree species. Its main advantage is availability and relatively low cost. However, because it generally has only a few bands with low spectral resolution (i.e. wide bands), it has less discriminative power when compared to hyperspectral imaging which provides a quasi continuous sampling of the spectra with many narrow bands. The interest in using the latter technology to identify single species comes from the fact that many species have subtle spectral features (a few nanometers wide) which can only be detected when considering adjacent narrow spectral bands. Some regions of the spectra in particular (cf. figure 5.1) provide information about biochemical properties such as leaf pigment concentration (e.g. chlorophyll, carotenes, xanthophylls, anthocyanins), water and dry matter contents (e.g. cellulose, lignin). The relative presence of these components can help discriminate species and evaluate tree health (Kumar et al., 2002). With multi/hyperspectral imaging, the spatial extent of pixels is generally larger than individual leafs or needles, so spectral signatures also contain information about canopy structural properties such a leaf area index and crown density (Schlerf and Atzberger, 2006b; Zheng and Moskal, 2009). Commonly used features to discriminate vegetation from multi/hyperspectral reflectance data include:

- **Raw reflectance values** of selected bands (based on known species traits);
- **Transformed reflectance values** after applying a Principal Component Analysis (PCA), Maximum Noise Fraction (MNF) (cf. (Green et al., 1988)), standard normal variate transform (Barnes et al., 1989) or other techniques to increase the signal to noise ratio.
- **First or second derivatives** of the raw or transformed reflectance signal (Demetriades-Shah et al., 1990; Tsai and Philpot, 1998);
- Narrow and broad band **vegetation indices** (Broge and Leblanc, 2001; Silleos et al., 2006; Stagakis et al., 2010; Garbulsky et al., 2011; Pettorelli, 2013);
- **Temporal change** of any of the above (due to phenology, growth, death, disease).

These features may be computed at pixel or object scales (e.g. tree crown, forest stand). In the latter case, a preliminary segmentation of the image is necessary (cf. chapter 3).

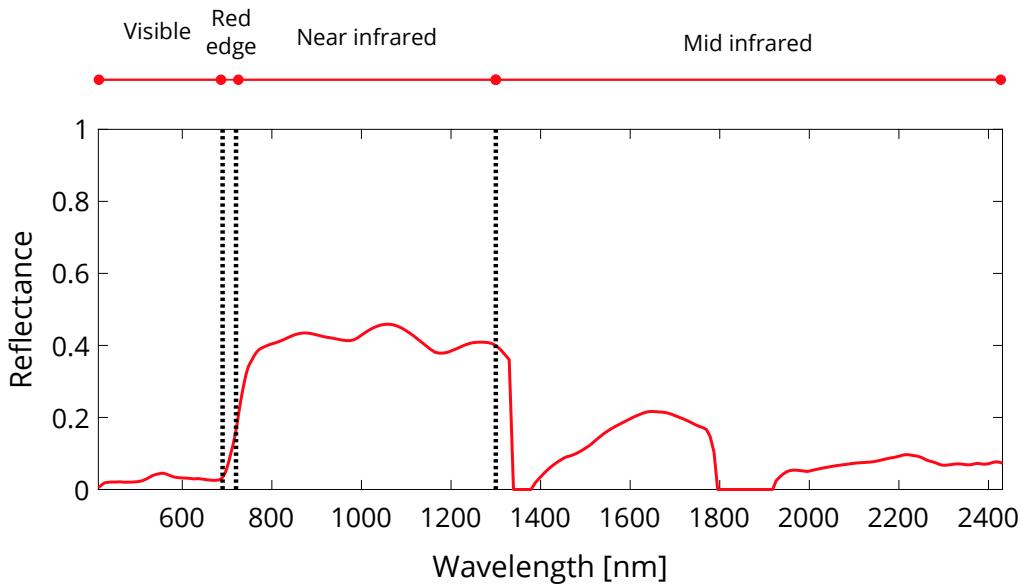


Figure 5.1: Important spectral regions for vegetation analysis. **Visible (400-700 nm)**: strong absorption by leaf pigments (chlorophyll a and b, carotenes, xanthophyll). **Red edge (690-720 nm)**: strong reflectance increase, **Near infrared (700-1300 nm)**: strong reflectance, **Mid infrared (1300-2500 nm)**: strong absorption by water and several compounds (e.g. cellulose, lignin).

Airborne laser scanning has also been used extensively for tree species mapping. Its ability to differentiate tree species is dependent on acquisition parameters such as phenological phase, instrument and flight parameters (e.g. pulse repetition rate, scan angle, flying height, aircraft speed) and processing of the full waveform signal (cf. chapter 1.3). There are several broad categories of descriptive features that can be derived from raw ALS data (Koenig and Höfle, 2016; Fassnacht et al., 2016; Lin and Hyppä, 2016; Shi et al., 2018b):

- **Spatial point pattern statistics** (e.g. density, spacing between points, height distribution).
- **Radiometric statistics** (e.g. intensity/amplitude distribution).
- **Opacity statistics** (e.g. echo ranks, counts and width distribution)
- **Temporal change** of any of the above (due to phenology, growth, death, disease).

Moreover, if individual tree segmentation (cf. chapter 3) is applied, the tree structure can be described explicitly in terms of:

- **External crown features** (e.g. height, volume, surface, surface-area, convexity, roughness, fractal dimension).
- **Internal crown features** (e.g. branch density, branch lengths, branch order, branch diameter, branching angles, lacunarity).
- **Temporal change** of any of the above (due to phenology, growth, death, disease).

External crown characteristics can be derived through volumetric representations of the point cloud such as alpha shapes (Edelsbrunner and Mücke, 1994) or voxels. Some internal crown characteristics may be estimated by examining the return intensity (amplitude), echo-width and fraction of different echo ranks or counts (e.g. fraction of last returns). Other internal crown features may be derived from a graph representation (Parkan and Tuia, 2015) or by further subdividing the tree segment into subcomponents (Harikumar et al., 2017b). However, it should be noted that computing internal crown features require extensive sampling of the branching structure (preferably in leaf-off conditions), thus restricting their use to very high density ALS.

The creation of thematic maps (e.g. genus, species, healthy/dead, etc) from ALS and/or AHS derived features involves the use of classification models. Of particular interest here is the use of statistical learning (also known as machine learning). Unlike deterministic methods which require an explicit definition of the classification model, statistical learning methods are able to learn a model and predict classes based on data derived features alone. The use of these statistical learning methods requires:

- The **definition of distinct thematic classes** of interest (e.g. genus, species, health). If they are known, these classes may be assigned explicitly to observations (i.e. labeling) by an expert prior to training the classification algorithm (i.e. supervised learning). Conversely, they may be assigned by the algorithm itself, if they are unknown (i.e. unsupervised learning). A less commonly used family of classification algorithms is able to train on a minimal set of labeled observations combined with unlabeled observations (i.e. semi-supervised learning).
- The **definition of a minimal set of discriminative features (i.e. descriptive variables)** that allows an unambiguous differentiation of classes. This set of features may be defined based on expert knowledge, it may be obtained from a list of predefined generic features using automatic feature selection techniques (Guyon and Elisseeff, 2003) or it may be learned from the raw data by feature learning algorithms.
- For some methods (e.g. support vector machines), the **definition of similarity metrics** (kernel functions) which are used to compare feature values is necessary.

Due to their superior performance and because the target classes are general known, supervised learning algorithms are by far the most commonly used for the classification of tree species. Technically speaking, the supervised learning process works by statistically evaluating the similarity between features characterizing the unlabeled and labeled data. A classification model is produced by assigning a weight to each of the features as a function of the discriminatory information they bring to distinguish labeled samples. For example, a feature related to green color intensity would likely be assigned a large weight, when attempting to discriminate deciduous and broadleaf trees in leaf off conditions. The main advantages of models produced with supervised learning is that they can be continuously improved by the addition of new training samples and they do not require any a priori assumptions about the weights of descriptive features. Currently, the main supervised learning algorithms applied to tree species classification are: Support Vector Machines (SVM), Random Forests (RF), Quadratic/Linear Discriminant Analysis (Q/LDA), Maximum Likelihood Estimation (MLE), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN). Example usage of these algorithms is provided in table 5.1. The respective suitability and performances of these algorithms are being actively researched (Vyas et al., 2011; Ørka et al., 2012; Dalponte et al., 2012; Féret and Asner, 2013; Ghosh et al., 2014b; Omer et al., 2015; Deng et al., 2016; Zhang et al., 2016b; Ballanti et al., 2016; Raczko and Zagajewski, 2017; Piiroinen et al., 2017; Tuominen et al., 2018). Thanks to their ability to handle non-linear classification problems and to their ease of use, Random forests and Support Vector Machine are currently the most popular and have been shown to have similar performances (Ørka et al., 2012; Dalponte et al., 2012; Ghosh et al., 2014b; Ballanti et al., 2016; Piiroinen et al., 2017) with SVM slightly surpassing RF in some cases (Deng et al., 2016; Raczko and Zagajewski, 2017). The use of convolutional neural networks for tree detection and species classification has not been extensively investigated, but initial results indicate that it is a promising approach (Guan et al., 2015; Li et al., 2016; Cheang et al., 2017; Zou et al., 2017; Han et al., 2017; Hamraz et al., 2018; Ayrey and Hayes, 2018; Trier et al., 2018; Ko et al., 2018).

Overall, a large part of the work load when using supervised statistical learning methods is the creation of relevant discriminant features from the raw data, the definition of adequate similarity metrics and the labeling of a large number of training samples. This last point is generally the

most tedious, since it is a manual process, often the result of repeated visual interpretation and/or field observations. For this purpose, the use of crowd sourced (collaborative) labeling approaches to produce very large numbers of labeled samples could be an interesting solution. A form of parsimonious crowd sourcing could also be employed, through the use of active learning methods which only require the labeling of a minimal set of examples (Persello et al., 2014). In addition to the critical characteristics of ALS and AHI data presented in chapter 1.3 and 1.4, several forest characteristics also influence the difficulty of classification including:

- The number of different thematic classes.
- The number of individuals for each thematic class and the fraction of observations used for calibrating the model (Baldeck and Asner, 2014).
- The number of different age classes which relates to the diversity of height and shape (Buddenbaum et al., 2005; Nordkvist et al., 2012; Hovi et al., 2016).
- The spectral and structural separability of the considered classes.
- Differences in reflectance characteristics within a class due to local environmental conditions, phenological phase, genetic differences, damage.
- Differences in structural characteristics within a class due to local environmental conditions, phenological phase, genetic differences, damage.

Due to these differences, comparing the results of tree species classification in different studies is difficult and should be done with attention to details.

Table 5.1: Main supervised classification algorithms used for tree species detection/classification.

<b>Algorithm</b>	<b>Usage references</b>
Random Forests (RF) (Breiman, 2001)	Ørka et al. (2012), Dalponte et al. (2012), Yu et al. (2014), Ghosh et al. (2014b), Ballanti et al. (2016), Piiroinen et al. (2017), Shen and Cao (2017), Dechesne et al. (2017)
Support Vector Machines (SVM) (Vapnik (1995), Schölkopf et al. (2002))	Liu et al. (2011), Ørka et al. (2012), Dalponte et al. (2012), Féret and Asner (2013), Ghosh et al. (2014b), Baldeck et al. (2015), Omer et al. (2015), Matsuki et al. (2015) Ballanti et al. (2016), Lin and Hyppä (2016), Piiroinen et al. (2017)
Artificial Neural Networks (ANN)	Féret and Asner (2013), Omer et al. (2015) Raczko and Zagajewski (2017)
Quadratic/Linear Discriminant Analysis (Q/LDA)	Holmgren and Persson (2004a), Brandtberg (2007), Ørka et al. (2007), Kim et al. (2009), Puttonen et al. (2009), Suratno et al. (2009), Ørka et al. (2012), Féret and Asner (2013), Lindberg et al. (2014), Hovi et al. (2016)
Convolutional Neural Networks (CNN) (LeCun et al., 2015)	Guan et al. (2015), Li et al. (2016), Cheang et al. (2017), Zou et al. (2017), Han et al. (2017), Hamraz et al. (2018), Ayrey and Hayes (2018), Trier et al. (2018)

Table 5.2: Selected publications covering tree classification based on airborne laser scanning and/or multi/hyperspectral imaging. The terrestrial ecoregion classification follows the global map of Olson et al. (2001). **Abbreviations:** **DR** = discrete return, **FW** = full waveform, **MS** = multispectral, **HS** = hyperspectral, **Sp.** = number of individually classified species, **Gr.** = number of classified groups (e.g. broadleaf/coniferous, genera, associations).

Ecozone	Ecoregion	Reference	ALS		Imaging		Classes	
			DR	FW	MS	HS	Sp.	Gr.
Palearctic	Scandinavian and Russian taiga	Korpela et al. (2007)	✓	✗	✓	✗	2	1
		Ørka et al. (2007)	✓	✗	✗	✗	3	-
		Liang et al. (2007)	✓	✗	✗	✗	-	2
		Säynäjoki et al. (2008)	✓	✗	✓	✗	1	1
		Vauhkonen et al. (2009)	✓	✗	✗	✗	2	1
		Tooke et al. (2009)	✓	✗	✓	✗	1	1
		Ørka et al. (2009)	✓	✗	✗	✗	-	2
		Puttonen et al. (2009)	✓	✗	✓	✗	3	-
		Heikkinen et al. (2011)	✗	✗	✓	✗	3	-
		Nordkvist et al. (2012)	✓	✗	✓	✗	-	7
		Ørka et al. (2012)	✓	✗	✓	✗	2	1
		Ørka et al. (2013)	✓	✗	✓	✓	2	1
		Dalponte et al. (2013)	✓	✗	✗	✓	3	1
		Dalponte et al. (2014)	✓	✗	✗	✓	2	1
		Schumacher and Nord-Larsen (2014)	✓	✗	✓	✗	-	2
		Yu et al. (2014)	✗	✓	✗	✗	3	-
		Lin and Hyppä (2016)	✓	✗	✗	✗	4	-
		Hovi et al. (2016)	✗	✓	✗	✗	3	-
		Yu et al. (2017)	✓	✗	✗	✗	3	-
		Tuominen et al. (2017)	✗	✗	✗	✓	14	-
		Blomley et al. (2017)	✗	✓	✗	✗	3	-
		Trier et al. (2018)	✗	✓	✗	✓	3	-

Table 5.3: Selected publications covering tree classification based on airborne laser scanning and/or multi/hyperspectral imaging. The terrestrial ecoregion classification follows the global map of Olson et al. (2001). **Abbreviations:** **DR** = discrete return, **FW** = full waveform, **MS** = multispectral, **HS** = hyperspectral, **Sp.** = number of individually classified species, **Gr.** = number of classified groups (e.g. broadleaf/coniferous, genera, associations).

Ecozone	Ecoregion	Reference	ALS		Imaging		Classes	
			DR	FW	MS	HS	Sp.	Gr.
Palearctic	Western European broadleaf forests	Buddenbaum et al. (2005)	×	×	×	✓	-	6
		Reitberger et al. (2006)	×	✓	×	×	-	2
		Reitberger et al. (2008b)	×	✓	×	×	-	2
		Waser et al. (2008)	×	×	✓	×	5	-
		Waser et al. (2010)	×	×	✓	×	8	-
		Heinzel and Koch (2011a)	×	✓	×	×	6	2
		Waser et al. (2011)	✓	×	✓	×	9	-
		Heinzel and Koch (2012)	×	✓	✓	✓	4	-
		Yao et al. (2012)	×	✓	×	×	2	-
		Engler et al. (2013)	×	×	✓	×	6	-
		Ghosh et al. (2014a)	✓	×	×	✓	5	-
		Fassnacht et al. (2014)	×	×	×	✓	7	-
		Torabzadeh (2016)	✓	×	×	✓	7	1
		Sommer et al. (2016)	×	✓	×	✓	13	-
		Bruggisser et al. (2017)	×	✓	×	×	3	-
Sarmatic mixed forests		Dechesne et al. (2017)	✓	×	✓	×	-	6
		Raczko and Zagajewski (2017)	×	×	×	✓	5	-
		Shi et al. (2018b)	×	✓	×	×	6	-
		Shi et al. (2018a)	×	✓	×	✓	5	-
		Holmgren and Persson (2004a)	✓	×	×	×	2	-
		Persson et al. (2004)	✓	×	✓	×	2	1
		Holmgren et al. (2008)	✓	×	✓	×	2	1
		Dinuls et al. (2012)	✓	×	✓	×	5	-
		Lindberg et al. (2014)	×	✓	×	×	5	1

Table 5.4: Selected publications covering tree classification based on airborne laser scanning and/or multi/hyperspectral imaging. The terrestrial ecoregion classification follows the global map of Olson et al. (2001). **Abbreviations:** **DR** = discrete return, **FW** = full waveform, **MS** = multispectral, **HS** = hyperspectral, **Sp.** = number of individually classified species, **Gr.** = number of classified groups (e.g. broadleaf/coniferous, genera, associations).

Ecozone	Ecoregion	Reference	ALS		Imaging		Classes	
			DR	FW	MS	HS	Sp.	Gr.
Alps conifer and mixed forests		Dalponte et al. (2009)	✓	✗	✗	✓	20	3
		Hollaus et al. (2009b)	✗	✓	✗	✗	3	-
		Dalponte et al. (2012)	✓	✗	✓	✓	6	2
		Kandare et al. (2017)	✓	✗	✗	✓	4	1
		Kukunda et al. (2018)	✓	✗	✓	✗	2	-
Atlantic mixed forests		Geerling et al. (2007)	✓	✗	✓	✗	-	8
		Hantson et al. (2012)	✓	✗	✗	✓	6	-
		Van Coillie et al. (2014)	✗	✗	✗	✓	7	-
		Laslier et al. (2017)	✓	✗	✗	✗	8	-
Palearctic	English Lowlands beech forests	Koukoulas and Blackburn (2005)	✓	✗	✓	✗	3	1
		Hill et al. (2010)	✗	✗	✓	✗	6	-
		Lee et al. (2016)	✓	✗	✗	✓	6	1
	Central European mixed forests	Tigges et al. (2013)	✗	✗	✓	✗	8	-
		Richter et al. (2016)	✗	✗	✗	✓	10	-
	Changjiang Plain evergreen forests	Cao et al. (2016)	✗	✓	✗	✗	6	-
		Shen and Cao (2017)	✓	✗	✗	✓	5	-
Manchurian mixed forests		Liu et al. (2011)	✗	✓	✓	✗	5	-
		Dian et al. (2015)	✗	✗	✗	✓	5	-
	Pannonian mixed forests	Immitzer et al. (2012)	✗	✗	✓	✗	10	-
Po Basin mixed forests		Verlič et al. (2014)	✓	✗	✓	✗	5	-
		Dalponte et al. (2008)	✓	✗	✗	✓	20	-
		Barilotti et al. (2009)	✓	✗	✗	✗	-	2

Table 5.5: Selected publications covering tree classification based on airborne laser scanning and/or multi/hyperspectral imaging. The terrestrial ecoregion classification follows the global map of Olson et al. (2001). **Abbreviations:** **DR** = discrete return, **FW** = full waveform, **MS** = multispectral, **HS** = hyperspectral, **Sp.** = number of individually classified species, **Gr.** = number of classified groups (e.g. broadleaf/coniferous, genera, associations).

Ecozone	Ecoregion	Reference	ALS		Imaging		Classes	
			DR	FW	MS	HS	Sp.	Gr.
Palearctic	Taiheiyo evergreen forests	Sasaki et al. (2012)	✓	✗	✓	✗	10	5
		Matsuki et al. (2015)	✓	✗	✗	✓	16	-
	Huang He Plain mixed forests	Li et al. (2015)	✗	✗	✓	✗	4	-
	Nihonkai montane deciduous forests	Deng et al. (2017)	✓	✗	✓	✗	13	-
	Northeast Spain and South France Mediter. forests	Dunford et al. (2009)	✗	✗	✓	✗	4	1
		Sheeren et al. (2011)	✗	✗	✗	✓	12	-
	Puget lowland forests	Kim et al. (2009)	✓	✗	✗	✗	-	2
		Jones et al. (2010)	✓	✗	✗	✓	11	-
		Jones et al. (2011)	✓	✗	✗	✓	9	-
		Vaughn et al. (2011)	✗	✓	✗	✗	3	-
		Vaughn et al. (2012)	✗	✓	✗	✗	5	-
		Zhang et al. (2016b)	✓	✗	✗	✓	7	-
		Liu et al. (2017)	✓	✗	✗	✓	15	-
Nearctic	Eastern forest-boreal transition	Li et al. (2013)	✓	✗	✗	✗	4	-
		Ko et al. (2013)	✓	✗	✗	✗	3	-
	Appalachian mixed mesophytic forests	Key et al. (2001b)	✗	✗	✓	✗	4	-
		Brandtberg (2007)	✓	✗	✗	✗	3	-
	California interior chaparral and woodlands	Ballanti et al. (2016)	✓	✗	✗	✓	8	-
		Zhou et al. (2017)	✗	✓	✗	✗	3	1
	Central tall grasslands	Sugumaran and Voss (2007)	✓	✗	✗	✓	9	-
		Voss and Sugumaran (2008)	✓	✗	✗	✓	7	-

Table 5.6: Selected publications covering tree classification based on airborne laser scanning and/or multi/hyperspectral imaging. The terrestrial ecoregion classification follows the global map of Olson et al. (2001). **Abbreviations:** **DR** = discrete return, **FW** = full waveform, **MS** = multispectral, **HS** = hyperspectral, **Sp.** = number of individually classified species, **Gr.** = number of classified groups (e.g. broadleaf/coniferous, genera, associations).

Ecozone	Ecoregion	Reference	ALS		Imaging		Classes	
			DR	FW	MS	HS	Sp.	Gr.
Nearctic	Southeastern conifer forests (Nearctic)	Pu and Landry (2012)	✗	✗	✓	✗	4	4
		Zhang et al. (2013)	✓	✗	✓	✗	-	7
	New England-Acadian forests	Anderson et al. (2011)	✗	✓	✗	✓	1	-
	Allegheny Highlands forests	Ke et al. (2010a)	✓	✗	✗	✓	4	1
	South Central Rockies forests	Suratno et al. (2009)	✗	✓	✗	✗	4	-
	Appalachian-Blue Ridge forests	Fang et al. (2018)	✓	✗	✓	✗	4	-
	California montane chaparral and woodlands	Alonzo et al. (2014)	✗	✓	✗	✓	29	-
	Canadian Aspen forests and parklands	Bork and Su (2007)	✓	✗	✓	✗	-	6
	Sierra Nevada forests	Swatantran et al. (2011)	✓	✗	✗	✓	4	2
Afrotropic	Zambezian and Mopane woodlands	Texas blackland prairies	✓	✗	✗	✓	40	-
		Upper Midwest forest-savanna transition	✓	✗	✗	✓	3	1
		Cho et al. (2010)	✗	✓	✗	✓	10	-
		Colgan et al. (2012)	✗	✗	✗	✓	15	-
		Cho et al. (2012)	✗	✓	✓	✓	5	1
		Naidoo et al. (2012)	✓	✗	✗	✓	8	1
		Sarrazin et al. (2012)	✗	✓	✗	✓	4	-
	Eastern Arc forests	Baldeck et al. (2014)	✓	✗	✗	✓	15	-
	Piiroinen et al. (2017)	Piiroinen et al. (2017)	✓	✗	✗	✓	31	-

Table 5.7: Selected publications covering tree classification based on airborne laser scanning and/or multi/hyperspectral imaging. The terrestrial ecoregion classification follows the global map of Olson et al. (2001). **Abbreviations:** **DR** = discrete return, **FW** = full waveform, **MS** = multispectral, **HS** = hyperspectral, **Sp.** = number of individually classified species, **Gr.** = number of classified groups (e.g. broadleaf/coniferous, genera, associations).

Ecozone	Ecoregion	Reference	ALS		Imaging		Classes	
			DR	FW	MS	HS	Sp.	Gr.
Neotropic	Isthmian-Atlantic moist forests	Clark (2005)	✓	✗	✗	✓	7	-
		Clark and Roberts (2012)	✗	✗	✗	✓	7	-
		Baldeck et al. (2015)	✓	✗	✗	✓	3	-
	Alto Paraná Atlantic forests	Ferreira et al. (2016)	✗	✗	✓	✓	8	-
	Isthmian-Pacific moist forests	Graves et al. (2016)	✗	✓	✗	✓	20	1
Australasia	Tumbes-Piura dry forests	Baena et al. (2017)	✗	✗	✓	✗	3	-
	Brigalow tropical savanna	Moffiet et al. (2005)	✓	✗	✗	✗	-	4
		Lucas et al. (2008)	✓	✗	✗	✓	8	-
	Southeast Australia temperate forests	Zhang and Liu (2013)	✓	✗	✗	✗	2	-
	Northland temperate kauri forests	Pham et al. (2016)	✓	✗	✓	✗	1	-
Oceania	Hawaii tropical moist forests	Asner et al. (2008)	✗	✓	✗	✓	4	-
		Féret and Asner (2012)	✓	✗	✗	✓	9	-
		Féret and Asner (2013)	✓	✗	✗	✓	17	12
Indo-Malaya	Himalayan subtropical broadleaf forests	Karna et al. (2015)	✓	✗	✓	✗	6	-

## 5.2 Error assessment framework

The classification error indicates how well an algorithm is able to correctly assign discrete classes (e.g. species, healthy/diseased, live/dead) to observations (Congalton and Green, 2008; Sokolova and Lapalme, 2009). The error can be evaluated at the point or tree scale. In both cases, the assessment procedure involves the construction of a confusion matrix (cf. figure 5.2) and the computation of performance metrics (cf. table 5.8). The interpretation of the confusion matrix is the primary mean of error assessment, while generic scores (such as overall accuracy and kappa scores) should be regarded as partial summary metrics that should not be interpreted independently.

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

Figure 5.2: Multi-class confusion matrix. Abbreviations: C: class, N: total number of classes,  $n_{obs}$ : number of observations,  $n_{i,j}$  the number of times point/tree with class  $i$  was assigned to class  $j$ .

Table 5.8: Multi-class classification performance metrics.

Metric	Formula	Interpretation
Overall Accuracy	$OA = \frac{\sum_{i=1}^N n_{i,i}}{n_{obs}} \in [0, 1]$	Overall performance ignoring class occurrence frequencies
Kappa coefficient	$K = \frac{p_o - p_e}{1 - p_e} \in [-1, 1]$ observed accuracy $p_o$ : $p_o = OA$ expected accuracy $p_e$ : $p_e = \frac{1}{n_{obs}^2} \sum_{i=1}^N \left( \sum_{j=1}^N n_{i,j} \sum_{j=1}^N n_{j,i} \right)$	Overall performance considering class occurrence frequencies (comparison to chance class assignment)
Recall (Producer's Accuracy) for class $i$	$r(i) = \frac{n_{i,i}}{\sum_{j=1}^N n_{i,j}} \in [0, 1]$	Tendency to detect (sensitivity) class $i$ among N classes
Precision (User's Accuracy) for class $i$	$p(i) = \frac{n_{i,i}}{\sum_{j=1}^N n_{j,i}} \in [0, 1]$	Fraction of the detected class $i$ occurrences that were correctly assigned
F score for class $i$	$F(i) = 2 \cdot \frac{p(i) \cdot r(i)}{p(i) + r(i)} \in [0, 1]$	Harmonic mean of $r$ and $p$ .
Average Precision (Average Accuracy)	$\bar{p} = \frac{1}{N} \sum_{i=1}^N p(i) \in [0, 1]$	Average of per class precisions

### 5.3 General tree characteristics

The following section presents some general structural and spectral separability characteristics for the main tree genus/species encountered in Swiss forests.

#### 5.3.1 Structural separability

Many trees have a distinct structure (cf. tables 5.9 to 5.10) which can vary significantly depending on age and growth conditions (i.e. shape plasticity). When measuring trees with ALS, their structural characteristics affect the laser beam by locally modifying the opacity of space. Depending on the size of the laser beam's cross-section (typically 10-30 cm for long range ALS) and the size of the intercepted structure, the beam may be partially or totally reflected. In essence, this means that large/opaque structures have a higher probability of generating a single return per laser pulse (or a last return) and a high return intensity (cf. figures 5.3 to 5.5). This is apparent for example with the dense branching and foliage of persistent coniferous trees or with large stems or branches. Conversely, the thin branches of a birch tree scanned in leaf off conditions are more likely to generate multiple returns per pulse with a low return intensity.

Table 5.9: General tree shapes characteristics for selected persistent foliage species encountered in Switzerland (Johnson, 2006; Fischesser et al., 2008). Note that these should be considered as trends rather than absolute identification criteria.

<b>Latin name</b>	<b>Vernacular name</b>	<b>Foliage</b>	<b>Characteristics</b>
<i>Pseudotsuga menziesii</i>	Douglas fir	Persistent	Crown: dense, radial symmetry Tip: pointy Stem: long, straight Branching: horizontal Max. Height: 60 m
<i>Abies alba</i>	Silver fir	Persistent	Crown: dense, radial symmetry Tip: pointy (flat on large trees) Stem: long, straight Branching: horizontal Max. Height: 60 m
<i>Picea abies</i>	Norway spruce	Persistent	Crown: dense, radial symmetry Tip: pointy (flat on large trees) Stem: long, straight Branching: horizontal Max. Height: 60 m
<i>Pinus sylvestris</i>	Scots pine	Persistent	Crown: sparse, radial assymetry Tip: broad Stem: curved Max. Height: 40 m

Table 5.10: General tree shapes characteristics for selected deciduous foliage species encountered in Switzerland (Johnson, 2006; Fischesser et al., 2008). Note that these should be considered as trends rather than absolute identification criteria.

<b>Latin name</b>	<b>Vernacular name</b>	<b>Foliage</b>	<b>Characteristics</b>
<i>Larix decidua</i>	Larch	Deciduous	Crown: dense Tip: pointy Stem: long, straight Branching: vertical Max. Height: 40 m
<i>Fagus sylvatica</i>	European beech	Deciduous	Crown: dense Tip: broad Stem: straight Branching: forked Max. Height: 40 m
<i>Quercus sp.</i>	Oak	Deciduous	Crown: dense Tip: broad Branching: sinuous Max. Height: 25-35 m
<i>Castanea sativa</i>	Chestnut	Deciduous	Crown: dense Tip: round dome Max. Height: 30 m
<i>Fraxinus excelsior</i>	European ash	Deciduous	Crown: open, sparse Tip: broad Stem: straight Branching: vertical, forked Max. Height: 35 m
<i>Acer sp.</i>	Maple	Deciduous	Crown: dense Tip: broad, pyramidal Max. Height: 20-30 m
<i>Betula pendula</i>	Birch	Deciduous	Crown: sparse Tip: pointy Stem: straight Branching: thin Max. Height: 25 m
<i>Tilia cordata</i>	Lime	Deciduous	Crown: dense Tip: round dome Stem: straight Branching: arched Max. Height: 35 m

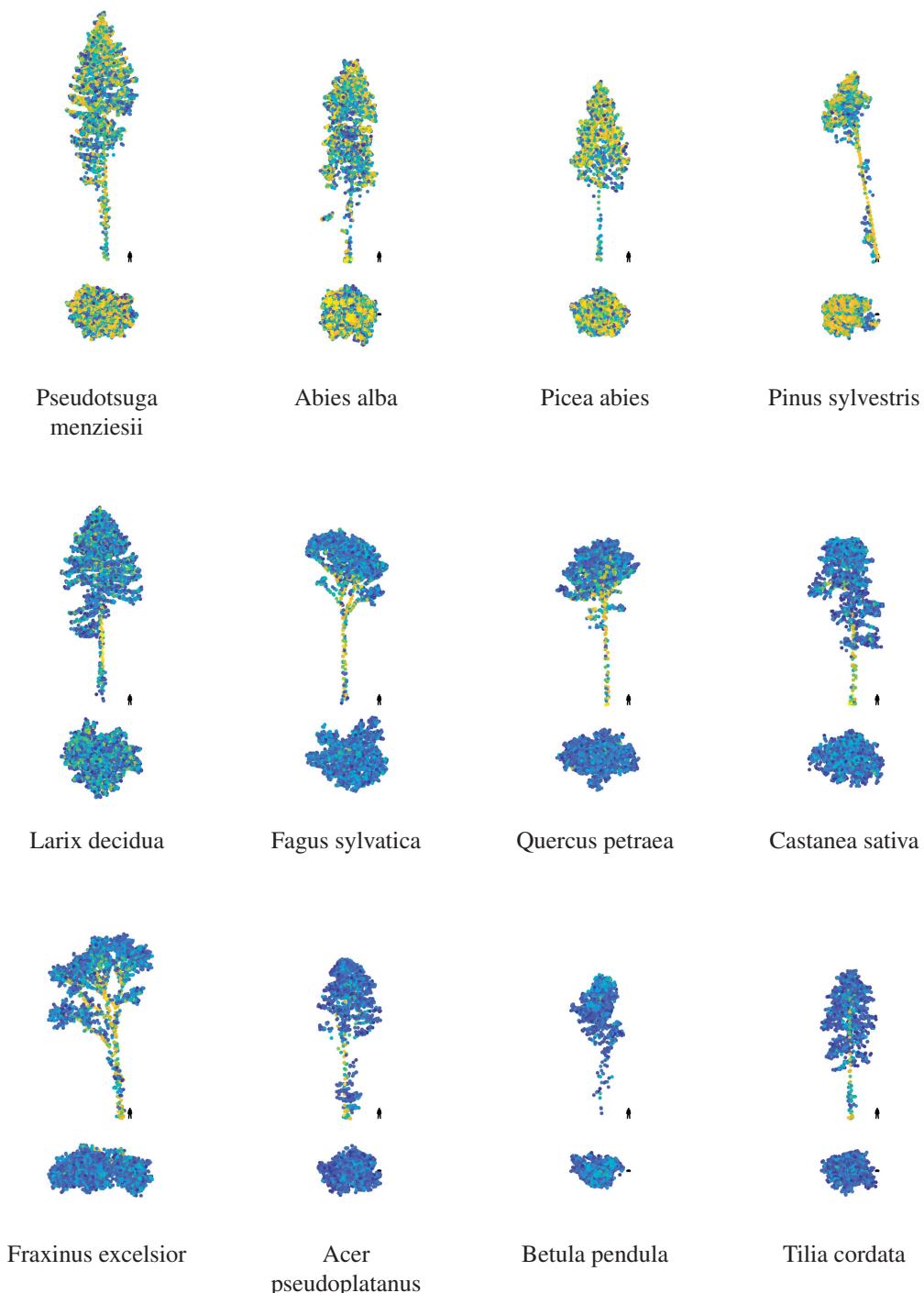


Figure 5.3: Examples (side and top view) of frequently encountered species. **The color represents laser return intensity (leaf-off acquisition).** Note that the crowns of deciduous species have a low return intensity and that opaque or large structures (stems, main branches) have a high return intensity. The relative height scale is given by the human silhouette (1.8 m high) next to each tree.

