



TECHNISCHE
UNIVERSITÄT
WIEN



Diplomarbeit

Ableitung von Forstparametern aus ALS Daten unter Verwendung von maschinellem Lernen und GIS

Ausgeführt zum Zwecke der Erlangung des akademischen Grades einer

Diplom-Ingenieurin

im Rahmen des Studiums

Masterstudium Geodäsie und Geoinformation

eingereicht von

Johanna Thiel, BSc

Matrikelnummer 01427107

ausgeführt am Department für Geodäsie und Geoinformation der
Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Betreuer:

Senior Scientist Dipl.-Ing. Dr.techn. Markus Hollaus

Univ.Prof. Dipl.-Ing. Dr.techn. Norbert Pfeifer

Wien, 10.10.2022

(Unterschrift Verfasserin)

(Unterschrift Betreuer)



Ich habe zur Kenntnis genommen, dass ich zur Drucklegung dieser Arbeit nur mit Bewilligung der Prüfungskommission berechtigt bin.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet. Die aus den Quellen wörtlich entnommenen Stellen, sind als solche kenntlich gemacht.

Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Arbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilten Arbeit überein.

Ich nehme zur Kenntnis, dass die vorgelegte Arbeit mit geeigneten und dem derzeitigen Stand der Technik entsprechenden Mitteln (Plagiat-Erkennungssoftware) elektronisch-technisch überprüft wird. Dies stellt einerseits sicher, dass bei der Erstellung der vorgelegten Arbeit die hohen Qualitätsvorgaben im Rahmen der geltenden Regeln zur Sicherung guter wissenschaftlicher Praxis „Code of Conduct“ an der TU Wien eingehalten wurden. Zum anderen werden durch einen Abgleich mit anderen studentischen Abschlussarbeiten Verletzungen meines persönlichen Urheberrechts vermieden.

Ort, Datum

Unterschrift

Abstract

Laser scanning is a technology to capture and display topographical data. It also offers several advantages for forest applications. Information about the size of the forest areas, the wood stock and tree-specific parameters such as location, height, crown area, crown diameter, trunk circumference and tree species are the basis for ecological forest management.

Aim of this work is the assessment and analysis of the forest population for the Illwerke vkw company. An individual tree detection is carried out in order to identify the risk potential of individual trees and to determine the corresponding parameters for each tree. A potential risk would be the damage of critical infrastructure from falling trees. The results of the forest analysis can be used for forest management and further analyses. The Airborne Laser Scanning (ALS) point clouds are the input data to determine tree positions. These are detected as local maxima on a surface model derived from the point clouds. By using digital terrain models, the tree height can be determined for each tree position. Existing forest inventories are used as reference data to calibrate models and derive timber stock and diameter at breast height. In addition, the main tree species should also be determined at the individual tree level. In addition to the ALS data, orthophotos (RGB and infrared channels) are used for the tree species classification. The classification is performed with different databases to assess how the results differ. The automation of the analysis is given by the implementation as a Python script. ArcGIS modules are accessed through the ArcGIS API for Python. The aim of this work is to use a machine learning tool to carry out the classification of tree species. The proof of concept areas are situated in such a way that several small areas of different altitudes and tree species are chosen to take forest areas with different characteristics into account. A choice of seven reference areas is made in which trees were surveyed and tree parameters documented. This data is used to validate the single tree detection and to build the tree species classification model.

The automatic and reusable analysis presented above creates a digitalization of the forest inventory and thus the management and preservation of the forests can be improved.

After validation of the method with the reference data, a detection rate of 75% was achieved for single tree detection. Better results were reached in coniferous forests. In

areas with many deciduous trees, an accuracy of less than 70% was achieved. When classifying the detected individual trees into the main tree species spruce, fir, deciduous and other conifers, a model accuracy of 71% was achieved.

Kurzfassung

Laser Scanning ist eine qualitative Technologie um topografische Daten zu erfassen und darzustellen und bietet auch für Forstanwendungen mehrere Vorteile. Informationen über die Größe der Waldflächen, den Holzvorrat und baumspezifische Parameter wie Lage, Höhe, Kronenfläche, Kronendurchmesser, Stammumfang und Baumart sind die Basis für ein ökologisches Waldmanagement.

Das Ziel dieser Arbeit ist die Abschätzung und Analyse des Forstbestandes der Illwerke vkw AG. Es wird eine Einzelbaumdetektion durchgeführt, um das Gefahrenpotential einzelner Bäume zu identifizieren und entsprechende Parameter für jeden Baum zu bestimmen. Ein mögliches Risiko ist die Beschädigung von kritischer Infrastruktur durch umfallende Bäume. Die Ergebnisse der Forstanalyse sollen als Input für den Waldmanagementplan und weitere Analysen dienen. Die mittels Airborne Laser Scanning (ALS) erfassten Punktwolken sind die Eingangsdaten um die Baumpositionen zu bestimmen. Diese werden als lokale Maxima auf einem aus den Punktwolken abgeleiteten Oberflächenmodell detektiert. Durch Verwendung von digitalen Geländemodellen kann mit der Baumposition auch die Baumhöhe bestimmt werden. Bestehende Forstinventuren werden als Referenzdaten eingesetzt um Modelle zu kalibrieren und den Holzvorrat und den Brusthöhendurchmesser abzuleiten. Zusätzlich sollen auch die Hauptbaumarten auf Einzelbaumniveau bestimmt werden. Für die Baumartenklassifizierung werden zusätzlich zu den ALS-Daten auch Orthophotos (RGB und Infrarot Kanäle) eingesetzt. Die Klassifizierung wird mit unterschiedlicher Datenbasis durchgeführt, um zu beurteilen, wie sich die Ergebnisse unterscheiden. Die Automatisierbarkeit der Analyse soll durch die Realisierung als Python Script gegeben werden. Es wird über die ArcGIS API for Python auf ArcGIS Module zugegriffen. Das Ziel der Arbeit ist es, die Klassifizierung nach Baumarten durch den Einsatz eines Machine Learning Tools durchzuführen. Die Proof of Concept Gebiete werden so situiert, dass mehrere kleine Gebiete unterschiedlicher Seehöhe und Baumarten gewählt werden, um Waldflächen unterschiedlicher Eigenschaften zu berücksichtigen. Für diese Arbeit wurde eine Wahl von sieben Referenzgebieten getroffen, in denen Bäume eingemessen und Baumparameter dokumentiert wurden. Diese Daten wurden für die Validierung der

Einzelbaumdetektion und die Erstellung des Modells zur Baumartenklassifizierung eingesetzt.

Durch die oben vorgestellte automatische und wiederverwendbare Analyse wird eine Digitalisierung der Forstinventur geschaffen und somit kann die Pflege und Erhaltung der Wälder verbessert werden.

Nach Validierung der Methode mit den Referenzdaten konnte für die Einzelbaumdetektion eine Detektionsrate von 75% erreicht werden. In Nadelwäldern konnten teilweise auch bessere Ergebnisse erreicht werden. In Gebieten mit vielen Laubbäumen konnte teilweise nur eine Genauigkeit unter 70% erlangt werden. Bei der Klassifizierung der detektierten Einzelbäume in die Hauptbaumartengruppen Fichte, Tanne, Laub und andere Nadelbäume konnte eine Modellgenauigkeit von 71% erreicht werden.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation und Ziel der Arbeit	1
1.2	Aufbau der Arbeit	2
2	Theoretische Grundlagen	3
2.1	ALS Systeme.....	3
2.2	Forstwirtschaftliche Grundlagen	4
2.2.1	Bedeutung des Waldes.....	4
2.2.2	Waldmanagement	5
2.3	Forstparameter.....	5
2.3.1	Forstinventur	6
2.4	Methoden digitaler Waldinventuren	7
2.4.1	Baumbasierter und flächenbasierter Ansatz.....	7
2.4.2	Ableitung von Einzelbaumparametern.....	8
2.4.3	Analyse von Baumarten.....	11
2.5	Waldkartierung.....	13
2.5.1	Kartierung mit optischen Fernerkundungssystemen	14
2.5.2	Kartierung mit Mikrowellenfernerkundung	15
2.5.3	Kartierung mit Airborne Laser Scanning.....	16
2.6	Maschinelles Lernen	18
2.6.1	Random Forest Classifier	19
2.6.2	Kurvenanpassung	20
3	Methoden und Implementierung	21
3.1	Datensätze	21
3.1.1	ALS Daten.....	21
3.1.2	Geländemodell	22
3.1.3	Orthophoto.....	23
3.1.4	Referenzdaten	24
3.2	Einzelbaumdetektion.....	26

3.2.1	Ableitung des Oberflächenmodells	26
3.2.2	Maximadetektion	26
3.3	Baumartenklassifizierung	28
3.3.1	Datensatz	28
3.3.2	Parameter	29
3.3.3	Modell.....	31
3.4	Holzvorrat	32
3.4.1	Datensatz	32
3.4.2	Modell.....	32
3.5	Validierung.....	32
3.5.1	Einzelbaumdetektion.....	33
3.5.2	Baumartenklassifizierung	34
4	Ergebnisse und Diskussion.....	35
4.1	Einzelbaumdetektion.....	35
4.1.1	Validierung.....	35
4.2	Baumartenklassifizierung	43
4.2.1	Validierung.....	43
4.3	Holzvorrat	47
4.4	Diskussion der Ergebnisse	49
5	Zusammenfassung	52
6	Abbildungsverzeichnis.....	54
7	Tabellenverzeichnis	56
8	Literaturverzeichnis.....	57

1 EINLEITUNG

1.1 Motivation und Ziel der Arbeit

Österreich ist in Mitteleuropa mit 48% Waldfläche eines der am dichtesten bewaldeten Länder. Von den über vier Millionen Hektar Wald in Österreich liegen ca. 100.000 ha in Vorarlberg, das entspricht einem Drittel der gesamten Fläche des Bundeslandes. Eine genaue Dokumentation der Wälder ist ein wichtiger Bestandteil einer effektiven Waldbewirtschaftung. Waldinventuren umfassen wichtige Parameter bezogen auf die Struktur der Bäume, wie die Höhe, den Stammdurchmesser, die Stammgrundfläche und das Volumen. Aufgrund der heutigen Datenverfügbarkeit und der technischen Analysemöglichkeiten kommen immer öfter luftgestützte Laserscandaten in Forstanalysen zum Einsatz. Für Inventuren über große Waldflächen, in abgelegenen, schwer zugänglichen oder auch gefährlichen Gebieten können luftgestützte Erhebungen die einzige Möglichkeit sein (Windrim und Bryson 2020).

Aufgrund der Datenverfügbarkeit in Vorarlberg und dem Mangel an vorhandenen Inventurdaten innerhalb der Waldflächen der Illwerke vkw AG wurde im Zuge dieser Arbeit ein digitales Forstanalysewerkzeug erstellt. Das Ziel war eine Analyse auf Einzelbaumniveau mit den resultierenden Parametern Baumposition, Höhe, Hauptbaumartengruppe und Holzvolumen. Nicht nur für die Ableitung von Maßnahmen für ein nachhaltiges Waldmanagement sondern auch zur Identifizierung des Gefahrenpotentials in der Nähe von technischen Anlagen sollen die Ergebnisse der Analyse dienen. Die Analyse kann in mehrere Abschnitte unterteilt werden. Ein erster Schritt war die eigenständige Aufnahme von Referenzdaten, um genügend Daten zur Erstellung und Validierung des Modells zur Verfügung zu haben. Die digitale Forstanalyse besteht aus drei Abschnitten. Das erste Ziel war die Detektion von Baumspitzen und die Ableitung der Höhe. Anschließend wurde eine Baumartenklassifizierung durchgeführt, auf Basis der Baumarten wurde als letzter Schritt der Holzvorrat abgeleitet.