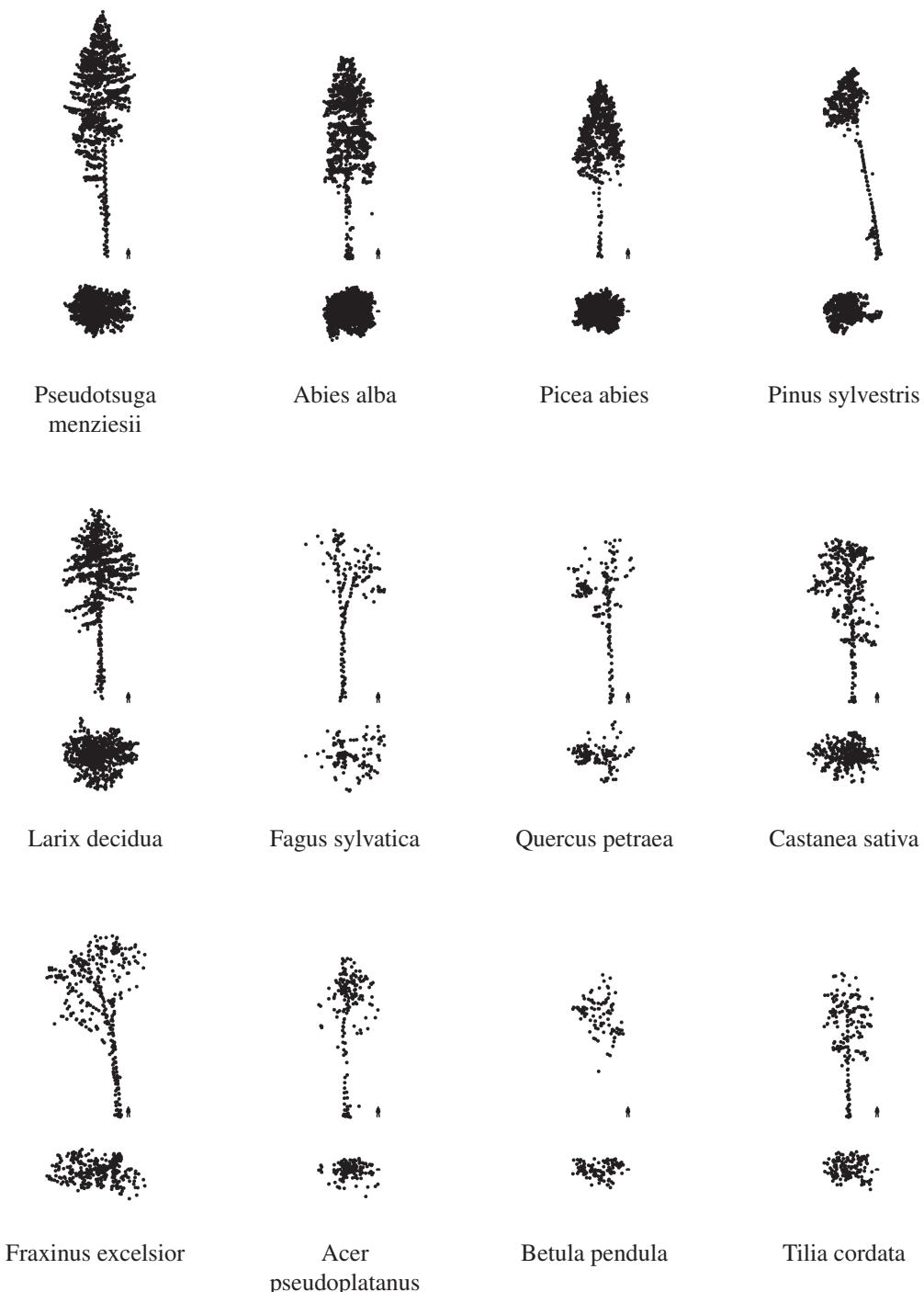


Figure 5.4: Examples (side and top view) of frequently encountered species. **Only last returns are represented (leaf-off acquisition)**. Note that the primary (large diameter) branches and stems of deciduous species are apparent. The relative height scale is given by the human silhouette (1.8 m high) next to each tree.

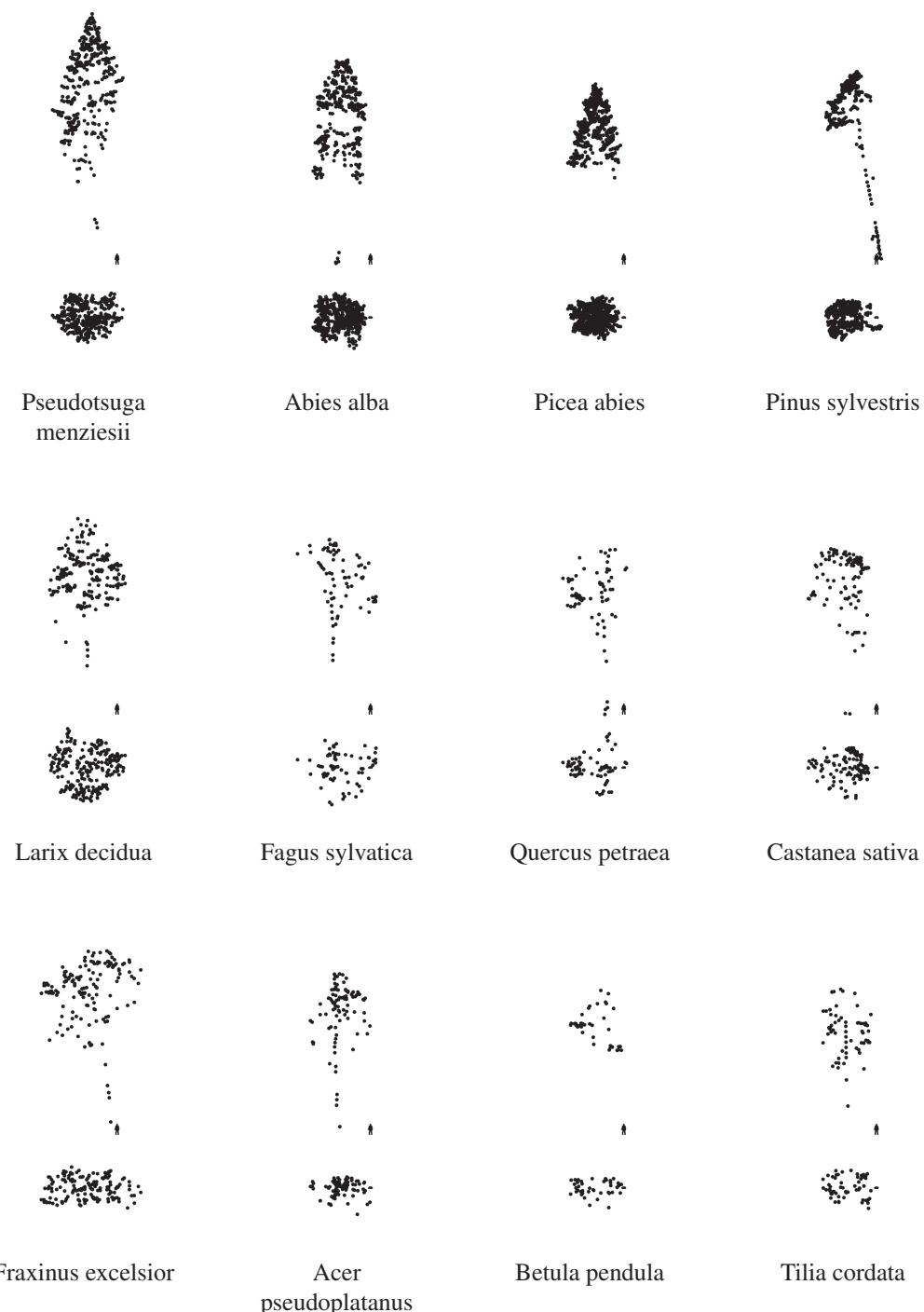


Figure 5.5: Examples (side and top view) of frequently encountered species. **Only single returns are represented (leaf-off acquisition).** Note that the primary (large diameter) branches and stems of deciduous species are apparent. The relative height scale is given by the human silhouette (1.8 m high) next to each tree.

### 5.3.2 Spectral separability

Spectral separability is a measure of dissimilarity between pure reflectance signatures (also called endmembers) that can help identify regions of the spectra useful for tree genus/species discrimination. Extracting the pure reflectance signatures of individual genus/species can be done with different approaches: the analysis of sampled leafs with a field or laboratory spectrometer, automatic endmember extraction methods (Plaza et al., 2002, 2004; Veganzones and Grana, 2008) or pure pixel isolation using a land cover map (i.e. where each cover class defines a unique spectral endmember, e.g. soil, water, tree species, unknown, etc) in a multi/hyperspectral image. Due to unavailability of field spectra and the uncertainty of automatic endmember extraction methods in the presence of highly similar endmembers, pure pixel isolation using a land cover map was used in this work. The extents of trees manually delineated from high density ALS point clouds (cf. chapter 2.2) served as a reference land cover map which was overlaid on the hyperspectral image. Based on this overlay, a pixel purity index indicating the fractional abundances of land cover classes in each pixel was computed and the median reflectance signatures of tree genus/species were subsequently determined from the purest pixels (cf. figure 5.7). The main steps of this procedure are summarized below and illustrated in figure 5.6:

1. Label individual tree crowns in the forest canopy with a genus/species class (cf. figure 5.6b) to create a reference map (cf. chapter 2.2). Tree crowns manually labeled in high density ALS data were used for this purpose (cf. chapter 2.2). The fraction of mixed class pixels depends on the shape (area, perimeter, branch density) of the tree crowns, the spacing between crowns and the homogeneity of species in the canopy layer. Thus, when working with coarse resolution pixels, to increase the fraction of pure class pixels, it is preferable to collect pixel samples in dense single species stands.
2. Create a coarse resolution label image (i.e. same resolution as the hyperspectral image) by rasterizing the labeled point cloud. Each pixel is assigned the label of the majority class of ALS points in its extent (cf. figure 5.6c).
3. Create a fine resolution label image (i.e. 1/4 resolution of the hyperspectral image) by rasterizing the labeled point cloud. Each pixel is assigned the label of the majority class of ALS points in its extent (cf. figure 5.6d).
4. Compute the class purity index defined in equation 5.1 for each pixel of the hyperspectral image.
5. Compute the mean or median reflectance signatures of each class of interest, using only pixels with a purity above a chosen level (85% was used here). The resulting median reflectance for nine genus/species of interest is present in figure 5.7.

$$P_i = \frac{C_i}{N_i} \quad (5.1)$$

where:

$C_i$  is the number of fine pixels located in the coarse pixel  $i$  that are labeled with the majority class;

$N_i$  is the number fine pixels located in the coarse pixel  $i$ .

This procedure assumes that the ALS derived genus/species reference map and hyperspectral image are spatially co-referenced. For this reason, it is imperative that the hyperspectral image be orthorectified using a surface modeled (preferably derived from the considered ALS data) and that it is co-registered with the reference land cover map. This was the case for the Airborne Prism Experiment (APEX) (Itten et al., 2008) imagery used in this study (cf. appendix D.1).

Looking at the spectral signature of individual genus/species in figure 5.7, two characteristics are apparent. First, deciduous species generally have a larger reflectance than coniferous species, in the near infrared region (Williams, 1991). Second, there is large inner species/genus reflectance variability (particularly in the near infrared region). In addition to measurement artifacts (illumination, shadowing, instrumental noise), the spectral reflectance within a species may vary considerably and depends on many biophysical factors including genetic differences, soil moisture, nutrient availability and health (Zhang et al., 2006; Castro-Esau et al., 2006; Jensen, 2007; Papeş et al., 2013; Danusevicius et al., 2014). The similar reflectance signatures of different species combined with the high within species variability result in a low spectral separability. In particular, it has been reported in previous studies that coniferous species generally have lower spectral separability than deciduous species (Roberts et al., 2004; Leckie et al., 2005; van Aardt and Wynne, 2007; Trier et al., 2018). One way of quantifying spectral separability is by computing a Separability Index (SI) which considers the ratio of the between and within class reflectance variabilities at different wavelengths (Somers and Asner, 2013):

$$SI = \frac{|\mu_{1,i} - \mu_{2,i}|}{1.96(\sigma_{1,i} + \sigma_{2,i})} \quad (5.2)$$

where:

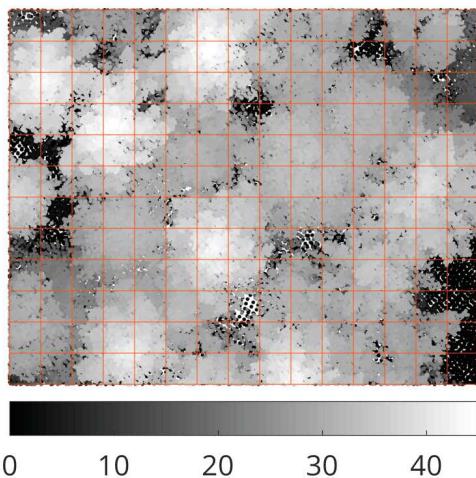
$\mu_{1,i}$  is the mean reflectance of class 1 at wavelength  $i$ ;

$\mu_{2,i}$  is the mean reflectance of class 2 at wavelength  $i$ ;

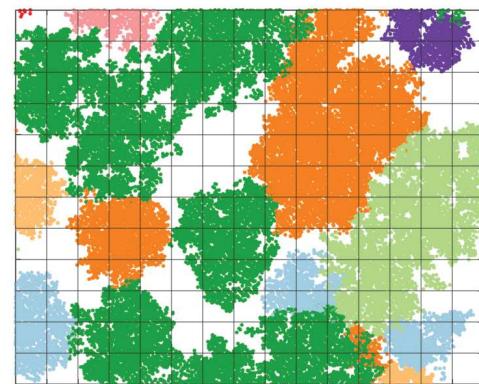
$\sigma_{1,i}$  is the standard deviation of the reflectance of class 1 at wavelength  $i$ ;

$\sigma_{2,i}$  is the standard deviation of the reflectance of class 2 at wavelength  $i$ .

This index assumes that the reflectance at given wavelength is normally distributed and that the one-sided 95% confidence interval (which explains the 1.96 factor) is representative of within class variability (Somers et al., 2010). It is conceptually similar to the signal to noise ratio or the inverse coefficient of variation. This index provides an initial indication about regions of the spectra which may be used to differentiate pairs of species. The spectral separability of the considered genus/species is illustrated in figure 5.9. As could be expected, deciduous have a large separability with coniferous species, but the separability of species within each of these groups is generally low ( $< 0.2$ ). The spectral regions with the best separability are in the blue ( $< 460$  nm), near-infrared and mid-infrared ranges, although there is a large variability in the separability index depending on the considered species pairs.



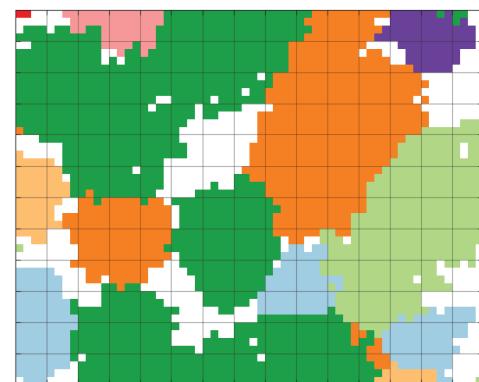
(a) ALS point cloud colored by height [m].



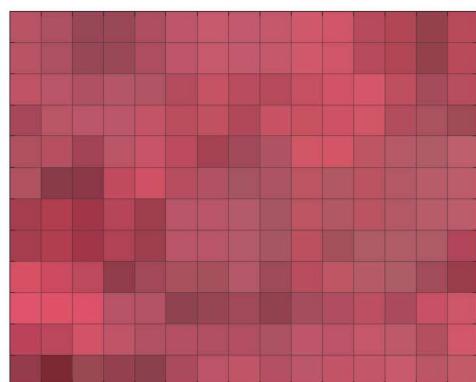
(b) ALS point cloud segments colored by genus/species class. Only labeled points are represented.



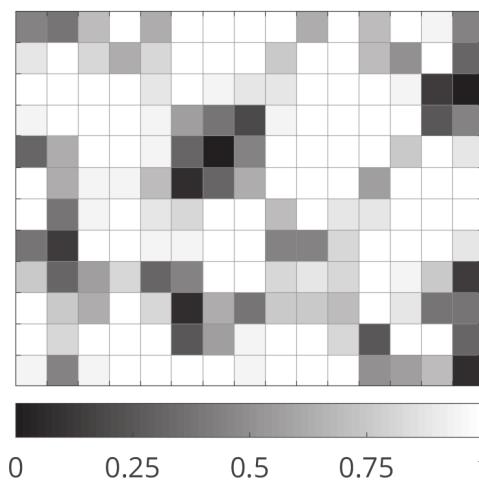
(c) Majority genus/species class at coarse raster resolution (= AHI resolution).



(d) Majority genus/species class at fine raster resolution (= 0.25·AHI resolution).



(e) False color AHI composite.



(f) Class purity index.

Figure 5.6: Pixel genus/species class purity index computation (pixel dimensions: 2.76 x 2.76 m).

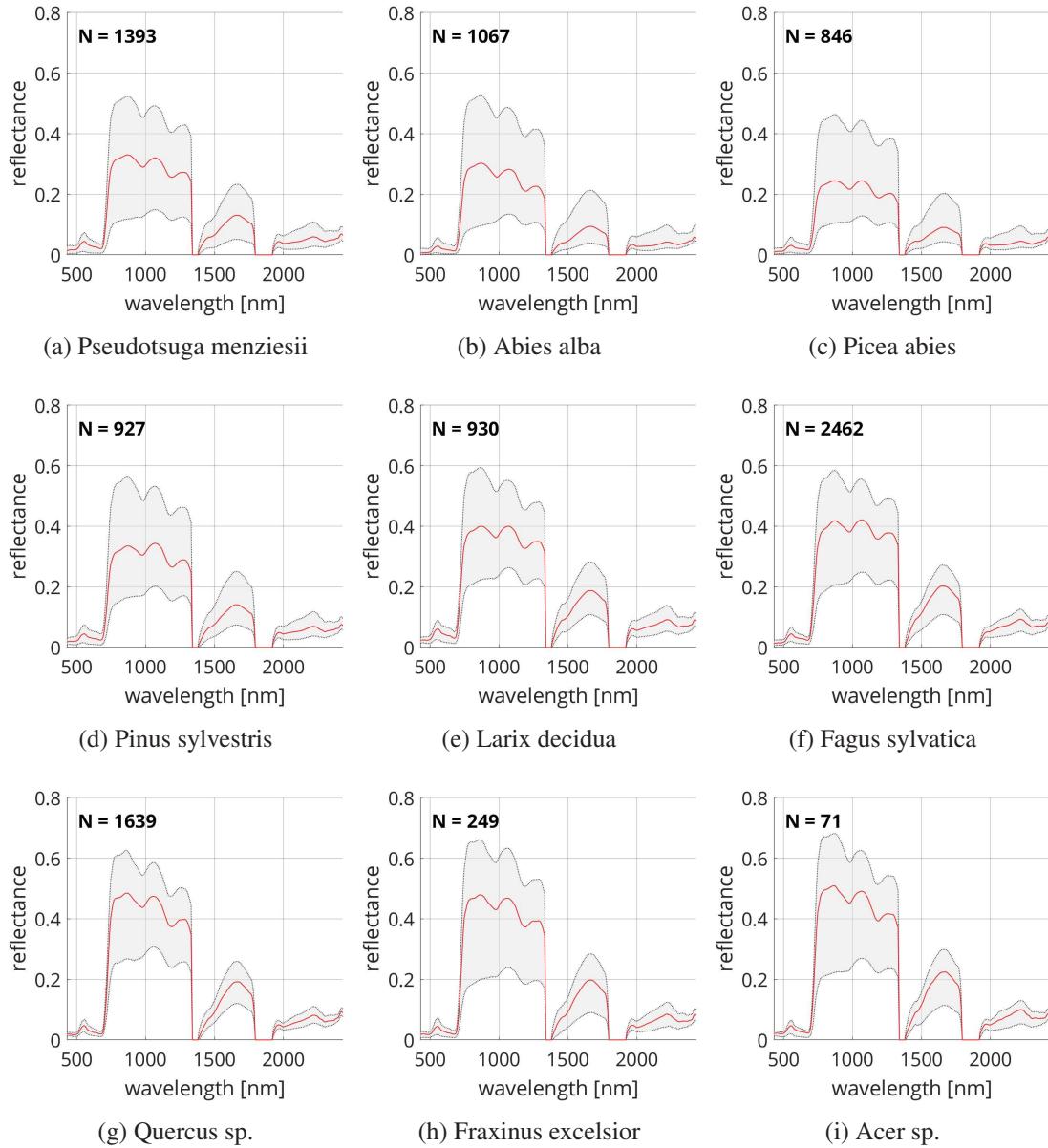


Figure 5.7: Spectral signatures for nine selected tree genus/species (mixed age classes) obtained with the APEX sensor in July 2014. The red line indicates the median reflectance for pixels with a class purity above 85%. The lower 5% and upper 95% quantile limits are represented by the gray areas. The number of pixel samples  $N$  used to produce the signature is indicated in the upper left corner of each graph.



Figure 5.8: Three different color composites for nine selected tree genus/species (mixed age classes) obtained with the APEX sensor in July 2014. The colors are based on the median values of pixels with a class purity above 85%. The last color composite (on the right) combines an estimation of carotenoid content  $CAR = (\rho_{515}^{-1} - \rho_{565}^{-1}) \cdot \rho_{780}$ , chlorophyll content  $CHL = (\rho_{550}^{-1} - \rho_{780}^{-1}) \cdot \rho_{780}$  and water content  $W = 1 - \frac{\rho_{1193}}{\rho_{1126}}$  (cf. Gitelson et al. (2006); Schneider et al. (2017)).

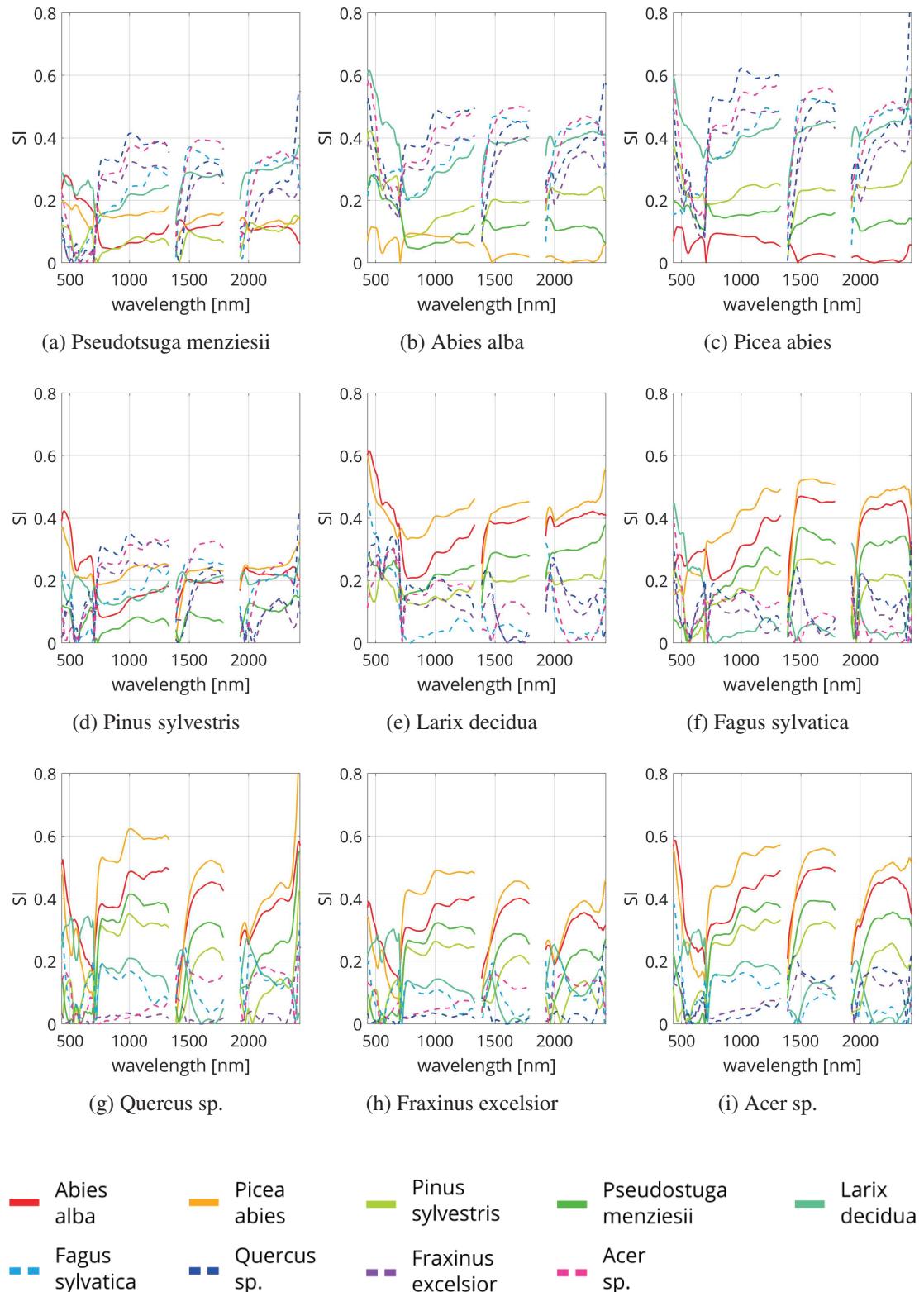


Figure 5.9: Spectral separability for nine selected tree genus/species (mixed age classes). Broadleaf and coniferous species (except larch) have mutually the highest separability index.

## 5.4 ALS based classification

This section present genus/species classification based only on features derived from ALS acquired in leaf-off ALS acquisitions (cf. Appendix E.1). It illustrate the classification of several commonly encountered genus/species (*Abies alba*, *Picea abies*, *Pseudotsuga menziesii*, *Pinus sylvestris*, *Larix decidua*, *Fagus sylvatica*, *Quercus petraea/robur*, *Fraxinus excelsior*, *Acer sp.*).

### 5.4.1 Description

As a first step, ALS measurements conducted in leaf-off conditions and trees with a height larger than 4 m were selected. Leaf-off acquisitions simplify the differentiation of deciduous and coniferous trees, and also provides a better sampling of the tree branching structure. Trees smaller than 4 m were not considered because they are typically sampled with too few points and/or do not have a distinct structure in the point cloud. To ensure reasonable training/test dataset sizes, the oaks (*Quercus petraea*, *Quercus robur*) and the maples (*Acer platanoides*, *Acer pseudoplatanus*) were grouped at the genus taxonomic level.

Because the ALS data was acquired with different sensors that produce different intensity ranges, the 1% and 99% quantile range for a given sensor were rescaled to the 0-1 range. Subsequently, the observations were grouped by species and partitioned using stratified sampling, so that the training (40%) and validation (60%) sets contain roughly the same height frequency distributions within each group (cf. figure 5.10). Two sets of descriptive features were used 5.11): one set contained only scale invariant (relative) features and the other set contained all features. This was done in a effort to reduce the risk of obtaining overly optimistic predictions due to the non-uniform tree size (height, volume) frequency distributions between species in the sample (cf. chapter 2.3). To avoid missing dependency issues, a directed graph representation of feature relations (e.g. the crown radius depends on the concave area which in turn depends on the concave hull) was created and the computation order was determined through topological sorting of the graph nodes (with depth first search).

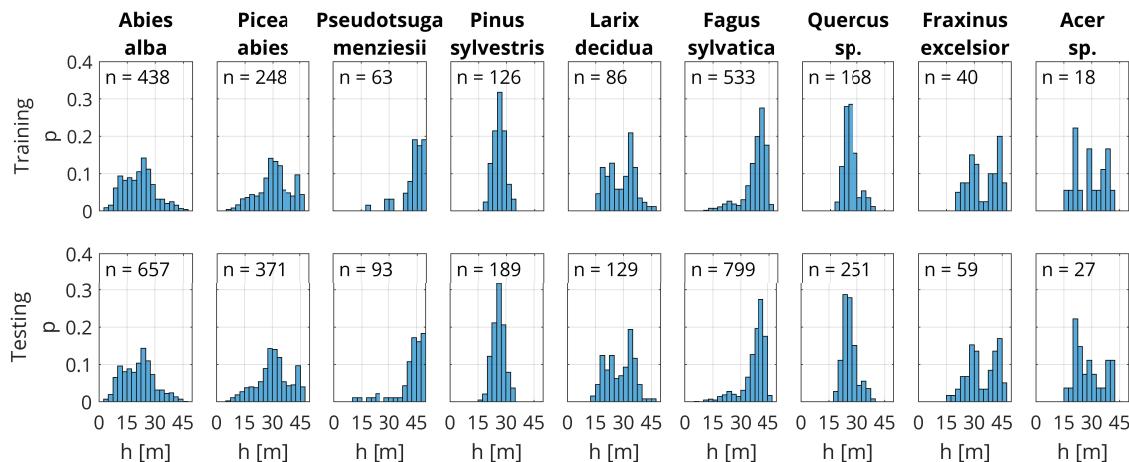


Figure 5.10: Height probability distributions for the 40% training (upper row) and 60% validation (lower row) sets, after class grouping and stratified sampling.

The Random Forest (RF) classification algorithm (Breiman, 2001; Belgiu and Drăguț, 2016) was chosen to predict species, because of its strong ability to handle non-linear classification problems, its embedded capacity to provide feature ranking, its ease of use and its efficient implementations. It has also been shown to provide similar performances than other state of the art methods such as

Support Vector Machines (SVM) (Ørka et al., 2012; Dalponte et al., 2012; Ghosh et al., 2014b; Ballanti et al., 2016; Piiroinen et al., 2017). The RF was configured to use 600 decision trees with a minimum leaf size of 1. To estimate feature importance, the values of each feature were randomly permuted in the out-of-bag observations and the average decrease in the classification margin (i.e. the difference between the predicted probability for the correct class and the maximum predicted probability for incorrect classes) was computed (cf. figure 5.14). A minimal set of features was then selected by using recursive feature elimination (i.e. recursively removing the feature with the least weight and retraining the RF model, until the prediction kappa score decreased significantly). The workflow described above is summarized in figure 5.11.

Table 5.11: ALS derived features used for species classification. **Abbreviations:**  $\sigma$ : Standard deviation,  $c_v$ : Coefficient of variation,  $\kappa$ : Kurtosis,  $\gamma_1$ : Skewness,  $Q(p)$ : Quantile for probability  $p$ .

Category	Features	Scale invariant
External shape	Total Height	×
	Variances of the three principal components (PCA)	×
	Convex ( $A_{conv}$ ) and concave ( $A_{conc}$ ) hull 2D areas	×
	Convex ( $V_{conv}$ ) and concave ( $V_{conc}$ ) hull volumes	×
	Convex ( $SA_{conv}$ ) and concave hull ( $SA_{conc}$ ) surface areas	×
	Convexity ( $= \frac{SA_{conc}}{SA_{conv}}$ )	✓
	Convex hull lacunarity ( $= \frac{V_{conv} - V_{conc}}{V_{conv}}$ )	✓
	Convex and concave specific surface ( $= \frac{SA_{conv}}{V_{conv}}$ and $= \frac{SA_{conc}}{V_{conc}}$ )	✓
	Aspect ratio $\frac{2}{height} \cdot \sqrt{\frac{A_{conv}}{\pi}}$	✓
	Number of points	×
Point pattern	Number of points on the convex ( $N_{conv}$ ) and concave ( $N_{conc}$ ) hulls	×
	Fraction of points on the convex ( $= \frac{N_{conv}}{N_{tot}}$ and concave $\frac{N_{conc}}{N_{tot}}$ ) hulls	✓
	Normalized height statistics $(\sigma, c_v, Q(0.25), Q(0.5), Q(0.75), Q(0.9), \kappa, \gamma_1)$	✓
	All and First return intensity statistics $(\sigma, c_v, Q(0.25), Q(0.5), Q(0.75), Q(0.9), \kappa, \gamma_1, max)$	✓
Intensity	Fraction of first/last/single returns	✓
	Convex and concave point density ( $= \frac{N_{tot}}{V_{conv}}$ and $\frac{N_{tot}}{V_{conc}}$ )	✓
	Median opacity (return number / number of returns)	✓

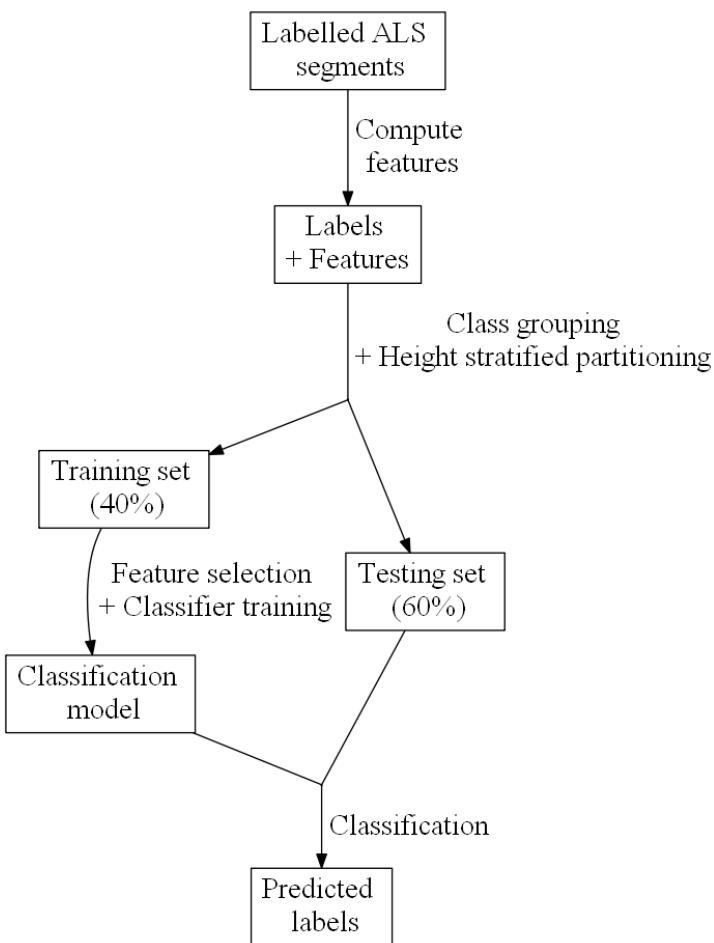


Figure 5.11: ALS classification workflow. Note that zhe same workflow was used

To evaluate the robustness of the classifier to segmentation error, it was also trained and validated on a set of tree segments with degraded quality. The following approach was used to simulate segmentation error (cf. figure 5.12):

1. Select an individual tree segment (cf. figure 5.12a). Here the manually delineated segments presented in chapter 2.2 are used.
2. Add some random noise to the points in the segment. The amount of segmentation error can be controlled by setting the amount of noise (cf. figure 5.12b). Five different levels of noise were used: 0 m, 1.5 m, 3 m, 4.5 m, 6 m.
3. Randomly select a fraction  $f$  (empirically set to 15% here) of the points in the segment (cf. figure 5.12c).
4. Compute the single region alpha shape (Edelsbrunner and Mücke, 1994) of the randomly selected points (cf. figure 5.12d).
5. Select the points of the complete point cloud that are located within the alpha shape (i.e. all points, not only those of the considered segment), cf. figure 5.12e. Examples of the simulated segmentation error are illustrated in figure 5.13.

To avoid overly optimistic predictions, all quality levels of a given segment were either assigned to the training or the test set.

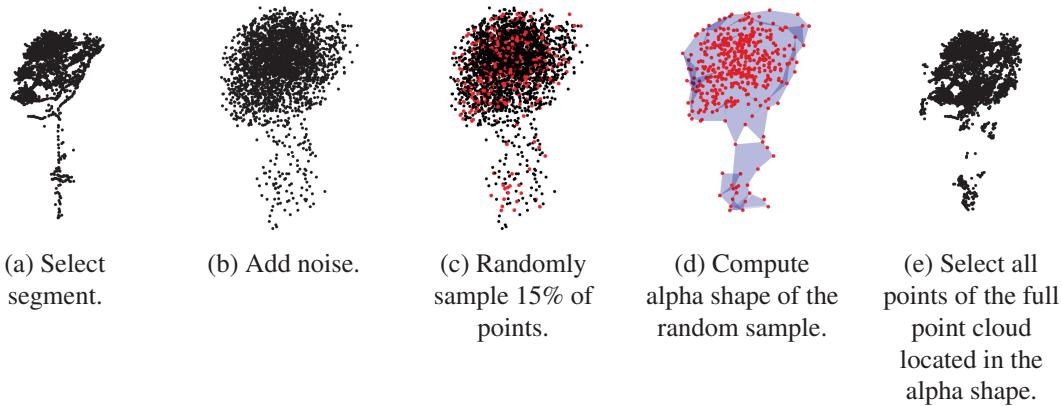


Figure 5.12: Simulation of the segmentation error.

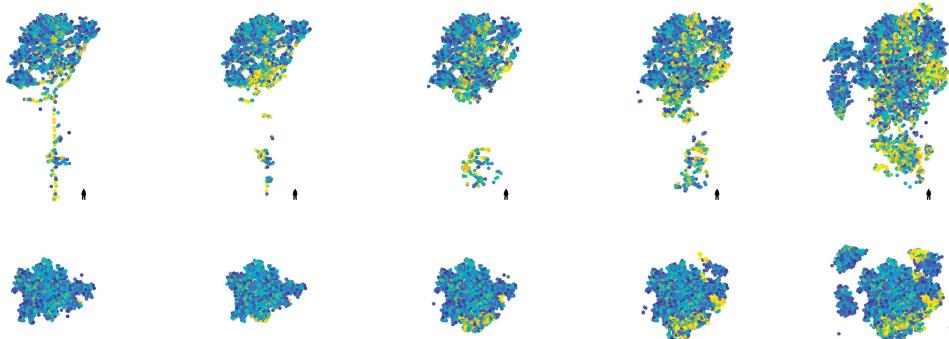
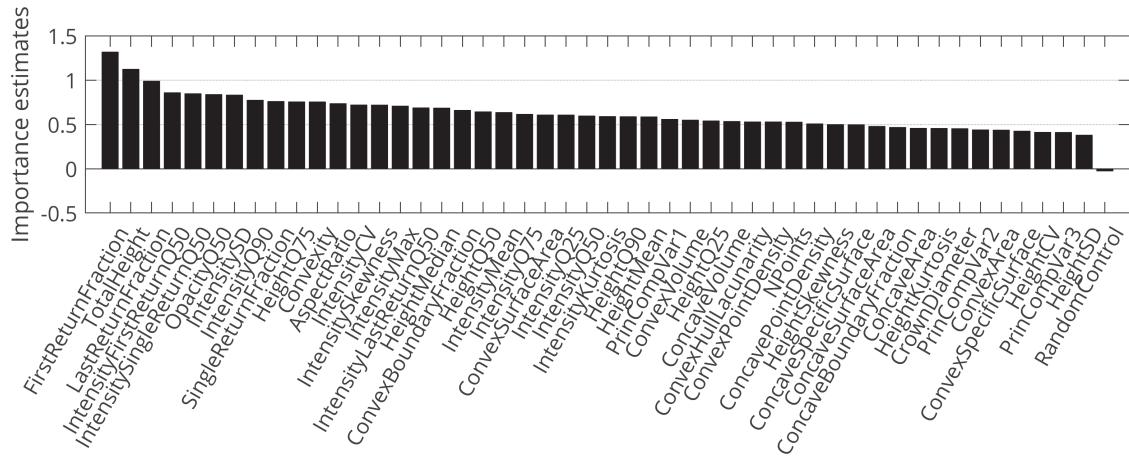


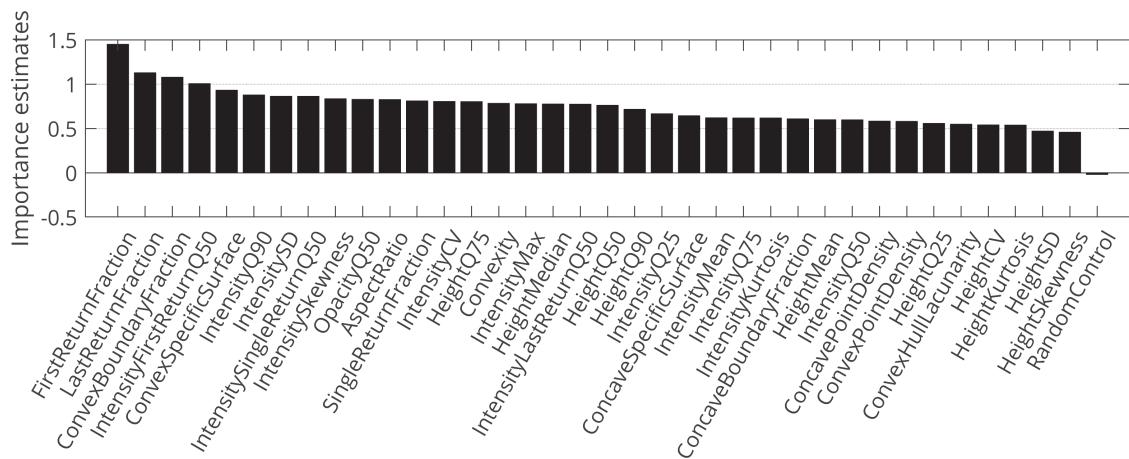
Figure 5.13: Side and top view of a tree with increasing amounts of simulated segmentation error. The color represents laser return intensity (leaf-off acquisition). The relative height scale is given by the human silhouette (1.8 m high) next to each tree.

## 5.4.2 Results

The classification was validated using the framework presented in section 5.2. The feature rankings and final sets of selected features are provided in figure 5.14. The class confusion matrices obtained with the Random Forest algorithm using all features and scale invariant features only are reported in tables 5.12 and 5.13. The performances of the classification when trained and evaluated with simulated segmentation error are reported in tables 5.14 to 5.15.



(a) Initial estimation of feature importance for all features. After recursive feature elimination and re-ranking at each iteration, the following set of 11 features was selected: Fraction of first returns, fraction of last returns, total height, 75% height quantile, 50% first returns intensity quantile, 50% single return intensity quantile, convexity, fraction of points on the convex hull, convex specific surface, intensity standard deviation, intensity skewness, intensity standard deviation, 90% intensity quantile.



(b) Initial estimation of feature importance for scale invariant features only. After recursive feature elimination and re-ranking at each iteration, the following set of 13 features was selected: Fraction of first returns, fraction of last returns, 50% intensity of first returns quantile, fraction of points on the convex hull, convex specific surface, intensity standard deviation, aspect ratio, 75% height quantile, 50% height quantile, 50% single return intensity quantile, intensity skewness, convexity, 90% intensity quantile.

Figure 5.14: To estimate feature importance, the values of each feature were randomly permuted in the out-of-bag observations and the average decrease in the classification margin (i.e. the difference between the predicted probability for the correct class and the maximum probability predicated for incorrect classes) was computed. Note that a random control feature (ranked last in the importance estimate) was introduced.

Table 5.12: Class confusion matrix and classification scores for the validation set (60%) with **no segmentation error** using **all ALS features**. Abbreviations: **A. a.**: *Abies alba*, **P. a.**: *Picea abies*, **P. m.**: *Pseudotsuga menziesii*, **P. s.**: *Pinus sylvestris*, **L. d.**: *Larix decidua*, **F. s.**: *Fagus sylvatica*, **Q.**: *Quercus* sp., **F. e.**: *Fraxinus excelsior*, **A.**: *Acer* sp.

	Prediction									<b>Recall</b>
	<b>A. a.</b>	<b>P. a.</b>	<b>P. m.</b>	<b>P. s.</b>	<b>L. d.</b>	<b>F. s.</b>	<b>Q.</b>	<b>F. e.</b>	<b>A.</b>	
<b>Reference</b>	<b>A. a.</b>	<b>605</b>	37	5	7	0	3	0	0	0.92
	<b>P. a.</b>	31	<b>329</b>	5	4	0	1	1	0	0.89
	<b>P. m.</b>	6	9	<b>78</b>	0	0	0	0	0	0.84
	<b>P. s.</b>	7	5	0	<b>177</b>	0	0	0	0	0.94
	<b>L. d.</b>	0	0	0	1	<b>122</b>	4	2	0	0.95
	<b>F. s.</b>	1	4	0	0	8	<b>759</b>	12	7	0.95
	<b>Q.</b>	0	2	0	0	3	2	<b>237</b>	4	0.94
	<b>F. e.</b>	1	1	0	0	1	3	13	<b>39</b>	0.66
	<b>A.</b>	0	0	0	0	0	6	5	6	<b>10</b>
<b>Precis.</b>										<b>OA = 0.91</b>
										<b>K = 0.89</b>

Table 5.13: Class confusion matrix and classification scores for the validation set (60%) with **no segmentation error** using **scale invariant ALS features only**.

	Prediction									<b>Recall</b>
	<b>A. a.</b>	<b>P. a.</b>	<b>P. m.</b>	<b>P. s.</b>	<b>L. d.</b>	<b>F. s.</b>	<b>Q.</b>	<b>F. e.</b>	<b>A.</b>	
<b>Reference</b>	<b>A. a.</b>	<b>604</b>	38	4	8	0	3	0	0	0.92
	<b>P. a.</b>	30	<b>328</b>	4	5	0	1	3	0	0.88
	<b>P. m.</b>	9	8	<b>75</b>	1	0	0	0	0	0.81
	<b>P. s.</b>	7	5	0	<b>177</b>	0	0	0	0	0.94
	<b>L. d.</b>	0	1	0	0	<b>121</b>	4	3	0	0.94
	<b>F. s.</b>	1	2	0	0	9	<b>759</b>	18	6	0.95
	<b>Q.</b>	0	4	0	0	4	4	<b>234</b>	3	0.93
	<b>F. e.</b>	0	2	0	0	1	4	15	<b>35</b>	0.59
	<b>A.</b>	0	0	0	0	0	7	5	5	<b>10</b>
<b>Precis.</b>										<b>OA = 0.91</b>
										<b>K = 0.89</b>

Table 5.14: Class confusion matrix and classification scores for the validation set (60%) **with simulated segmentation error using all ALS features**. Abbreviations: **A. a.**: *Abies alba*, **P. a.**: *Picea abies*, **P. m.**: *Pseudotsuga menziesii*, **P. s.**: *Pinus sylvestris*, **L. d.**: *Larix decidua*, **F. s.**: *Fagus sylvatica*, **Q.**: *Quercus* sp., **F. e.**: *Fraxinus excelsior*, **A.**: *Acer* sp.

		Prediction									Recall
Reference		A. a.	P. a.	P. m.	P. s.	L. d.	F. s.	Q.	F. e.	A.	
	<b>A. a.</b>	<b>3612</b>	243	1	24	0	31	16	0	2	0.92
	<b>P. a.</b>	192	<b>1891</b>	26	45	11	41	16	3	1	0.85
	<b>P. m.</b>	41	37	<b>470</b>	9	1	0	0	0	0	0.84
	<b>P. s.</b>	39	38	0	<b>1027</b>	19	4	6	0	1	0.91
	<b>L. d.</b>	2	13	0	11	<b>719</b>	12	8	9	0	0.93
	<b>F. s.</b>	26	52	0	16	54	<b>4474</b>	68	70	34	0.93
	<b>Q.</b>	10	12	0	9	35	33	<b>1371</b>	29	7	0.91
	<b>F. e.</b>	11	9	0	5	11	35	71	<b>200</b>	12	0.56
	<b>A.</b>	2	7	0	4	4	58	28	35	<b>24</b>	0.15
<b>Precis.</b>											<b>OA = 0.89</b>
											<b>K = 0.87</b>

Table 5.15: Class confusion matrix and classification scores for the validation set (60%) **with simulated segmentation error using scale invariant ALS features only**.

		Prediction									Recall
Reference		A. a.	P. a.	P. m.	P. s.	L. d.	F. s.	Q.	F. e.	A.	
	<b>A. a.</b>	<b>3607</b>	239	6	29	1	28	18	0	1	0.92
	<b>P. a.</b>	185	<b>1889</b>	25	52	11	37	25	2	0	0.85
	<b>P. m.</b>	48	46	<b>454</b>	9	1	0	0	0	0	0.81
	<b>P. s.</b>	40	33	2	<b>1026</b>	23	4	5	0	1	0.9
	<b>L. d.</b>	2	11	1	13	<b>712</b>	17	17	1	0	0.92
	<b>F. s.</b>	28	55	0	15	54	<b>4456</b>	97	56	33	0.93
	<b>Q.</b>	17	11	0	9	44	33	<b>1371</b>	20	1	0.91
	<b>F. e.</b>	9	13	0	3	13	31	99	<b>175</b>	11	0.49
	<b>A.</b>	2	7	0	3	4	60	30	30	<b>26</b>	0.16
<b>Precis.</b>											<b>OA = 0.89</b>
											<b>K = 0.86</b>

### 5.4.3 Discussion

Overall high classification scores were obtained ( $OA = 0.91$ ,  $K = 0.89$ ) and the classifier performed well for all genus/species except ash (*Fraxinus Excelsior*) and maple (*Acer sp.*). The poor scores obtained for these two classes is at least partially imputable to the low number of available training samples. The classification of ash trees bears an additional difficulty, because due to their open crown (which lets light penetrate) and their preference for humid areas, they are frequently colonized by ivy (*Hedera helix*) which changes the shape and opacity characteristics of the tree (e.g. the persistent foliage of Ivy increases the amount of high intensity echos). The confusion matrices (cf. tables 5.12 and 5.13) indicate that the separation of deciduous and persistent foliage with leaf-off ALS is highly successful, with OA and K scores above 0.98. This indicates that preliminary separation of deciduous foliage before segmentation or integration of information on foliage persistence during the segmentation process could be a potential approach to improve individual tree delineation (cf. chapter 3).

The classification of segments with simulated error produced remarkably good results, only slightly lower ( $OA = 0.89$ ,  $K = 0.86-0.87$ ) than what is obtained when using error free segments. It has been shown indirectly by Ko et al. (2016) that the Random Forest classifier can effectively handle a moderate amount of tree segmentation error, if representative examples of different segmentation quality are used in the training phase. The results obtained here confirm and enhance this finding independently, using a simulation approach which allows explicit control over segmentation error. By allowing the creation of many erroneous training examples, the approach may be used to train genus/species classifiers that are more robust to segmentation error.

The feature ranking shows that intensity and opacity (echo rank distributions) metrics are among the most important features to differentiate species and no difference in performance was observed when using only scale invariant features. Shape related features are less important, which can be explained by the shape variability observed within a species. In particular, due to crown plasticity (Purves et al., 2007; Pretzsch, 2014; Jucker et al., 2015), shape variability tends to increase as a function of age (heteroscedasticity), making it a less distinctive feature. The importance of non-shape features is in accordance with findings previously reported by Holmgren and Persson (2004b), Korpela et al. (2009), Ørka et al. (2009), Suratno et al. (2009), Shi et al. (2018b). This fact can also be linked to the importance of amplitude and echo width features reported in studies which use full waveform LiDAR for species classification (e.g. Reitberger et al. (2008a); Heinzel and Koch (2011b)). One possible interpretation of this result is that the intensity and opacity features are related to the branch size (diameter) distribution within a tree. Stems and large branches (or opaque structures like evergreen coniferous branches) have a higher probability of completely intercepting the laser beam (thus generating more last returns) than small branches.

Comparison of these results with existing work is difficult, because of the small number of studies which consider the same species, employ data with similar characteristics and report the same error metrics. Moreover, a significant number of studies do not report performances per species and/or combine multiple types of data (e.g. multi/hyperspectral imagery and ALS) without analyzing the performance of single sensor classification. In this regard, the most comparable work found in the corpus was by Heinzel and Koch (2011b) (6 classes,  $OA = 59\%$ ), Lindberg et al. (2014) (6 classes,  $OA = 0.71$ ), Torabzadeh (2016) (8 classes,  $K = 0.76$ ) and Shi et al. (2018b) (6 classes,  $OA = 0.6-0.62$ ,  $K = 0.49-0.51$ ).

Possible improvements in the analysis include the addition and evaluation of other features, adding more observations in poorly represented species and/or age classes, evaluating other classification algorithms.

## 5.5 AHI based classification

In this section, the nine genus/species of interest presented previously are identified by analyzing hyperspectral imagery.

### 5.5.1 Description

The analysis presented here is based on hyperspectral imagery acquired in July 2015 (leaf-on) with the Airborne Prism Experiment (APEX) system (cf. appendix D.1). The dataset covers study sites around the Boudry (Neuchâtel) and Lausanne (Vaud) sectors. It is composed of several orthorectified images corresponding to different flight lines with a variable 2.6 to 3 m Ground Sampling Distance (GSD). It has 285 narrow bands (< 5 nm VIS/NIR, < 10 nm SWIR) covering the 400 to 2500 nm spectral range.

Pixels corresponding to individual tree crowns were sampled by overlaying the manually delineated tree crown extents (only for canopy trees) presented in chapter 2.2 to the hyperspectral images. Then, a pixel purity index (cf. 5.3.2) was computed and pixels with a class purity larger than 0.5, a Normalized Difference Vegetation Index (NDVI) larger than 0.5 and part of crowns with at least 4 pixels were selected. Moreover, shaded pixels were detected and removed with the simple blue-green indicator (i.e. shadow if  $\rho_{446} > \rho_{544}$ ) proposed in Trier et al. (2018). Erroneous reflectance bands in the dark blue (< 430 nm) and null reflectance bands in the atmospheric water vapor absorption regions (around 1320-1390 nm and 1770-1930 nm) were also removed. Then, several sets of descriptive features were computed from the reflectance signatures at the pixel and crowns scales (i.e. average of pixel scale values in each crown):

- The Principal Component Analysis (PCA) scores of the raw and standard normal variate transformation of reflectance (i.e. for each reflectance pixel  $i$ , subtract the mean of  $i$  and divide by the standard deviation of  $i$ , cf. Barnes et al. (1989)). The PCA is a linear transformation which produces a new set of bands which are decorrelated and ordered by decreasing explained variance.
- The Maximum Noise Fraction (MNF) scores (Green et al., 1988) of the raw and standard normal variate transformation of reflectance. The MNF is a linear transformation where the variables (features) are first decorrelated with a first PCA, followed by a statistical noise whitening transformation. A second PCA is then applied to the noise-whitened variables. The result is a new set of bands ordered by decreasing signal to noise ratio.
- Vegetation Indices (VI) from established lists (Sims and Gamon, 2002; Stagakis et al., 2010) and custom simple band ratio indices based on apparent separability of the reflectance signatures (cf. figure 5.7)

Two alternatives were tested to aggregate pixel level classification at the crown scale: averaging feature values (i.e. before classification) or averaging class probabilities (i.e. after classification). Both approaches resulted in similar performances and only results for the former approach are detailed in the following subsections.

Subsequently, the observations were grouped by species and partitioned at the crown level using stratified sampling, so that the training (50%) and validation (50%) sets contain roughly the same crown size (i.e. number of pixels per crown) frequency distributions within each group (cf. figure 5.15). This partitioning was done at the crown level to avoid model over-fitting and to ensure different age categories were represented in both sets. Thus, the pixels of a given tree crown were either all in the training set or all in the validation set, but not in both.

Then, two different statistical classifiers were evaluated to predict the genus/species: the Random Forest (Breiman, 2001) and the Support Vector Machine (Vapnik, 1995; Schölkopf et al., 2002). In a first stage, for each complete set of features, an initial classification was conducted with the random forest algorithm using 600 decision trees and a minimum leaf size of 1. To

estimate feature importance, the values of each feature were randomly permuted in the out-of-bag observations and the average decrease in the classification margin (i.e. the difference between the predicted probability for the correct class and the maximum predicted probability for incorrect classes) was computed. Based on this initial feature ranking, recursive feature elimination was used to discard non-informative features in each of the feature sets. It was determined that the first 30 (out of 255) components of PCA, the 20 (out of 255) first components of the MNF transform and 15 (out of 96) vegetation indices (cf. table 5.16) were sufficient to model class differences. The final selection of vegetation indices covered all major leaf pigments (chlorophyll, carotenoids, anthocyanin) and water content. Then, the random forest classifier was reapplied to the training data, using each of the pruned feature sets. Using each of the features sets, separate models were trained at the pixel and crown scales. Finally, classification performances were evaluated using the framework presented in section 5.2 (cf. table 5.18). The workflow described above is summarized in figure 5.16.

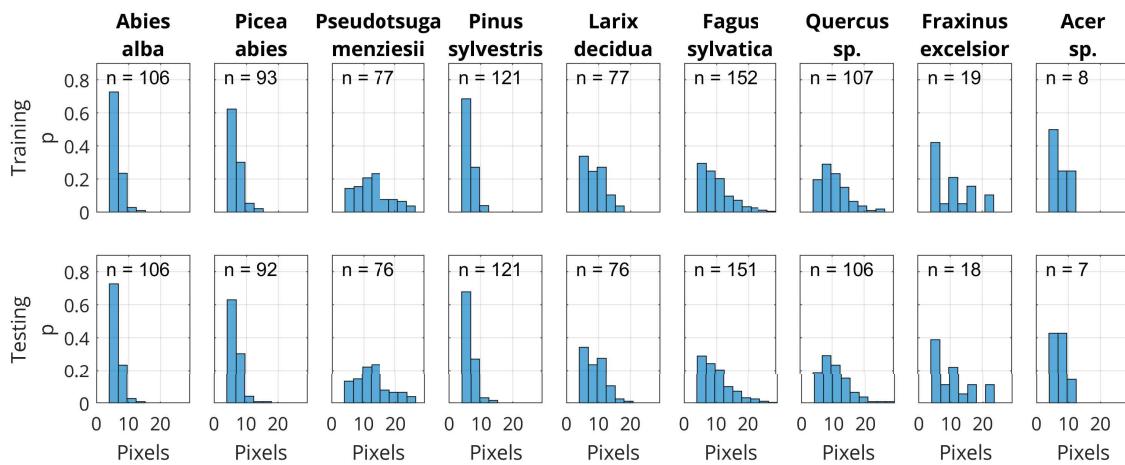


Figure 5.15: Number of pixels per crown probability distributions for the 50% training (upper row) and 50% validation (lower row) sets, after class grouping and stratified sampling.

In a second stage, the SVM classifier (using the libsvm implementation by Chang and Lin (2011)) was evaluated on the two best performing sets of pruned features. The SVM is a binary classifier which uses functions called kernels to quantify similarity between observations. Depending on its form, the kernel function can also serve to map variables to a new space where they potentially become linearly separable. Within this new space, the SVM then attempts to find a hyperplane that maximizes the distance (allowing some slack) between the two labeled classes. Both the linear and Radial Basis Function (RBF) kernels were tested. The first kernel produces a linear classification boundary while the second produces a non-linear classification boundary. The hyperparameters of the SVM (i.e. the regularization parameter  $C$  that penalizes misclassification and the  $\gamma$  parameter which defines the width of the RBF kernel) were optimized using a grid search where different parameter value combinations were evaluated with 5-fold cross-validation on the training data. Moreover, to mitigate the effect of class imbalance, misclassification costs inversely proportional to the class occurrence frequencies were imposed. Since SVM is a binary classifier, a pairwise (one-versus-one) classification setup was used to distinguish all classes. As previously, class predictions were made on validation data at the pixel and crown scales and the performances were evaluated.

Table 5.16: Vegetation indices automatically selected from the list in Stagakis et al. (2010) and custom indices with recursive feature elimination using Random Forest. The ranking was obtained with the Random Forest by randomly permuting the out-of-bag observations and computing the average decrease in the classification margin (i.e. the difference between the predicted probability for the correct class and the maximum predicted probability for incorrect classes).

Category	Name	Formula	Source	Ranking
Leaf pigments	Green Vegetation Index (GVI)	$\frac{\rho_{682} - \rho_{553}}{\rho_{682} + \rho_{553}}$	Gandia et al. (2004)	15
	Greenness Index (GI)	$\frac{\rho_{554}}{\rho_{677}}$	Zarco-Tejada et al. (2005)	14
	Simple Ratio	$\frac{\rho_{685}}{\rho_{655}}$	Zarco-Tejada et al. (2003)	2
	Simple Ratio	$\frac{\rho_{750}}{\rho_{700}}$	Gitelson and Merzlyak (1997)	5
	Simple Ratio	$\frac{\rho_{860}}{\rho_{780}}$	-	10
	Plant Pigment Ratio (PPR)	$\frac{\rho_{550} - \rho_{450}}{\rho_{550} + \rho_{450}}$	Metternicht (2003)	13
	Red Edge Inflection Point (REIP)	$700 + 40 \frac{0.5(\rho_{670} + \rho_{780}) - \rho_{700}}{\rho_{740} - \rho_{700}}$	Guyot et al. (1988)	9
	Carotenoid Reflectance Index (CRI)	$\frac{1}{\rho_{510}} - \frac{1}{\rho_{550}}$	Gitelson et al. (2002)	3
	Anthocyanin Reflectance Index (ARI)	$\rho_{800} \frac{1}{\rho_{550}} - \frac{1}{\rho_{700}}$	Merzlyak et al. (2003)	7
Water content	Photochemical Reflectance Ratio (PRR)	$\frac{\rho_{531}}{\rho_{570}}$	Zheng and Chen (2017)	4
	Max reflectance wavelength between 980 and 1165 nm	$\text{argmax}_{\rho}(\rho_{980-1165})$	-	6
	Max reflectance wavelength between 1165 and 1330 nm	$\text{argmax}_{\rho}(\rho_{1165-1330})$	-	8
	EWT	$1 - \frac{\rho_{1193}}{\rho_{1126}}$	Underwood et al. (2003)	1
	Floating Water Band Index (fWBI)	$\frac{\rho_{900}}{\min_{\rho}(\rho_{920-980})}$	Peñuelas et al. (1993)	11
	Water Band Index (WBI)	$\frac{\rho_{900}}{\rho_{970}}$	Peñuelas et al. (1993)	12

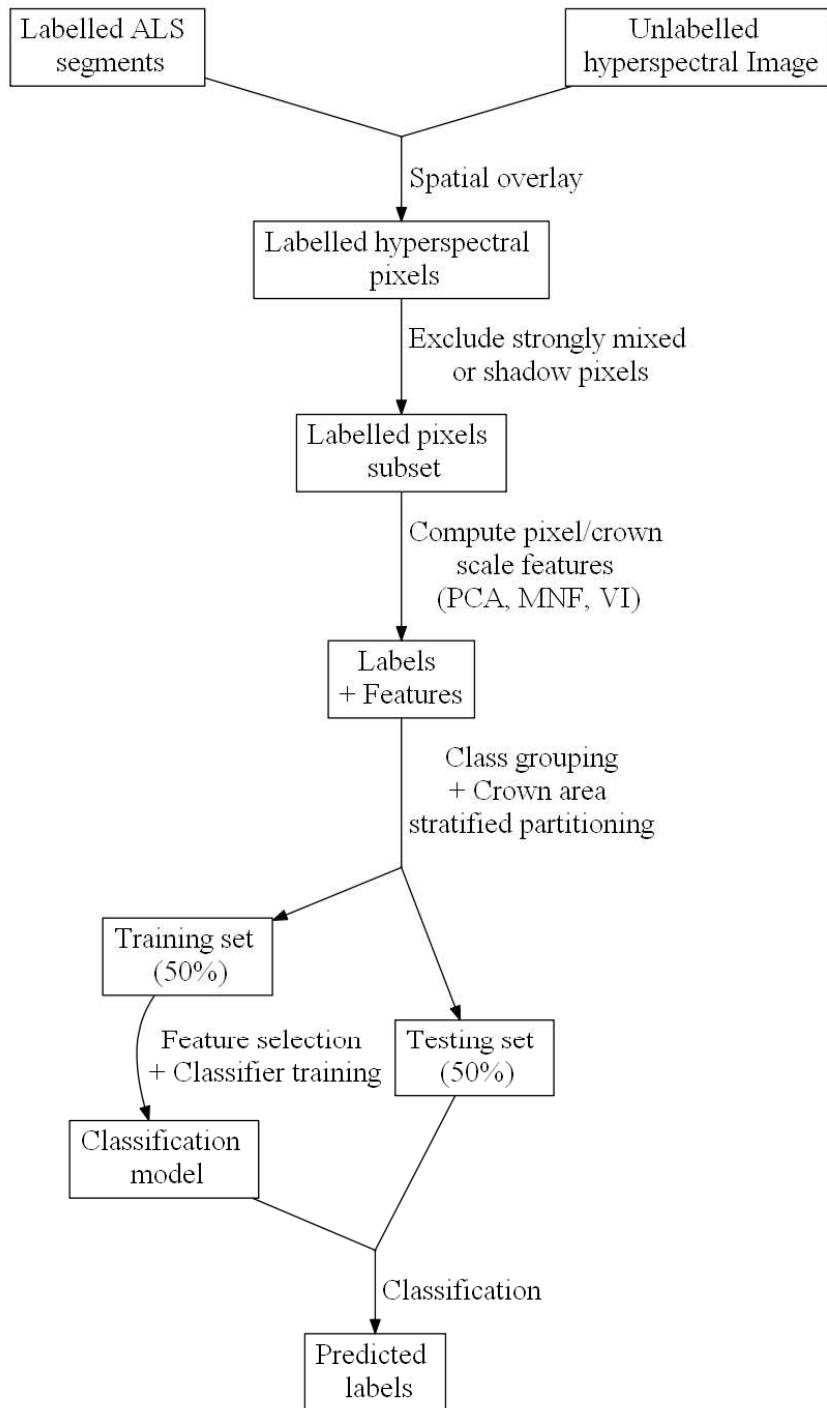


Figure 5.16: AHI classification workflow. Note that the same workflow was used at the pixel and crown scales.

### 5.5.2 Results

The overall classification scores obtained with the RF and SVM classifiers at the pixel and crown scales are reported in table 5.17. Detailed per class scores and confusion matrices for the best classifier and feature set are reported in tables 5.18 (pixel scale) and 5.19 (crown scale).

Table 5.17: Overall genus/species classification scores using AHI derived features (after feature selection at the pixel scale (6418 observations) and crown scale (753 observations). **Abbreviations:** OA: Overall Accuracy, K: kappa,  $\bar{p}$ : average of per class precisions,  $\bar{r}$ : average of per class recalls,  $\bar{F}$ : average of per class F-scores.