1.2 Aufbau der Arbeit

In Kapitel 2 wird auf die theoretischen Grundlagen eingegangen. Neben dem Aufnahmesystem der ALS Daten und den forstwirtschaftlichen Grundlagen, werden auch Methoden zur digitalen Waldinventur und Waldkartierung vorgestellt. Im letzten Teil des Kapitels wird auf die zur Analyse eingesetzten Methoden von maschinellem Lernen eingegangen. In Kapitel 3 werden die Datensätze und die implementierten Methoden näher beschrieben. Nach der Vorstellung der Methoden zur Einzelbaumdetektion und zur Baumartenklassifizierung wird auch auf die Validierung eingegangen. Die Ergebnisse der Analyse werden in Kapitel 4 erläutert und diskutiert. Nachdem in Kapitel 4.4 Probleme und mögliche Verbesserungen vorgestellt werden, folgt als Abschluss in Kapitel 5 eine Zusammenfassung der Arbeit gemeinsam mit einem Ausblick auf den weiterführenden Einsatz der gewonnenen Daten.

2 THEORETISCHE GRUNDLAGEN

2.1 ALS Systeme

Airborne Laserscanning (ALS), auch LiDAR (light detection and ranging) genannt, ist eine Technik im Bereich der aktiven Fernerkundung. Der Fernerkundungssensor sendet Laserpulse aus, diese werden von der Erdoberfläche oder darauf befindlichen Objekten zurück zum Sensor reflektiert und dort von einem Teleskop empfangen. Messgröße ist die Signallaufzeit, durch welche man mithilfe der Ausbreitungsgeschwindigkeit die Distanz, in der Fernerkundung als Range bezeichnet, berechnen kann. Ein Puls kann ein oder mehrere rückkehrende Signale haben. Man unterscheidet zwischen diskreten Systemen, welche nur einzelne Echos (z.B. erstes und letztes) aufzeichnen, und Full-Waveform Systemen, welche das gesamte zurückgestreute Signal digitalisieren. Zusätzlich wird auch basierend auf der Größe der beleuchteten Fläche auf der Erdoberfläche unterschieden. Bei Systemen mit einem kleinen Durchmesser des Lichtkegels ist eine Größenordnung von wenigen Zentimetern gemeint, von einem großen Durchmesser spricht man bei mehreren Metern (Dubayah und Drake 2000).

In den meisten Fällen befinden sich Lidar-Systeme auf Flugzeugplattformen und bestehen aus dem Laserscanner, einem globalen Positionierungssystem und einer inertialen Messeinheit (IMU), um auch die Position des Laserscanners und die Neigung des Flugzeuges zu bestimmen. Abbildung 2.1 zeigt den Aufbau eines Lidar Systems. Die Bestimmung von Position und Richtung ist notwendig, um auch die Richtung des ausgesendeten Pulses zu bestimmen. Zusätzlich zur Position wird auch die Intensität des zurückgestreuten Signales aufgezeichnet. Moderne Lidarsysteme können bis zu 800.000 Pulse pro Sekunde aussenden, welche wie schon oben erwähnt jeweils mehrere Echos haben können. Die Positionen aller Echos bilden eine Punktwolke (Markus Melin et al. 2017).

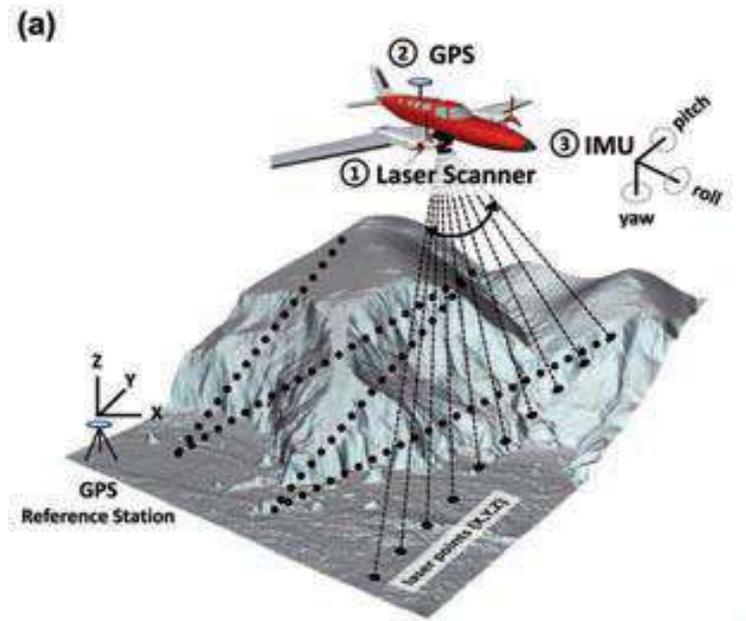


Abbildung 2.1: Aufbau eines Lidar-Systems. Abbildung aus Höfle und Rutzinger (2011)

2.2 Forstwirtschaftliche Grundlagen

Der Klimawandel ist heute die größte Herausforderung, er ist nicht nur ein Umweltproblem, sondern auch problematisch in den Bereichen Wirtschaft, Sicherheit, Moral und Sozialleben. Für das Wohlergehen der Bevölkerung und der Erde selbst ist die Zukunft der Wälder von zentraler Bedeutung (Freer-Smith et al. 2007).

Heutzutage ist eine effektive Waldbewirtschaftung besonders wichtig und in vielen Bereichen ein Thema. Die Wälder des Landes sind eine der wichtigsten Ressourcen um negative Klimaveränderungen zu verhindern (Avdeeva et al. 2020).

2.2.1 Bedeutung des Waldes

Die Wälder in Österreich dienen nicht nur zur Erholung oder als Holzlieferant, sie haben auch noch viele andere wichtige Aufgaben. Wald kann als Schutz vor Lärm oder Lawinen dienen und schützt außerdem vor Bodenerosionen. Bei Regenfällen sorgt der Waldboden für einen langsamen Abfluss, speichert und reinigt das Wasser und leistet so auch einen Beitrag zur Trinkwassergewinnung. Aufgrund der genannten Punkte und vor allem durch

ihre Rolle als Kohlenstoffspeicher haben Wälder eine wichtige Rolle im Kampf gegen den Klimawandel.

Wälder machen rund 30 % der Landoberfläche der Erde aus und Waldökosysteme einschließlich ihrer Böden speichern mit rund 1200 Gigatonnen Kohlenstoff deutlich mehr als in der Atmosphäre vorhanden ist. Ein großer anthropogen verursachter Klimawandel ist im Gange, und dieser wird größtenteils durch Änderungen des globalen Kohlenstoffkreislaufs durch die Emission von Kohlendioxid und Methan verursacht. Die Wechselwirkungen zwischen Waldökosystemen und der Atmosphäre sind eindeutig von entscheidender Bedeutung für die Einschätzung und Bekämpfung des Klimawandels und somit für die Bewertung von Maßnahmen gegen den Klimawandel (Freer-Smith et al. 2007).

2.2.2 Waldmanagement

In Evens et al (2001) wird der Begriff Waldbewirtschaftungssystem beschrieben als die Techniken und Regeln, die entwickelt wurden um Waldflächen in einem gewünschten Zustand zu erhalten. Diese Regeln und Techniken werden so angepasst, dass auch mit veränderten Gegebenheiten umgegangen werden kann. Bestandteil des Managements sind die Identifizierung und Festlegung der Waldgrenzen, Festlegung und Kontrolle des Ertragsniveaus des Waldes und auch die Kontrolle der Verfügbarkeit und Kosten der Rohstoffe. Die andauernde Verfügbarkeit der Rohstoffe kann durch eine aktive Bewirtschaftung garantiert werden.

Schon kleine Unterschiede in Bewirtschaftungssystemen oder waldbaulichen Praktiken können erhebliche Auswirkungen auf Wälder und ihre biologische Vielfalt haben. Die Auswirkungen der Waldbewirtschaftung, insbesondere der Holzproduktion, auf die Biodiversität haben große öffentliche Aufmerksamkeit erregt (Julian Evens 2001).

2.3 Forstparameter

Basis für Bewirtschaftungssysteme sind Daten über den jeweiligen Wald. Zu diesen Daten zählen unter anderem Parameter wie Baumhöhe, Baumart, Parameter des Baumstammes oder der Baumkrone, Altersstrukturen, Wachstum und auch

Informationen über die Gesundheit der Bäume. Da es vor allem bei großen Waldflächen unmöglich ist, Informationen über alle Bäume des Waldes zu haben, wird von ausgewählten Stichproben auf den gesamten Wald geschlossen. Der Vorgang einer solchen Datenerhebung wird in sogenannten Forstinventuren durchgeführt und wird im nächsten Absatz erklärt.

2.3.1 Forstinventur

Eine Forstinventur oder auch Waldinventur dient der Aufnahme des Zustandes des Waldes. Die meisten Forstbetriebe führen regelmäßig Inventuren durch, um den Zustand und Veränderungen zu dokumentieren und daraus Bewirtschaftungskonzepte abzuleiten. Solche Inventuren gibt es jedoch nicht nur auf regionaler Ebene, viele Länder führen auch landesweite Inventuren durch. In Österreich werden seit 1961 Daten über den österreichischen Wald erhoben. Die österreichische Waldinventur besteht aus Stichprobenstandorten, welche systematisch über das Land verteilt sind. Insgesamt sind es über 11000 Probeflächen, an denen alle fünf bis sieben Jahre Daten erfasst werden (Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft (BFW), 2022).

Rund um jeden Standort einer Stichprobe werden ein kleiner und ein großer Probekreis gezogen. Innerhalb der kleinen Probefläche mit einer Größe von 21 m² werden nur Bäume mit einem Brusthöhendurchmesser (BHD) zwischen 50 und 104 mm aufgenommen. Der BHD ist der Stammdurchmesser auf einer Höhe von 1.3 m. Der große Probekreis mit einer Größe von 300 m² dient zur Erfassung von Flächenmerkmalen, hier werden alle Bäume mit einem BHD ab 105 mm aufgenommen. Im Datensatz jeder Stichprobe sind Informationen über den Waldanteil, Holzvorrat, BHD der Probebäume, Baumart und Baumhöhe enthalten. Die beschriebenen Daten der Probeflächen werden in der Bestandesgrundfläche zusammengefasst. Dies ist in der Forstwirtschaft ein Maß zur Einschätzung von stehenden Holzvorräten. Die Bestandesgrundfläche bezeichnet die Dichte und wird aus den Grundflächen der Einzelbäume berechnet. Ein häufig angewendetes Stichprobenverfahren der Forstinventur ist die Winkelzählprobe nach Bitterlich, welche dazu dient, die Probestämme innerhalb eines Plots auszuwählen und die Bestandesgrundfläche zu erfassen um anschließend Bestockungsgrad und

Holzvolumen daraus zu berechnen. Die Bäume werden mit einem Spiegelrelaskop erfasst. Im Zuge der Inventur werden von jedem Probebaum die Polarkoordinaten ausgehend vom Probeflächenmittelpunkt, der BHD, die Baumart, die Baumhöhe, Kronengrenzhöhe und die Baumklasse erfasst. Aus diesen Parametern wird der Waldanteil, in Zehntel-Flächenanteilen abgeleitet. Der Holzvorrat wird durch statistische Beziehungen von BHD, Baumhöhe und Formzahl in m^3/ha berechnet. Die Formzahl wird in Kapitel 2.4.2.3 erklärt.