Scores							
Classifier	Features	Scale	OA	K	$\bar{r} \pm \sigma$	$\bar{p} \pm \sigma$	$\bar{F} \pm \sigma$
RF	PCA of raw reflectance	Pixel	0.77	0.73	$0.67 \pm 0.24$	$0.7 \pm 0.18$	$0.67 \pm 0.22$
		Crown	0.82	0.8	$0.73 \pm 0.29$	$0.73 \pm 0.29$	$0.72 \pm 0.28$
	PCA of norm. reflectance	Pixel	0.78	0.74	$0.67 \pm 0.26$	$0.7 \pm 0.2$	$0.67 \pm 0.23$
		Crown	0.82	0.79	$0.72 \pm 0.28$	$0.72 \pm 0.28$	$0.72 \pm 0.28$
	MNF of raw reflectance	Pixel	0.82	0.79	$0.72 \pm 0.26$	$0.73 \pm 0.23$	$0.72 \pm 0.24$
		Crown	0.83	0.81	$0.75 \pm 0.25$	$0.76 \pm 0.24$	$0.75 \pm 0.23$
	MNF of norm. reflectance	Pixel	0.83	0.8	$0.74 \pm 0.22$	$0.75 \pm 0.2$	$0.74 \pm 0.2$
		Crown	0.85	0.82	$0.78 \pm 0.21$	$0.79 \pm 0.16$	$0.78 \pm 0.18$
SVM (linear)	VI	Pixel	0.74	0.69	$0.61 \pm 0.26$	$0.61 \pm 0.26$	$0.61 \pm 0.26$
		Crown	0.75	0.71	$0.65 \pm 0.24$	$0.65 \pm 0.21$	$0.65 \pm 0.22$
	MNF of norm. refl. and VI	Pixel	0.83	0.8	$0.74 \pm 0.21$	$0.75 \pm 0.19$	$0.75 \pm 0.19$
		Crown	0.85	0.83	$0.79 \pm 0.21$	$0.80 \pm 0.14$	$0.78 \pm 0.17$
	MNF of norm. reflectance	Pixel	0.8	0.76	$0.70 \pm 0.24$	$0.71 \pm 0.20$	$0.70 \pm 0.22$
		Crown	0.84	0.81	$0.76 \pm 0.24$	$0.77 \pm 0.18$	$0.76 \pm 0.22$
	MNF of norm. refl. and VI	Pixel	0.8	0.77	$0.69 \pm 0.25$	$0.71 \pm 0.22$	$0.70 \pm 0.24$
		Crown	0.83	0.8	$0.75 \pm 0.24$	$0.77 \pm 0.18$	$0.76 \pm 0.22$
SVM (RBF)	MNF of norm. reflectance	Pixel	0.82	0.79	$0.72 \pm 0.23$	$0.74 \pm 0.19$	$0.73 \pm 0.21$
		Crown	0.85	0.82	$0.76 \pm 0.25$	$0.77 \pm 0.22$	$0.76 \pm 0.24$
	MNF of norm. refl. and VI	Pixel	0.81	0.78	$0.72 \pm 0.21$	$0.73 \pm 0.21$	$0.72 \pm 0.2$
		Crown	0.84	0.82	$0.74 \pm 0.29$	$0.74 \pm 0.29$	$0.74 \pm 0.29$

Table 5.18: Class confusion matrix and classification scores for the validation set (50%) using **AHI features (20 first MNF of normalized reflectance) at the pixel scale (6442 observations) and RF classifier**. Abbreviations: **A.** *a.*: *Abies alba*, **P.** *a.*: *Picea abies*, **P.** *m.*: *Pseudotsuga menziesii*, **P.** *s.*: *Pinus sylvestris*, **L.** *d.*: *Larix decidua*, **F.** *s.*: *Fagus sylvatica*, **Q.**: *Quercus* sp., **F.** *e.*: *Fraxinus excelsior*, **A.**: *Acer* sp.

Table 5.19: Class confusion matrix and classification scores for the validation set (50%) using **AHI features (20 first MNF of normalized reflectance) at the crown scale (753 observations)** and RF classifier.

Prediction										
Reference	A. a.	P. a.	P. m.	P. s.	L. d.	F. s.	Q.	F. e.	A.	Recall
	<b>81</b>	2	0	8	1	5	9	0	0	0.76
	<b>64</b>	0	2	17	6	1	0	1	1	0.7
	<b>72</b>	0	3	1	0	0	0	0	0	0.95
	<b>114</b>	1	1	0	3	1	1	0	0	0.94
	<b>69</b>	0	3	2	2	0	0	0	0	0.91
	<b>128</b>	0	3	4	5	2	7	1	1	0.85
	<b>93</b>	5	0	1	3	2	2	0	0	0.88
	<b>14</b>	0	0	1	0	2	0	0	1	0.78
<b>A.</b>	<b>2</b>	0	0	1	1	0	0	3	<b>2</b>	0.29

Table 5.20: Class confusion matrix and classification scores for the validation set (50%) using **AHI features (20 first MNF of normalized reflectance) at the pixel scale (6442 observations) and SVM RBF classifier**. Abbreviations: **A.** *a.*: *Abies alba*, **P.** *a.*: *Picea abies*, **P.** *m.*: *Pseudotsuga menziesii*, **P.** *s.*: *Pinus sylvestris*, **L.** *d.*: *Larix decidua*, **F.** *s.*: *Fagus sylvatica*, **Q.**: *Quercus* sp., **F.** *e.*: *Fraxinus excelsior*, **A.**: *Acer* sp.

Table 5.21: Class confusion matrix and classification scores for the validation set (50%) using AHI features (20 first MNF of normalized reflectance) at the crown scale (753 observations) and SVM RBF classifier.

### 5.5.3 Discussion

Overall, the best set of features to distinguish genus/species was the 20 first MNF components. There was no clear performance gain of adding VI to MNF components. This is consistent with multiple studies on tree species classification which report that dimensional reduction with PCA or MNF is beneficial and sometimes sufficient to produce good classification results (Ghosh et al., 2014b; Fassnacht et al., 2014; Torabzadeh, 2016; Lee et al., 2016; Dabiri and Lang, 2018) without relying on more complex procedures.

Predictions at the crown scale were systematically better than at the pixel scale, highlighting the importance of taking into account the spatial structure of the image and the usefulness of individual tree segmentation. Although the higher scores obtained at object level do not indicate a better performance of the classification algorithm per se, but rather are a consequence of the averaging effect when aggregating multiple observations.

Consistent with previous studies (Ørka et al., 2012; Dalponte et al., 2012; Ghosh et al., 2014b; Ballanti et al., 2016; Piiroinen et al., 2017), the RF and SVM classifiers produced similar results, with the highest kappa scores (at the crown scale) ranging from 0.82 (RF) to 0.84 (SVM-RBF).

Oak (pedunculate and sessile), European beech and Douglas fir systematically obtained the highest classification scores with precision and recall in the 0.8-0.95 range. Norway spruce and silver fir were frequently misidentified as other species (silver fir with Oak and spruce with Scots pine). This is consistent with other studies that report difficulties in separating coniferous species (Roberts et al., 2004; Leckie et al., 2005; van Aardt and Wynne, 2007; Trier et al., 2018). Since observations were collected in mixed forests, part of the misidentification may also be imputed to corrupt labeling related to image orthorectification and/or georeferencing errors. This hypothesis however cannot be easily verified because of the relatively coarse resolution of the imagery (2.5-3 m) which precludes fine coregistration with a DSM or high resolution RGB image. To reduce the problems related to mis-registration of ALS derived crown extents and AHI, a solution would be to collect pixel samples in pure forest stands (which may be problematic for some species that do not occur in large groups). The low scores obtained for ash and maple should not be considered reliable, as they are based on very small training samples and also because other studies (Torabzadeh, 2016; Dabiri and Lang, 2018) have reported high success rates in identifying these species.

As discussed previously, comparing the results obtained here with those in other studies is complicated due to differences in data characteristics, considered species and analytical approach. Similar studies on multispectral or hyperspectral based species classification are listed in table 5.22 for reference. Interestingly, some of these studies (Engler et al., 2013; Immitzer et al., 2012; Waser et al., 2014, 2010, 2008) that use high spatial resolution (< 1m) multispectral data were able to achieve similar performances than those using hyperspectral data. This may indicate that, at least for a limited set of tree species, the lack of spectral information may be compensated by a high spatial resolution (which provides more information on texture). Keeping in mind that the comparison is not fully relevant, the performances obtained here ( $OA = 86$ ,  $K = 0.84$ ) can be considered on par with other recent studies. This confirms the ability of the APEX system to differentiate common European trees at least at the genus level on a limited number of genus/species. However, additional testing on a larger dataset (also including different geographic regions) would be necessary for conclusive results on a more extensive number of species representative of the true diversity of European forests.

Basic improvements to the analysis presented here include adding observations in under represented classes (ash and maple) and using finer spatial resolution hyperspectral imagery. More advanced improvements could include the use of spectral unmixing procedures (Somers and Asner, 2013, 2014) and at the pixel scale smoothing (regularisation) the genus/species class map with a conditional random field (cf. Dechesne et al. (2017)) using the class probabilities from the RF or SVM classification.

Table 5.22: Selected multispectral and hyperspectral based tree species classification studies similar (northern European palearctic, 5-13 species) to the one presented in this section. The performances reported here are based on the use of the multi/hyperspectral imagery only, even though some of the studies also use ALS. If multiple sites were present in the study, scores for the one with the most species was selected. The generic OA and K scores are the highest values reported in the studies and indicate general trends; they are not sufficient to reliably compare the performances of studies.

Sensor	Bands	GSD	Classes	OA	K	Reference
Leica ADS40-SH52	4	0.5 m	6	0.72	0.65	Engler et al. (2013)
Leica RC30 + Leica ADS40	4	0.25 m	5	-	0.86	Waser et al. (2008)
Z/I Imaging DMC	4	0.2 m	8	0.88	0.86	Waser et al. (2010)
WorldView-2	8	2 m	10	0.84	0.81	Immitzer et al. (2012)
	8	2 m	7	0.83	0.79	Waser et al. (2014)
Daedalus 1268 ATM	11	2 m	6	0.71	0.63	Hill et al. (2010)
AISA Eagle	126	1 m	8	0.74	0.66	Dalponte et al. (2012)
HyMAP	125	4 m	5	0.82	0.77	Ghosh et al. (2014b)
HySpex VNIR-1600	160	1.6 m	13 (6)	0.8	-	Sommer et al. (2016)
HySpex (VNIR + SWIR)	416	1-2 m	5	0.69	0.59	Shi et al. (2018a)
AISA DUAL (Hawk + Eagle)	367	3 m	7	0.92	-	Fassnacht et al. (2014)
	367	2 m	10	0.75	-	Richter et al. (2016)
APEX	286	-	7	0.7	0.61	Van Coillie et al. (2014)
	285	2 m	8	-	0.75	Torabzadeh (2016)
	288	3.35 m	5	0.77	0.72	Raczko and Zagajewski (2017)
	288	2.5 m	6	0.85	0.8	Dabiri and Lang (2018)
	288	2.6 - 3 m	9	0.86	0.84	This study (2018)

## 5.6 Combined ALS and AHI classification

In this section, the ALS and AHI features presented in sections 5.4 and 5.5 are combined at the tree scale to identify genus/species. An overall performance comparison of different ALS/AHI feature combinations and classifiers is also conducted (cf. table 5.23).

### 5.6.1 Description

The combination of ALS and AHI allows an integral description of the tree in terms of structure and foliage biophysical characteristics. Here, different combinations of ALS and AHI features are tested to determine their relative value for species identification at the tree scale. The tested descriptive features were described in chapters 5.4 and 5.5. They include:

- the 10 scale variant ALS features presented in chapter 5.4.
- the 36 scale invariant ALS features presented in chapter 5.4.
- the 20 first components (averaged at the crown scale) of the Maximum Noise Fraction transform applied to the normalized AHI reflectance spectra (cf. chapter 5.5).
- the 15 vegetation indices (averaged at the crown scale) obtained by recursive feature elimination on a list of indices found in literature (cf. chapter 5.5).

The number of observations covered by AHI limits the size of the combined ALS-AHI dataset to 1513 observations (out of 6699 covered by ALS). The same grouping and stratified partitioning used in section 5.5 is applied here. That is, the observations are grouped by species and partitioned using stratified sampling, so that the training (50%) and validation (50%) sets contain roughly the same crown size (i.e. number of pixels per crown) frequency distributions within each group (cf. figure 5.15).

Both the Random Forest (RF) and the Support Vector Machine (SVM) classifiers are evaluated. The former is evaluated on all feature set combinations and the later is only evaluated on the single sensor and combined sensor features which obtained the best results with RF.

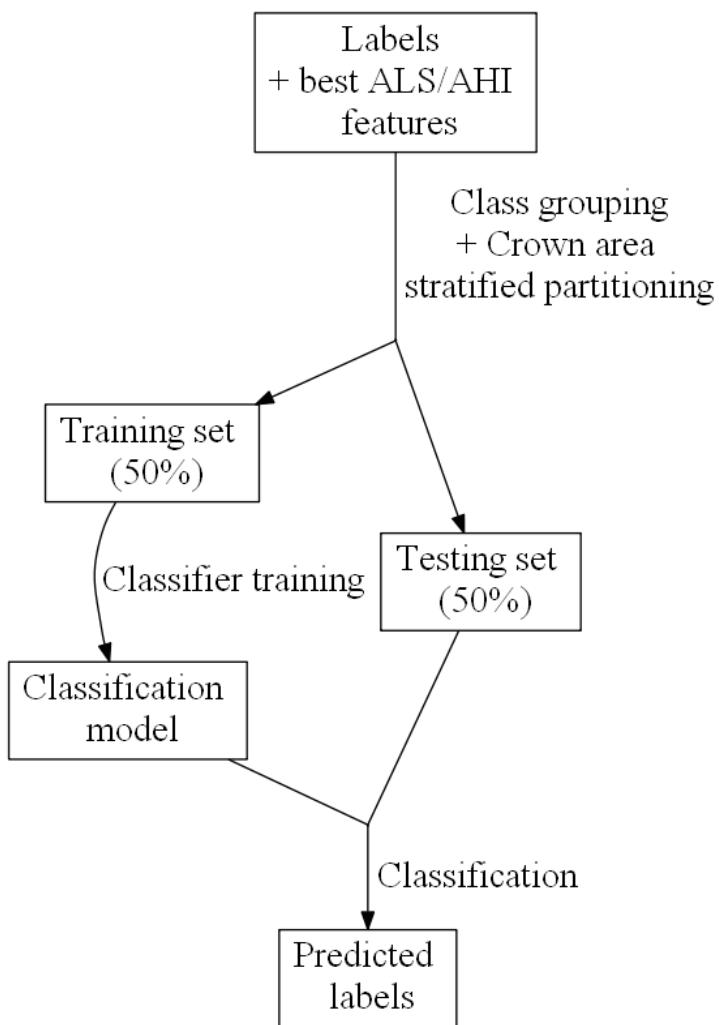


Figure 5.17: Combined ALS and AHI classification workflow.

### 5.6.2 Results

The classification was validated using the framework presented in section 5.2. The overall classification scores obtained with the RF and SVM classifiers for different combinations of ALS and AHI features at the crown scales are reported in table 5.23. Detailed per class scores and confusion matrices for the best classifier and feature setw are reported in tables 5.24 and 5.25.

Table 5.23: Comparison of overall genus/species classification scores using different combinations of ALS and AHI derived features. **Abbreviations:** **OA:** Overall Accuracy, **K:** kappa,  $\bar{p}$ : average of per class precisions,  $\bar{r}$ : average of per class recalls,  $\bar{F}$ : average of per class F-scores, **SI:** Scale Invariant, **SV:** Scale Variant, **MNF:** Maximum Noise Fraction, **VI:** Vegetation Indices, #: Number of features.

Classifier	Features			Scores				
	AHI	ALS	#	OA	K	$\bar{r} \pm \sigma$	$\bar{p} \pm \sigma$	$\bar{F} \pm \sigma$
RF	-	SV + SI	46	0.89	0.88	$0.82 \pm 0.19$	$0.82 \pm 0.17$	$0.81 \pm 0.18$
	-	SI	36	0.9	0.89	$0.82 \pm 0.22$	$0.84 \pm 0.18$	$0.81 \pm 0.21$
	MNF	-	20	0.85	0.82	$0.78 \pm 0.21$	$0.79 \pm 0.16$	$0.78 \pm 0.18$
	MNF + VI	-	35	0.85	0.83	$0.79 \pm 0.21$	$0.80 \pm 0.14$	$0.78 \pm 0.17$
	MNF	SV + SI	46	0.93	0.92	$0.90 \pm 0.09$	$0.88 \pm 0.16$	$0.89 \pm 0.13$
	MNF + VI	SV + SI	81	0.93	0.92	$0.90 \pm 0.09$	$0.89 \pm 0.15$	$0.89 \pm 0.12$
	MNF	SI	56	0.95	0.94	$0.93 \pm 0.06$	$0.90 \pm 0.14$	$0.91 \pm 0.10$
	MNF + VI	SI	71	0.95	0.94	$0.93 \pm 0.06$	$0.89 \pm 0.17$	$0.90 \pm 0.12$
SVM (linear)	-	SI	36	0.90	0.88	$0.72 \pm 0.41$	$0.70 \pm 0.40$	$0.71 \pm 0.41$
	MNF	-	20	0.84	0.81	$0.76 \pm 0.24$	$0.77 \pm 0.18$	$0.76 \pm 0.22$
	MNF	SI	56	0.95	0.94	$0.88 \pm 0.18$	$0.88 \pm 0.18$	$0.88 \pm 0.18$
SVM (RBF)	-	SI	36	0.91	0.89	$0.74 \pm 0.40$	$0.82 \pm 0.31$	$0.73 \pm 0.39$
	MNF	-	20	0.85	0.82	$0.76 \pm 0.25$	$0.77 \pm 0.22$	$0.76 \pm 0.24$
	MNF	SI	56	0.95	0.94	$0.91 \pm 0.10$	$0.90 \pm 0.15$	$0.90 \pm 0.13$

Table 5.24: Best class confusion matrix and classification scores for the validation set (50%) using **AHI (MNF) and ALS (SI) features at the crown scale with the RF classifier**. Abbreviations: A. a.: *Abies alba*, P. a.: *Picea abies*, P. m.: *Pseudotsuga menziesii*, P. s.: *Pinus sylvestris*, L. d.: *Larix decidua*, F. s.: *Fagus sylvatica*, Q.: *Quercus* sp., F. e.: *Fraxinus excelsior*, A.: *Acer* sp.

Reference	Prediction									Recall
	A. a.	P. a.	P. m.	P. s.	L. d.	F. s.	Q.	F. e.	A.	
A. a.	102	1	0	3	0	0	0	0	0	0.96
P. a.	4	80	4	4	0	0	0	0	0	0.87
P. m.	0	5	71	0	0	0	0	0	0	0.93
P. s.	4	0	0	117	0	0	0	0	0	0.97
L. d.	0	0	0	0	75	0	1	0	0	0.99
F. s.	0	0	0	0	1	144	3	1	2	0.95
Q.	0	0	0	0	1	1	102	1	1	0.96
F. e.	0	0	0	0	0	0	1	15	2	0.83
A.	0	0	0	0	0	1	0	0	6	0.86
Precis.	0.93	0.93	0.95	0.94	0.97	0.99	0.95	0.88	0.55	OA = 0.95 K = 0.94

Table 5.25: Best class confusion matrix and classification scores for the validation set (50%) using **AHI (MNF) and ALS (SI) features at the crown scale (753 observations) with the SVM (RBF) classifier**

Reference	Prediction									Recall
	A. a.	P. a.	P. m.	P. s.	L. d.	F. s.	Q.	F. e.	A.	
A. a.	103	0	0	3	0	0	0	0	0	0.97
P. a.	6	80	3	3	0	0	0	0	0	0.87
P. m.	1	3	71	1	0	0	0	0	0	0.93
P. s.	3	3	0	115	0	0	0	0	0	0.95
L. d.	0	0	0	0	76	0	0	0	0	1
F. s.	0	0	0	0	2	145	1	1	2	0.96
Q.	0	0	0	0	0	1	105	0	0	0.99
F. e.	0	0	0	0	0	1	0	14	3	0.78
A.	0	0	0	0	0	1	1	0	5	0.71
Precis.	0.91	0.93	0.96	0.94	0.97	0.98	0.98	0.93	0.5	OA = 0.95 K = 0.94

### 5.6.3 Discussion

By combining the AHI and ALS feature sets, most of the genus/species could be identified reliably, with a maximum overall accuracy of 0.95, kappa of 0.94 and average F-score of 0.9. The combination of ALS and AHI features systematically produced better predictions than any single sensor feature set. The gain in prediction performance brought by the combination was ~4-5% in terms of overall accuracy and kappa, ~10% in terms of average precision, recall and F-score. The RF and SVM (RBF) produced similar performances, here again confirming what has been reported in previous studies (Ørka et al., 2012; Dalponte et al., 2012; Ghosh et al., 2014b; Ballanti et al., 2016; Piiroinen et al., 2017).

The fir-oak and spruce-pine confusions that occurred when using only AHI, were suppressed by adding ALS features to the model. The only class that remained poorly identified was maple, as reported in previous sections, likely due to the insufficient number of available observations rather than poor structural/spectral separability.

ALS features alone produced better predictions than AHI features alone, in terms of overall accuracy and kappa scores (~5% higher), but had similar (~3-4% higher) performances in terms of average precision, recall and F-score. These results are consistent with the analysis conducted in chapter 5.4.2 which used ALS features only on a larger dataset and which obtained very similar results. As reported in chapter 5.4.3, the addition of scale dependent ALS features did not improve the classification. At this point, it should also be noted, that ALS has a major advantage over AHI with regard to the number of trees that can be detected/characterized because AHI does not cover the forest understory, potentially missing a large fraction of the trees.

Regarding the AHI feature sets, the addition of vegetation indices did not improve the classification and simply using the first 20 components of the maximum noise transform of normalized reflectance produced the best results, as noted previously in chapter 5.5.3.

Summary scores and basic metadata of selected similar studies are provided for reference in table 5.26. They should be interpreted carefully as the analytical approach and data characteristics are different across studies. Nonetheless, the comparison indicate that the results obtained here are in line with recent similar studies.

An alternative option to analysis at the tree scale is to conduct the entire analysis using AHI and ALS features at the pixel level (e.g. Torabzadeh (2016)), avoiding the need for individual tree segmentation. This simplified approach, which was not investigated here, may provide species distribution information at the canopy level and answer some operational requirements.

An improvement could be to simultaneously conduct segmentation and classification, by integrating the AHI features into the individual tree segmentation process. As noted in chapter 5.5, an other improvement could be the application of probabilistic smoothing to the class map as proposed in Dechesne et al. (2017).

Table 5.26: Selected combined multi/hyperspectral and ALS based tree species classification studies similar (northern European palearctic, 5-13 species) to the one presented in this section. If multiple sites were present in the study, scores for the one with the most species was selected. The generic OA and K scores are the highest values reported in the studies and indicate general trends; they are not sufficient to reliably compare the performances of studies.

M/HI			ALS			Score			Reference
Sensor	Bands	GSD	Sensor	Density	Classes	OA	K		
Leica ADS40 SH52 Leica RC30	4	0.25-0.5 m	Various	0.8 m <sup>-2</sup>	7	0.76	0.7	Waser et al. (2011)	
AISA Eagle	126	1 m	Optech ALTM 3100C	0.5 m <sup>-2</sup>	8	0.83	0.77	Dalponte et al. (2012)	
CASI-1500	13	0.5	Optech Gemini	4 m <sup>-2</sup>	5	0.98	0.97	Dinuls et al. (2012)	
HySpex VNIR-1600	160	1.6 m	Riegl LMS-Q680i	25-30 m <sup>-2</sup>	13 (6)	0.91	-	Sommer et al. (2016)	
AISA Fenix	361	1.2 m	Leica ALS-50 II	6 m <sup>-2</sup>	7	0.92	0.9	Lee et al. (2016)	
HySpex (VNIR + SWIR)	416	1-2 m	Riegl LMS-Q680i	70 m <sup>-2</sup>	5	0.84	0.74	Shi et al. (2018a)	
APEX	285	2 m	Riegl LMS-Q560 LMS-Q680i	20, 40 m <sup>-2</sup>	8	-	0.93	Torabzadeh (2016)	
	288	2.6 - 3 m	Optech Gemini Riegl LMS Q1560	30-70 m <sup>-2</sup>	9	0.95	0.94	This study (2018)	

## 5.7 Synthesis

In this chapter, a state of the art on tree species classification from ALS and AHI data was conducted. It revealed that most studies have focused on palearctic and nearctic ecozones and have usually considered less than 10 genus or species. Studies that attempt to map forest species composition at the individual tree scale using both ALS and AHI in environments similar to Switzerland remain relatively scarce and are often based on small datasets with uncertain segmentation quality. Based on this premise, three approaches to identify nine commonly encountered genus/species were investigated:

- An approach to classify tree genus/species at the crown scale using discrete return ALS features (from different sensors) was evaluated and good performances ( $OA = 0.91$ ,  $K = 0.89$ ), in par with other state of the art studies, were obtained. Lower performances were obtained for ash and maple and it was suspected that this was at least in part due to the small number of observations available for these classes. Both scale variant and scale invariant features were tested with no conclusive evidence that the inclusion/exclusion of scale variant features affected the performance. Metrics related to structural opacity, such as echo ranking distribution and intensity were found to be the most discriminative. A novel method to simulate individual tree segmentation error and evaluate its effect on species classification was also developed and evaluated. The approach demonstrated that it is possible to train the Random Forest classifier to identify species from segments containing significant delineation error and still obtain good predictions.
- An approach to classify tree genus/species using AHI features (from the APEX sensor) at the pixel and crown scale was evaluated. SVM and RF classifiers were evaluated and produced similar results. Relatively good performances were achieved at both scales (Pixel:  $OA = 0.83$ ,  $K = 0.8$ , Crown:  $OA = 0.85$ ,  $K = 0.82$ ). Higher performances were systematically obtained at the crown scale underlining the importance of taking into account information about the extent of individual tree crowns (and the utility of segmentation). Different sets of feature combinations (PCA, MNF, VI) were tested and the best performing model simply employed the 20 first components of the MNF transform.
- The combination of ALS and AHI derived feature for tree genus/species at the crown scale was evaluated. As previously, SVM and RF classifiers were evaluated and produced similar results. It was found that combining the two sets of features systematically improved the performances (by about ~5% in terms of OA and K, ~10% in terms of average F-score) and provided a high overall performance ( $OA = 0.95$ ,  $K = 0.94$ ) in line with similar state of the art studies. The best performances were obtained by combining the 20 first MNF of the AHI and the scale invariant features from ALS.





## 6. Conclusion

### Insights

Throughout this thesis, the capacity of Airborne Laser Scanning (ALS) augmented with Hyperspectral Imaging (AHI) to describe forests at the individual tree scale was investigated. Going back to the research questions and technical objectives formulated in the introduction, the following conclusions can be drawn:

#### **How and under which conditions can Airborne Laser Scanning (ALS) and Hyperspectral Imaging (AHI) be used to map forest characteristics at the individual tree scale?**

These research questions were investigated by attempting to fulfill three fundamental requirements of forest inventories, using ALS and AHI: locating trees, measuring their diameter and identifying their genus/species. Technical solutions to address each of these requirements - some built upon existing work and others entirely novel - were proposed and validated. In chapter 2, it was shown how visual interpretation of ALS point clouds can be used to accurately determine the position and characteristics of individual trees. An interactive application was developed and used to extract several thousand detailed 3D tree models which subsequently served to calibrate and validate segmentation, regression and classification algorithms. This approach can also be employed to prepare detailed maps of small plots (e.g. long term research plots, growth and yield plots, training plots) without any surveying skills. In chapter 3, two new algorithms for individual tree delineation and stem detection in ALS data were presented. Special emphasis was put on pure deciduous broadleaf forests in which it is more difficult to locate individual trees due to their relatively flat canopy. In these types of forests, good stem detection and tree delineation performances were obtained with high density ALS ( $> 70 \text{ per m}^2$ ) acquired in leaf-off conditions. However, the accurate localization and delineation of tree shapes across all environments and in particular in multi-layered mixed forest remains an unsolved problem with currently no universal solution. It can also be noted that delineating the exact shape of trees may not always be required and that just detecting the stems may be sufficient for some applications. In chapter 4, methods to estimate stem diameter through allometry and direct measurement in the ALS point cloud were presented. The allometric approach is based on a modified state of the art method and the direct

measurement approach (which also estimates taper) is novel. Both approaches are able to predict diameter with about 10% error. However, the allometric approach requires that individual tree height and crown extent be determined, while the direct measurement approach only requires stem detection. It also has the advantage of not relying on knowledge about the species. With the increasing availability of high density ALS (including from unmanned systems), it can be expected that direct measurement of stem geometry will become more relevant in the near future. In the meantime, indirect allometric methods can provide valuable information on DBH, but is conditioned by the delineation performances of individual tree segmentation algorithms. Finally, in chapter 5, eight out of nine tree species commonly found in Switzerland could be successfully identified using a combination of ALS and AHI derived features. Using ALS alone a maximum overall accuracy of 91% and a Kappa score of 89% were obtained. It was found that non-structural ALS features (such as echo ranking and intensity distributions) played the most important role in discriminating genus/species, confirming previously reported observations. It was also determined, through a novel error simulation approach, that a moderate amount of individual tree segmentation error did not affect the ALS based classification performance very much if the classifier was trained on error-containing observations. Using moderate resolution (2.5-3 m) AHI acquired in leaf-on (July) conditions, a maximum overall accuracy of 0.85% and a kappa score of 0.82% were obtained at the crown scale. Systematically higher scores were obtained at the crown scale versus the pixel scale, once again highlighting the importance of individual tree segmentation. The combination of ALS and AHI data increased the classification scores by about ~5%. It was suspected that the two poorly identified species, ash and maple, had an insufficient number of observations to reliably calibrate the classification models and the reported performances for these two species should not be considered reliable. Based on the work conducted here, to support forest inventories at the individual tree scale, it is recommended to acquire high density ( $> 100$  points per  $m^2$ ) ALS in leaf-off conditions and to ensure echo intensity is calibrated/corrected adequately. Under these conditions, the probabilities of locating, determining the diameter and identifying the species of individual trees can be considered high for trees in the upper and intermediate canopy layers. It is also recommended that any future work should focus on improving individual tree delineation and/or stem detection in complex multi-layered forests which is a prerequisite for tree scale inventories and which has not yet reached operational readiness.

### **How can the reliability of remote sensing derived forest inventory maps be quantified?**

It is important that practitioners using ALS derived products for inventory purposes be provided with a reasonable estimate of their reliability. However, reading through the vast number of publications on remote sensing of forests, a lack of standardization in the reporting of errors and the quantification of problem difficulty (e.g. detecting trees in dense tropical forests is much harder than in a sparse woodland pasture) is apparent. In particular, the performance metrics used to characterize tree segmentation algorithms do not necessarily coincide with reality and are hardly comparable between studies. This thesis attempted to address part of the problem by describing and using a rigorous error assessment framework which compares automatic segmentation to a reliable manual segmentation in terms of 3D shape. It is suggested that an indication of 3D shape (delineation) quality, which has been mostly ignored in segmentation studies, should be included in future studies and/or forest maps. Using a simulated forest dataset, a simple method to quantify the difficulty of individual tree segmentation problems in terms of spatial adjacency was also proposed. A novel method (called ensemble filtering) to estimate segmentation error without field surveys and/or manually delineated tree crowns was also developed and provides a practical alternative to estimate the quality of segmentation in coniferous forests.

### **What is the relative value of ALS and AHI to support current forest inventory needs?**

For the classification of the nine considered species, the benefit of adding descriptive features derived from the APEX AHI to ALS was about 5% in terms of overall accuracy and kappa, about 10% in terms of F-score. The addition of AHI features also helped resolve several species confusion cases (ash in particular). In this regard AHI has a significant value, but it was determined that high density ALS (> 70 points per m<sup>2</sup>) acquired in leaf-off conditions already provides sufficient information to reliably identify seven of the nine considered species. Moreover, for the determination of basic forest inventory metrics at the tree scale (in particular stem location and DBH), the AHI does not bring any added value. Thus, if a single sensor solution had to be chosen (e.g. for economic or availability reasons), the author would suggest ALS. If a multi-solution was to be recommended in the current state of affairs, high spatial resolution multispectral imagery (down to 10 cm), which is already available operationally, may be a more adequate and lightweight solution than AHI. Such imagery is capable of providing detailed information on canopy texture which may be used in place of detailed spectral information. It is also important to note that the validity of this conclusion could be shorted lived, as new low cost hyperspectral sensors may become available in the near future.

### **Perspectives**

The short time span of this thesis has been a pivotal period, with data acquisition, analysis and interpretation means evolving significantly; bringing new opportunities and challenges along the way. Several expected developments should be underlined:

- ALS is becoming a mainstream product in national land surveys. Because of its ability to acquire high resolution data over large areas (several hundred km<sup>2</sup> per flight), it will continue to play a leading role in the characterization of forests, in the years to come. Data from new multispectral LiDAR sensors (Optech Titan, Riegl VQ-1560i-DW) will also be more common in forestry applications, for example to improve tree species identification (Budei et al., 2018; Axelsson et al., 2018; Dalponte et al., 2018a; Yu et al., 2017; Ahokas et al., 2016). Within the next decade, single flight coverage and point density may be further extended when photon counting and Geiger mode LiDAR reach commercial maturity (Harding et al., 2011; Swatantran et al., 2016; Stoker et al., 2016; Wästlund et al., 2018).
- Unmanned aerial vehicles already provide a practical and low cost solution to support forest monitoring up to several hundreds of hectares (Torresan et al., 2017; Zhang et al., 2016a; Tang and Shao, 2015; Paneque-Gálvez et al., 2014). A series of current or near-future developments will further increase their ubiquity in forest surveys: improvements in battery energy storage capacity, availability of vertical take-off and landing (VTOL) fixed wing airframes, better navigation with multi-constellation (GPS, GLONASS, Galileo, BeiDou) and dual-frequency GNSS chips (e.g. Broadcom BCM47755) previously restricted to high grade surveying equipment, small and lightweight active 3D sensing capabilities using waveform LiDAR or time-of-flight cameras (flash LiDAR).
- In late 2018, two spaceborne LiDAR missions, the Ice, Cloud, and land Elevation Satellite 2 (IceSat-2) and the Global Ecosystem Dynamics Investigation (GEDI) will start mapping canopy height at the global scale. These systems will complete the capabilities of passive optical satellites (e.g. Landsat-8, Sentinel-2) and will help monitor global forest carbon stocks. Maps based on this type of data will play a central role in substantiating forest degradation reports and verifying the effectiveness of conservation measures as required by the United Nations framework for Reducing Emissions from Deforestation and forest Degradation (REDD+).