2.4 Methoden digitaler Waldinventuren

Luftgestützte Laserscandaten haben großes Potential in der Ableitung von Baumparametern. Meist handelt es sich bei den gesuchten Parametern um Baumhöhe und Holzvolumen. Diese können auf Bestandsebene, aber auch auf Einzelbaumniveau abgeleitet werden.

2.4.1 Baumbasierter und flächenbasierter Ansatz

Ansätze auf Bestandsebene basieren auf empirischen Modellen und sind je nach Waldzusammensetzung, Stichprobengröße, Stichprobenanzahl und Gebietsgröße nicht immer aussagekräftig.

Hollaus et al. (2009c) geht zu Beginn des Papers auf die verschiedenen Modellarten ein, die bei einem flächenbasierten Ansatz zur Ableitung von Forstparametern eingesetzt werden können. Beim Einsatz von physikalischen Modellen wird ein dreidimensionales Modell benötigt, welches die Ausbreitung der Echos innerhalb der Baumschichten beschreibt. Der Vorteil der physikalischen Modelle ist es, dass mehrere Forstparameter zeitgleich abgeleitet werden können, jedoch müssen auch genaue Informationen über die Waldzusammensetzung, die Kronenform und die Reflexion der Waldoberfläche bekannt sein, um Fehler bei der Invertierung des Modells zu vermeiden. Da solche Informationen meist nicht vorhanden sind, können oft keine physikalischen Modelle zur Ableitung der Forstparameter eingesetzt werden.

Ein wichtiger Bestandteil der flächenbasierten Holzvorratsabschätzungen sind in-situ Forstinventur Daten. Um die Stichprobendaten einzusetzen, muss eine Co-Registrierung

zwischen den Baumpositionen der Probeflächen und den Laserscandaten durchgeführt werden. Das Volumen zwischen der Waldoberfläche und der Geländeoberfläche stellt das umhüllende Baumvolumen dar und wird dem Baumvolumen aus der Inventur der Referenzfläche gegenübergestellt. Mithilfe einer linearen Regressionsanalyse wird eine Beziehung zwischen den beiden Volumens Parametern abgeleitet (Schaudauer K., BFW Web).

Bei einem flächenbasierten Ansatz wird die hohe Auflösung der Laserscandaten nicht vollständig ausgeschöpft. In vielen Gebieten stehen Laserscandaten mit einer ausreichenden Punktdichte für eine Analyse auf Einzelbaumniveau zur Verfügung. Parameter, wie die Baumspitzen und die Kronenflächen können direkt aus der Punktwolke abgeleitet werden (Shardt et al. 2002).

Vor allem bei einem Waldmanagement für kleinere Gebiete kann eine Analyse auf Einzelbaumniveau interessant sein und bei der Planung Abhilfe schaffen. Außerdem ist für Abschätzungen des Gefahrenpotentials eines Waldes oft Information auf Einzelbaumniveau notwendig, wenn es beispielsweise darum geht einzelne Bäume zu identifizieren, welche aufgrund ihrer Höhe eine Leitung beschädigen könnten. Im nachfolgenden Absatz werden verschiedene baumbasierte Ansätze beschrieben.

2.4.2 Ableitung von Einzelbaumparametern

2.4.2.1 Baumposition

In Eysn et al. (2015) werden verschiedene Methoden zur Ableitung der Baumpositionen miteinander verglichen. Dafür wurden acht Referenzgebiete in fünf unterschiedlichen alpinen Ländern gewählt. Somit konnten Aussagen bezüglich der Genauigkeiten der verschiedenen Methoden in Gebieten unterschiedlicher Waldzusammensetzungen getroffen werden. Es wurden insgesamt acht Methoden aus Forschungsprojekten unterschiedlicher Länder getestet. Fast alle Teilnehmer setzten eine Lokale Maxima Suche auf dem Oberflächenmodell ein. Nur bei einer der Methoden werden die Baumpositionen direkt in der Punktwolke detektiert (Eysn et al. 2015).

Bei sechs der getesteten Methoden wurde eine Lokale Maxima Suche auf dem DSM durchgeführt, sie unterscheiden sich einerseits in der Auflösung des Oberflächenmodells,

welche zwischen 0.2 m und 1 m variiert, und andererseits in der gewählten Kernelgröße zur Maxima Suche, welche zwischen 1.5 x 1.5 m und 7 x 7 m liegt. Bei einer der Methoden wurde eine Detektion in mehreren Höhenmodellen verschiedener Ebenen durchgeführt, um auch Information in unteren Höhengschichten zu nutzen. Bei zwei weiteren Methoden wurden zusätzlich zur Maxima Detektion auch die Baumkronen segmentiert. Hierfür wurden Region Growing und die Watershed Methode eingesetzt. Eine weitere vorgestellte Möglichkeit, beschrieben in Lindberg et al. (2014), ist die Polynomische Anpassung. Hierbei werden für jeden Pixel des Höhenmodells mit einer Nachbarschaft von 3.5 x 3.5 m die Polynome zweiten Grades gefittet, um anschließend konkave Nachbarschaften als potentielle Baumspitzen zu detektieren. Es hat sich gezeigt, dass vor allem in niedrigeren Höhengschichten mit fortgeschrittenen Methoden, wie der Baumdetektion direkt in der Punktwolke, bessere Detektionsraten erreicht wurden. Bei der in Lindberg et al. (2014) vorgestellten Methode für die Baumdetektion in der Punktwolke erfolgt die Abgrenzung der Baumkronen durch Segmentierung eines Korrelationsoberflächenmodells. Die Segmentierung basiert auf geometrischen Baumkronenmodellen und Rasterdaten. Es wird für jede Rasterzelle eine Ellipsoidfläche berechnet, dessen Höhe wird dann durch Korrelation den Höhen der ALS-Daten gegenübergestellt. Das Korrelationsoberflächenmodell enthält die Korrelationskoeffizienten und zeigt somit, wie gut jede Rasterzelle eine Baumspitze repräsentiert. Anschließend wird ein dreidimensionales ellipsoidisches Baummodell-Clustering in den Laserscandaten durchgeführt. (Eysn et al. 2015).

2.4.2.2 Brusthöhendurchmesser

Einige Einzelbaumparameter, wie einzelne Kronenparameter können aus dichten Punktwolken direkt abgeleitet werden, weitere Parameter können anschließend über mathematische Beziehungen berechnet werden.

In Windrim und Bryson (2020) können aufgrund der hohen Punktdichte von 300 bis 700 Punkten pro Quadratmeter, aufgenommen mit einem Riegl VUX-1 LiDAR Scanner, nach der Lokalisierung der Baumpositionen noch weitere Parameter aus der Punktwolke abgeleitet werden. Die Analyse besteht aus mehreren Stufen. Nach der Identifizierung und Entfernung von Bodenpunkten werden die Einzelbäume abgegrenzt. Anschließend werden die Baumpunkte in Stamm- und Kronenpunkte segmentiert. Auf die Stammpunkte wird mittels Deep Learning, einer Methode von maschinellem Lernen

durch künstliche neuronale Netze, eine Modellanpassung durchgeführt. Dafür wird auf die Stammpunkte eine Hauptkomponentenanalyse angewandt um die Vektorrichtung der Punkte und somit die ungefähre Richtung der Baummittellinie zu bestimmen. Aus dem Richtungsvektor wird anschließend eine Koordinatentransformation berechnet um die z-Achse entlang der Baummittellinie auszurichten und den höchsten Punkt der Baumspitze zu identifizieren. Durch Berechnung von horizontalen Kreisen entlang der Baumachse und anschließendem Modell Fitting mittels der Methode der kleinsten Quadrate wird ein Stammmodell erstellt. Aus den Stammmodellen der Bäume wird der Brusthöhendurchmesser abgeleitet (Windrim und Bryson 2020).

Auch für die Beziehungen der Forstparameter untereinander gibt es einige gängige Formeln. In Shardt et al. (2002) werden empirische Modelle vorgestellt, auf Basis welcher der Brusthöhendurchmesser berechnet werden kann. Das in dem Paper abgeleitete Modell wurde mit Daten von Fichten in einem Gebiet in den Hohentauern erstellt. Für ähnliche Waldgebiete kann die untenstehende Formel (2.1) zu guten Ergebnissen führen. Mit der Baumhöhe $h(B)$ und der Kronenfläche $A(K)$ kann der Brusthöhendurchmesser berechnet werden aus:

$$BHD[mm] = -31,96 + 1,33 * h(B) + 5,19 * A(K) \quad (2.1)$$

Es kann auch sinnvoll sein, in die Formel des Brusthöhendurchmessers andere Parameter, wie z.B. die Seehöhe, mit einfließen zu lassen.

2.4.2.3 Holzvorrat

Wenn Baumhöhe und BHD erfasst wurden, wird für die Volumensberechnung noch ein dritter Parameter, die Formzahl, benötigt. Sie beschreibt die Form des Baumstammes, welcher nach oben hin meist dünner wird. Für die Berechnung einer echten Formzahl, müsste der Stammdurchmesser auch 0.9 m unter der Baumspitze gemessen werden, deswegen wird bei Inventuren eine „unechte“ Formzahl eingesetzt. Dafür werden die in Pollanschütz (1974) vorgestellten Schätzfunktionen eingesetzt, die den BHD, die Baumhöhe und zusätzlich Regressionskoeffizienten beinhalten. Zur Berechnung dieser Regressionskoeffizienten wurden 17000 Probestämme unterschiedlicher Baumarten,

Altersklassen und Größen eingemessen. Aus den aufgenommenen Parametern von BHD, Stammlänge, Kronendurchmesser, Kronenbreite, Stammdurchmesser in mehreren Sektionen und Rindenstärke, wurde das Schaftholzvolumen abgeleitet. Mit einem Regressionsprogramm wurden daraus jeweils sieben Regressionskoeffizienten der Baumarten Fichte, Tanne, Lärche, Kiefer, Buche und Eiche berechnet. Die Schätzfunktion zur Ermittlung von Schaftholzformzahlen wurde aus Pollanschütz (1974) entnommen. Die unechte Schaftholzformzahl (f_{SmR}) wird mit der Formel (2.2) berechnet. Hierbei sind BHD und Baumhöhe ($h(B)$) in Dezimetern anzugeben.

$$f_{SmR} = b_1 + b_2 * \ln^2(BHD) + b_3 * \frac{1}{h(B)} + b_4 * \frac{1}{BHD} + b_5 * \frac{1}{BHD^2} + b_6 * \frac{1}{BHD * h(B)} + b_7 * \frac{1}{BHD^2 * h(B)} \quad (1.2)$$

In der obigen Formel steht \ln für den natürlichen Logarithmus und b_1 bis b_7 sind die Regressionskoeffizienten einer Baumart. Mit der Formzahl kann anschließend das Holzvolumen berechnet werden:

$$HV[m^3] = \frac{(BHD^2/4 * \pi * h(B) * FZ)}{1000} \quad (2.3)$$

Die Formzahl und somit auch das berechnete Holzvolumen beziehen sich auf das Schaftholz, welches bei Nadelbäumen den gesamten Baumstamm samt Rinde bis zur Baumspitze, jedoch ohne Astholz bezeichnet. Auch der Verlust aufgrund der Fällung mit Motorsägen wird berücksichtigt. Da es bei Laubbäumen keinen bis zum Kronendach durchgehenden Baumstamm gibt, bezieht sich hier das Holzvolumen auf den Baumstamm und zusätzlich einem bis zur höchsten Stelle der Krone durchgehenden Ast (Pollanschütz 1974).