- Improvements in the navigation, depth sensing and computation capabilities of consumer grade mobile devices will also benefit operations during local forest inventories. More accurate and faster positioning under canopies will help correlate field observations with remote sensing data. The addition of embedded or external depth sensors to mobile devices will also allow live measurement of 3D structures and preliminary studies have already demonstrated possible applications of such sensors in forest inventories (Hyyppä et al., 2017; Tomaštík et al., 2017; Huang et al., 2018). Finally, advanced image processing and classification capabilities embedded in the mobile devices will help with automatic identification of species and disease diagnosis from pictures (Joly et al., 2016).
- Mobile terrestrial surveying instruments that use integrated navigation (GNSS, IMU, LiDAR) for Simultaneous Localization and Mapping (SLAM) may eventually replace the more bulky static terrestrial laser scanners (Tang et al., 2015; Ryding et al., 2015; Qian et al., 2016; Bauwens et al., 2016).
- Digital interactive representations of forests will allow for easier communication, planning and execution of maintenance and exploitation tasks. Immersive virtual or augmented reality devices could for example be used for training purposes (e.g. virtual timber marking).
- high performance object detection and classification algorithms based on the deep learning paradigm have emerged and their application is diffusing in all research domains including forest remote sensing. In particular, 2D and 3D convolutional neural networks are showing promising performances for individual tree detection and might bring a decisive answer to a problem which has been tackled for two decades without any universal solution (Ayrey and Hayes, 2018; Windrim and Bryson, 2018).

Overall, effective remote sensing solutions are being developed to answer forest mapping requirements at the local, regional and global scales. However, in this constantly expanding and diversifying technological landscape, transmitting remote sensing skills or providing simple turn key solutions to forest practitioners remains problematic. With thousands of scientific publications investigating the benefits of remote sensing for forestry, but few practical software tools and tutorials available to field foresters, the adoption pace is relatively slow. Perhaps, the single most important point to improve this situation is a better integration of the forestry and remote sensing (surveying) communities, so that knowledge and skills can be shared.



## Bibliography

- Agutter, P. S., Wheatley, D. N., Nov. 2004. Metabolic scaling: Consensus or controversy? *Theoretical Biology and Medical Modelling* 1, 13.
- Ahokas, E., Hyppä, J., Yu, X., Liang, X., Matikainen, L., Karila, K., Litkey, P., Kukko, A., Jaakkola, A., Kaartinen, H., Holopainen, M., Vastaranta, M., Jun. 2016. Towards Automatic Single-Sensor Mapping by Multispectral Airborne Laser Scanning. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 49B3, 155–162.
- Alexander, C., Jul. 2009. Delineating tree crowns from airborne laser scanning point cloud data using Delaunay triangulation. *International Journal of Remote Sensing* 30 (14), 3843–3848.
- Allouis, T., Durrieu, S., Véga, C., Couteron, P., Apr. 2013. Stem Volume and Above-Ground Biomass Estimation of Individual Pine Trees From LiDAR Data: Contribution of Full-Waveform Signals. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (2), 924–934.
- Alonso, M., Bookhagen, B., Roberts, D. A., May 2014. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sensing of Environment* 148, 70–83.
- American Society for Photogrammetry and Remote Sensing, Jul. 2013. LAS Specification Version 1.4 - R13. Format specification, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, USA.
- Amiri, N., Polewski, P., Yao, W., Krzystek, P., Skidmore, A., Li, D., Sep. 2017. Detection of single tree stems in forested areas from high density ALS point clouds using 3D shape descriptors. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. IV-2-W4. Wuhan, China, pp. 35–42.
- Andersen, H.-E., Reutebuch, S. E., McGaughey, R. J., Oct. 2006. A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. *Canadian Journal of Remote Sensing* 32 (5), 355–366.

- Anderson, J. E., Ducey, M. J., Fast, A., Martin, M. E., Lepine, L., Smith, M.-L., Lee, T. D., Dubayah, R. O., Hofton, M. A., Hyde, P., Peterson, B. E., Blair, J. B., 2011. Use of waveform lidar and hyperspectral sensors to assess selected spatial and structural patterns associated with recent and repeat disturbance and the abundance of sugar maple (*Acer saccharum* Marsh.) in a temperate mixed hardwood and conifer forest. *Journal of Applied Remote Sensing* 5 (1), 053504–053504–18.
- Anderson, J. E., Plourde, L. C., Martin, M. E., Braswell, B. H., Smith, M.-L., Dubayah, R. O., Hofton, M. A., Blair, J. B., Apr. 2008. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment* 112 (4), 1856–1870.
- Aplin, P., Jun. 2006. On scales and dynamics in observing the environment. *International Journal of Remote Sensing* 27 (11), 2123–2140.
- Aschoff, T., Spiecker, H., 2004. Algorithms for the automatic detection of trees in laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 8), W2.
- Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Jones, M. O., Martin, R. E., Boardman, J., Hughes, R. F., 2008. Invasive species detection in Hawaiian rainforests using airborne imaging spectroscopy and LiDAR. *Remote Sensing of Environment* 112 (5), 1942–1955.
- Avery, T. E., 1969. Forester's guide to aerial photo interpretation.
- Axelsson, A., Lindberg, E., Olsson, H., Jan. 2018. Exploring Multispectral ALS Data for Tree Species Classification. *Remote Sensing* 10 (2), 183.
- Ayrey, E., Fraver, S., Kershaw Jr, J. A., Kenefic, L. S., Hayes, D., Weiskittel, A. R., Roth, B. E., 2017. Layer stacking: A novel algorithm for individual forest tree segmentation from LiDAR point clouds. *Canadian Journal of Remote Sensing* 43 (1), 16–27.
- Ayrey, E., Hayes, D. J., Apr. 2018. The Use of Three-Dimensional Convolutional Neural Networks to Interpret LiDAR for Forest Inventory. *Remote Sensing* 10 (4), 649.
- Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., Houghton, R. A., Mar. 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change* 2 (3), 182–185.
- Baena, S., Moat, J., Whaley, O., Boyd, D. S., Nov. 2017. Identifying species from the air: UAVs and the very high resolution challenge for plant conservation. *PLOS ONE* 12 (11), e0188714.
- Baldeck, C. A., Asner, G. P., Mar. 2014. Improving Remote Species Identification through Efficient Training Data Collection. *Remote Sensing* 6 (4), 2682–2698.
- Baldeck, C. A., Asner, G. P., Martin, R. E., Anderson, C. B., Knapp, D. E., Kellner, J. R., Wright, S. J., Jul. 2015. Operational Tree Species Mapping in a Diverse Tropical Forest with Airborne Imaging Spectroscopy. *PLOS ONE* 10 (7), e0118403.
- Baldeck, C. A., Colgan, M. S., Féret, J.-B., Levick, S. R., Martin, R. E., Asner, G. P., Jan. 2014. Landscape-scale variation in plant community composition of an African savanna from airborne species mapping. *Ecological Applications* 24 (1), 84–93.
- Ballanti, L., Blesius, L., Hines, E., Kruse, B., May 2016. Tree Species Classification Using Hyperspectral Imagery: A Comparison of Two Classifiers. *Remote Sensing* 8 (6), 445.

- Barilotti, A., Crosilla, F., Sepic, F., 2009. Curvature analysis of LiDAR data for single tree species classification in alpine latitude forests. *Laser scanning*, 1–2.
- Barnes, R. J., Dhanoa, M. S., Lister, S. J., May 1989. Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy* 43 (5), 772–777.
- Barrett, F., McRoberts, R. E., Tomppo, E., Cienciala, E., Waser, L. T., Mar. 2016. A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. *Remote Sensing of Environment* 174, 279–289.
- Bauwens, S., Bartholomeus, H., Calders, K., Lejeune, P., Jun. 2016. Forest Inventory with Terrestrial LiDAR: A Comparison of Static and Hand-Held Mobile Laser Scanning. *Forests* 7 (6), 127.
- Bayer, D., Seifert, S., Pretzsch, H., Aug. 2013. Structural crown properties of Norway spruce (*Picea abies* L. Karst.) and European beech (*Fagus sylvatica* L.) in mixed versus pure stands revealed by terrestrial laser scanning. *Trees* 27 (4), 1035–1047.
- Bejan, A., Oct. 2000. *Shape and Structure, from Engineering to Nature*. Cambridge University Press.
- Belgiu, M., Drăguț, L., Apr. 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24–31.
- Bentley, L. P., Stegen, J. C., Savage, V. M., Smith, D. D., von Allmen, E. I., Sperry, J. S., Reich, P. B., Enquist, B. J., Jul. 2013. An empirical assessment of tree branching networks and implications for plant allometric scaling models. *Ecology Letters* 16 (8), 1069–1078.
- Bienert, A., Scheller, S., Keane, E., Mohan, F., Nugent, C., 2007. Tree detection and diameter estimations by analysis of forest terrestrial laserscanner point clouds. In: ISPRS Workshop on Laser Scanning. Vol. 2007. pp. 50–55.
- Blomley, R., Hovi, A., Weinmann, M., Hinz, S., Korpela, I., Jutzi, B., Nov. 2017. Tree species classification using within crown localization of waveform LiDAR attributes. *ISPRS Journal of Photogrammetry and Remote Sensing* 133, 142–156.
- Bock, J., Piboule, A., Jolly, A., 2017. TidALS: Trunk identification in dense Airborne LiDAR Scanner to estimate tree diamter.
- Borchert, R., Slade, N. A., 1981. Bifurcation Ratios and the Adaptive Geometry of Trees. *Botanical Gazette* 142 (3), 394–401.
- Bork, E. W., Su, J. G., Nov. 2007. Integrating LIDAR data and multispectral imagery for enhanced classification of rangeland vegetation: A meta analysis. *Remote Sensing of Environment* 111 (1), 11–24.
- Boutin, J., Rey, P., Luzu, G., Miguet, J. M., Chabrol, P., Plaisance, G., 1953. La photographie aérienne et à ses applications forestières. *Revue Forestière Française* (11), 746.
- Bouvier, M., Durrieu, S., Fournier, R. A., Renaud, J.-P., Jan. 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment* 156, 322–334.
- Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS Journal of Photogrammetry and Remote Sensing* 61 (5), 325–340.

- Brandtberg, T., Warner, T. A., Landenberger, R. E., McGraw, J. B., May 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment* 85 (3), 290–303.
- Brede, B., Lau, A., Bartholomeus, H. M., Kooistra, L., Oct. 2017. Comparing RIEGL RiCOPTER UAV LiDAR Derived Canopy Height and DBH with Terrestrial LiDAR. *Sensors* 17 (10), 2371.
- Breiman, L., Aug. 1996. Bagging predictors. *Machine Learning* 24 (2), 123–140.
- Breiman, L., Oct. 2001. Random Forests. *Machine Learning* 45 (1), 5–32.
- Broge, N. H., Leblanc, E., May 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment* 76 (2), 156–172.
- Brolly, G., Király, G., 2009. Algorithms for stem mapping by means of terrestrial laser scanning. *Acta Silvatica et Lignaria Hungarica* 5, 119–130.
- Bruggisser, M., Roncat, A., Schaeppman, M. E., Morsdorf, F., Jul. 2017. Retrieval of higher order statistical moments from full-waveform LiDAR data for tree species classification. *Remote Sensing of Environment* 196, 28–41.
- Bucksch, A., Lindenbergh, R., Menenti, M., 2010. SkelTre. *The Visual Computer* 26 (10), 1283–1300.
- Bucksch, A., Lindenbergh, R., Menenti, M., Rahman, M. Z., Aug. 2009. Skeleton-based botanic tree diameter estimation from dense LiDAR data. In: Lidar Remote Sensing for Environmental Monitoring X. Vol. 7460. International Society for Optics and Photonics, p. 746007.
- Bucksch, A., Lindenbergh, R., Rahman, M. Z. A., Menenti, M., Jun. 2014. Breast Height Diameter Estimation From High-Density Airborne LiDAR Data. *IEEE Geoscience and Remote Sensing Letters* 11 (6), 1056–1060.
- Buddenbaum, H., Schlerf, M., Hill, J., Dec. 2005. Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *International Journal of Remote Sensing* 26 (24), 5453–5465.
- Budei, B. C., St-Onge, B., Hopkinson, C., Audet, F.-A., Jan. 2018. Identifying the genus or species of individual trees using a three-wavelength airborne lidar system. *Remote Sensing of Environment* 204, 632–647.
- Burt, A. P., Sep. 2017. New 3D measurements of forest structure. Doctoral, UCL (University College London).
- Cabo, C., Ordóñez, C., López-Sánchez, C. A., Armesto, J., Jul. 2018. Automatic dendrometry: Tree detection, tree height and diameter estimation using terrestrial laser scanning. *International Journal of Applied Earth Observation and Geoinformation* 69, 164–174.
- Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., Kaasalainen, M., Feb. 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution* 6 (2), 198–208.

- Calders, K., Origo, N., Burt, A., Disney, M., Nightingale, J., Raumonen, P., Åkerblom, M., Malhi, Y., Lewis, P., Jun. 2018. Realistic Forest Stand Reconstruction from Terrestrial LiDAR for Radiative Transfer Modelling. *Remote Sensing* 10 (6), 933.
- Cao, L., Coops, N. C., Innes, J. L., Dai, J., Ruan, H., She, G., Jul. 2016. Tree species classification in subtropical forests using small-footprint full-waveform LiDAR data. *International Journal of Applied Earth Observation and Geoinformation* 49, 39–51.
- Cardone, A., Gupta, S. K., Karnik, M., 2003. A survey of shape similarity assessment algorithms for product design and manufacturing applications. *Journal of Computing and Information Science in Engineering* 3 (2), 109–118.
- Castro-Esau, K. L., Sánchez-Azofeifa, G. A., Rivard, B., Wright, S. J., Quesada, M., Jan. 2006. Variability in leaf optical properties of Mesoamerican trees and the potential for species classification. *American Journal of Botany* 93 (4), 517–530.
- Chang, C.-C., Lin, C.-J., May 2011. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* 2 (3), 27:1–27:27.
- Cheang, E. K., Cheang, T. K., Tay, Y. H., 2017. Using Convolutional Neural Networks to Count Palm Trees in Satellite Images. arXiv preprint arXiv:1701.06462.
- Chen, M., Wan, Y., Wang, M., Xu, J., 2018. Automatic Stem Detection in Terrestrial Laser Scanning Data With Distance-Adaptive Search Radius. *IEEE Transactions on Geoscience and Remote Sensing* PP (99), 1–12.
- Chen, Q., Baldocchi, D., Gong, P., Kelly, M., Aug. 2006. Isolating Individual Trees in a Savanna Woodland Using Small Footprint Lidar Data. *Photogrammetric Engineering & Remote Sensing* 72 (8), 923–932.
- Chen, Q., Gong, P., Baldocchi, D., Tian, Y. Q., 2007. Estimating basal area and stem volume for individual trees from lidar data. *Photogrammetric Engineering & Remote Sensing* 73 (12), 1355–1365.
- Chisholm, R. A., Cui, J., Lum, S. K., Chen, B. M., 2013. UAV LiDAR for below-canopy forest surveys. *Journal of Unmanned Vehicle Systems* 1 (01), 61–68.
- Cho, M., Debba, P., Mathieu, R., Naidoo, L., Van Aardt, J., Asner, G., 2010. Improving Discrimination of Savanna Tree Species Through a Multiple-Endmember Spectral Angle Mapper Approach: Canopy-Level Analysis. *IEEE Transactions on Geoscience and Remote Sensing* 48 (11), 4133–4142.
- Cho, M. A., Mathieu, R., Asner, G. P., Naidoo, L., van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I. P., 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment* 125, 214–226.
- Clark, M. L., 2005. An Assessment of Hyperspectral and Lidar Remote Sensing for the Monitoring of Tropical Rain Forest Trees.
- Clark, M. L., Roberts, D. A., 2012. Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sensing* 4 (6), 1820–1855.
- Clark, M. L., Roberts, D. A., Clark, D. B., Jun. 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment* 96 (3), 375–398.

- Colgan, M. S., Baldeck, C. A., Féret, J.-B., Asner, G. P., 2012. Mapping Savanna Tree Species at Ecosystem Scales Using Support Vector Machine Classification and BRDF Correction on Airborne Hyperspectral and LiDAR Data. *Remote Sensing* 4 (11), 3462–3480.
- Congalton, R. G., Green, K., Dec. 2008. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Second Edition. CRC Press.
- Coomes, D. A., Nov. 2006. Challenges to the generality of WBE theory. *Trends in Ecology & Evolution* 21 (11), 593–596.
- Coomes, D. A., Lines, E. R., Allen, R. B., Apr. 2011. Moving on from Metabolic Scaling Theory: Hierarchical models of tree growth and asymmetric competition for light. *Journal of Ecology* 99 (3), 748–756.
- Côté, J.-F., Widlowski, J.-L., Fournier, R. A., Verstraete, M. M., May 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. *Remote Sensing of Environment* 113 (5), 1067–1081.
- Dabiri, Z., Lang, S., 2018. Spectral-Spatial Dimensionality Reduction of APEX Hyperspectral Imagery for Tree Species Classification; a Case Study of Salzach Riparian Mixed Forest.
- Dalponte, M., Bruzzone, L., Gianelle, D., 2008. Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas. *IEEE Transactions on Geoscience and Remote Sensing* 46 (5), 1416–1427.
- Dalponte, M., Bruzzone, L., Gianelle, D., Jul. 2011. A System for the Estimation of Single-Tree Stem Diameter and Volume Using Multireturn LIDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* 49 (7), 2479–2490.
- Dalponte, M., Bruzzone, L., Gianelle, D., Aug. 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.
- Dalponte, M., Bruzzone, L., Vescovo, L., Gianelle, D., 2009. The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sensing of Environment* 113 (11), 2345–2355.
- Dalponte, M., Ene, L., Gobakken, T., Næsset, E., Gianelle, D., Dalponte, M., Ene, L. T., Gobakken, T., Næsset, E., Gianelle, D., Apr. 2018a. Predicting Selected Forest Stand Characteristics with Multispectral ALS Data. *Remote Sensing* 10 (4), 586.
- Dalponte, M., Frizzera, L., Ørka, H. O., Gobakken, T., Næsset, E., Gianelle, D., Feb. 2018b. Predicting stem diameters and aboveground biomass of individual trees using remote sensing data. *Ecological Indicators* 85, 367–376.
- Dalponte, M., Ørka, H. O., Ene, L. T., Gobakken, T., Næsset, E., Jan. 2014. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sensing of Environment* 140, 306–317.
- Dalponte, M., Ørka, H. O., Gobakken, T., Gianelle, D., Næsset, E., 2013. Tree species classification in boreal forests with hyperspectral data.
- Dalponte, M., Reyes, F., Kandare, K., Gianelle, D., 2015. Delineation of Individual Tree Crowns from ALS and Hyperspectral data: A comparison among four methods. *European Journal of Remote Sensing* 48, 365–382.

- Danusevicius, D., Masaitis, G., Mozgeris, G., Dec. 2014. Visible and near infrared hyperspectral imaging reveals significant differences in needle reflectance among Scots pine provenances. *Silvae Genetica* 63 (1-6), 169–180.
- Dassot, M., Colin, A., Santenoise, P., Fournier, M., Constant, T., Nov. 2012. Terrestrial laser scanning for measuring the solid wood volume, including branches, of adult standing trees in the forest environment. *Computers and Electronics in Agriculture* 89, 86–93.
- Dechesne, C., Mallet, C., Le Bris, A., Gouet-Brunet, V., Apr. 2017. Semantic segmentation of forest stands of pure species combining airborne lidar data and very high resolution multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 126, 129–145.
- Delagrange, S., Jauvin, C., Rochon, P., Mar. 2014. PypeTree: A Tool for Reconstructing Tree Perennial Tissues from Point Clouds. *Sensors* 14 (3), 4271–4289.
- Demetriades-Shah, T. H., Steven, M. D., Clark, J. A., Jul. 1990. High resolution derivative spectra in remote sensing. *Remote Sensing of Environment* 33 (1), 55–64.
- Deneubourg, J. L., Goss, S., Franks, N., Pasteels, J. M., Sep. 1989. The blind leading the blind: Modeling chemically mediated army ant raid patterns. *Journal of Insect Behavior* 2 (5), 719–725.
- Deng, S., Katoh, M., Takenaka, Y., Cheung, K., Ishii, A., Fujii, N., Gao, T., 2017. Tree Species Classification of Broadleaved Forests in Nagano, Central Japan, Using Airborne Laser Data and Multispectral Images. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33–38.
- Deng, S., Katoh, M., Yu, X., Hyppä, J., Gao, T., Dec. 2016. Comparison of Tree Species Classifications at the Individual Tree Level by Combining ALS Data and RGB Images Using Different Algorithms. *Remote Sensing* 8 (12), 1034.
- Dennison, P. E., Roberts, D. A., Oct. 2003. The effects of vegetation phenology on endmember selection and species mapping in southern California chaparral. *Remote Sensing of Environment* 87 (2), 295–309.
- Dian, Y., Li, Z., Pang, Y., Mar. 2015. Spectral and Texture Features Combined for Forest Tree species Classification with Airborne Hyperspectral Imagery. *Journal of the Indian Society of Remote Sensing* 43 (1), 101–107.
- Dijkstra, E. W., 1959. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik* 1, 269–271.
- Dinuls, R., Erins, G., Lorencs, A., Mednieks, I., Sinica-Sinavskis, J., 2012. Tree Species Identification in Mixed Baltic Forest Using LiDAR and Multispectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5 (2), 594–603.
- Disney, M. I., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S., Pfeifer, M., Jul. 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.
- Dong, T., Zhou, Q., Gao, S., Shen, Y., May 2018. Automatic Detection of Single Trees in Airborne Laser Scanning Data through Gradient Orientation Clustering. *Forests* 9 (6), 291.
- Dubois, E., 1897. Sur le rapport du poids de l'encéphale avec la grandeur du corps chez les mammifères. *Bulletins et Mémoires de la Société d'Anthropologie de Paris* 8 (1), 337–376.

- Düggelin, C., Keller, M., 2017. Schweizerisches Landesforstinventar - Feldaufnahme-Anleitung.
- Duncanson, L., Rourke, O., Dubayah, R., Nov. 2015. Small Sample Sizes Yield Biased Allometric Equations in Temperate Forests. *Scientific Reports* 5, 17153.
- Duncanson, L. I., Cook, B. D., Hurt, G. C., Dubayah, R. O., Nov. 2014. An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. *Remote Sensing of Environment* 154, 378–386.
- Dunford, R., Michel, K., Gagnage, M., Piégay, H., Trémelo, M.-L., Sep. 2009. Potential and constraints of Unmanned Aerial Vehicle technology for the characterization of Mediterranean riparian forest. *International Journal of Remote Sensing* 30 (19), 4915–4935.
- Edelsbrunner, H., Mücke, E. P., Jan. 1994. Three-dimensional Alpha Shapes. *ACM Trans. Graph.* 13 (1), 43–72.
- Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., Glennie, C. L., Joerg, P. C., LeWinter, A. L., Magney, T. S., Mandlburger, G., Morton, D. C., Müller, J., Vierling, K. T., Dec. 2016. Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences. *Remote Sensing of Environment* 186, 372–392.
- Elzinga, C., Shearer, R. C., Elzinga, G., 2005. Observer variation in tree diameter measurements. *Western Journal of Applied Forestry* 20 (2), 134–137.
- Ene, L., Næsset, E., Gobakken, T., Aug. 2012. Single tree detection in heterogeneous boreal forests using airborne laser scanning and area-based stem number estimates. *International Journal of Remote Sensing* 33 (16), 5171–5193.
- Engler, R., Waser, L. T., Zimmermann, N. E., Schaub, M., Berdos, S., Ginzler, C., Psomas, A., Dec. 2013. Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *Forest Ecology and Management* 310, 64–73.
- Enquist, B. J., 2002. Universal scaling in tree and vascular plant allometry: Toward a general quantitative theory linking plant form and function from cells to ecosystems. *Tree Physiology* 22, 1045–1064.
- Enquist, B. J., Niklas, K. J., Apr. 2001. Invariant scaling relations across tree-dominated communities. *Nature* 410 (6829), 655–660.
- Enquist, B. J., West, G. B., Charnov, E. L., Brown, J. H., Oct. 1999. Allometric scaling of production and life-history variation in vascular plants. *Nature* 401 (6756), 907–911.
- Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M., Lingua, E., Mongus, D., Pfeifer, N., May 2015. A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space. *Forests* 6 (5), 1721–1747.
- Falkowski, M. J., Smith, A. M. S., Hudak, A. T., Gessler, P. E., Vierling, L. A., Crookston, N. L., Apr. 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. *Canadian Journal of Remote Sensing* 32, 153–161.
- Fang, F., McNeil, B. E., Warner, T. A., Maxwell, A. E., Aug. 2018. Combining high spatial resolution multi-temporal satellite data with leaf-on LiDAR to enhance tree species discrimination at the crown level. *International Journal of Remote Sensing* 0 (0), 1–19.

- FAO, 2016. Global Forest Resources Assessment 2015: How Are the World's Forests Changing?, 2nd Edition. Rome, oCLC: 1003587002.
- Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., Ghosh, A., Dec. 2016. Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment* 186, 64–87.
- Fassnacht, F. E., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P. K., Koch, B., Jun. 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.
- Féret, J.-B., Asner, G. P., Aug. 2012. Semi-Supervised Methods to Identify Individual Crowns of Lowland Tropical Canopy Species Using Imaging Spectroscopy and LiDAR. *Remote Sensing* 4 (12), 2457–2476.
- Féret, J.-B., Asner, G. P., Jan. 2013. Tree Species Discrimination in Tropical Forests Using Airborne Imaging Spectroscopy. *IEEE Transactions on Geoscience and Remote Sensing* 51 (1), 73–84.
- Ferraro, P., Godin, C., 2000. A distance measure between plant architectures. *Annals of Forest Science* 57 (5), 445–461.
- Ferraz, A., Bretar, F., Jacquemoud, S., Gon\ccalves, G., Pereira, L., Tomé, M., Soares, P., 2012. 3-D mapping of a multi-layered Mediterranean forest using ALS data. *Remote Sensing of Environment* 121, 210–223.
- Ferraz, A., Saatchi, S., Mallet, C., Meyer, V., Sep. 2016. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment* 183, 318–333.
- Ferreira, M. P., Zortea, M., Zanotta, D. C., Shimabukuro, Y. E., de Souza Filho, C. R., Jun. 2016. Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sensing of Environment* 179, 66–78.
- Finer, M., Novoa, S., Weisse, M. J., Petersen, R., Mascaro, J., Souto, T., Stearns, F., Martinez, R. G., Jun. 2018. Combating deforestation: From satellite to intervention. *Science* 360 (6395), 1303–1305.
- Fischesser, B., Campan, J., Campan, M., Apr. 2008. Reconnaître facilement les arbres. Identifier, soigner, planter, protéger. Delachaux, Paris.
- Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24 (6), 381–395.
- Freeman, L. C., Mar. 1977. A Set of Measures of Centrality Based on Betweenness. *Sociometry* 40 (1), 35–41.
- Fritz, A., Kattenborn, T., Koch, B., Aug. 2013. UAV-based photogrammetric point clouds - Tree stem mapping in open stands in comparison to terrestrial laser scanner point clouds. In: ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XL-1-W2. Copernicus GmbH, Rostock, Germany, pp. 141–146.
- Gandia, S., Fernández, G., García, J. C., Moreno, J., 2004. Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign. *Esa Sp* 578, 40–48.

- Gao, B.-C., Montes, M. J., Davis, C. O., Goetz, A. F. H., Sep. 2009. Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment* 113, S17–S24.
- Garbulsky, M. F., Peñuelas, J., Gamon, J., Inoue, Y., Filella, I., Feb. 2011. The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies: A review and meta-analysis. *Remote Sensing of Environment* 115 (2), 281–297.
- Gastellu-Etchegorry, J.-P., Yin, T., Lauret, N., Cajgfinger, T., Gregoire, T., Grau, E., Feret, J.-B., Lopes, M., Guilleux, J., Dedieu, G., Malenovský, Z., Cook, B. D., Morton, D., Rubio, J., Durrieu, S., Cazanave, G., Martin, E., Ristorcelli, T., Feb. 2015. Discrete Anisotropic Radiative Transfer (DART 5) for Modeling Airborne and Satellite Spectroradiometer and LiDAR Acquisitions of Natural and Urban Landscapes. *Remote Sensing* 7 (2), 1667–1701.
- Gatziolis, D., Andersen, H.-E., others, 2008. A guide to LiDAR data acquisition and processing for the forests of the Pacific Northwest.
- Gatziolis, D., Popescu, S., Sheridan, R., Ku, N., 2010. Evaluation of terrestrial lidar technology for the development of local tree volume equations. In: Proceedings of SilviLaser. Vol. 1. Freiburg.
- Gayon, J., Oct. 2000. History of the Concept of Allometry. *Integrative and Comparative Biology* 40 (5), 748–758.
- Geerling, G. W., Labrador-Garcia, M., Clevers, J. G. P. W., Ragas, A. M. J., Smits, A. J. M., Oct. 2007. Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. *International Journal of Remote Sensing* 28 (19), 4263–4284.
- Ghosh, A., Fassnacht, F. E., Joshi, P., Koch, B., Feb. 2014a. 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, 49–63.
- Ghosh, A., Fassnacht, F. E., Joshi, P. K., Koch, B., Feb. 2014b. 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, 49–63.
- Ginzler, C., Bärtschi, H., Bedolla, A., Brassel, P., Hägeli, M., Hauser, M., Kamphues, M., Laranjeiro, L., Mathys, L., Uebersax, D., Weber, E., Wicki, P., Zulliger, D., 2005. Luftbildinterpretation LFI3 - Interpretationsanleitung zum dritten Landesforstinventar.
- Gitelson, A. A., Keydan, G. P., Merzlyak, M. N., Jun. 2006. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Research Letters* 33 (11).
- Gitelson, A. A., Merzlyak, M. N., Aug. 1997. Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing* 18 (12), 2691–2697.
- Gitelson, A. A., Zur, Y., Chivkunova, O. B., Merzlyak, M. N., Mar. 2002. Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. *Photochemistry and Photobiology* 75 (3), 272–281.
- Gobakken, T., Næsset, E., Dec. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scandinavian Journal of Forest Research* 19 (6), 529–542.

- Gobakken, T., Næsset, E., Dec. 2005. Weibull and percentile models for lidar-based estimation of basal area distribution. *Scandinavian Journal of Forest Research* 20 (6), 490–502.
- Godin, C., Ferraro, P., 2010. Quantifying the degree of self-nestedness of trees: Application to the structural analysis of plants. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on* 7 (4), 688–703.
- Gonzalez-Benecke, C. A., Gezan, S. A., Samuelson, L. J., Cropper, W. P., Leduc, D. J., Martin, T. A., Mar. 2014. Estimating *Pinus palustris* tree diameter and stem volume from tree height, crown area and stand-level parameters. *Journal of Forestry Research* 25 (1), 43–52.
- Gorte, B., Pfeifer, N., 2004. Structuring laser-scanned trees using 3D mathematical morphology. *International Archives of Photogrammetry and Remote Sensing* 35 (B5), 929–933.
- Gorte, B., Winterhalder, D., 2004a. Reconstruction of laser-scanned trees using filter operations in the 3D raster domain. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 8), W2.
- Gorte, B., Winterhalder, D., 2004b. Reconstruction of laser-scanned trees using filter operations in the 3D raster domain. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (8).
- Gould, S. J., 1966. Allometry and Size in Ontogeny and Phylogeny. *Biological Reviews* 41 (4), 587–638.
- Goulden, T., Hopkinson, C., 2010. The forward propagation of integrated system component errors within airborne lidar data. *Photogrammetric Engineering & Remote Sensing* 76 (5), 589–601.
- Graves, S., Asner, G., Martin, R., Anderson, C., Colgan, M., Kalantari, L., Bohlman, S., Graves, S. J., Asner, G. P., Martin, R. E., Anderson, C. B., Colgan, M. S., Kalantari, L., Bohlman, S. A., Feb. 2016. Tree Species Abundance Predictions in a Tropical Agricultural Landscape with a Supervised Classification Model and Imbalanced Data. *Remote Sensing* 8 (2), 161.
- Green, A. A., Berman, M., Switzer, P., Craig, M. D., Jan. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing* 26 (1), 65–74.
- Guan, H., Yu, Y., Ji, Z., Li, J., Zhang, Q., Nov. 2015. Deep learning-based tree classification using mobile LiDAR data. *Remote Sensing Letters* 6 (11), 864–873.
- Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. *Journal of machine learning research* 3 (Mar), 1157–1182.
- Guyot, G., Baret, F., Major, D. J., 1988. High spectral resolution: Determination of spectral shifts between the red and the near infrared. *International Archives of Photogrammetry and Remote Sensing* 11 (750-760).
- Habib, A., Bang, K. I., Kersting, A. P., Lee, D.-C., 2009. Error budget of LiDAR systems and quality control of the derived data. *Photogrammetric Engineering & Remote Sensing* 75 (9), 1093–1108.
- Hackenberg, J., Morhart, C., Sheppard, J., Spiecker, H., Disney, M., May 2014. Highly Accurate Tree Models Derived from Terrestrial Laser Scan Data: A Method Description. *Forests* 5 (5), 1069–1105.

- Halle, F., Oldeman, R. A. A., Tomlinson, P. B., 1978. Tropical Trees and Forests: An Architectural Analysis. Springer-Verlag, Berlin Heidelberg.
- Hamraz, H., Contreras, M. A., Zhang, J., Oct. 2016. A robust approach for tree segmentation in deciduous forests using small-footprint airborne LiDAR data. *International Journal of Applied Earth Observation and Geoinformation* 52, 532–541.
- Hamraz, H., Contreras, M. A., Zhang, J., Aug. 2017. Vertical stratification of forest canopy for segmentation of understory trees within small-footprint airborne LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 130, 385–392.
- Hamraz, H., Jacobs, N. B., Contreras, M. A., Clark, C. H., 2018. Deep learning for conifer/deciduous classification of airborne LiDAR 3D point clouds representing individual trees. *arXiv preprint arXiv:1802.08872*.
- Han, L. T., Merry, M., Gee, T., Strozzi, A. G., Delmas, P., Gimel'farb, G., Dec. 2017. Automated Kauri trees detection in high resolution aerial images. In: 2017 International Conference on Image and Vision Computing New Zealand (IVCNZ). pp. 1–6.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., Townshend, J. R. G., Nov. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342 (6160), 850–853.
- Hantson, W., Kooistra, L., Slim, P. A., Oct. 2012. Mapping invasive woody species in coastal dunes in the Netherlands: A remote sensing approach using LiDAR and high-resolution aerial photographs. *Applied Vegetation Science* 15 (4), 536–547.
- Harding, D. J., Dabney, P. W., Valett, S., Oct. 2011. Polarimetric, two-color, photon-counting laser altimeter measurements of forest canopy structure. In: International Symposium on Lidar and Radar Mapping 2011: Technologies and Applications. Vol. 8286. International Society for Optics and Photonics, p. 828629.
- Harikumar, A., Bovolo, F., Bruzzone, L., Oct. 2017a. An approach to conifer stem localization and modeling in high density airborne LiDAR data. In: Image and Signal Processing for Remote Sensing XXIII. Vol. 10427. International Society for Optics and Photonics, p. 104270Q.
- Harikumar, A., Bovolo, F., Bruzzone, L., 2017b. An Internal Crown Geometric Model for Conifer Species Classification With High-Density LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* PP (99), 1–17.
- Heidemann, H. K., Dec. 2014. Lidar base specification (version 1.2). In: U.S. Geological Survey Techniques and Methods. Vol. 11. U.S. Geological Survey.
- Heikkilä, V., Korpela, I., Tokola, T., Honkavaara, E., Parkkinen, J., 2011. An SVM classification of tree species radiometric signatures based on the Leica ADS40 sensor. *Geoscience and Remote Sensing, IEEE Transactions on* 49 (11), 4539–4551.
- Heinzel, J., Huber, M. O., Dec. 2016. Detecting Tree Stems from Volumetric TLS Data in Forest Environments with Rich Understory. *Remote Sensing* 9 (1), 9.
- Heinzel, J., Koch, B., Feb. 2011a. Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation* 13 (1), 152–160.

- Heinzel, J., Koch, B., Feb. 2011b. Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation* 13 (1), 152–160.
- Heinzel, J., Koch, B., Aug. 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, 101–110.
- Heinzel, J. N., Weinacker, H., Koch, B., Sep. 2011. Prior-knowledge-based single-tree extraction. *International Journal of Remote Sensing* 32 (17), 4999–5020.
- Heller, R. C., Doverspike, G. E., Aldrich, R. C., 1964. Identification of Tree Species on Large-Scale Panchromatic and Color Aerial Photographs. US Department of Agriculture, Forest Service.
- Hengl, T., Nov. 2006. Finding the right pixel size. *Computers & Geosciences* 32 (9), 1283–1298.
- Henning, J. G., Radtke, P. J., 2006. Detailed stem measurements of standing trees from ground-based scanning lidar. *Forest Science* 52 (1), 67–80.
- Henry, M., Bombelli, A., Trotta, C., Alessandrini, A., Birigazzi, L., Sola, G., Vieilledent, G., Santenoise, P., Longuetaud, F., Valentini, R., Picard, N., Saint-André, L., 2013. GlobAllomeTree: International platform for tree allometric equations to support volume, biomass and carbon assessment. *iForest - Biogeosciences and Forestry* 6 (6), 326.
- Hershey, R. R., Befort, W. A., 1995. Aerial photo guide to New England forest cover types. Gen. Tech. Rep. NE-195. Radnor, PA: US Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 70 p. 195, 1–70.
- Hesketh, M., Sánchez-Azofeifa, G. A., Mar. 2012. The effect of seasonal spectral variation on species classification in the Panamanian tropical forest. *Remote Sensing of Environment* 118, 73–82.
- Hill, R. A., Wilson, A., George, M., Hinsley, S., 2010. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science* 13 (1), 86–99.
- Höfle, B., Pfeifer, N., Dec. 2007. Correction of laser scanning intensity data: Data and model-driven approaches. *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (6), 415–433.
- Hollaus, M., Dorigo, W., Wagner, W., Schadauer, K., Höfle, B., Maier, B., Sep. 2009a. Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data. *International Journal of Remote Sensing* 30 (19), 5159–5175.
- Hollaus, M., Mücke, W., Höfle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., Regner, B., Oct. 2009b. Tree species classification based on full-waveform airborne laser scanning data. In: *Silvilaser 2009*. College Station, Texas, USA.
- Hollaus, M., Wagner, W., Maier, B., Schadauer, K., Aug. 2007. Airborne Laser Scanning of Forest Stem Volume in a Mountainous Environment. *Sensors* 7 (8), 1559–1577.
- Holmgren, J., Lindberg, E., Jun. 2013. Tree crown segmentation based on a geometric tree crown model for prediction of forest variables. *Canadian Journal of Remote Sensing* 39 (sup1), S86–S98.
- Holmgren, J., Nilsson, M., Olsson, H., Jun. 2003a. Estimation of tree height and stem volume on plots using airborne laser scanning. *Forest Science; Bethesda* 49 (3), 419–428.

- Holmgren, J., Nilsson, M., Olsson, H., Jun. 2003b. Estimation of Tree Height and Stem Volume on Plots Using Airborne Laser Scanning. *Forest Science* 49 (3), 419–428.
- Holmgren, J., Persson, A., Apr. 2004a. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment* 90 (4), 415–423.
- Holmgren, J., Persson, A., Apr. 2004b. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment* 90 (4), 415–423.
- Holmgren, J., Persson, A., Söderman, U., 2008. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *International Journal of Remote Sensing* 29 (5), 1537–1552.
- Honda, H., Fisher, J. B., Feb. 1978. Tree Branch Angle: Maximizing Effective Leaf Area. *Science* 199 (4331), 888–890.
- Honda, H., Fisher, J. B., Jan. 1979. Ratio of tree branch lengths: The equitable distribution of leaf clusters on branches. *Proceedings of the National Academy of Sciences* 76 (8), 3875–3879.
- Hopkinson, C., Chasmer, L., Gynan, C., Mahoney, C., Sitar, M., Sep. 2016. Multisensor and Multispectral LiDAR Characterization and Classification of a Forest Environment. *Canadian Journal of Remote Sensing* 42 (5), 501–520.
- Hopkinson, C., Chasmer, L., Young-Pow, C., Treitz, P., Mar. 2004. Assessing forest metrics with a ground-based scanning lidar. *Canadian Journal of Forest Research; Ottawa* 34 (3), 573–583.
- Horton, R. E., 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological society of America bulletin* 56 (3), 275–370.
- Hovi, A., Korhonen, L., Vauhkonen, J., Korpela, I., Feb. 2016. LiDAR waveform features for tree species classification and their sensitivity to tree- and acquisition related parameters. *Remote Sensing of Environment* 173, 224–237.
- Hu, S., Li, Z., Zhang, Z., He, D., Wimmer, M., Oct. 2017a. Efficient tree modeling from airborne LiDAR point clouds. *Computers & Graphics* 67, 1–13.
- Hu, X., Chen, W., Xu, W., Feb. 2017b. Adaptive Mean Shift-Based Identification of Individual Trees Using Airborne LiDAR Data. *Remote Sensing* 9 (2), 148.
- Huang, Y., Hou, S., Ling, H., Xu, H., Dec. 2018. Trunk volume prediction of individual *Populus euphratrica* trees based on point clouds analysis. *Ecological Indicators* 95, 964–971.
- Huxley, J., Strauss, R. E., Churchill, F. B., 1932. Problems of relative growth. Methuen And Company Limited.
- Hyyppä, H., Yu, X., Hyyppä, J., Kaartinen, H., Kaasalainen, S., Honkavaara, E., Rönnholm, P., 2005. Factors affecting the quality of DTM generation in forested areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (3/W19), 85–90.
- Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M., May 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing* 39 (5), 969–975.
- Hyyppä, J., Virtanen, J.-P., Jaakkola, A., Yu, X., Hyyppä, H., Liang, X., Dec. 2017. Feasibility of Google Tango and Kinect for Crowdsourcing Forestry Information. *Forests* 9 (1), 6.

- Immitzer, M., Atzberger, C., Koukal, T., Sep. 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing* 4 (9), 2661–2693.
- Ioki, K., Imanishi, J., Sasaki, T., Morimoto, Y., Kitada, K., Jan. 2010. Estimating stand volume in broad-leaved forest using discrete-return LiDAR: Plot-based approach. *Landscape and Ecological Engineering* 6 (1), 29.
- Itten, K. I., Dell'Endice, F., Hueni, A., Kneubühler, M., Schläpfer, D., Odermatt, D., Seidel, F., Huber, S., Schopfer, J., Kellenberger, T., Bühler, Y., D'Odorico, P., Nieke, J., Alberti, E., Meuleman, K., Oct. 2008. APEX - the Hyperspectral ESA Airborne Prism Experiment. *Sensors* 8 (10), 6235–6259.
- Jaakkola, A., Hyppä, J., Yu, X., Kukko, A., Kaartinen, H., Liang, X., Hyppä, H., Wang, Y., Jul. 2017. Autonomous Collection of Forest Field Reference—The Outlook and a First Step with UAV Laser Scanning. *Remote Sensing* 9 (8), 785.
- Jakubowski, M. K., Guo, Q., Kelly, M., Mar. 2013a. Tradeoffs between lidar pulse density and forest measurement accuracy. *Remote Sensing of Environment* 130, 245–253.
- Jakubowski, M. K., Li, W., Guo, Q., Kelly, M., Aug. 2013b. Delineating Individual Trees from Lidar Data: A Comparison of Vector and Raster-based Segmentation Approaches. *Remote Sensing* 5 (9), 4163–4186.
- Jensen, J. L. R., Humes, K. S., Conner, T., Williams, C. J., DeGroot, J., May 2006. Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar. *Canadian Journal of Forest Research* 36 (5), 1129–1138.
- Jensen, J. R., 2007. *Remote Sensing of the Environment: An Earth Resource Perspective*. Pearson Prentice Hall.
- Johnson, O., 2006. *Collins Tree Guide*. Collins.
- Joly, A., Bonnet, P., Goëau, H., Barbe, J., Selmi, S., Champ, J., Dufour-Kowalski, S., Affouard, A., Carré, J., Molino, J.-F., Boujema, N., Barthélémy, D., Nov. 2016. A look inside the PI@ntNet experience. *Multimedia Systems* 22 (6), 751–766.
- Jones, T. G., Coops, N. C., Sharma, T., Dec. 2010. Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment* 114 (12), 2841–2852.
- Jones, T. G., Coops, N. C., Sharma, T., 2011. Exploring the Utility of Hyperspectral Imagery and LiDAR Data for Predicting *Quercus garryana* Ecosystem Distribution and Aiding in Habitat Restoration. *Restoration Ecology* 19 (201), 245–256.
- Jucker, T., Bouriaud, O., Coomes, D. A., Aug. 2015. Crown plasticity enables trees to optimize canopy packing in mixed-species forests. *Functional Ecology* 29 (8), 1078–1086.
- Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K. Y., Forrester, D. I., Haeni, M., Higgins, S. I., Holdaway, R. J., Iida, Y., Lorimer, C., Marshall, P. L., Momo, S., Moncrieff, G. R., Ploton, P., Poorter, L., Rahman, K. A., Schlund, M., Sonké, B., Sterck, F. J., Trugman, A. T., Usoltsev, V. A., Vanderwel, M. C., Waldner, P., Wedeux, B. M. M., Wirth, C., Wöll, H., Woods, M., Xiang, W., Zimmermann, N. E., Coomes, D. A., Jul. 2016. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology*, n/a–n/a.

- Junod, P., Ammann, P., 2018. Liste des marteloscopes. Tech. rep., Centre de compétence en sylviculture (CCS), Centre forestier de formation Lyss Hardernstrasse 20, CH-3250 Lyss, <http://www.waldbau-sylviculture.ch>.
- Kaartinen, H., Hyppä, J., Yu, X., Vastaranta, M., Hyppä, 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., Mar. 2012a. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing* 4 (4), 950–974.
- Kaartinen, H., Hyppä, J., Yu, X., Vastaranta, M., Hyppä, 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., Mar. 2012b. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing* 4 (4), 950–974.
- Kandare, K., Dalponte, M., Ganelle, D., Chan, J.-W., Jul. 2014. A new procedure for identifying single trees in understory layer using discrete LiDAR data. In: Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International. pp. 1357–1360.
- Kandare, K., Dalponte, M., Ørka, H., Frizzera, L., Næsset, E., Kandare, K., Dalponte, M., Ørka, H. O., Frizzera, L., Næsset, E., Apr. 2017. Prediction of Species-Specific Volume Using Different Inventory Approaches by Fusing Airborne Laser Scanning and Hyperspectral Data. *Remote Sensing* 9 (5), 400.
- Kandare, K., Ørka, H. O., Chan, J. C.-W., Dalponte, M., 2016. Effects of forest structure and airborne laser scanning point cloud density on 3D delineation of individual tree crowns. *European Journal of Remote Sensing* 49 (1), 337–359.
- Kankare, V., Liang, X., Vastaranta, M., Yu, X., Holopainen, M., Hyppä, J., Oct. 2015. Diameter distribution estimation with laser scanning based multisource single tree inventory. *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 161–171.
- Karna, Y. K., Hussin, Y. A., Gilani, H., Bronsveld, M. C., Murthy, M. S. R., Qamer, F. M., Karky, B. S., Bhattacharai, T., Aigong, X., Baniya, C. B., Jun. 2015. Integration of WorldView-2 and airborne LiDAR data for tree species level carbon stock mapping in Kayar Khola watershed, Nepal. *International Journal of Applied Earth Observation and Geoinformation* 38, 280–291.
- Kashani, A. G., Olsen, M. J., Parrish, C. E., Wilson, N., Nov. 2015. A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. *Sensors* 15 (11), 28099–28128.
- Ke, Y., Quackenbush, L. J., Im, J., Jun. 2010a. Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment* 114 (6), 1141–1154.
- Ke, Y., Zhang, W., Quackenbush, L. J., 2010b. Active contour and hill climbing for tree crown detection and delineation. *Photogrammetric Engineering & Remote Sensing* 76 (10), 1169–1181.
- Key, T., Warner, T. A., McGraw, J. B., Fajvan, M. A., Jan. 2001a. A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest. *Remote Sensing of Environment* 75 (1), 100–112.
- Key, T., Warner, T. A., McGraw, J. B., Fajvan, M. A., Jan. 2001b. A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest. *Remote Sensing of Environment* 75 (1), 100–112.

- Kim, S., McGaughey, R. J., Andersen, H.-E., Schreuder, G., Aug. 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.
- Kleiber, M., 1932. Body size and metabolism. *ENE* 1 (9).
- Ko, C., Kang, J., Sohn, G., 2018. Deep Multi-Task Learning for Tree Genera Classification. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 153–159.
- Ko, C., Sohn, G., Remmel, T. K., Jun. 2013. Tree genera classification with geometric features from high-density airborne LiDAR. *Canadian Journal of Remote Sensing* 39 (sup1), S73–S85.
- Ko, C., Sohn, G., Remmel, T. K., Miller, J. R., Aug. 2016. Maximizing the Diversity of Ensemble Random Forests for Tree Genera Classification Using High Density LiDAR Data. *Remote Sensing* 8 (8), 646.
- Koch, B., Heyder, U., Weinacker, H., Apr. 2006. Detection of Individual Tree Crowns in Airborne Lidar Data. *Photogrammetric Engineering & Remote Sensing* 72 (4), 357–363.
- Koch, B., Kattenborn, T., Straub, C., Vauhkonen, J., 2014. Segmentation of Forest to Tree Objects. In: *Forestry Applications of Airborne Laser Scanning. Managing Forest Ecosystems*. Springer, Dordrecht, pp. 89–112.
- Koenig, K., Höfle, B., Sep. 2016. Full-Waveform Airborne Laser Scanning in Vegetation Studies—A Review of Point Cloud and Waveform Features for Tree Species Classification. *Forests* 7 (9), 198.
- Koreň, M., Mokroš, M., Bucha, T., Dec. 2017. Accuracy of tree diameter estimation from terrestrial laser scanning by circle-fitting methods. *International Journal of Applied Earth Observation and Geoinformation* 63, 122–128.
- Korpela, I., Dahlin, B., Schäfer, H., Bruun, E., Haapaniemi, F., Honkasalo, J., Ilvesniemi, S., Kuutti, V., Linkosalmi, M., Mustonen, J., 2007. Single-tree forest inventory using lidar and aerial images for 3D treetop positioning, species recognition, height and crown width estimation. In: *Proceedings of ISPRS Workshop on Laser Scanning*. pp. 227–233.
- Korpela, I., Heikkilä, V., Honkavaara, E., Rohrbach, F., Tokola, T., Aug. 2011. Variation and directional anisotropy of reflectance at the crown scale — Implications for tree species classification in digital aerial images. *Remote Sensing of Environment* 115 (8), 2062–2074.
- Korpela, I., Tokola, T., Ørka, H. O., Koskinen, M., 2009. Small-footprint discrete-return LIDAR in tree species recognition. In: *Proceedings of ISPRS Hannover Workshop*. pp. 2–5.
- Koukoulas, S., Blackburn, G. A., 2005. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LIDAR and multi-spectral remotely sensed data. *International Journal of Remote Sensing* 26 (3), 431–455.
- Kozłowski, J., Konarzewski, M., Apr. 2004. Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant? *Functional Ecology* 18 (2), 283–289.
- Kozłowski, J., Konarzewski, M., Aug. 2005. West, Brown and Enquist's model of allometric scaling again: The same questions remain. *Functional Ecology* 19 (4), 739–743.
- Kukkonen, M., Korhonen, L., Maltamo, M., Suvanto, A., Packalen, P., Oct. 2018. How much can airborne laser scanning based forest inventory by tree species benefit from auxiliary optical data? *International Journal of Applied Earth Observation and Geoinformation* 72, 91–98.

- Kukunda, C. B., Duque-Lazo, J., González-Ferreiro, E., Thaden, H., Kleinn, C., Mar. 2018. Ensemble classification of individual Pinus crowns from multispectral satellite imagery and airborne LiDAR. *International Journal of Applied Earth Observation and Geoinformation* 65, 12–23.
- Kumar, L., Schmidt, K., Dury, S., Skidmore, A., 2002. Imaging Spectrometry and Vegetation Science. In: van der Meer, F. D., Jong, S. M. D. (Eds.), *Imaging Spectrometry*. No. 4 in Remote Sensing and Digital Image Processing. Springer Netherlands, pp. 111–155.
- Kuncheva, L. I., Aug. 2004. Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons.
- Kwak, D.-A., Lee, W.-K., Lee, J.-H., Biging, G. S., Gong, P., Oct. 2007. Detection of individual trees and estimation of tree height using LiDAR data. *Journal of Forest Research* 12 (6), 425–434.
- Lahivaara, T., Seppanen, A., Kaipio, J. P., Vauhkonen, J., Korhonen, L., Tokola, T., Maltamo, M., May 2014. Bayesian Approach to Tree Detection Based on Airborne Laser Scanning Data. *IEEE Transactions on Geoscience and Remote Sensing* 52 (5), 2690–2699.
- Lalonde, J.-F., Vandapel, N., Hebert, M., 2006. Automatic three-dimensional point cloud processing for forest inventory. Robotics Institute, 334.
- Lamprecht, S., Stoffels, J., Dotzler, S., Haßs, E., Udelhoven, T., 2015. aTrunk—An ALS-Based Trunk Detection Algorithm. *Remote Sensing* 7 (8), 9975–9997.
- Lapicque, L., 1907. Tableau général des poids somatique et encéphalique dans les espèces animales. *Bulletins et Mémoires de la Société d'Anthropologie de Paris* 8 (1), 248–270.
- Larsen, D. R., 2017. Simple taper: Taper equations for the field forester. In: 20th Central Hardwood Forest Conference. US Department of Agriculture, Forest Service, Northern Research Station, Columbia, pp. 265–278.
- Larsen, M., Eriksson, M., Descombes, X., Perrin, G., Brandtberg, T., Gougeon, F. A., Oct. 2011. Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions. *International Journal of Remote Sensing* 32 (20), 5827–5852.
- Laslier, M., Ba, A., Hubert-Moy, L., Dufour, S., Nov. 2017. Comparison of leaf-on and leaf-off ALS data for mapping riparian tree species. In: *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIX*. Vol. 10421. International Society for Optics and Photonics, p. 104210X.
- Leckie, D. G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F. A., Cloney, E., Paradine, D., Jan. 2005. Issues in species classification of trees in old growth conifer stands. *Canadian Journal of Remote Sensing* 31 (2), 175–190.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *nature* 521 (7553), 436.
- Lee, H., Slatton, K. C., Roth, B. E., JR, W. P. C., Jan. 2010. Adaptive clustering of airborne LiDAR data to segment individual tree crowns in managed pine forests. *International Journal of Remote Sensing* 31 (1), 117–139.
- Lee, J., Cai, X., Lellmann, J., Dalponte, M., Malhi, Y., Butt, N., Morecroft, M., Schönlieb, C. B., Coomes, D. A., Jun. 2016. Individual Tree Species Classification From Airborne Multisensor Imagery Using Robust PCA. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (6), 2554–2567.

- Lehnebach, R., Beyer, R., Letort, V., Heuret, P., Apr. 2018. The pipe model theory half a century on: A review. *Annals of Botany* 121 (5), 773–795.
- Leopold, L. B., 1971. Trees and streams: The efficiency of branching patterns. *Journal of Theoretical Biology* 31 (2), 339–354.
- Li, D., Ke, Y., Gong, H., Li, X., Li, D., Ke, Y., Gong, H., Li, X., Dec. 2015. Object-Based Urban Tree Species Classification Using Bi-Temporal WorldView-2 and WorldView-3 Images. *Remote Sensing* 7 (12), 16917–16937.
- Li, H.-T., Han, X.-G., Wu, J.-G., Oct. 2005. Lack of Evidence for 3/4 Scaling of Metabolism in Terrestrial Plants. *Journal of Integrative Plant Biology* 47 (10), 1173–1183.
- Li, J., Hu, B., Noland, T. L., Apr. 2013. Classification of tree species based on structural features derived from high density LiDAR data. *Agricultural and Forest Meteorology* 171–172, 104–114.
- Li, W., Fu, H., Yu, L., Cracknell, A., Dec. 2016. Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. *Remote Sensing* 9 (1), 22.
- Li, W., Guo, Q., Jakubowski, M. K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering & Remote Sensing* 78 (1), 75–84.
- Liang, X., Hyppä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., Holopainen, M., Brolly, G., Francesco, P., Hackenberg, J., Huang, H., Jo, H.-W., Katoh, M., Liu, L., Mokroš, M., Morel, J., Olofsson, K., Poveda-Lopez, J., Trochta, J., Wang, D., Wang, J., Xi, Z., Yang, B., Zheng, G., Kankare, V., Luoma, V., Yu, X., Chen, L., Vastaranta, M., Saarinen, N., Wang, Y., Oct. 2018. International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing* 144, 137–179.
- Liang, X., Hyppä, J., Matikainen, L., 2007. Deciduous-coniferous tree classification using difference between first and last pulse laser signatures. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (3/W52), 253–257.
- Liang, X., Litkey, P., Hyppä, J., Kaartinen, H., Vastaranta, M., Holopainen, M., Feb. 2012. Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning. *IEEE Transactions on Geoscience and Remote Sensing* 50 (2), 661–670.
- Lin, Y., Hyppä, J., Apr. 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.
- Lindberg, E., Eysn, L., Hollaus, M., Holmgren, J., Pfeifer, N., Jul. 2014. Delineation of Tree Crowns and Tree Species Classification From Full-Waveform Airborne Laser Scanning Data Using 3-D Ellipsoidal Clustering. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (7), 3174–3181.
- Lindberg, E., Hollaus, M., Apr. 2012. Comparison of Methods for Estimation of Stem Volume, Stem Number and Basal Area from Airborne Laser Scanning Data in a Hemi-Boreal Forest. *Remote Sensing* 4 (4), 1004–1023.
- Lindberg, E., Holmgren, J., Mar. 2017. Individual Tree Crown Methods for 3D Data from Remote Sensing. *Current Forestry Reports* 3 (1), 19–31.

- Lindberg, E., Holmgren, J., Olofsson, K., Olsson, H., Nov. 2012. Estimation of stem attributes using a combination of terrestrial and airborne laser scanning. *European Journal of Forest Research* 131 (6), 1917–1931.
- Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., Olsson, H., Mar. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *International Journal of Remote Sensing* 31 (5), 1175–1192.
- Lines, E. R., Zavala, M. A., Purves, D. W., Coomes, D. A., Sep. 2012. Predictable changes in aboveground allometry of trees along gradients of temperature, aridity and competition. *Global Ecology and Biogeography* 21 (10), 1017–1028.
- Liu, C., Xing, Y., Duanmu, J., Tian, X., Mar. 2018. Evaluating Different Methods for Estimating Diameter at Breast Height from Terrestrial Laser Scanning. *Remote Sensing* 10 (4), 513.
- Liu, D., Kelly, M., Gong, P., Mar. 2006. A spatial-temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery. *Remote Sensing of Environment* 101 (2), 167–180.
- Liu, H., Wu, C., Jun. 2018. Crown-level tree species classification from AISA hyperspectral imagery using an innovative pixel-weighting approach. *International Journal of Applied Earth Observation and Geoinformation* 68, 298–307.
- Liu, J., Shen, J., Zhao, R., Xu, S., Aug. 2013. Extraction of individual tree crowns from airborne LiDAR data in human settlements. *Mathematical and Computer Modelling* 58 (3), 524–535.
- Liu, L., Coops, N. C., Aven, N. W., Pang, Y., Oct. 2017. Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sensing of Environment* 200, 170–182.
- Liu, L., Pang, Y., Fan, W., Li, Z., Li, M., 2011. Fusion of airborne hyperspectral and LiDAR data for tree species classification in the temperate forest of northeast China. In: *Geoinformatics, 2011 19th International Conference On*. pp. 1–5.
- Liu, T., Im, J., Quackenbush, L. J., Dec. 2015. A novel transferable individual tree crown delineation model based on Fishing Net Dragging and boundary classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 110, 34–47.
- Livny, Y., Yan, F., Olson, M., Chen, B., Zhang, H., El-Sana, J., 2010. Automatic Reconstruction of Tree Skeletal Structures from Point Clouds. In: *ACM SIGGRAPH Asia 2010 Papers. SIGGRAPH ASIA '10*. ACM, New York, NY, USA, pp. 151:1–151:8.
- Lo, C. S., Lin, C., Apr. 2013. Growth-Competition-Based Stem Diameter and Volume Modeling for Tree-Level Forest Inventory Using Airborne LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* 51 (4), 2216–2226.
- Lorimer, N. D., Haight, R. G., Leary, R. A., 1994. The fractal forest: Fractal geometry and applications in forest science.
- Lu, X., Guo, Q., Li, W., Flanagan, J., Aug. 2014. A bottom-up approach to segment individual deciduous trees using leaf-off lidar point cloud data. *ISPRS Journal of Photogrammetry and Remote Sensing* 94, 1–12.

- Lucas, R., Bunting, P., Paterson, M., Chisholm, L., 2008. Classification of Australian forest communities using aerial photography, CASI and HyMap data. *Remote Sensing of Environment* 112 (5), 2088–2103.
- Luoma, V., Saarinen, N., Wulder, M. A., White, J. C., Vastaranta, M., Holopainen, M., Hyppä, J., Feb. 2017. Assessing Precision in Conventional Field Measurements of Individual Tree Attributes. *Forests* 8 (2), 38.
- Luther, J. E., Skinner, R., Fournier, R. A., Lier, V., R. O., Bowers, W. W., Coté, J.-F., Hopkinson, C., Moulton, T., Apr. 2014. Predicting wood quantity and quality attributes of balsam fir and black spruce using airborne laser scanner data. *Forestry: An International Journal of Forest Research* 87 (2), 313–326.
- Luzum, B., Starek, M., Slatton, K. C., 2004. Normalizing ALSM intensities. Tech. Rep. Rep\_2004-07-01, Civil and Coastal Engineering Department, University of Florida, Gainesville, FL, USA.
- Maas, H.-G., Bienert, A., Scheller, S., Keane, E., Mar. 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. *International Journal of Remote Sensing* 29 (5), 1579–1593.
- Magnussen, S., Næsset, E., Gobakken, T., Frazer, G., Apr. 2012. A fine-scale model for area-based predictions of tree-size-related attributes derived from LiDAR canopy heights. *Scandinavian Journal of Forest Research* 27 (3), 312–322.
- Maltamo, M., Eerikäinen, K., Packalén, P., Hyppä, J., Apr. 2006. Estimation of stem volume using laser scanning-based canopy height metrics. *Forestry: An International Journal of Forest Research* 79 (2), 217–229.
- Maltamo, M., Næsset, E., Bollandsås, O. M., Gobakken, T., Packalén, P., Dec. 2009. Non-parametric prediction of diameter distributions using airborne laser scanner data. *Scandinavian Journal of Forest Research* 24 (6), 541–553.
- Mascaro, J., Asner, G. P., Davies, S., Dehgan, A., Saatchi, S., Sep. 2014. These are the days of lasers in the jungle. *Carbon Balance and Management* 9 (1), 7.
- Matérn, B., 1956. On the geometry of the cross-section of a stem. *Meddelanden från Statens skogsforskningsinstitut* (0369-2167) 46 (16).
- Matsuki, T., Yokoya, N., Iwasaki, A., May 2015. Hyperspectral Tree Species Classification of Japanese Complex Mixed Forest With the Aid of Lidar Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (5), 2177–2187.
- McDaniel, M. W., Nishihata, T., Brooks, C. A., Salesses, P., Iagnemma, K., 2012. Terrain classification and identification of tree stems using ground-based LiDAR. *Journal of Field Robotics* 29 (6), 891–910.
- McGlone, J. C., George Y. G. Lee, American Society for Photogrammetry and Remote Sensing, 2013. Manual of Photogrammetry, sixth edition Edition. American Society for Photogrammetry and Remote Sensing, Bethesda Md.
- McRoberts, R. E., Tomppo, E. O., Oct. 2007. Remote sensing support for national forest inventories. *Remote Sensing of Environment* 110 (4), 412–419.

- Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L., Hendrix, C. J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering and Remote Sensing* 66 (11), 1367–1372.
- Mehtätalo, L., Maltamo, M., Packalen, P., 2007. Recovering plot-specific diameter distribution and height-diameter curve using ALS based stand characteristics. In: Proceedings of ISPRS Workshop Laser Scanning. pp. 12–14.
- Méndez, V., Rosell-Polo, J. R., Pascual, M., Escolà, A., Aug. 2016. Multi-tree woody structure reconstruction from mobile terrestrial laser scanner point clouds based on a dual neighbourhood connectivity graph algorithm. *Biosystems Engineering* 148, 34–47.
- Merzlyak, M. N., Solovchenko, A. E., Gitelson, A. A., 2003. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. *Postharvest Biology and Technology* 27 (2), 197–211.
- Metternicht, G., Jan. 2003. Vegetation indices derived from high-resolution airborne videography for precision crop management. *International Journal of Remote Sensing* 24 (14), 2855–2877.
- Meyer, F., Jul. 1994. Topographic distance and watershed lines. *Signal Processing* 38 (1), 113–125.
- Meyer, F., Beucher, S., Sep. 1990. Morphological segmentation. *Journal of Visual Communication and Image Representation* 1 (1), 21–46.
- Moffiet, T., Mengersen, K., Witte, C., King, R., Denham, R., Aug. 2005. Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. *ISPRS Journal of Photogrammetry and Remote Sensing* 59 (5), 289–309.
- Mongus, D., Žalik, B., Oct. 2015. An efficient approach to 3D single tree-crown delineation in LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 219–233.
- Monnet, J. M., Chanussot, J., Berger, F., May 2011. Support Vector Regression for the Estimation of Forest Stand Parameters Using Airborne Laser Scanning. *IEEE Geoscience and Remote Sensing Letters* 8 (3), 580–584.
- Morsdorf, F., Meier, E., Kötz, B., Itten, K. I., Dobbertin, M., Allgöwer, B., Aug. 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.
- Moskal, L. M., Zheng, G., Dec. 2011. Retrieving Forest Inventory Variables with Terrestrial Laser Scanning (TLS) in Urban Heterogeneous Forest. *Remote Sensing* 4 (1), 1–20.
- Muller-Landau, H. C., Condit, R. S., Chave, J., Thomas, S. C., Bohlman, S. A., Bunyavejchewin, S., Davies, S., Foster, R., Gunatilleke, S., Gunatilleke, N., Harms, K. E., Hart, T., Hubbell, S. P., Itoh, A., Kassim, A. R., LaFrankie, J. V., Lee, H. S., Losos, E., Makana, J.-R., Ohkubo, T., Sukumar, R., Sun, I.-F., Supardi, M. N. N., Tan, S., Thompson, J., Valencia, R., Muñoz, G. V., Wills, C., Yamakura, T., Chuyong, G., Dattaraja, H. S., Esufali, S., Hall, P., Hernandez, C., Kenfack, D., Kiratiprayoon, S., Suresh, H. S., Thomas, D., Vallejo, M. I., Ashton, P., Mar. 2016. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecology Letters* 9 (5), 575–588.
- Muller-Landau, H. C., Condit, R. S., Harms, K. E., Marks, C. O., Thomas, S. C., Bunyavejchewin, S., Chuyong, G., Co, L., Davies, S., Foster, R., Gunatilleke, S., Gunatilleke, N., Hart, T., Hubbell,

- S. P., Itoh, A., Kassim, A. R., Kenfack, D., LaFrankie, J. V., Lagunzad, D., Lee, H. S., Losos, E., Makana, J.-R., Ohkubo, T., Samper, C., Sukumar, R., Sun, I.-F., Supardi, M. N. N., Tan, S., Thomas, D., Thompson, J., Valencia, R., Vallejo, M. I., Muñoz, G. V., Yamakura, T., Zimmerman, J. K., Dattaraja, H. S., Esufali, S., Hall, P., He, F., Hernandez, C., Kiratiprayoon, S., Suresh, H. S., Wills, C., Ashton, P., Mar. 2006. Comparing tropical forest tree size distributions with the predictions of metabolic ecology and equilibrium models. *Ecology Letters* 9 (5), 589–602.
- Næsset, E., Apr. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* 80 (1), 88–99.
- Næsset, E., Dec. 2004. Accuracy of forest inventory using airborne laser scanning: Evaluating the first nordic full-scale operational project. *Scandinavian Journal of Forest Research* 19 (6), 554–557.
- Næsset, E., Oct. 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment* 98 (2–3), 356–370.
- Næsset, E., Bollandsås, O. M., Gobakken, T., Feb. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sensing of Environment* 94 (4), 541–553.
- Nagendra, H., Rocchini, D., 2008. High resolution satellite imagery for tropical biodiversity studies: The devil is in the detail. *Biodiversity and Conservation* 17 (14), 3431–3442.
- Naidoo, L., Cho, M. A., Mathieu, R., Asner, G., Apr. 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.
- Naveed, F., Hu, B., 2017. Individual Tree Crown Delineation Using Multi-Wavelength Titan LiDAR Data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 42.
- Nijland, W., Addink, E. A., De Jong, S. M., Van der Meer, F. D., Apr. 2009. Optimizing spatial image support for quantitative mapping of natural vegetation. *Remote Sensing of Environment* 113 (4), 771–780.
- Niklas, K. J., Oct. 1994. *Plant Allometry: The Scaling of Form and Process*. University of Chicago Press.
- Nikopensius, M., Pisek, J., Raabe, K., Dec. 2015. Spectral reflectance patterns and seasonal dynamics of common understory types in three mature hemi-boreal forests. *International Journal of Applied Earth Observation and Geoinformation* 43, 84–91.
- Nordkvist, K., Granholm, A.-H., Holmgren, J., Olsson, H., Nilsson, M., 2012. Combining optical satellite data and airborne laser scanner data for vegetation classification. *Remote Sensing Letters* 3 (5), 393–401.
- Nyland, R. D., May 2016. *Silviculture: Concepts and Applications*, Third Edition. Waveland Press.
- Oester, B., 2003. *Einführung in die Fernerkundung*. Tech. rep., Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft (WSL).

- Olofsson, K., Holmgren, J., Olsson, H., May 2014. Tree Stem and Height Measurements using Terrestrial Laser Scanning and the RANSAC Algorithm. *Remote Sensing* 6 (5), 4323–4344.
- Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., Kassem, K. R., Nov. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* 51 (11), 933–938.
- Omer, G., Mutanga, O., Abdel-Rahman, E. M., Adam, E., 2015. Performance of support vector machines and artificial neural network for mapping endangered tree species using WorldView-2 data in Dukuduku forest, South Africa. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (10), 4825–4840.
- Omule, S. A. Y., Oct. 1980. Personal Bias in Forest Measurements. *The Forestry Chronicle* 56 (5), 222–224.
- Oohata, S.-i., Shidei, T., 1971. Studies on the branching structure of trees: I. Bifurcation ratio of trees in Horton's law. *Japanese Journal of Ecology* 21 (1-2), 7–14.
- Ørka, H., Næsset, E., Bollandsås, O., Sep. 2007. Utilizing airborne laser intensity for tree species classification. In: ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007. Finland.
- 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* (ahead-of-print), 1–12.
- 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 (02), 125–138.
- Ørka, H. O., Næsset, E., Bollandsås, O. M., Jun. 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.
- Ørka, H. O., Næsset, E., Bollandsås, O. M., Jul. 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.
- Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., Jul. 2011. A Large and Persistent Carbon Sink in the World's Forests. *Science*, 1201609.
- Paneque-Gálvez, J., McCall, M. K., Napoletano, B. M., Wich, S. A., Koh, L. P., Jun. 2014. Small Drones for Community-Based Forest Monitoring: An Assessment of Their Feasibility and Potential in Tropical Areas. *Forests* 5 (6), 1481–1507.
- Papeş, M., Tupayachi, R., Martínez, P., Peterson, A. T., Asner, G. P., Powell, G. V. N., Apr. 2013. Seasonal Variation in Spectral Signatures of Five Genera of Rainforest Trees. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (2), 339–350.

- Paris, C., Bruzzone, L., Jul. 2016. A data-driven identification of growth-model classes for the adaptive estimation of single-tree stem diameter in LiDAR data. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 6918–6921.
- Paris, C., Valduga, D., Bruzzone, L., 2016. A Hierarchical Approach to Three-Dimensional Segmentation of LiDAR Data at Single-Tree Level in a Multilayered Forest. *IEEE Transactions on Geoscience and Remote Sensing* PP (99), 1–14.
- Parkan, M., Aug. 2017a. Digital-Forestry-Toolbox: A collection of digital forestry tools for Matlab [<https://github.com/mparkan/Digital-Forestry-Toolbox>].
- Parkan, M., Mar. 2017b. Données LiDAR aériennes: Pertinence de l’interprétation visuelle pour la foresterie. *Schweizerische Zeitschrift fur Forstwesen* 168 (3), 127–133.
- Parkan, M., Junod, P., Lugrin, R., Ginzler, C., 2018. A reference airborne LiDAR dataset for forest research. *Remote Sensing* (Submitted).
- Parkan, M., Tuia, D., 2015. Individual tree segmentation in deciduous forests using geodesic voting. In: Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, pp. 637–640.
- Parkan, M., Tuia, D., Feb. 2018. Estimating Uncertainty of Point-Cloud Based Single-Tree Segmentation with Ensemble Based Filtering. *Remote Sensing* 10 (2), 335.
- Patenaude, G., Milne, R., Dawson, T. P., 2005. Synthesis of remote sensing approaches for forest carbon estimation: Reporting to the Kyoto Protocol. *Environmental Science & Policy* 8 (2), 161–178.
- Peñuelas, J., Filella, I., Biel, C., Serrano, L., Save, R., 1993. The reflectance at the 950–970 nm region as an indicator of plant water status. *International journal of remote sensing* 14 (10), 1887–1905.
- Persello, C., Boulaïas, A., Dalponte, M., Gobakken, T., Næsset, E., Schölkopf, B., Oct. 2014. Cost-Sensitive Active Learning With Lookahead: Optimizing Field Surveys for Remote Sensing Data Classification. *IEEE Transactions on Geoscience and Remote Sensing* 52 (10), 6652–6664.
- Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing* 68 (9), 925–932.
- Persson, A., Holmgren, J., Söderman, U., Olsson, H., 2004. Tree species classification of individual trees in Sweden by combining high resolution laser data with high resolution near-infrared digital images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (8), 204–207.
- Pettorelli, N., Oct. 2013. The Normalized Difference Vegetation Index. OUP Oxford.
- Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J., Packalén, P., Dec. 2007. Preharvest Measurement of Marked Stands Using Airborne Laser Scanning. *Forest Science*; Bethesda 53 (6), 653–661.
- Peuhkurinen, J., Mehtätalo, L., Maltamo, M., Feb. 2011. Comparing individual tree detection and the area-based statistical approach for the retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands. *Canadian Journal of Forest Research* 41 (3), 583–598.

- Pfeifer, N., Gorte, B., Winterhalder, D., 2004. Automatic reconstruction of single trees from terrestrial laser scanner data. In: Proceedings of 20th ISPRS Congress. ISPRS Istanbul, pp. 114–119.
- Pham, L. T. H., Brabyn, L., Ashraf, S., Aug. 2016. Combining QuickBird, LiDAR, and GIS topography indices to identify a single native tree species in a complex landscape using an object-based classification approach. International Journal of Applied Earth Observation and Geoinformation 50, 187–197.
- Picard, N., Saint-André, L., Matieu, H., Aug. 2012. Manual for Building Tree Volume and Biomass Allometric Equations: From Field Measurement to Prediction. Food and Agriculture Organization of the United Nations (FAO) and Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD).
- Piironen, R., Heiskanen, J., Maeda, E., Viinikka, A., Pellikka, P., Aug. 2017. Classification of Tree Species in a Diverse African Agroforestry Landscape Using Imaging Spectroscopy and Laser Scanning. Remote Sensing 9 (9), 875.
- Pirotti, F., Oct. 2010. Assessing a Template Matching Approach for Tree Height and Position Extraction from Lidar-Derived Canopy Height Models of Pinus Pinaster Stands. Forests 1 (4), 194–208.
- Pirotti, F., Kobal, M., Roussel, J. R., Sep. 2017. A Comparison of Tree Segmentation Methods Using Very High Density Airborne Laser Scanner Data. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 42W7, 285–290.
- Pitkänen, J., Maltamo, M., Hyppä, J., Yu, X., 2004. Adaptive methods for individual tree detection on airborne laser based canopy height model. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (8), 187–191.
- Plaza, A., Martinez, P., Perez, R., Plaza, J., Sep. 2002. Spatial/spectral endmember extraction by multidimensional morphological operations. IEEE Transactions on Geoscience and Remote Sensing 40 (9), 2025–2041.
- Plaza, A., Martinez, P., Perez, R., Plaza, J., Mar. 2004. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 42 (3), 650–663.
- Popescu, S. C., Wynne, R. H., 2004. Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering and Remote Sensing 70 (5), 589–604.
- Pretzsch, H., Jan. 2006. Species-specific allometric scaling under self-thinning: Evidence from long-term plots in forest stands. Oecologia 146 (4), 572–583.
- Pretzsch, H., 2009. Forest Dynamics, Growth, and Yield. In: Forest Dynamics, Growth and Yield. Springer, Berlin, Heidelberg, pp. 1–39.
- Pretzsch, H., Sep. 2014. Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures. Forest Ecology and Management 327, 251–264.
- Preuksakarn, C., Boudon, F., Ferraro, P., Durand, J.-B., Nikinmaa, E., Godin, C., 2010. Reconstructing plant architecture from 3D laser scanner data. In: 6th International Workshop on Functional-Structural Plant Models. pp. 12–17.

- Pu, R., Landry, S., Sep. 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment* 124, 516–533.
- Pueschel, P., Newnham, G., Rock, G., Udelhoven, T., Werner, W., Hill, J., Mar. 2013. The influence of scan mode and circle fitting on tree stem detection, stem diameter and volume extraction from terrestrial laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing* 77, 44–56.
- Purves, D. W., Lichstein, J. W., Pacala, S. W., Sep. 2007. Crown Plasticity and Competition for Canopy Space: A New Spatially Implicit Model Parameterized for 250 North American Tree Species. *PLOS ONE* 2 (9), e870.
- Puttonen, E., Litkey, P., Hyppä, J., Dec. 2009. Individual Tree Species Classification by Illuminated-Shaded Area Separation. *Remote Sensing* 2 (1), 19–35.
- Qian, C., Liu, H., Tang, J., Chen, Y., Kaartinen, H., Kukko, A., Zhu, L., Liang, X., Chen, L., Hyppä, J., Dec. 2016. An Integrated GNSS/INS/LiDAR-SLAM Positioning Method for Highly Accurate Forest Stem Mapping. *Remote Sensing* 9 (1), 3.
- Raczko, E., Zagajewski, B., Jan. 2017. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing* 50 (1), 144–154.
- Radkau, J., 2012. *Wood: A History*. Polity.
- Rahman, A. F., Gamon, J. A., Sims, D. A., Schmidts, M., Feb. 2003. Optimum pixel size for hyperspectral studies of ecosystem function in southern California chaparral and grassland. *Remote Sensing of Environment* 84 (2), 192–207.
- Rahman, M. Z. A., Gorte, B., 2008. Individual tree detection based on densities of high points of high resolution airborne LiDAR. *GEOBIA*, 350–355.
- Rahman, M. Z. A., Gorte, B. G. H., 2009. Tree crown delineation from high resolution airborne lidar based on densities of high points. *Int Arch Photogramm Remote Sens Spat Inf Sci XXXVIII part 3, W8*.
- Rahman, M. Z. A., Gorte, B. G. H., Bucksch, A. K., 2009. A new method for individual tree measurement from airborne LiDAR. *Proc. SilviLaser*, 14–16.
- Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P., Jan. 2013. Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sensing* 5 (2), 491–520.
- Reich, P. B., Oct. 2011. Biogeochemistry: Taking stock of forest carbon. *Nature Climate Change* 1 (7), 346–347.
- Reitberger, J., Krzystek, P., Stilla, U., 2006. Analysis of full waveform lidar data for tree species classification. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 3), 228–233.
- Reitberger, J., Krzystek, P., Stilla, U., 2007. Combined tree segmentation and stem detection using full waveform lidar data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36, 332–337.

- Reitberger, J., Krzystek, P., Stilla, U., Mar. 2008a. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing* 29 (5), 1407–1431.
- Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., 2008b. 3D segmentation of full waveform LIDAR data for single tree detection using normalized cut. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 37, 77–83.
- Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., Nov. 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (6), 561–574.
- Richter, R., Reu, B., Wirth, C., Doktor, D., Vohland, M., Oct. 2016. The use of airborne hyperspectral data for tree species classification in a species-rich Central European forest area. *International Journal of Applied Earth Observation and Geoinformation* 52, 464–474.
- Riegl Laser Measurement Systems, Dec. 2017. LAS Extrabytes Implementation in RIEGL Software. Tech. rep., Riegl Laser Measurement Systems GmbH, Horn, Riedenburgstrasse 48, Austria.
- Roberts, D. A., Ustin, S. L., Ogunjemiyo, S., Greenberg, J., Dobrowski, S. Z., Chen, J., Hinckley, T. M., Aug. 2004. Spectral and Structural Measures of Northwest Forest Vegetation at Leaf to Landscape Scales. *Ecosystems* 7 (5), 545–562.
- Roth, K. L., Roberts, D. A., Dennison, P. E., Peterson, S. H., Alonso, M., Dec. 2015. The impact of spatial resolution on the classification of plant species and functional types within imaging spectrometer data. *Remote Sensing of Environment* 171, 45–57.
- Rouchdy, Y., Cohen, L. D., Oct. 2013. Geodesic voting for the automatic extraction of tree structures. Methods and applications. *Computer Vision and Image Understanding* 117 (10), 1453–1467.
- Rüger, N., Condit, R., May 2012. Testing metabolic theory with models of tree growth that include light competition. *Functional Ecology* 26 (3), 759–765.
- Russo, S. E., Wiser, S. K., Coomes, D. A., Jul. 2007. Growth-size scaling relationships of woody plant species differ from predictions of the Metabolic Ecology Model. *Ecology Letters* 10 (10), 889–901.
- Ryding, J., Williams, E., Smith, M. J., Eichhorn, M. P., Jan. 2015. Assessing Handheld Mobile Laser Scanners for Forest Surveys. *Remote Sensing* 7 (1), 1095–1111.
- Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., Morel, A., Jun. 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences* 108 (24), 9899–9904.
- Sarrazin, M. J. D., van Aardt, J. A. N., Asner, G. P., McGlinchy, J., Messinger, D. W., Wu, J., Jan. 2012. Fusing small-footprint waveform LiDAR and hyperspectral data for canopy-level species classification and herbaceous biomass modeling in savanna ecosystems. *Canadian Journal of Remote Sensing* 37 (6), 653–665.
- Sasaki, T., Imanishi, J., Ioki, K., Morimoto, Y., Kitada, K., Jul. 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.

- Sayn-Wittgenstein, L., 1978. Recognition of tree species on aerial photographs. Information Report FMR-X-118, Canadian Forestry Service, Department of the Environment, Ottawa, Ontario, Canada.
- Säynäjoki, R., Packalén, P., Maltamo, M., Vehmas, M., Eerikäinen, K., Aug. 2008. Detection of Aspens Using High Resolution Aerial Laser Scanning Data and Digital Aerial Images. *Sensors* 8 (8), 5037–5054.
- Schapire, R. E., Jun. 1990. The Strength of Weak Learnability. *Machine Learning* 5 (2), 197–227.
- Schaub, M., Dobbertin, M., Kräuchi, N., Dobbertin, M. K., Mar. 2011. Preface—long-term ecosystem research: Understanding the present to shape the future. *Environmental Monitoring and Assessment* 174 (1-4), 1–2.
- Schlerf, M., Atzberger, C., Feb. 2006a. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sensing of Environment* 100 (3), 281–294.
- Schlerf, M., Atzberger, C., Feb. 2006b. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sensing of Environment* 100 (3), 281–294.
- Schneider, F. D., Morsdorf, F., Schmid, B., Petchey, O. L., Hueni, A., Schimel, D. S., Schaepman, M. E., Nov. 2017. Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nature Communications* 8 (1), 1441.
- Schölkopf, B., Smola, A. J., Scholkopf, M. D. o. t. M. P. I. f. B. C. i. T. G. P. B., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press.
- Schumacher, J., Nord-Larsen, T., May 2014. Wall-to-wall tree type classification using airborne lidar data and CIR images. *International Journal of Remote Sensing* 35 (9), 3057–3073.
- Sexton, J. O., Bax, T., Siqueira, P., Swenson, J. J., Hensley, S., Feb. 2009. A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America. *Forest Ecology and Management* 257 (3), 1136–1147.
- Sheeren, D., Fauvel, M., Ladet, S., Jacquin, A., Bertoni, G., Gibon, A., 2011. Mapping ash tree colonization in an agricultural mountain landscape: Investigating the potential of hyperspectral imagery. In: Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International. pp. 3672–3675.
- Shen, X., Cao, L., Nov. 2017. Tree-Species Classification in Subtropical Forests Using Airborne Hyperspectral and LiDAR Data. *Remote Sensing* 9 (11), 1180.
- Shendryk, I., Broich, M., Tulbure, M. G., Alexandrov, S. V., Feb. 2016. Bottom-up delineation of individual trees from full-waveform airborne laser scans in a structurally complex eucalypt forest. *Remote Sensing of Environment* 173, 69–83.
- Shi, Y., Skidmore, A. K., Wang, T., Holzwarth, S., Heiden, U., Pinnel, N., Zhu, X., Heurich, M., Dec. 2018a. Tree species classification using plant functional traits from LiDAR and hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation* 73, 207–219.
- Shi, Y., Wang, T., Skidmore, A. K., Heurich, M., Mar. 2018b. Important LiDAR metrics for discriminating forest tree species in Central Europe. *ISPRS Journal of Photogrammetry and Remote Sensing* 137, 163–174.

- Shinozaki, K., Yoda, K., Hozumi, K., Kira, T., 1964. A quantitative analysis of plant form-the pipe model theory: I. Basic analyses. *Japanese Journal of Ecology* 14 (3), 97–105.
- Silleos, N. G., Alexandridis, T. K., Gitas, I. Z., Perakis, K., Dec. 2006. Vegetation Indices: Advances Made in Biomass Estimation and Vegetation Monitoring in the Last 30 Years. *Geocarto International* 21 (4), 21–28.
- Simonse, M., Aschoff, T., Spiecker, H., Thies, M., 2003. Automatic determination of forest inventory parameters using terrestrial laser scanning. In: Proceedings of the Scandlaser Scientific Workshop on Airborne Laser Scanning of Forests. Vol. 2003. pp. 252–258.
- Sims, D. A., Gamon, J. A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote sensing of environment* 81 (2-3), 337–354.
- Snyder, G. I., 2012. National Enhanced Elevation Assessment at a Glance. US Department of the Interior, US Geological Survey.
- Soille, P., 2013. Morphological Image Analysis: Principles and Applications. Springer Science & Business Media, Berlin, Germany.
- Sokolova, M., Lapalme, G., Jul. 2009. A systematic analysis of performance measures for classification tasks. *Information Processing & Management* 45 (4), 427–437.
- Solberg, S., Naesset, E., Bollandsas, O. M., 2006. Single tree segmentation using airborne laser scanner data in a structurally heterogeneous spruce forest. *Photogrammetric Engineering & Remote Sensing* 72 (12), 1369–1378.
- Somers, B., Asner, G. P., Sep. 2013. Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests. *Remote Sensing of Environment* 136, 14–27.
- Somers, B., Asner, G. P., Sep. 2014. Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: Wavelength adaptive spectral mixture analysis. *International Journal of Applied Earth Observation and Geoinformation* 31, 57–66.
- Somers, B., Delalieux, S., Verstraeten, W. W., van Aardt, J. A. N., Albrigo, G. L., Coppin, P., Oct. 2010. An automated waveband selection technique for optimized hyperspectral mixture analysis. *International Journal of Remote Sensing* 31 (20), 5549–5568.
- Sommer, C., Holzwarth, S., Heiden, U., Heurich, M., Müller, J., Mauser, W., 2016. Feature-based tree species classification using hyperspectral and Lidar data in the Bavarian Forest National Park. *EARSeL eProceedings* 14 (S2), 49–70.
- St-Onge, B., Vega, C., Fournier, R. A., Hu, Y., Jun. 2008. Mapping canopy height using a combination of digital stereo-photogrammetry and lidar. *International Journal of Remote Sensing* 29 (11), 3343–3364.
- Stagakis, S., Markos, N., Sykioti, O., Kyparissis, A., 2010. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sensing of Environment* 114 (5), 977–994.
- Stoker, J. M., Abdullah, Q. A., Nayegandhi, A., Winehouse, J., Sep. 2016. Evaluation of Single Photon and Geiger Mode Lidar for the 3D Elevation Program. *Remote Sensing* 8 (9), 767.

- Stoy, P. C., Williams, M., Spadavecchia, L., Bell, R. A., Prieto-Blanco, A., Evans, J. G., van Wijk, M. T., Jun. 2009. Using Information Theory to Determine Optimum Pixel Size and Shape for Ecological Studies: Aggregating Land Surface Characteristics in Arctic Ecosystems. *Ecosystems* 12 (4), 574–589.
- Strahler, A. N., 1952. Hypsometric (area-altitude) analysis of erosional topography. *Geological Society of America Bulletin* 63 (11), 1117–1142.
- Straub, C., Dees, M., Weinacker, H., Koch, B., Jul. 2009. Using Airborne Laser Scanner Data and CIR Orthophotos to Estimate the Stem Volume of Forest Stands. <https://www.ingentaconnect.com/content/schweiz/pfg/2009/00002009/00000003/art00010>.
- Strîmbu, V. F., Strîmbu, B. M., Jun. 2015. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing* 104, 30–43.
- Sugumaran, R., Voss, M., 2007. Object-Oriented Classification of LIDAR-Fused Hyperspectral Imagery for Tree Species Identification in an Urban Environment. In: *Urban Remote Sensing Joint Event*, 2007. pp. 1–6.
- Suratno, A., Seielstad, C., Queen, L., Nov. 2009. Tree species identification in mixed coniferous forest using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (6), 683–693.
- Swatantran, A., Dubayah, R., Roberts, D., Hofton, M., Blair, J. B., 2011. Mapping biomass and stress in the Sierra Nevada using lidar and hyperspectral data fusion. *Remote Sensing of Environment* 115 (11), 2917–2930.
- Swatantran, A., Tang, H., Barrett, T., DeCola, P., Dubayah, R., Jun. 2016. Rapid, High-Resolution Forest Structure and Terrain Mapping over Large Areas using Single Photon Lidar. *Scientific Reports* 6, 28277.
- Tang, J., Chen, Y., Kukko, A., Kaartinen, H., Jaakkola, A., Khoramshahi, E., Hakala, T., Hyppä, J., Holopainen, M., Hyppä, H., Dec. 2015. SLAM-Aided Stem Mapping for Forest Inventory with Small-Footprint Mobile LiDAR. *Forests* 6 (12), 4588–4606.
- Tang, L., Shao, G., Dec. 2015. Drone remote sensing for forestry research and practices. *Journal of Forestry Research* 26 (4), 791–797.
- Tang, S., Dong, P., Buckles, B. P., Feb. 2013. Three-dimensional surface reconstruction of tree canopy from lidar point clouds using a region-based level set method. *International Journal of Remote Sensing* 34 (4), 1373–1385.
- Tao, S., Wu, F., Guo, Q., Wang, Y., Li, W., Xue, B., Hu, X., Li, P., Tian, D., Li, C., Yao, H., Li, Y., Xu, G., Fang, J., Dec. 2015. Segmenting tree crowns from terrestrial and mobile LiDAR data by exploring ecological theories. *ISPRS Journal of Photogrammetry and Remote Sensing* 110, 66–76.
- Tero, A., Kobayashi, R., Nakagaki, T., Feb. 2007. A mathematical model for adaptive transport network in path finding by true slime mold. *Journal of Theoretical Biology* 244 (4), 553–564.
- Thies, M., Pfeifer, N., Winterhalder, D., Gorte, B. G. H., Dec. 2004. Three-dimensional reconstruction of stems for assessment of taper, sweep and lean based on laser scanning of standing trees. *Scandinavian Journal of Forest Research* 19 (6), 571–581.

- Thomas, V., Oliver, R. D., Lim, K., Woods, M., Dec. 2008. LiDAR and Weibull modeling of diameter and basal area. *The Forestry Chronicle* 84 (6), 866–875.
- Thomas, V., Treitz, P., McCaughey, J. H., Morrison, I., 2006. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: An examination of scanning density. *Canadian Journal of Forest Research* 36 (1), 34–47.
- Thompson, D. W., 1917. On Growth and Form. University press, Cambridge [Eng.], open Library ID: OL6604798M.
- Tiede, D., Hochleitner, G., Blaschke, T., 2005. A full GIS-based workflow for tree identification and tree crown delineation using laser scanning. In: ISPRS Workshop CMRT. Vol. 5. p. 2005.
- Tigges, J., Lakes, T., Hostert, P., Sep. 2013. Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. *Remote Sensing of Environment* 136, 66–75.
- Tittmann, P., Shafii, S., Hartsough, B., Hamann, B., 2011. Tree Detection and Delineation from LiDAR point clouds using RANSAC. In: Proceedings of Silvilaser 2011. University of Tasmania, Australia.
- Tomaštík, J., Saloň, Š., Tunák, D., Chudý, F., Kardoš, M., Sep. 2017. Tango in forests – An initial experience of the use of the new Google technology in connection with forest inventory tasks. *Computers and Electronics in Agriculture* 141, 109–117.
- Tooke, T. R., Coops, N. C., Goodwin, N. R., Voogt, J. A., Feb. 2009. Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. *Remote Sensing of Environment* 113 (2), 398–407.
- Torabzadeh, H., 2016. Forest characterization by fusion of imaging spectroscopy and airborne laser scanning. Dissertation, University of Zürich, Zürich, Switzerland.
- Torabzadeh, H., Morsdorf, F., Schaeppman, M. E., Nov. 2014. Fusion of imaging spectroscopy and airborne laser scanning data for characterization of forest ecosystems – A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 97, 25–35.
- Torresan, C., Berton, A., Carotenuto, F., Gennaro, S. F. D., Gioli, B., Matese, A., Miglietta, F., Vagnoli, C., Zaldei, A., Wallace, L., May 2017. Forestry applications of UAVs in Europe: A review. *International Journal of Remote Sensing* 38 (8-10), 2427–2447.
- Trier, O. D., Salberg, A.-B., Kermit, M., Rudjord, O., Gobakken, T., Næsset, E., Aarsten, D., Jan. 2018. Tree species classification in Norway from airborne hyperspectral and airborne laser scanning data. *European Journal of Remote Sensing* 51 (1), 336–351.
- Tsai, F., Philpot, W., Oct. 1998. Derivative Analysis of Hyperspectral Data. *Remote Sensing of Environment* 66 (1), 41–51.
- Tuominen, S., Näsi, R., Honkavaara, E., Balazs, A., Hakala, T., Viljanen, N., Pölönen, I., Saari, H., Ojanen, H., May 2018. Assessment of Classifiers and Remote Sensing Features of Hyperspectral Imagery and Stereo-Photogrammetric Point Clouds for Recognition of Tree Species in a Forest Area of High Species Diversity. *Remote Sensing* 10 (5), 714.
- Tuominen, S., Näsi, R., Honkavaara, E., Balazs, A., Hakala, T., Viljanen, N., Pölönen, I., Saari, H., Reinikainen, J., 2017. Tree species recognition in species rich area using UAV-borne hyperspectral imagery and stereo-photogrammetric point cloud. *International archives of the photogrammetry, remote sensing and spatial information sciences; Volume XLII-3/W3*.

- Underwood, E., Ustin, S., DiPietro, D., 2003. Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment* 86 (2), 150–161.
- Vain, A., Kaasalainen, S., 2011. Correcting Airborne Laser Scanning Intensity Data. INTECH Open Access Publisher.
- van Aardt, J. A. N., Wynne, R. H., Jan. 2007. Examining pine spectral separability using hyperspectral data from an airborne sensor: An extension of field-based results. *International Journal of Remote Sensing* 28 (2), 431–436.
- Van Coillie, F., Liao, W., Devriendt, F., Gautama, S., De Wulf, R., Vandekerckhove, K., 2014. Effect of hyperspectral image denoising with PCA and total variation on tree species mapping using Apex Data. In: *SOUTH-EASTERN EUROPEAN JOURNAL OF EARTH OBSERVATION AND GEOMATICS*. Vol. 3. pp. 281–286.
- Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag New York, Inc.
- Vastaranta, M., Wulder, M. A., White, J. C., Pekkarinen, A., Tuominen, S., Ginzler, C., Kankare, V., Holopainen, M., Hyppä, J., Hyppä, H., Nov. 2013. Airborne laser scanning and digital stereo imagery measures of forest structure: Comparative results and implications to forest mapping and inventory update. *Canadian Journal of Remote Sensing* 39 (5), 382–395.
- Vaughn, N. R., Moskal, L. M., Turnblom, E. C., 2011. Fourier transformation of waveform Lidar for species recognition. *Remote Sensing Letters* 2 (4), 347–356.
- Vaughn, N. R., Moskal, L. M., Turnblom, E. C., Feb. 2012. Tree Species Detection Accuracies Using Discrete Point Lidar and Airborne Waveform Lidar. *Remote Sensing* 4 (2), 377–403.
- Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K. M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., Maltamo, M., Jan. 2012. Comparative testing of single-tree detection algorithms under different types of forest. *Forestry: An International Journal of Forest Research* 85 (1), 27–40.
- Vauhkonen, J., Hakala, T., Suomalainen, J., Kaasalainen, S., Nevalainen, O., Vastaranta, M., Holopainen, M., Hyppä, J., Sep. 2013. Classification of Spruce and Pine Trees Using Active Hyperspectral LiDAR. *IEEE Geoscience and Remote Sensing Letters* 10 (5), 1138–1141.
- Vauhkonen, J., Korpela, I., Maltamo, M., Tokola, T., Jun. 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.
- Vauhkonen, J., Tokola, T., Packalén, P., Maltamo, M., 2009. Identification of Scandinavian Commercial Species of Individual Trees from Airborne Laser Scanning Data Using Alpha Shape Metrics. *Forest Science* 55 (1), 37–47.
- Vega, C., Hamrouni, A., El Mokhtari, S., Morel, J., Bock, J., Renaud, J. P., Bouvier, M., Durrieu, S., Dec. 2014. PTrees: A point-based approach to forest tree extraction from lidar data. *International Journal of Applied Earth Observation and Geoinformation* 33, 98–108.
- Veganzones, M. A., Grana, M., 2008. Endmember extraction methods: A short review. In: *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. Springer, pp. 400–407.