2.4.3 Analyse von Baumarten

Es gibt viele Studien zur flächenbasierten Analyse der Baumarten. Dabei werden Laserscandaten oft gemeinsam mit Orthophotos eingesetzt. Durch die immer größere Verfügbarkeit von Fernerkundungsdaten mit sehr hoher räumlicher Auflösung ist es

möglich, auch bei der Baumartenklassifizierung von einem flächenbasierten Ansatz auf einen einzelbaumbasierten Ansatz überzugehen. Beim Einsatz von passiven Fernerkundungsdaten können die Abhängigkeit des Sonnenlichts und der starke Einfluss der wechselnden Beleuchtungsbedingungen zu Schwierigkeiten führen. Es kommt oft zu sehr dunklen Schattenbereichen und auch zu überbelichteten Bereichen in Luftbildern. Beim Einsatz von passiven Bilddaten ist die Unterscheidung verschiedener Nadelbaumarten aufgrund der spektralen Ähnlichkeit kaum möglich. Vor allem mit Full-Waveform Laserscan Daten können die strukturellen Unterschiede dieser Arten jedoch besser erfasst werden. Mit dem alleinigen Einsatz von Orthophotos sind noch keine Klassifizierungen auf Einzelbaumniveau bekannt (Nevalainen et al. 2017).

Nevalainen et al. (2017) führte eine Klassifizierung durch Kombination der hyperspektralen Bilder mit hochaufgelösten luftgestützten Punktwolken durch. Es wird bei der Klassifizierung vor allem Wert auf die Bedeutung verschiedener spektraler und struktureller Merkmale der Einzelbäume gelegt. Die Ergebnisse zeigten, dass die Klassifizierung durch Einsatz der hyperspektralen Bilder durch die Hinzunahme von Merkmalen der Punktwolke nicht verbessert wurde (Nevalainen et al. 2017).

Sehr gute Genauigkeiten wiederum brachte die Klassifizierung mit Multi Wavelength ALS Daten, welche in Amiri et al. (2019) diskutiert wird.

Auch maschinelles Lernen findet seinen Einsatz in der Baumartenklassifizierung. Mustafic und Schardt (2019) testeten den Einsatz von Deep Learning zur Klassifizierung basierend auf Laserscandaten. Dafür wurden Rasterdatensätze, wie Histogramme, Perzentile und Profile, aus den Punktwolken abgeleitet. Die Rasterbilder der Baumsegmente mit bestimmtem Radius wurden an ein Convolutional Neural Network übergeben, welches mit 500 Bäumen als Referenz trainiert wurde um eine Klassifizierung zwischen Kiefern, Fichten und Laubbäumen zu erzeugen. Es konnte durch die Kombination der abgeleiteten Rasterdatensätze eine mittlere Klassifizierungsgenauigkeit von bis zu 74% erreicht werden (Mustafic und Schardt 2019).

2.5 Waldkartierung

Die in Kapitel 2.3.1 beschriebene Inventur von Wäldern ist nur ein Stichprobenverfahren. Sie kann nicht für den gesamten Wald flächendeckend durchgeführt werden und ist vor allem in heterogenen Waldflächen stark von der Wahl der Stichprobenstandpunkte abhängig.

Der digitale Fortschritt bringt viele Möglichkeiten zur Analyse von Wäldern mit sich. Durch Einsatz von flächendeckenden Datensätzen können Analysen über die gesamte Fläche durchgeführt werden.

Nitoslawski et al. (2021) führte eine Literaturdatenbanksuche durch, um für die Zeitspanne von 2010 bis 2020 festzustellen, welche Technologien für die Kartierung der Wälder eingesetzt wurden. In dem Beitrag werden die Unterschiede der Tools und Technologien zur Datengewinnung, die Analysemethoden, deren Anwendungen und deren räumliche und zeitliche Auflösung analysiert. In der Datenbank wird eine Stichwortsuche durchgeführt um festzustellen, welche Technologien über die letzten Jahre an Bedeutung gewonnen haben. In Abbildung 2.2 ist zu sehen, dass maschinelles Lernen in den letzten Jahren erheblich angewachsen ist. Viele Vorkommnisse beziehen sich auf die Sammlung, Verarbeitung und Analyse von Fernerkundungsdaten, auch Algorithmen zur Klassifizierung und Segmentierung kommen vor. Die Gruppe der unbemannten Fahrzeuge, bestehend vor allem aus Flugzeugen, aber auch terrestrischen Fahrzeugen, gewinnt seit 2015 stetig an Interesse. Auch bei Schlagwörtern bezüglich der Nutzung von LiDAR Daten ist eine Steigerung des Interesses und der Nutzung in den letzten 10 Jahren zu erkennen. Der Abfall im Jahr 2020 entsteht dadurch, dass die Suche nur bis August 2020 durchgeführt wurde (Nitoslawski et al. 2021).

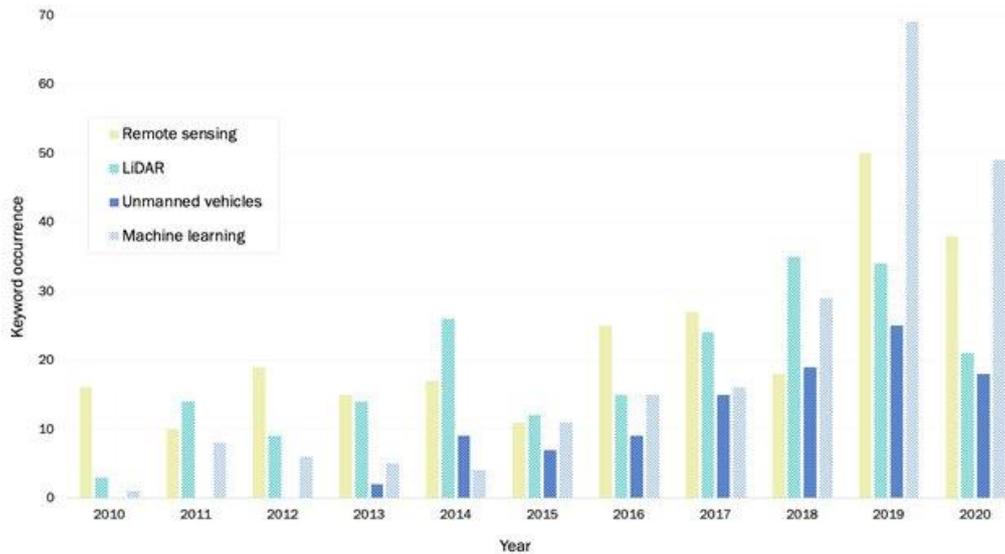


Abbildung 2.2: Die häufigsten Schlüsselwörter in Artikeln bezogen auf Technologien zur Datengewinnung für die Kartierung von Wäldern im Zeitraum 2010 bis 2020. Abbildung aus Nitoslawski et al. (2021)

Im folgenden Kapitel wird auf die Anwendungen der unterschiedlichen Datengewinnungstechnologien eingegangen.

2.5.1 Kartierung mit optischen Fernerkundungssystemen

Daten von optischen Erdbeobachtungssatelliten werden vor allem eingesetzt, um die Veränderungen der Waldflächen zu detektieren.

Pitkänen et al. (2020) beschäftigt sich mit der Erstellung eines Tools zur Überwachung von Managementmaßnahmen. Dazu wurde eine Veränderungsdetektion (engl. Change Detection) unter Verwendung von multi-temporalen Sentinel-2 Satellitenbildern durchgeführt. Diese wurde als automatisierte Modellierung in nahezu Echtzeit umgesetzt.

Die Studie von Pourghasemi et al. (2020) bewertet die Anfälligkeit für Waldbrände im Iran durch Einsatz von Algorithmen des maschinellen Lernens basierend auf geographischen Informationssystemen. Die Informationssysteme enthalten neben wichtigen Einflussfaktoren wie Höhe, Neigung, topografischer Nässeindex, Ausrichtung, Entfernung von städtischen Gebieten, Jahresmitteltemperatur, Landnutzung und Entfernung von Straßen auch Landsat-8 OLI und MODIS-Satellitenbilder. Die

resultierenden Karten zur Anfälligkeit für Waldbrände können die Planung und Bewirtschaftung von Waldressourcen in der analysierten Provinz verbessern.

Kislov und Korznikov (2020) beschäftigten sich mit der Erkennung und Kartierung von Windwurfflächen, welche aus Sicht der Waldbewirtschaftung und des Naturschutzes von hoher Bedeutung ist. Hierbei wird das Potenzial neuronaler Netze bei der Lösung von Bildklassifizierungsproblemen genutzt. Es wurde ein Algorithmus entwickelt, der auf einem U-Net-ähnlichen CNN basiert, das darauf trainiert wurde, Windwurfgebiete auf der Insel Kunaschir, in Russland, zu erkennen. Als Datenbasis wurden Satellitenbilder mit einer räumlichen Auflösung von 0.5 x 0.5 m eingesetzt. Für die Erkennung von Windwurfstellen wurde eine Gesamtgenauigkeit von bis zu 94% für bewaldete Landschaften mit Nadel- und Nadelmischwäldern erreicht. Um die Genauigkeit zu erhöhen wurden zusätzlich eine Waldmaske und weitere Informationen über blattlose Waldbestände eingesetzt.

2.5.2 Kartierung mit Mikrowellenfernerkundung

Im Gegensatz zu optischen Satellitenbildern, bei denen es zu Wolkenbedeckungen kommen kann, ist der große Vorteil bei Mikrowellenfernerkundung die Wetterunabhängigkeit. Daten des Radars mit synthetischer Apertur (SAR) können eingesetzt werden, um zwischen unterschiedlichen Arten von Waldbedeckung zu unterscheiden oder auch um die Biomasse der Wälder abzuschätzen. Jüngste Studien zeigen den Einsatz von multitemporalen Sentinel-1 Daten für die Waldkartierung, zur Klassifizierung zwischen Laub- und Nadelwald sowie zur Ableitung phänologischer Variablen (Dostálová et al. 2021).

Kaushik und Kumar (2020) versuchten mit dem C-Band Datensatz der Sentinel 1A/B-Mission Informationen über Störungen im Wald abzuleiten und diese zu lokalisieren. Dabei wird die Waldverteilung und deren zeitliche Veränderung analysiert. Der große Vorteil von Radardatensätzen ist die Fähigkeit, Informationen nicht nur von der Oberfläche abzuleiten. Die Wellen des C-Bandes dringen in die oberen Zentimeter des Bodens ein, langwellige Radarsysteme sogar mehrere Meter tief.

In der Studie von Dostálová et al. (2021) wurde der Einsatz einer kürzlich veröffentlichten multitemporalen SAR-Klassifizierungsmethode für die Waldkartierung auf kontinentaler Ebene getestet. Die Klassifizierung wurde auf ganz Europa angewandt und die daraus resultierenden Datensätze zur Dichte der Baumbedeckung und der Waldtypen wurden mit dem Walddatensatz der Copernicus High Resolution Layer (HRL) abgeglichen. Die Auswertung zeigte, dass die Genauigkeiten der erhaltenen Sentinel-1 Waldkarten sehr nahe bei den HRL-Datensätzen liegen.

Piermattei et al. (2017) nutzten die Fähigkeit von Radarsensoren, in das Laub einzudringen, um vertikale Waldstrukturen zu erfassen. Dabei wurden Daten eines hubschraubergestützten Profiling-Radarsystems im KU-Band analysiert. In dieser Studie werden die Profilwellenformen und das Rückstreusignal von borealen Wäldern analysiert. Zusätzlich wurden auch LiDAR Daten eingesetzt, mit welchen die Daten des hubschraubergestützten Radarsystems verglichen wurden. Baumkronen und Waldbodenoberfläche konnten im Datensatz erkannt werden. Das entwickelte System eignet sich zur Erstellung von Baumkronenprofilen und zur Schätzung der Geländehöhen.

2.5.3 Kartierung mit Airborne Laser Scanning

Datensätze, welche durch luftgestütztes Laser Scanning gewonnen wurden, spielen vor allem in der Klassifizierung von Wäldern und Bäumen eine wichtige Rolle. ALS Daten haben sich in der Vergangenheit zur Abgrenzung von Waldflächen und der Klassifizierung dieser, als auch zur Einzelbaumsegmentierung als effizient erwiesen. Die Abgrenzung von Waldflächen mit ALS Daten wurde schon vielfach eingesetzt. In Gebirgsregionen gestaltet sich diese Analyse schwieriger. Nicht nur die Erfassung und Weiterverarbeitung der Daten ist komplexer als in flachen Regionen, einen wesentlichen Einfluss haben auch die unterschiedlichen Strukturen der alpinen Wälder. Hollaus et al. (2006) testete bereits eingesetzte Datensätze und Methoden zur Kartierung der Baumkronenhöhen und der Abschätzung der Baumhöhen in einem Gebiet der westlichen österreichischen Alpen. Durch den Einsatz einer Großzahl an Bodenkontrollpunkten zeigte sich, dass im abgeleiteten Geländemodell bei einer Hangneigung von mehr als 60° ein Höhenfehler von über 50 cm entsteht. Jedoch bleibt der Fehler bei geringeren Neigungen von unter 20° auch unter 20 cm. Für die Ableitung der Baumhöhen wurden

verschiedene Methoden getestet. Mit dem Einsatz dreidimensionaler erster Pulse konnte eine Korrelation von 0.73-0.84 zu den eingesetzten Referenzen erzeugt werden. Auch unter Verwendung eines gitterbasierten Kronenhöhenmodells konnte eine ähnlich gute Korrelation erreicht werden (Hollaus et al. 2006).

Auf dasselbe Gebiet wurde in Hollaus et al. (2009c) ein neues flächenbasiertes semiempirisches Modell zur Schätzung des Holzvorrates angewandt. Hier wurden LiDAR Daten mit einer hohen Punktdichte eingesetzt, aus welchen das Oberflächenvolumen der Wälder abgeleitet wurde. Das hier vorgestellte Modell nimmt eine lineare Beziehung zwischen diesem Volumen und dem Holzvorrat an. Es zeigte sich, dass das Modell robust gegenüber Änderungen des Aufnahmezeitpunktes und der Intensität der Laserpunkte ist.

Das in Hollaus et al. (2009c) vorgestellte Modell zur Abschätzung des Holzvorrates wurde in Hollaus et al. (2009a) getestet und bewertet. Dafür wurde das Modell mit allen verfügbaren Waldbestandsdaten des Gebiets kalibriert. Das Resultat war eine Stammvolumenkarte des gesamten Landes Vorarlberg, welche mit der Validierung mit lokalen Waldinventurdaten eine Genauigkeit von 75% hat (Hollaus et al. 2009a).

Die vorgestellten Studien zeigen, dass Parameter wie Baumhöhe und flächenbasierter Holzvorrat auch aus ALS Daten mit geringer Punktdichte abgeleitet werden können.

Auch das Potential von Full-Waveform-Laserscandaten wurde schon vielseitig getestet. In Hollaus et al. (2009b) wurde mithilfe von Echobreite, Rückstreuquerschnitt und der vertikalen Verteilung der Echos eine Baumartenklassifizierung in einem Mischwald durchgeführt. Mithilfe der Attribute der Laserpunkte konnten Parameter berechnet werden, um eine Unterscheidung zwischen Nadel- und Laubbäumen mit einer Genauigkeit von 83% durchzuführen. Anschließend wurden noch weitere Unterscheidungen getroffen und es konnte final für die Klassifizierung zwischen Fichte, Lärche und Rotbuche eine Genauigkeit von 75% erreicht werden (Hollaus et al. 2009b).

Die zwei folgenden Beiträge zeigen noch weitere Möglichkeiten der Ableitung von Baumparametern auf Einzelbaumniveau.

Yan et al. (2020) nutzt das Potential der ALS Daten um vertikale Waldstrukturen zu erfassen. Da die auf festen Bandbreiten, hier ist der Bereich zwischen minimaler und maximaler Segmentgröße gemeint, basierenden Segmentierungsverfahren nicht für unterschiedliche Kronengrößen geeignet sind, wird eine selbstadaptive Schätzung der

Bandbreite durchgeführt. Auch in komplexen Waldgebieten konnten damit 80% der Bäume korrekt segmentiert werden.

Im Beitrag von Apostol et al. (2020) werden ALS Daten mit hochaufgelösten Luftbildern kombiniert. Durch den Einsatz objektbasierter Bildanalyse werden die Bäume automatisch detektiert und die Baumarten Fichte und Buche voneinander unterschieden. Basierend auf den ALS Daten, kombiniert mit Referenzdaten, wurde auch der Brusthöhendurchmesser der einzelnen Bäume geschätzt.

2.6 Maschinelles Lernen

Im Gegensatz zum klassischen Programmieren, bei welchem bekannte Algorithmen an den Computer übergeben werden, wird bei maschinellem Lernen das Wissen bzw. der Algorithmus als Beziehung zwischen Ein- und Ausgangsdaten künstlich erzeugt und muss dem Programmierer davor nicht bekannt sein. Dafür werden Erfahrungen in Form von Trainingsdaten benötigt. Die Maschine sucht und erkennt darin durch die Verwendung von Computeralgorithmen automatisch Gesetzmäßigkeiten und Verhaltensmuster. Durch den Einsatz dieser automatisch abgeleiteten Gesetzmäßigkeiten kann anschließend je nach Art der Daten z.B. eine Klassifizierung nach unterschiedlichen Kategorien durchgeführt werden (Frochte 2021; Bishop 2006).

Beim maschinellen Lernen unterscheidet man zwischen überwachtem und unüberwachtem Lernen. Der große Unterschied dieser beiden Kategorien, ist der Zustand der Trainingsdaten. Bei ersterem sind die Daten gelabelt, man kennt den korrekten Funktionswert, beim unüberwachten Lernen geht es im Gegensatz dazu um das Erkennen unbekannter Strukturen im Datensatz. Bei der dazwischen liegenden Klasse des bestärkenden Lernens sind die Daten ebenfalls nicht gelabelt, jedoch erhält der Computer Information darüber, ob die abgeleiteten Ergebnisse gut oder schlecht sind (Frochte 2021).

Näheres über die unterschiedlichen Klassen ist Frochte (2021) zu entnehmen, hier wird nur auf überwachtes Lernen näher eingegangen.

Es wird eine genügend große Menge an Trainingsdaten benötigt, von denen nicht nur die Eingangsfunktion, zum Beispiel einzelne Merkmale, sondern auch die Ausgangsfunktion bekannt sind. Mit Ausgangsfunktion ist hier eine Klasse oder ein Wert gemeint, auch als

Label der Daten bezeichnet. Unterschieden werden beim überwachten Lernen zwei Kategorien. Bei beiden Kategorien geht es darum, einem Element der Eingangsfunktion einen Wert der Zielmenge zuzuordnen. Bei der Klassifikation besteht die Zielmenge aus Klassen, dies sind diskrete Werte. Bei der Regression ist die Zielmenge ein Wertebereich und somit ein kontinuierliches Spektrum (Frochte 2021).

Sowohl für die Klassifikation als auch für die Regression gibt es unterschiedliche Verfahren. Im nächsten Absatz werden nur die im Zuge dieser Arbeit eingesetzten Verfahren näher behandelt.

2.6.1 Random Forest Classifier

Der Random Forest Classifier ist ein Klassifikationsverfahren auf Basis mehrerer Entscheidungsbäume. Der Name kommt von dem in Abbildung 2.3 dargestellten Aufbau der Datenstruktur eines Entscheidungsbaumes. Der Baum hat am Beginn eine Wurzel, von welcher zwei Stränge zu den nächsten Knoten gehen. An jedem Knoten wird eine Entscheidung getroffen. Am Ende jedes Stranges steht ein Blatt, welches die resultierende Klasse oder den jeweiligen Regressionswert angibt (Frochte 2021).