- Veltkamp, R. C., 2001. Shape matching: Similarity measures and algorithms. In: Shape Modeling and Applications, SMI 2001 International Conference On. IEEE, pp. 188–197.
- Verlič, A., DJurić, N., Kokalj, Ž., Marsetič, A., Simončič, P., Oštir, K., 2014. Tree species classification using WorldView-2 satellite images and laser scanning data in a natural urban forest. *Šumarski list* 138 (9-10), 477–488.
- Vonderach, C., Voegtle, T., Adler, P., 2012. Voxel-based approach for estimating urban tree volume from terrestrial laser scanning data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 39, 451–456.
- Voss, M., Sugumaran, R., May 2008. Seasonal Effect on Tree Species Classification in an Urban Environment Using Hyperspectral Data, LiDAR, and an Object- Oriented Approach. *Sensors* 8 (5), 3020–3036.
- Vyas, D., Krishnayya, N., Manjunath, K., Ray, S., Panigrahy, S., Apr. 2011. Evaluation of classifiers for processing Hyperion (EO-1) data of tropical vegetation. *International Journal of Applied Earth Observation and Geoinformation* 13 (2), 228–235.
- Wang, D., Brunner, J., Ma, Z., Lu, H., Hollaus, M., Pang, Y., Pfeifer, N., May 2018. Separating Tree Photosynthetic and Non-Photosynthetic Components from Point Cloud Data Using Dynamic Segment Merging. *Forests* 9 (5), 252.
- Wang, D., Hollaus, M., Puttonen, E., Pfeifer, N., Nov. 2016a. Automatic and Self-Adaptive Stem Reconstruction in Landslide-Affected Forests. *Remote Sensing* 8 (12), 974.
- Wang, L., Birt, A. G., Lafon, C. W., Cairns, D. M., Coulson, R. N., Tchakerian, M. D., Xi, W., Popescu, S. C., Guldin, J. M., 2011. Computer-based synthetic data to assess the tree delineation algorithm from airborne LiDAR survey. *Geoinformatica* 17 (1), 35–61.
- Wang, Y., Hyppä, J., Liang, X., Kaartinen, H., Yu, X., Lindberg, E., Holmgren, J., Qin, Y., Mallet, C., Ferraz, A., Torabzadeh, H., Morsdorf, F., Zhu, L., Liu, J., Alho, P., Sep. 2016b. International Benchmarking of the Individual Tree Detection Methods for Modeling 3-D Canopy Structure for Silviculture and Forest Ecology Using Airborne Laser Scanning. *IEEE Transactions on Geoscience and Remote Sensing* 54 (9), 5011–5027.
- Wang, Y., Weinacker, H., Koch, B., Jun. 2008a. A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis And 3D Single Tree Modelling in Forest. *Sensors* 8 (6), 3938–3951.
- Wang, Y., Weinacker, H., Koch, B., Jun. 2008b. A Lidar Point Cloud Based Procedure for Vertical Canopy Structure Analysis And 3D Single Tree Modelling in Forest. *Sensors* 8 (6), 3938–3951.
- Waser, L., Küchler, M., Jütte, K., Stampfer, T., Waser, L. T., Küchler, M., Jütte, K., Stampfer, T., May 2014. Evaluating the Potential of WorldView-2 Data to Classify Tree Species and Different Levels of Ash Mortality. *Remote Sensing* 6 (5), 4515–4545.
- Waser, L. T., Boesch, R., Wang, Z., Ginzler, C., 2017a. Towards Automated Forest Mapping. In: *Mapping Forest Landscape Patterns*. Springer, New York, NY, pp. 263–304.
- Waser, L. T., Eisenbeiss, H., Kuechler, M., Baltsavias, B., 2008. Potential and limits of airborne remote sensing data for extraction of fractional canopy cover and forest stands and detection of tree species. *The international archives of the photogrammetry, remote sensing and spatial information sciences* 37, 1405–1412.

- Waser, L. T., Ginzler, C., Kuechler, M., Baltsavias, E., Hurni, L., Jan. 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.
- Waser, L. T., Ginzler, C., Rehush, N., Jul. 2017b. Wall-to-Wall Tree Type Mapping from Country-wide Airborne Remote Sensing Surveys. *Remote Sensing* 9 (8), 766.
- Waser, L. T., Klonus, S., Ehlers, M., Kuchler, M., Jung, A., 2010. Potential of Digital Sensors for Land Cover and Tree Species Classifications A Case Study in the Framework of the DGPF-Project. *Photogrammetrie-Fernerkundung-Geoinformation* 2010 (2), 141–156.
- Wästlund, A., Holmgren, J., Lindberg, E., Olsson, H., Sep. 2018. Forest Variable Estimation Using a High Altitude Single Photon Lidar System. *Remote Sensing* 10 (9), 1422.
- Welsh, D. J. A., Powell, M. B., Jan. 1967. An upper bound for the chromatic number of a graph and its application to timetabling problems. *The Computer Journal* 10 (1), 85–86.
- West, G. B., Brown, J. H., Enquist, B. J., Apr. 1997. A General Model for the Origin of Allometric Scaling Laws in Biology. *Science* 276 (5309), 122–126.
- West, G. B., Brown, J. H., Enquist, B. J., Apr. 1999a. The Fourth Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms. *Science* 284 (5420), 1677–1679.
- West, G. B., Brown, J. H., Enquist, B. J., Aug. 1999b. A general model for the structure and allometry of plant vascular systems. *Nature* 400 (6745), 664–667.
- West, P., Mar. 2013. *Tree and Forest Measurement*. Springer Science & Business Media.
- White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., Tompalski, P., Sep. 2016. Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Canadian Journal of Remote Sensing* 42 (5), 619–641.
- White, J. C., Wulder, M. A., Varhola, A., Vastaranta, M., Coops, N. C., Cook, B. D., Pitt, D. G., Woods, M., 2013a. A Best Practices Guide for Generating Forest Inventory Attributes from Airborne Laser Scanning Data Using an Area-Based Approach.
- White, J. C., Wulder, M. A., Vastaranta, M., Coops, N. C., Pitt, D., Woods, M., Jun. 2013b. The Utility of Image-Based Point Clouds for Forest Inventory: A Comparison with Airborne Laser Scanning. *Forests* 4 (3), 518–536.
- Whitney, G. G., 1976. The bifurcation ratio as an indicator of adaptive strategy in woody plant species. *Bulletin of the Torrey Botanical Club*, 67–72.
- Wieser, M., Mandlburger, G., Hollaus, M., Otepka, J., Glira, P., Pfeifer, N., Nov. 2017. A Case Study of UAS Borne Laser Scanning for Measurement of Tree Stem Diameter. *Remote Sensing* 9 (11), 1154.
- Williams, D. L., Feb. 1991. A comparison of spectral reflectance properties at the needle, branch, and canopy level for selected Conifer species. *Remote Sensing of Environment* 35 (2), 79–93.
- Windrim, L., Bryson, M., Oct. 2018. Forest Tree Detection and Segmentation using High Resolution Airborne LiDAR. arXiv:1810.12536 [cs].

- Wu, B., Yu, B., Wu, Q., Huang, Y., Chen, Z., Wu, J., Oct. 2016. Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests. *International Journal of Applied Earth Observation and Geoinformation* 52, 82–94.
- Wu, J., Yao, W., Choi, S., Park, T., Myneni, R. B., 2015. A Comparative Study of Predicting DBH and Stem Volume of Individual Trees in a Temperate Forest Using Airborne Waveform LiDAR. *IEEE Geosci. Remote Sensing Lett.* 12 (11), 2267–2271.
- Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W., Gobakken, T., Jun. 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment* 121, 196–209.
- Xia, S., Wang, C., Pan, F., Xi, X., Zeng, H., Liu, H., Oct. 2015. Detecting Stems in Dense and Homogeneous Forest Using Single-Scan TLS. *Forests* 6 (11), 3923–3945.
- Xiao, W., Xu, S., Elberink, S. O., Vosselman, G., Aug. 2016. Individual Tree Crown Modeling and Change Detection From Airborne Lidar Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (8), 3467–3477.
- Xu, C., Morgenroth, J., Manley, B., Sep. 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.
- Xu, H., Gossett, N., Chen, B., 2007. Knowledge and heuristic-based modeling of laser-scanned trees. *ACM Transactions on Graphics (TOG)* 26 (4), 19.
- Xu, L., Mould, D., Dec. 2012. A procedural method for irregular tree models. *Computers & Graphics* 36 (8), 1036–1047.
- Xu, S., Ye, N., Xu, S., Zhu, F., Jun. 2018. A supervoxel approach to the segmentation of individual trees from LiDAR point clouds. *Remote Sensing Letters* 9 (6), 515–523.
- Yan, D.-M., Wintz, J., Mourrain, B., Wang, W., Boudon, F., Godin, C., 2009. Efficient and robust reconstruction of botanical branching structure from laser scanned points. In: *Computer-Aided Design and Computer Graphics, 2009. CAD/Graphics' 09. 11th IEEE International Conference On*. IEEE, pp. 572–575.
- Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Lovell, J., Culvenor, D., Newnham, G., Ni-Meister, W., Schaaf, C., Woodcock, C., Wang, J., Li, X., Strahler, A., Nov. 2011. Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. *Remote Sensing of Environment* 115 (11), 2965–2974.
- Yao, W., Krzystek, P., Heurich, M., Aug. 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.
- Yao, W., Krzystek, P., Heurich, M., 2013. Enhanced detection of 3D individual trees in forested areas using airborne full-waveform LiDAR data by combining normalized cuts with spatial density clustering. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* 1 (2), 349–354.
- Yin, D., Wang, L., Oct. 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.

- Yu, X., Hyppä, J., Holopainen, M., Vastaranta, M., Jun. 2010. Comparison of Area-Based and Individual Tree-Based Methods for Predicting Plot-Level Forest Attributes. *Remote Sensing* 2 (6), 1481–1495.
- Yu, X., Hyppä, J., Litkey, P., Kaartinen, H., Vastaranta, M., Holopainen, M., Jan. 2017. Single-Sensor Solution to Tree Species Classification Using Multispectral Airborne Laser Scanning. *Remote Sensing* 9 (2), 108.
- Yu, X., Hyppä, J., Vastaranta, M., Holopainen, M., Viitala, R., Jan. 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (1), 28–37.
- Yu, X., Litkey, P., Hyppä, J., Holopainen, M., Vastaranta, M., May 2014. Assessment of Low Density Full-Waveform Airborne Laser Scanning for Individual Tree Detection and Tree Species Classification. *Forests* 5 (5), 1011–1031.
- Zarco-Tejada, P. J., Berjón, A., López-Lozano, R., Miller, J. R., Martín, P., Cachorro, V., González, M. R., De Frutos, A., 2005. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment* 99 (3), 271–287.
- Zarco-Tejada, P. J., Pushnik, J. C., Dobrowski, S., Ustin, S. L., 2003. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sensing of Environment* 84 (2), 283–294.
- Zeide, B., Dec. 1991. Fractal geometry in forestry applications. *Forest Ecology and Management* 46 (3–4), 179–188.
- Zhang, C., Qiu, F., 2012. Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery. *Photogrammetric Engineering & Remote Sensing* 78, 1079–1087.
- Zhang, C., Xie, Z., Selch, D., Oct. 2013. Fusing lidar and digital aerial photography for object-based forest mapping in the Florida Everglades. *GIScience & Remote Sensing* 50 (5), 562–573.
- Zhang, J., Hu, J., Lian, J., Fan, Z., Ouyang, X., Ye, W., Jun. 2016a. Seeing the forest from drones: Testing the potential of lightweight drones as a tool for long-term forest monitoring. *Biological Conservation* 198, 60–69.
- Zhang, J., Rivard, B., Sánchez-Azofeifa, A., Castro-Esau, K., Nov. 2006. Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. *Remote Sensing of Environment* 105 (2), 129–141.
- Zhang, J., Sohn, G., Brédif, M., Dec. 2014. A hybrid framework for single tree detection from airborne laser scanning data: A case study in temperate mature coniferous forests in Ontario, Canada. *ISPRS Journal of Photogrammetry and Remote Sensing* 98, 44–57.
- Zhang, Z., Kazakova, A., Moskal, L. M., Styers, D. M., Jun. 2016b. Object-Based Tree Species Classification in Urban Ecosystems Using LiDAR and Hyperspectral Data. *Forests* 7 (6), 122.
- Zhang, Z., Liu, X., Jul. 2013. Support vector machines for tree species identification using LiDAR-derived structure and intensity variables. *Geocarto International* 28 (4), 364–378.
- Zhao, Y., Hao, Y., Zhen, Z., Quan, Y., Oct. 2017a. A Region-Based Hierarchical Cross-Section Analysis for Individual Tree Crown Delineation Using ALS Data. *Remote Sensing* 9 (10), 1084.

- Zhao, Y., Hao, Y., Zhen, Z., Quan, Y., Oct. 2017b. A Region-Based Hierarchical Cross-Section Analysis for Individual Tree Crown Delineation Using ALS Data. *Remote Sensing* 9 (10), 1084.
- Zhen, Z., Quackenbush, L. J., Zhang, L., Apr. 2016. Trends in Automatic Individual Tree Crown Detection and Delineation—Evolution of LiDAR Data. *Remote Sensing* 8 (4), 333.
- Zheng, G., Moskal, L. M., Apr. 2009. Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. *Sensors* 9 (4), 2719–2745.
- Zheng, T., Chen, J. M., 2017. Photochemical reflectance ratio for tracking light use efficiency for sunlit leaves in two forest types. *ISPRS Journal of Photogrammetry and Remote Sensing* 123, 47–61.
- Zhou, T., Popescu, S., Lawing, A., Eriksson, M., Strimbu, B., Bürkner, P., Zhou, T., Popescu, S. C., Lawing, A. M., Eriksson, M., Strimbu, B. M., Bürkner, P. C., Dec. 2017. Bayesian and Classical Machine Learning Methods: A Comparison for Tree Species Classification with LiDAR Waveform Signatures. *Remote Sensing* 10 (1), 39.
- Zou, X., Cheng, M., Wang, C., Xia, Y., Li, J., Dec. 2017. Tree Classification in Complex Forest Point Clouds Based on Deep Learning. *IEEE Geoscience and Remote Sensing Letters* 14 (12), 2360–2364.



An aerial photograph showing a coastal area. On the left, there's a paved path or road with some yellow markings. The terrain is covered in dense, tall grasses and low-lying plants. In the background, there are more trees and possibly a body of water under a clear sky.

## Appendix

## A Site geography, topography and forest characteristics

Table A.1: Site geography, topography and forest characteristics

ID	Name	Lat., Lon.	Alt.	Slope	Aspect	Structure	Management
1	Versoix	46.29685 N, 6.13021 E	445 m	0-5°	-	High forest (mixed)	N/A
2	Sauvabelin	46.53656 N, 6.63655 E	640 m	15-20°	WSW	High forest (deciduous)	Safety maintenance
3	Benenté	46.58351 N, 6.65808 E	808 m	0-5°	-	High forest (deciduous)	Preserve
4	Jorat south	46.57971 N, 6.65589 E	790 m	5-15°	W	High Forest (coniferous)	N/A
5	Gésiaux	46.55536 N, 6.65342 E	760 m	0-5°	-	High forest (coniferous)	Preserve
6	La Brévine	46.96035 N, 6.50852 E	1153 m	5-15°	ESE	High forest (mixed)	Selection cutting
7	Couvet	46.92006 N, 6.64495 E	846 m	10-15°	NNW	High Forest/Uneven-aged (mixed)	Selection cutting
8	Cortaillod	46.94523 N, 6.81218 E	536 m	5-10°	E	High forest (mixed)	Selection cutting
9	Boudry D20	46.95635 N, 6.81163 E	598 m	5-10°	E	Uneven-aged stand (mixed)	Selection cutting
10	Boudry D19	46.95906 N, 6.81268 E	554 m	10-20°	WNW	Uneven-aged stand (mixed)	Selection cutting
11	Boudry D1	46.96568 N, 6.82346 E	580 m	0-5°	-	High forest (mixed)	Group-selection cutting
12	Chambrelien	46.96679 N, 6.81613 E	650 m	5-10°	SE	High forest (mixed)	Group-selection cutting
13	Rochefort	46.97115 N, 6.82152 E	598 m	30°	NE / SW	High forest (mixed)	Selection cutting
14	Bevaix	46.93305 N, 6.80317 E	610 m	10-15°	SE	High Forest (mixed)	Selection cutting
15	Grosszinggibrunn	47.51003 N, 7.66426 E	410 m	~10°	NNW	High forest (deciduous)	Selection cutting
16	Ottmarsingen	47.39887 N, 8.22677 E	500 m	15-20°	SSE	High Forest (mixed)	Selection cutting
17	Sihlwald	47.25096 N, 8.55908 E	630 m	5-15°	NE	High Forest/Uneven-aged (mixed)	Preserve
18	Oberaegeri	47.15412 N, 8.695 E	940 m	10-15°	S	High forest (coniferous)	Selection cutting
19	Dischma	46.77061 N, 9.8769 E	1832 m	35-40°	NE	High Forest (coniferous)	Selection cutting

## B Field survey metadata

Table B.1: Field survey metadata.

**Abbreviations:** **AbAl:** *Abies alba*, **Ac:** *Acer sp.*, **AcPs:** *Acer pseudoplatanus*, **CaBe:** *Carpinus betulus*, **FaSy:** *Fagus sylvatica*, **FrEx:** *Fraxinus excelsior*, **LaDe:** *Larix decidua*, **PiAb:** *Picea abies*, **PrAv:** *Prunus avium*, **PiSy:** *Pinus sylvestris*, **PsMe:** *Pseudotsuga menziesii*, **Qu:** *Quercus sp.*, **QuPe:** *Quercus petraea*, **QuRo:** *Quercus robur*, **SoAu:** *Sorbus aucuparia*, **TiCo:** *Tilia cordata*.

**CCS:** Centre de Compétence en Sylviculture (<http://www.waldbau-sylviculture.ch>).    **WSL:** Eidg. Forschungsanstalt für Wald, Schnee und Landschaft

ID	Name	Date	Density [ha <sup>-1</sup> ]	Basal area [m <sup>2</sup> /ha]	Main species	Source
1	Versoix	2012	337	31.8	QuPe (73.3 %), PiAb (15.7 %), CaBe (8 %)	CCS
2	Sauvabelin	Feb. 2015	168	40	FaSy (57.3 %), QuRo (29.9 %), PrAv (5.7 %)	CCS
3	Benenté	Mar. 2015	359	45.2	AbAl (55.7 %), FaSy (39.4 %), PiAb (4.4 %)	WSL
4	Jorat south	-	-	-	-	-
5	Gésiaux	-	-	-	-	-
6	La Brévine	Oct. 2015	382	28.4	AbAl (37.9 %), PiAb (33 %), FaSy (28.4 %)	CCS
7	Couvet	Dec. 2015	263	33	AbAl (74.1 %), PiAb (10.3 %), FaSy (7.6 %), Ac. (4.2 %))	CCS
8	Cortaillod	May 2016	270	23.3	FaSy (14.9 %), LaDe (% 14.6), PiAb (12.2 %), Ac. (12 %), FrEx (10.9 %), Qu. (9.3 %), AbAl (7.7 %), TiCo (6.6 %), PiSy (5.9 %)	CCS
9	Boudry D20	Sep. 2013	449	44.3	FaSy (35.1 %), PsMe (21.6 %), AcPs (13.7 %), AbAl (4.7 %), LaDe (4.3 %)	CCS
10	Boudry D19	Mar. 2012	348	40.8	AbAl (25.8 %), FaSy (19.6 %), FrEx (15.7 %), PiAb (11.5 %), PiSy (9.4 %), QuPe (5.4 %), AcPs (4.5 %)	CCS
11	Boudry D1	Jan. 2013	321	29.9	AbAl (52.2 %), QuPe (29.5 %), FaSy (15.9 %)	CCS
12	Chambrelien	-	-	-	-	-
13	Rochefort	-	-	-	-	-
14	Bevaix	-	-	-	-	-
15	Grosszinggibrunn	Oct. 2014	271	28.2	FaSy (46.4 %), CaBe (22 %), Ac (13 %), Qu (6.4 %), FrEx (6.1 %)	WSL
16	Ottmarsingen	Mar. 2015	184	45.6	FaSy (88.2 %), PiAb (9.6 %)	WSL
17	Sihlwald	Nov. 2015	301	36.9	FaSy (79.9 %), AbAl (8.6 %), PiAb (6.6 %)	WSL
18	Oberaegeri	Nov. 2009	684	56.7	PiAb (40.9 %), AbAl (40.7 %), FaSy (16.3 %)	WSL
19	Dischma	Sep. 2016	387	74.5	PiAb (77.2 %), LaDe (14 %), SoAu (8.6 %)	WSL

## C Study site pictures



(a) Versoix (Genève), Photo from April 7, 2018.



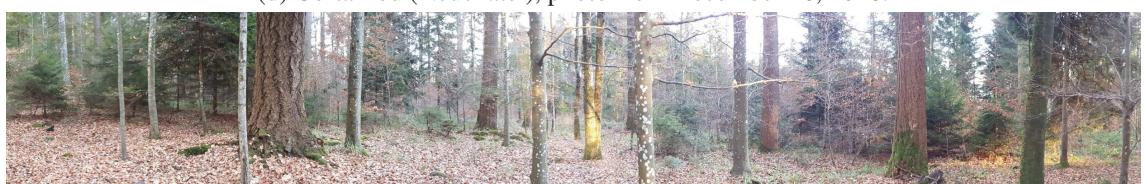
(b) Sauvabelin, (Vaud) photo from December 10, 2016.



(c) Benenté (Vaud), photo from December 4, 2016.



(d) Cortaillod (Neuchâtel), photo from December 26, 2016.



(e) Boudry division 20 (Neuchâtel), photo from December 26, 2016.



(f) Boudry division 19 (Neuchâtel), photo from December 26, 2016.



(g) Boudry division 1 (Neuchâtel), photo from December 26, 2016.

Figure C.1: Pictures of survey sites.

## D AHI survey metadata

Table D.1: APEX airborne hyperspectral imaging metadata. **Abbreviations:** RF: Random Forest, SVM: Support Vector Machine.

<b>Sensor</b>	Type	Dispersive push broom
	Bands	285
	Range	400 - 2500 nm
	Spectral resolution	< 10 nm (SWIR), 5 nm (VIS/NIR)
<b>Acquisition</b>	Dates	18.07.2014
	Flight time (local)	12h38 - 13h01
	Flying height (AGL)	4880 m
	Flight lines	Two lines per sector (Boudry and Lausanne)
	Aircraft	Dornier Do 228
<b>Processing</b>	Geometric correction	Orthorectification based on direct georeferencing with LiDAR DSM (Boudry, 2010 / Lausanne 2012)
	Atmospheric correction	Default settings ENVI software which uses the MODTRAN4 radiation transfer model
	Spectral smoothing	Default settings in Colibri software
<b>Delivery</b>	Provider	VITO (Belgium)
	Date	23.03.2015
	Format	IMG (BSQ), ENVI header files (HDR)
	Spatial resolution*	2.6 - 3.0 m
	Geolocation accuracy	< 2 m

\* The contract requested a spatial resolution of 1.75 m. However, due to poor flight planning by Vito, the optimal flying altitude was not granted by air traffic control and the aircraft had to fly higher than anticipated resulting in a coarser resolution product.

## E ALS survey metadata

Table E.1: ALS survey metadata. **AGIS**: Aargauische Geografische Informationssystem, **OIT**: Office d'Informations sur le Territoire (Vaud), **SITN**: Système d'Information du Territoire Neuchâtelois, **ARE**: Amt für Raumentwicklung (Zürich), **AGBL**: Amt für Geoinformation (Basel-Landschaft), **GKZ**: Geoportal Kanton Zug, **WSL**: Eidg. Forschungsanstalt für Wald, Schnee und Landschaft, **SITG**: Système d'Information du Territoire à Genève .

ID	Name	Sensor	Flight date(s) (phenology)	Point density [m <sup>-2</sup> ]	Intensity norm.	RGB color	Segments (Total)	Segments (Incl. diameter)	Segments (Incl. species)	Source
1	Versoix	Riegl LMS-Q1560	February 19, 2017 (leaf off)	46	Yes	No	375	261	262	SITG
2	Sauvabelin	Optech ALTM Gemini	March 10 / 11, 2012 (leaf off)	39	Yes	No	133	103	103	OIT
3	Benenté	Optech ALTM Gemini	March 10 / 11, 2012 (leaf off)	33	Yes	No	700	421	453	OIT
4	Jorat south	Optech ALTM Gemini	March 10 / 11, 2012 (leaf off)	33	Yes	No	357	0	350	OIT
5	Gésiaux	Optech ALTM Gemini	March 10 / 11, 2012 (leaf off)	33	Yes	No	126	0	95	OIT
6	La Brévine	Riegl LMS-Q1560	May 5, 2016 (leaf off)	31	Yes	Yes	896	371	370	SITN
7	Couvet	Riegl LMS-Q1560	May 4-5, 2016 (leaf off)	31	Yes	Yes	238	159	159	SITN
8	Cortaillod	Riegl LMS-Q1560	March 26 / May 4, 2016 (leaf off / on)	33	Yes	Yes	158	78	89	SITN
9	Boudry D20	Riegl LMS-Q1560	March 26 / May 4, 2016 (leaf off / on)	70	Yes	Yes	319	124	229	SITN

Table E.2: ALS survey metadata (continued from previous page). **AGIS**: Aargauische Geografische Informationssystem, **OIT**: Office d'Informations sur le Territoire (Vaud), **SITN**: Système d'Information du Territoire Neuchâtelois, **ARE**: Amt für Raumentwicklung (Zürich), **AGBL**: Amt für Geoinformation (Basel-Landschaft), **GKZ**: Geoportal Kanton Zug, **WSL**: Eidg. Forschungsanstalt für Wald, Schnee und Landschaft, **SITG**: Système d'Information du Territoire à Genève .

ID	Name	Sensor	Flight date(s) (phenology)	Point density [ $\text{m}^{-2}$ ]	Intensity norm.	RGB color	Segments (Total)	Segments (Incl. diameter)	Segments (Incl. species)	Source
10	Boudry D19	Riegl LMS-Q1560	March 26 / May 4, 2016 (leaf off / on)	71	Yes	Yes	321	117	168	SITN
11	Boudry D1	Riegl LMS-Q1560	March 26, 2016 (leaf off)	40	Yes	Yes	509	344	381	SITN
12	Chambrelien	Riegl LMS-Q1560	March 26 / May 4, 2016 (leaf off / on)	77	Yes	Yes	219	0	145	SITN
13	Rochefort	Riegl LMS-Q1560	March 26 / May 4, 2016 (leaf off / on)	35	Yes	Yes	372	0	71	SITN
14	Bevaix	Riegl LMS-Q1560	March 26 / May 4, 2016 (leaf off / on)	70	Yes	Yes	176	0	176	SITN
15	Grosszinggibrunn	Optech ALTM Gemini	June 12, 2012 (leaf on)	11	Yes	No	328	182	181	AGBL
16	Ottmarsingen	Riegl LMS-Q680i	March 27 / 28, 2014 (leaf off)	29	N/A	No	123	92	92	AGIS
17	Sihlwald	Trimble AX60	March 10, 2014 (leaf off)	41	N/A	No	1317	1166	1161	ARE
18	Oberaegeri	Optech ALTM Gemini	November 13, 2012 (leaf off)	20	Yes	No	302	204	204	GKZ
19	Dischma	Riegl LMS-Q780	August 5 / 6, 2015 (leaf on)	28	N/A	No	312	207	207	WSL

## F ALS point classification

Table F.1: Standard LAS 1.4 point classification for point data record formats 6-10. Adapted from the American Society for Photogrammetry and Remote Sensing (2013).

Class number	Description	Note
0	Created, never classified	Removed
1	Unclassified	Removed
2	Ground	Preserved
3	Low vegetation	$0 < \text{height} \leq 0.5 \text{ m}$
4	Medium vegetation	$0.5 < \text{height} \leq 3 \text{ m}$
5	High vegetation	$\text{height} > 3\text{m}$
6	Building	Removed
7	Low point (noise)	Removed
8	Reserved	Removed
9	Water	Removed
10	Rail	Removed
11	Road surface	Removed
12	Reserved for ASPRS Definition	Removed
13	Wire – Guard (Shield)	Removed
14	Wire – Conductor (Phase)	Removed
15	Transmission Tower	Removed
16	Wire-structure Connector (e.g. Insulator)	Removed
17	Bridge Deck	Removed
18	High Noise	Removed
19-63	Reserved for ASPRS Definition	Removed
64-255	User definable	Removed

## G Custom forestry EVLR specification

Table G.1: Custom forestry EVLR specification for ASPRS LAS files. Only items with an asterisk (\*) are filled in the benchmark dataset.

Item	Description	Record ID	Format (bits)	Unit	Example
LUID*	Locally Unique Identifier	5000	single (32)	-	42
UUID*	Universally Unique Identifier	5001	quadruple (128)	-	44c190e694384c7b ba7ca0e59396c58c
Stand ID	Stand Identifier	5002	uint16 (16)	-	10
Reserved	Reserved for identifiers	5003-5009	-	-	-
X*	X coordinate of stem proxy	5010	double (64)	m	2552299.30
Y*	Y coordinate of stem proxy	5011	double (64)	m	1200754.62
Z*	Altitude of root	5012	double (64)	m	595
Location Proxy*	Tree location proxy: root (1), centroid (2) or apex (3)	5013	uint8 (8)	-	1
Diameter*	Diameter at breast height (1.3 m)	5014	uint16 (16)	cm	11
Total height*	Distance from root to apex point	5015	uint16 (16)	cm	1613
Bole height	Distance from root to crown base	5016	uint16 (16)	cm	-
Bole angle	Bole angle rounded to nearest degree	5017	uint8 (8)	degree	-
Bole straightness	Linearity indicator (sinuose / straight length))	5018	uint8 (8)	-	-
Total projected area*	Area of single region (no holes) concave hull containing all xy points	5019	single (32)	$m^2$	-
Total Volume*	Volume of single region (no holes) concave hull containing all xyz points	5020	single (32)	$m^3$	-
Crown projected area	Area of single region (no holes) concave hull containing crown points	5021	single (32)	$m^2$	-
Crown volume	Volume of single region (no holes) concave hull containing crown points	5022	single (32)	$m^3$	-

Table G.2: Custom forestry EVLR specification for ASPRS LAS files (continued). Only items with an asterisk (\*) are filled in the benchmark dataset.

Item	Description	Record ID	Format (bits)	Unit	Example
Reserved	Reserved for geometric attributes	5023-5040	-	-	-
IPNI*	International Plant Name Index ID	5041	char(12) (96)	-	609009-1
Social rank	Social rank according to Nyland (2016): dominant (1), codominant (2), intermediate (3), overtopped (4)	5042	uint8(8)	-	3
Ivy flag*	Ivy presence flag	5043	uint8 (8)	-	1
Dead flag*	Dead tree flag	5044	uint8 (8)	-	0
Reserved	Reserved for biological and ecological attributes	5045-5060	-	-	-
Ambiguity flag*	Flag indicating if tree delineation is uncertain/unreliable	5061	uint8 (8)	-	0
Time stamp*	Initial tree creation time stamp in serial date format	5062	double (64)	s	736810.833413958
ColorIndex	Color index	5063	uint8 (8)	-	2
Reserved	Reserved for survey or display attributes	5064-5070	-	-	-

# **Matthew Parkan**

## **Environmental engineer**

---

### **Address**

Avenue du 24 Janvier 34  
CH-1020 Renens  
Switzerland

### **Core technical skills**

Geospatial data analysis, environmental monitoring and modelling, remote sensing, cartography, machine learning

### **Nationality**

Swiss

### **Working skills**

Independant, structured, proficient in written and graphic communication, strong work ethics

### **Email**

[matthew.parkan@gmail.com](mailto:matthew.parkan@gmail.com)

### **Language**

French (native), English (native), German (A2)

### **Telephone**

(+41) 79 789 81 51

### **Other**

Driver's licence (cat. B), ultralight aircraft pilot license

## **Experience**

---

### **Geographic Information Systems Lab (LASIG, EPFL) / Collaborator & PhD researcher (100%)**

JUN. 2013 - DEC. 2018, LAUSANNE

- Research on airborne laser scanning and hyperspectral imaging for forest mapping.
- Teaching, exercise preparation and student project supervision at the Bachelor/Master levels for the remote sensing, GIS, decision support and spatial analysis courses.
- Technical collaborations with forest services in western Switzerland and coordination of the “Forestry and LiDAR” workgroup.
- Intervention in continuing education workshops for geomatics and forestry specialists.
- Drone mapping of the Kuzikus natural preserve (Namibia) and preparation of maps and orthoimagery. Requirements analysis for the integration of geospatial information in the management of Kalahari ecosystems (SAVMAP project).

### **Association de la Grande Cariçaie / Civil service (100%)**

OCT. 2012 - APR. 2013, CHESEAUX-NORÉAZ

- Setup of the PostgreSQL/PostGIS database and the interactive web map presenting different themes related to the natural preserve. Preparation and analysis of bathymetric maps.

### **Etat de Vaud, Forest Service / Civil service (100%)**

AUG. 2012 - OCT. 2012, CHALET-À-GOBET

- Development of a method and software to delineate and characterize forest edges with airborne laser scanning.

## **EPFL, Planning and Logistics Services / Writer & translator (100%)**

JUN. 2012 - AUG. 2012, LAUSANNE

- Writing and translating (French-English) online documentation.

## **Amedis UE / Warehouseman (30%)**

JUL. 2007 - DEC. 2011, PUIDOUX

- Handling, packing and shipping of medical supplies (team work).

## **Education**

---

### **École Polytechnique Fédérale de Lausanne (EPFL) / PhD**

FEB. 2014 - OCT. 2018, LAUSANNE

Forest mapping with airborne laser scanning and hyperspectral imagery.

### **École Polytechnique Fédérale de Lausanne (EPFL) / Master in environmental engineering**

SEP. 2010 - JUN. 2012, LAUSANNE

Specialization in geomatics.

### **École Polytechnique Fédérale de Lausanne (EPFL) / Bachelor in environmental engineering**

OCT. 2006 - JUN. 2010, LAUSANNE

### **Swiss Armed Forces / Fire rescue specialist, non-commissioned officer**

JUL. 2005 - APR. 2006, PAYERNE & SION

## **Software skills**

---

<b>Office</b>	MS Word, MS Excel, Latex, Open Office
<b>Graphics</b>	Gimp***, Inkscape***, Blender*
<b>Generic</b>	Matlab/Octave***, R***, Python**, C++*
<b>Web / Mobile</b>	Html**, CSS3**, Javascript**, Jahia** (CMS), ThreeJS*, Cordova*, Phonegap*, php*, JQuery Mobile*
<b>Geospatial</b>	Quantum GIS***, SAGA***, GDAL***, Manifold**, Openlayers**, Leaflet**, Google Maps API**, Pix4Dmapper**, Agisoft Photoscan**, GeoServer*, ArcGIS*
<b>Database</b>	PostgreSQL/PostGIS*, SQLite*

**Level:** \* beginner, \*\* intermediate, \*\*\* advanced