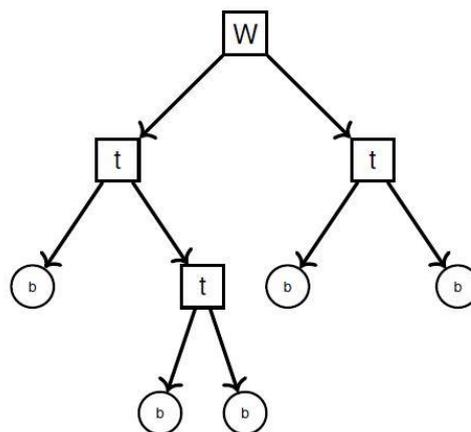


Abbildung 2.3: Datenstruktur des Entscheidungsbaumes zur Klassifikation aus Frochte (2021)

An jedem Knoten wird eine boolesche oder binäre Entscheidung getroffen, zum Beispiel, ob der Wert eines Merkmals über oder unter einer definierten Schranke liegt. Je nach

Entscheidung wird die linke oder rechte Abzweigung gewählt. Jeder Eingangsdatensatz durchläuft den Baum, bis er ein Blatt erreicht und somit ein Label zugewiesen bekommt. Wie schon erwähnt, besteht der Random Forest aus einer beliebigen Anzahl an Entscheidungsbäumen. Ergebnisklasse ist diejenige, die innerhalb der Bäume des Random Forest am häufigsten klassifiziert wurde. Dabei ist es wichtig, dass die Entscheidungsbäume untereinander eine niedrige Korrelation haben. Um zu erreichen, dass die Entscheidungsbäume innerhalb eines Random Forest möglichst nicht korrelieren, wird Bagging, kurz für Bootstrap Aggregation, angewandt. Bagging kann man als „Zurücklegen“ bezeichnen, es bedeutet hier, dass jeder Entscheidungsbaum eine zufällige Menge des Trainingsdatensatzes durch ebenfalls darin vorhandene Daten ersetzen darf. Da Entscheidungsbäume sehr empfindlich gegenüber Änderungen im Trainingsdatensatz sind, wird so eine große Korrelation zwischen den einzelnen Entscheidungsbäumen innerhalb des Random Forest vermieden. Um für zusätzliche Variation zu sorgen, kann jeder Baum zur Unterteilung an den Knoten nur eine zufällige Teilmenge der Merkmale verwenden. Weit verbreitet ist die Wahl von 100 Entscheidungsbäumen in einem Random Forest. Um die optimale Anzahl an Entscheidungsbäumen zu ermitteln, kann jedoch auch der Parameter „Out-of-Bag-Error“ eingesetzt werden, eine Methode, um den Vorhersagefehler zu berechnen. Der Out-of-Bag-Error ist der mittlere Vorhersagefehler für einen Datensatz der Trainingsdaten. Er wird berechnet, indem der Datensatz mit jedem Entscheidungsbaum ausgewertet wird, der diesen Datensatz nicht zum Training benutzt hat. Dieser Error wird für alle Datensätze berechnet, um anschließend auf dessen Basis die Anzahl der Bäume zu wählen (Frochte 2021).

2.6.2 Kurvenanpassung

Bei der Kurvenanpassung geht es darum, eine Funktion zu erhalten, welche die Werte bestmöglich vorhersagt. Je nach Beziehung der Merkmale der Referenzdaten untereinander können nicht nur lineare Funktionen, sondern auch nicht-lineare Funktionen angepasst werden. Hierfür werden die Abweichungen der Eingangsdaten von der angepassten Funktion minimiert, um die bestmögliche Funktion zu erhalten.

3 METHODEN UND IMPLEMENTIERUNG

Da die Ergebnisse der Forstanalyse nicht nur für forstwirtschaftliche Zwecke, sondern auch zur Analyse des Gefahrenpotentials eingesetzt werden sollen, wurde ein einzelbaumbasierter Ansatz gewählt.

3.1 Datensätze

3.1.1 ALS Daten

Vorarlberg ist den anderen Bundesländern mit der zeitlichen und räumlichen Auflösung der ALS Daten voraus. In Vorarlberg werden die luftgestützten Laserscandaten in einem Abstand von sechs Jahren flächendeckend über das ganze Bundesland aufgezeichnet. Die in dieser Analyse verwendeten Daten wurden vom Subunternehmer Flycom d.o.o. (Slowenien) aufgezeichnet. Aufgrund der gewählten Flughöhe von 1000 m/800 m, der daraus resultierenden Streifenbreite von 1155 m/924 m und einer gewählten Überlappung von 60%, ergab sich eine Gesamtanzahl von 1147 Streifen für das gesamte Bundesland. Aus den gewählten Flug- und Scanparametern ergab sich eine Punktdichte von 6.2 bzw. 7.7 Punkten/m² je Streifen. Die Datenaufzeichnung erfolgte im Zeitraum zwischen Februar 2017 und November 2017. Die Flüge wurden mit einem EuroCopter EC 120 Hubschrauber durchgeführt. Zusätzlich zu dem Laserscanner von Riegl (LMS-Q780, Full-Wave-Scanner, 400KHz) waren noch ein IGI IMU-IId Inertialsystem (256 Hz), ein Septentrio AsteRx4 GNSS und eine digitale Kamera mit an Bord. Die Wahl der Parameter ergab sich durch die Forderung einer Punktdichte von 8 Punkten/m² an den gesamten Datensatz. Der Zeitraum der Datenaufnahme der einzelnen Streifen wurde jeweils so gewählt, dass Schnee und Belaubung vermieden wurden (Würländer 2019b).

Die Flug- und Scanparameter wurden in Tabelle 3.1 nochmal zusammengefasst.

Tabelle 3.1: Parameter und daraus berechnete Kenngrößen des ALS Fluges, aus Würländer (2019b)

Parameter/Kenngrößen	Wert
Flughöhe über Grund	1000 m / 800 m
Fluggeschwindigkeit	70-75 ktn
Streifenabstand	450 m / 370 m
Scanfrequenz	266.667 Hz (netto)
Scanwinkel	$\pm 30^\circ$
Streifenbreite	1.155 m / 924 m
Überlappung	jeweils 60%
Mittlere Punktdichte je Streifen	6.2 / 7.7 Punkte je m ²

Die Flugfirma führte eine direkte Georeferenzierung durch, welche im Anschluss extern mit der Software ORIENT-LIDAR verfeinert wurde. In ORIENT-LIDAR wurde mithilfe von zur Verfügung stehender Passflächenmessungen eine Passflächeneinpassung eingesetzt. Hierfür wurde das Gebiet in 12 Blöcke unterteilt. Für jeden Block ergaben sich eigene Genauigkeitsparameter. Der hier relevante Parameter ist das Sigma gemittelt für alle Punkte pro Koordinatenachse. Dieser Parameter erstreckt sich in einem Bereich von 1.4 cm bis 5.6 cm. Für die Klassifizierung der Laserscandaten des gesamten Projektgebietes wurden automatische Softwaremethoden eingesetzt. Es wurde eine Klassifizierung in 16 Klassen durchgeführt, welche anschließend kontrolliert und manuell korrigiert wurde. Die für die Analyse dieser Arbeit relevantesten Klassen sind die Klassen 3-5 für niedrige, mittlere und hohe Vegetation (Würländer 2019a).

3.1.2 Geländemodell

Zusätzlich zu den ALS Daten werden vom Land Vorarlberg auch daraus abgeleitete digitale Höhenmodelle zur Verfügung gestellt. Ein Oberflächenmodell wird im Zuge dieser Arbeit selbst erstellt. Als Geländemodell zur Normalisierung der Punktwolke und des Oberflächenmodells wird das digitale Geländemodell vom Land Vorarlberg eingesetzt. Hierbei handelt es sich um einen Datensatz flächendeckend für ganz Vorarlberg mit einer Pixelgröße von 0.5 x 0.5 m.

3.1.3 Orthophoto

Gemeinsam mit der Datenaufnahme der ALS Flüge wurden auch die Luftbilder aufgenommen. Daraus wurde ein technisches Orthophoto erzeugt, woraus in weiteren Arbeitsschritten ein digitales Orthophotomosaik erstellt wurde. Bei der Aufnahme der Luftbilder gab es keine Anforderungen an die Beleuchtungssituation, wodurch auch bewölkte Gebiete und große schattige Bereiche vorkommen können.

3.1.3.1 True Orthophoto

Im Orthophoto des Landes Vorarlberg werden die Baumkronen nicht lagerichtig dargestellt, da es auf die Erdoberfläche entzerrt wurde und auf der Oberfläche befindliche Objekte aufgrund des Aufnahmewinkels der Kamera verzerrt bzw. von der Seite dargestellt sind. Um im Zuge der Analyse zu testen, ob der Einsatz von True Orthophotos bei der Baumartenklassifizierung im Vergleich zu Orthophotos Unterschiede zeigt, wurden für vier der sieben Referenzgebiete an der TU Wien True Orthophotos erstellt.



Abbildung 3.1: Vergleich von Orthophoto (links) und True Orthophoto (rechts), mit den detektierten Baumspitzen in Rot dargestellt

In Abbildung 3.1 ist gut erkennbar, dass die in Rot dargestellten Baumspitzen aus den Laserscandaten über den Baumspitzen im True Orthophoto liegen. In dem hier gezeigten Referenzgebiet ergab sich im Orthophoto aufgrund der Kameraposition und der Baumhöhe ein Versatz von ungefähr 2.5 m an den Baumspitzen.

3.1.4 Referenzdaten

Zur Validierung der durchgeführten Analyse und auch zur Parameterwahl werden Referenzdaten benötigt. Voraussetzung für diese Daten ist bei einer Analyse auf Einzelbaumniveau eine hohe Lagegenauigkeit. Trotz der zur Verfügung stehenden Daten eines Schulprojektes WaldChecks in Vorarlberg, wurde auch ein eigener Datensatz aufgezeichnet, um genügend Daten zur Validierung und zur Klassifizierung der Baumarten zu haben.

3.1.4.1 Schulprojekt

Hierbei handelt es sich um einen Punktdatensatz bestehend aus Baumpositionen und zusätzlich gemessener und berechneter Baumparameter. Dokumentiert wurden die Koordinaten, die Baumart, der Brusthöhendurchmesser, die Höhe, die Formzahl und das Volumen jedes Baumes. Im Zuge des Schulprojektes WaldChecks entstanden zwei Datensätze. Einer wurde von den Schülern und der andere vom Fachpersonal aufgezeichnet. Die beiden Datensätze enthalten überlappende, aber auch unterschiedliche Bereiche. In überlappenden Gebieten wurde der Datensatz des Fachpersonals verwendet, die Daten der Schüler wurden für Gebiete verwendet, welche nicht Teil des anderen Datensatzes waren. In Abbildung 3.2 sind die Bereiche der Datenaufnahme im Zuge des Schulprojektes dargestellt.

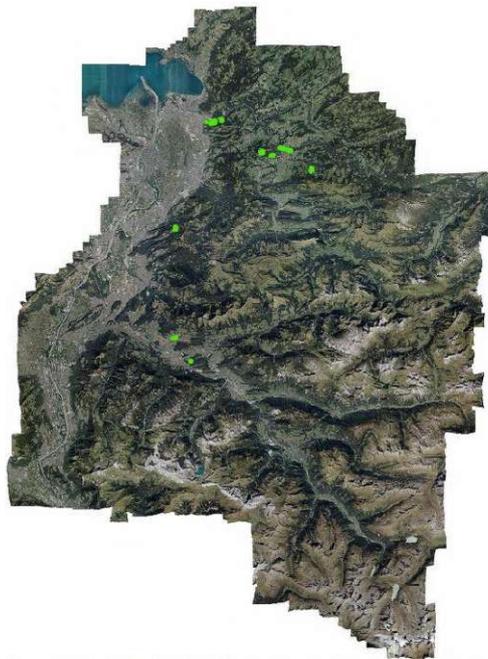


Abbildung 3.2: Aufnahmegebiete der Schulinventur in Grün. Orthophoto Land Vorarlberg

3.1.4.2 Datenaufnahme

Da die Referenzdaten möglichst gut über Vorarlberg verteilt sein sollen und die drei Hauptbaumartengruppen der Klassifizierung im Datensatz gut vertreten sein müssen, wurden auch eigens aufgenommene Referenzdaten eingesetzt. Die Datenaufnahme erfolgte in sieben über das Bundesland verteilten Gebieten, welche sich aufgrund der Höhenlage und der Waldzusammensetzung unterscheiden. Abbildung 3.3 zeigt die Lage der aufgenommenen Referenzdaten. Innerhalb der definierten Bereiche wurden alle Baumpositionen mit einem Leica TS15 Tachymeter aufgenommen. Zusätzlich wurde der Brusthöhdurchmesser mit einer Messkluppe gemessen und auch die Baumart dokumentiert. Überschattete Bäume wurden als solche gekennzeichnet, um dies bei der Validierung der digitalen Analyse mit einzubeziehen.



Abbildung 3.3: In den orangemarkierten Gebieten wurden eigenständig Referenzdaten aufgenommen.
Orthophoto Land Vorarlberg

3.2 Einzelbaumdetektion

Eine Analyse mit allen Punkten der Laserscandaten würde zu einem sehr hohen Rechenaufwand führen. Für die Bestimmung von Baumpositionen hat es sich bewährt, die Detektion nicht direkt in der Punktwolke, sondern in einem daraus abgeleiteten Oberflächenmodell durchzuführen.

3.2.1 Ableitung des Oberflächenmodells

Bei dem vorhandenen ALS Datensatz handelt es sich um eine gut klassifizierte Punktwolke, daher wurde die Punktwolke nach der Klasse 5, hohe Vegetation, gefiltert. Nur die Punkte dieser Klasse wurden verwendet, um ein digitales Oberflächenmodell mit einer Auflösung von 0.5×0.5 m zu erstellen. Jede Rasterzelle bekommt den Maximalwert der Punkte innerhalb der Zelle. Rasterzellen ohne Punkte enthalten mittels Interpolation über natürliche Nachbarn ihre Werte. Somit entsteht ein lückenfreies DSM, welches anschließend noch mittels einem Tiefpassfilter geglättet wird, um Ausreißer zu entfernen und Unebenheiten der Waldoberfläche herauszufiltern. Anschließend wird durch Subtraktion des digitalen Geländemodells des Landes Vorarlberg, welches ebenfalls eine Pixelgröße von 0.5×0.5 m hat, ein normalisiertes Oberflächenmodell abgeleitet.

3.2.2 Maximadetektion

Die Suche nach lokalen Maxima Werten wird durch einen dynamischen Kernel auf dem Oberflächenmodell durchgeführt. Alle lokalen Maxima innerhalb des Kernels mit einer bestimmten Größe sollen Baumspitzen und somit die Positionen der einzelnen Bäume darstellen. Der dynamische Kernel wird umgesetzt, indem basierend auf der normalisierten Oberflächenhöhe und der Geländehöhe unterschiedliche Kernelgrößen eingesetzt werden. Um die richtigen Werte der Kernelgrößen zu finden, wird der Validierungsalgorithmus, bestehend aus Matching Algorithmus und Kennwert Berechnungen, eingesetzt. Mit dem Algorithmus werden die statistischen Parameter der Einzelbaumdetektion, verglichen mit den Referenzdaten, berechnet. Durch Tests verschiedener Kombinationen von Kernelgrößen werden so die Parameter für die Einzelbaumdetektion gewählt.

Für das eingesetzte Geoverarbeitungswerkzeug von ArcGIS Pro stehen verschiedene Nachbarschaftstypen zur Verfügung. Hier wurden die Nachbarschaftstypen Kreis und Rechteck verwendet. Die folgende Abbildung 3.4 zeigt Beispiele von Kreis- und Rechtecknachbarschaften.

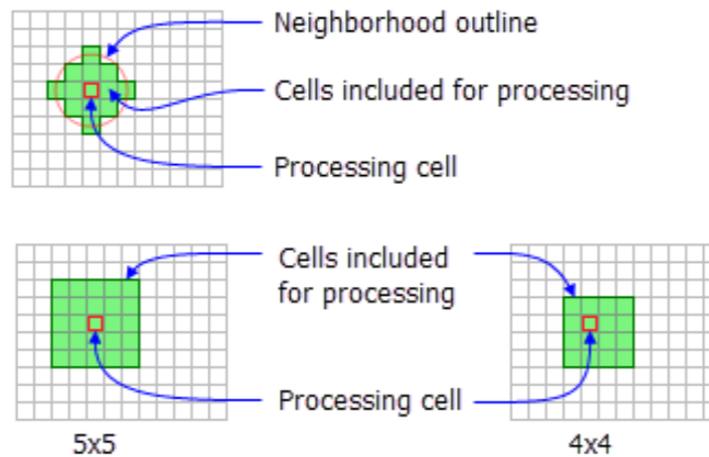


Abbildung 3.4: Beispiele von Kreis- und Rechtecknachbarschaften für die lokale Maxima Suche (Funktionsweise von „Focal Statistics“—ArcGIS Pro | Dokumentation, o. D.)

Rechtecknachbarschaften wurden nur in Form von Quadraten eingesetzt, da die Verarbeitungszelle längs und quer denselben Abstand zur Nachbarschaftsgrenze haben soll. In der obigen Abbildung ist zu sehen, dass die Verarbeitungszelle in der Rechtecknachbarschaft, nur bei ungeraden Werten für die Breite und Höhe des Rechtecks, im Zentrum des Kerns liegt. Deswegen wurde statt einem 4 x 4 Zellen Rechteck eine Kreisnachbarschaft mit einem Radius von 2 Zellen gewählt. Diese Kreisnachbarschaft beinhaltet weniger Zellen als eine 5 x 5 Rechtecknachbarschaft, jedoch mehr Zellen als ein 3 x 3 Kernel.

Die nachfolgende Tabelle (3.2) zeigt die eingesetzten Nachbarschaftstypen und Größen basierend auf der Baumhöhe, bzw. der Höhe der Rasterzelle im normalisierten DSM, und der Geländehöhe. Die Nachbarschaftsgrößen sind in Zelleneinheiten angegeben. Aufgrund der Auflösung der abgeleiteten Rasterdaten von 0.5 m ergibt sich für eine 5 x 5 Nachbarschaft eine Fenstergröße von 2.5 x 2.5 m.

Tabelle 3.2: Kernelgrößen der Einzelbaumdetektion abhängig von der Geländehöhe (Zellenwert im DTM) und der Baumhöhe (Zellenwert im nDSM). Ein $n \times n$ Kernel, bedeutet eine Maxima Suche innerhalb einer Rechtecknachbarschaft, bei einer Kreisnachbarschaft wird der Radius angegeben, beides wird jeweils in Zelleneinheiten angegeben

DTM < 480 m ü.A.	DTM < 800 m ü.A.	DTM > 800 m ü.A.
nDSM \geq 4: Kreis R=1	nDSM \geq 4: 3 x 3	nDSM \geq 4: 3 x 3
nDSM \geq 20: 3 x 3	nDSM \geq 25: Kreis R = 2	nDSM \geq 15: Kreis R = 2
nDSM > 30: 3 x 3	nDSM > 30: 5 x 5	nDSM > 35: 5 x 5

Da innerhalb der Waldflächen der Illwerke vkw unterhalb einer Geländehöhe von 480 m ü.A. hauptsächlich Misch- und Laubwälder liegen, wurde in diesen Bereichen auch mit einem DSM gearbeitet, auf welches nach der Tiefpassfilterung eine weitere Glättung angewandt wurde. Diese Glättung konnte Abhilfe bei der stark verrauschten Waldoberfläche in Laubwäldern verschaffen.

3.3 Baumartenklassifizierung

Als Basis für die Baumartenklassifizierung wird die gesamte Punktwolke herangezogen. Umgesetzt wird die Klassifizierung mithilfe von maschinellem Lernen.

3.3.1 Datensatz

Da für eine Baumartenklassifizierung die Referenzdaten aus der Natur mit den detektierten Baumspitzen übereinstimmen müssen, werden die eigens aufgenommenen Referenzdaten und die Daten des Schulprojektes den in den jeweiligen Gebieten detektierten Bäumen paarweise zugeordnet. Alle Bäume, für welche ein Partner gefunden wird, können für die Erstellung des Klassifizierungsmodells verwendet werden.

Um das Modell mit den oben genannten Daten mit einer Klassifizierung unter Hinzunahme eines Orthophotos zu vergleichen, stehen für vier der Testgebiete zusätzlich auch True Orthophotos zur Verfügung. Somit können drei Klassifizierungsmodelle mit unterschiedlicher Datenbasis verglichen werden.

3.3.2 Parameter

Für das Modell werden Merkmale der Bäume benötigt, welche sich aufgrund der Hauptbaumartengruppe voneinander unterscheiden. Dafür wird um jeden Baum ein Zylinder, mit der Baumposition als Mittelpunkt, gelegt. Nach dem Testen verschiedener Zylinderradien wurde ein Radius von 3 Metern gewählt. Alle in diesem Zylinder befindlichen Punkte werden als zum Baum zugehörig betrachtet und für die Berechnung der Merkmale verwendet. Es werden Geometrie- und Intensitätsmerkmale berechnet, indem nicht nur die Koordinaten aller Punkte, sondern auch deren Attribute, wie die Intensität, die Klassifizierung, die Nummer des Echos und die Anzahl der Echos verwendet werden. Je nach Merkmal werden nicht alle Punkte benötigt, sondern nur der nach einem bestimmten Attribut gefilterte Anteil der Punkte. Aus den ungefähr 63 berechneten Merkmalen werden jene herausgefiltert, die keinen Bezug zur Baumart haben.

3.3.2.1 Baummerkmale

Es wurden Merkmale der Einzelbäume aus dem Orthophoto, dem True Orthophoto und aus der Punktwolke berechnet. Für die beiden Orthophotos wurden jeweils in dem Baumsegment, realisiert als Kreis mit 3 m Radius um die Baumspitze, die Mittelwerte der 4 Kanäle Rot, Grün, Blau und Infrarot berechnet. Daraus wurde mit der Formel (3.1) der Normalized Difference Vegetation Index abgeleitet.

$$NDVI = \frac{NIR - Rot}{NIR + Rot} \quad (3.1)$$

Die aus den Punktwolken der Bäume berechneten Merkmale sollen die Geometrie und die Intensitätsverteilung der Punkte darstellen. Abbildung 3.5 zeigt Punktwolken der Baumarten Fichte, Tanne und Laub. Ziel ist es nun, Merkmale zu berechnen, aufgrund welcher sich die Punktwolken unterscheiden.

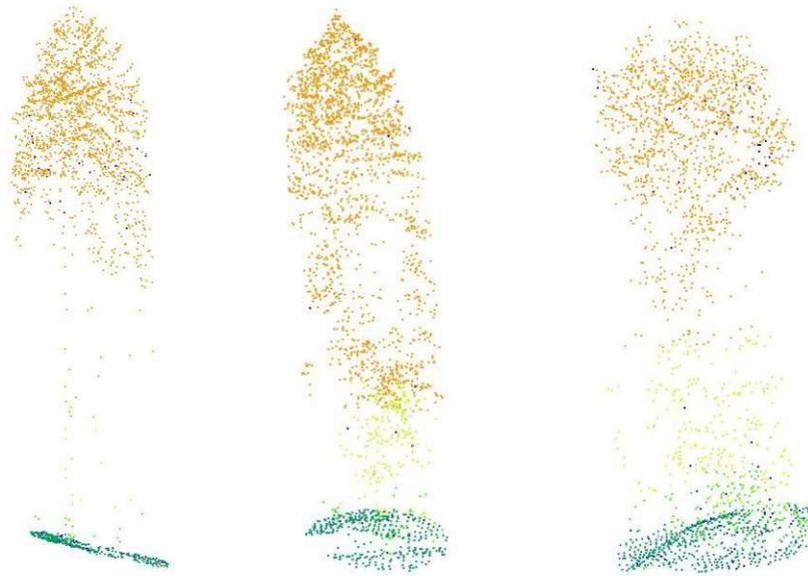


Abbildung 3.5: Punktwolken der drei Baumarten Fichte, Tanne, Laubbaum (von links nach rechts)

Für die Berechnung der Merkmale wurden die Punktwolken zuvor einerseits mittels dem Geländemodell normalisiert und zusätzlich auch mit dem Oberflächenmodell, um auch Parameter bezüglich der Höhen von der Kronenoberfläche ausgehend, berechnen zu können. Im Folgenden sind unter Vegetationspunkten alle Punkte der Klasse „Hohe Vegetation“ zu verstehen. Es wurden Merkmale zur Beschreibung der Höhenverteilung der Punkte, wie der Median der Z-Werte und die Standardabweichung, Schiefe und Wölbung der Z-Verteilung berechnet. Außerdem wurde das Verhältnis des Medians der Höhen aller Vegetationspunkte zur Baumhöhe und jeweils das Verhältnis des 75% und 90% Perzentil zur Baumhöhe berechnet. Auch über die Anzahl der Punkte im Baumsegment wurden Parameter, wie das Verhältnis der Bodenpunkte zur Gesamtanzahl und das Verhältnis der Vegetationspunkte zur Gesamtanzahl der Punkte berechnet. Um die Verteilung der Intensitäten der Vegetationspunkte zu beschreiben, wurde ebenso wie bei der Höhe ein Median, ein Maximum und eine Standardabweichung, Schiefe und Wölbung der Verteilung berechnet. Außerdem wurde für Vegetationspunkte das Verhältnis des Medians zum Maximum berechnet und der Median der Vegetationspunkte dem Median der Bodenpunkte gegenübergestellt. Zusätzlich wurde auch der Median der Intensitäten im obersten Meter des Baumsegments dem Median aller anderen Punkte gegenübergestellt. Mit dem Einsatz der mittels Oberflächenmodell normalisierten Punkte

wurden die Punkte bis 3 Meter unter der Baumhöhe als Kronenoberfläche bezeichnet und alle Vegetationspunkte darunter als Kronenmitte. Mit diesen beiden Punktlisten wurde das Verhältnis der Punktzahl dargestellt und das Verhältnis des Intensitätsmedians der beiden Klassen. Außerdem wurden die Intensitätssummen des Kronendaches und des gesamten Baumes einander gegenübergestellt. Für weitere Merkmale wurde der Baum in zehn gleich große Höhengschichten unterteilt um für jede Höhengschicht jeweils den mittleren Abstand der Punkte zum Baumstamm, das Verhältnis der Punktzahl der Schicht zur Gesamtzahl und den Median der Intensität der Schicht zu berechnen. Zusätzlich wurde der Baum auch noch in Höhengschichten mit gleicher Punktzahl unterteilt. Die Höhe der einzelnen Perzentile wurde der Baumhöhe gegenübergestellt. Abschließend wurden noch Merkmale berechnet, für die das Punkattribut „Number of Returns“, welches für den jeweiligen Punkt angibt ob der ausgesandte Puls nur dieses oder noch weitere Echos erzeugt hat, verwendet wurde. Somit konnte die Anzahl der Single Echos mit der Anzahl der First Echos verglichen werden, ebenso auch die Anzahl der Middle und Last Returns mit der Anzahl der First Echos und die Anzahl der Mehrfachreflexionen mit der Anzahl der Einfachreflexionen. Mit dem Punkattribut „Return Number“, welches angibt das wievielte Echo des ausgesendeten Pulses der Punkt ist, wurde auch die Intensitätssumme der einzelnen Return Numbers der Summe aller Intensitäten gegenübergestellt. Auf die Merkmale, die schlussendlich auch für die Baumartenklassifizierung eingesetzt wurden, wird im Ergebniskapitel näher eingegangen.

3.3.3 Modell

Nicht nur die Klassifizierung mit unterschiedlicher Datenbasis wurde verglichen, sondern auch der Einsatz verschiedener Klassifizierungsmethoden wurde gegenübergestellt. Als beste Methode zur Unterscheidung der Hauptbaumartengruppen erwies sich der Random Forest Classifier. Durch Einsatz von Python Funktionen wie der Feature Score, Grid Search CV und Kreuzvalidierung wurden die berechneten Merkmale aus der Punktwolke und die Parameter des Modells gewählt.

3.4 Holzvorrat

Mit den Ergebnissen der Einzelbaumdetektion und der Baumartenklassifizierung wird anschließend noch das Holzvolumen des Baumstammes abgeleitet.

3.4.1 Datensatz

Zur Erstellung eines Modells zur Ableitung des Stammvolumens werden die Daten des Schulprojekts Waldcheck eingesetzt. Für diesen Teil sind nur die baumspezifischen Parameter relevant, die Baumposition ist nicht notwendig. Hier relevante Parameter sind die Baumart, die Höhe des Baumes und das im Zuge des Schulprojekts abgeleitete Holzvolumen der Einzelbäume.

3.4.2 Modell

Zur Abschätzung des Holzvorrates wird eine Exponentialfunktion für den Zusammenhang zwischen Baumhöhe und Holzvolumen abgeleitet. Es wird für jede der Hauptbaumartengruppen des Klassifizierungsmodells eine eigene Funktion abgeleitet.

3.5 Validierung

Wie auch in (Eysn et al. 2015), zum Vergleich verschiedener Algorithmen zur Baumdetektion, wurde auch hier ein Validierungsalgorithmus, bestehend aus Matching Algorithmus und Kennwert Berechnungen eingesetzt, um statistische Kennwerte der Validierung zu erhalten. Der Matching Algorithmus wurde in python implementiert. Darin wurde für jeden Referenzbaum in einem bestimmten Radius der beste Partner durch Vergleich der Höhendifferenzen und der Distanzen gesucht. Für Referenzdaten ohne Information der Baumhöhe wurde der Brusthöhendurchmesser eingesetzt, um Bäume mit größerem Stammumfang und somit im Normalfall auch größerer Baumhöhe immer vor den dünneren Bäumen als jeweiliges potenzielles Match zu testen. Für das Matching ohne bekannter Baumhöhe wurde ein festes Distanzkriterium von zwei Metern gewählt. Bei bekannter Baumhöhe wurden die Werte der maximalen Distanz und

des maximalen Höhenunterschiedes abhängig von der Baumhöhe bestimmt. Die Werte sind der nachstehenden Tabelle (3.3) zu entnehmen.

Tabelle 3.3: Werte der maximalen Distanz und der maximalen Höhendifferenz zwischen detektierten Bäumen und Referenzbäumen im Matching Prozess, mit der Baumhöhe $h(B)$

	$h(B) \leq 10$ m	$h(B) \leq 15$ m	$h(B) \leq 25$ m	$h(B) > 25$ m
Max. Distanz	2 m	3 m	3 m	3 m
Max. Höhendiff.	3 m	3 m	4 m	4 m

Der Algorithmus wurde sowohl zur Wahl der Parameter der Einzelbaumdetektion, als auch zur Bewertung des finalen Detektionsalgorithmus eingesetzt. Im Zuge des Validierungsprozesses wurden die detektierten Bäume in die beiden Gruppen „true positives“ und „false positives“ klassifiziert, bei den Referenzbäumen blieb nach der Eliminierung der Matches noch die Gruppe der „false negatives“ übrig. Für jedes Gebiet einzeln konnten anschließend qualitative und quantitative statistische Parameter berechnet werden.

3.5.1 Einzelbaumdetektion

Zur Validierung der Einzelbaumdetektion wurden die in Abbildung 3.3 dargestellten Gebiete der eigenständigen Datenaufnahme gewählt. Durch einen visuellen Abgleich der eingemessenen Baumpositionen mit den Orthophotos aus dem Jahr der ALS Datenaufnahme wurden Veränderungen des Waldes, zum Beispiel durch Abholzungen, ausgeschlossen.

Für jedes der sieben Validierungsgebiete wurden die statistischen Parameter Users accuracy, Producers accuracy und F1 Score berechnet. Die Users accuracy, auch Precision genannt, stellt das Verhältnis der Matches zu den detektierten Bäumen dar:

$$\text{Users accuracy} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \quad (3.2)$$

Die Producers accuracy, oder auch Recall, wird mithilfe der Anzahl der Referenzbäume berechnet, welche sich aus true positives und false negatives zusammensetzt, aus:

$$\text{Producers accuracy} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \quad (3.3)$$

Der F1-Score ist eine Genauigkeitsangabe des Testsets und eine Kombination der Users und Producers accuracy. Er berechnet sich aus:

$$\text{F1 Score} = 2 * \frac{\text{Users accuracy} * \text{Producers accuracy}}{\text{Users accuracy} + \text{Producers accuracy}} \quad (3.4)$$

Der F1-Score wird auch als harmonisches Mittel zwischen Precision und Recall bezeichnet.

3.5.2 Baumartenklassifizierung

Zur Validierung der Baumartenklassifizierung wurden die selbst eingemessenen Referenzdaten und die Daten des Schulprojektes Waldcheck eingesetzt. Mithilfe des bereits erwähnten Matching Algorithmus wurden die Referenzdaten in den sieben Gebieten nach allen Bäumen gefiltert, für die mit der implementierten Einzelbaumdetektion Partner zu den Referenzbäumen gefunden werden konnten. Somit blieben insgesamt 665 Einzelbäume, die für die Klassifizierung genutzt wurden. 20% der Eingangsdaten für das Klassifizierungsmodell dienten im Zuge der Modellwahl als Validierungsdatensatz. Es wurde jedoch kein fester Anteil der Daten gewählt, da durch den Einsatz der Kreuzvalidierung verschiedene Validierungsdatensätze verwendet werden.

4 ERGEBNISSE UND DISKUSSION

4.1 Einzelbaumdetektion

4.1.1 Validierung

Im Folgenden werden die Ergebnisse der Validierung der sieben Testgebiete dargestellt und analysiert. Nicht nur die Waldzusammensetzung, sondern auch die Neigung des Geländes hat eine große Auswirkung darauf, wie gut die Ergebnisse der Einzelbaumdetektion sind. In Tabelle 4.1 sind die Parameter der einzelnen Referenzgebiete zusammengefasst. Die Gebiete liegen auf unterschiedlicher Geländehöhe in einem Bereich zwischen 400 m und 1000 m ü.A. Innerhalb einige Gebiete gibt es auch große Schwankungen der Geländeneigung. Vor allem in den Gebieten 1 und 4 spielt die Neigung eine große Rolle bezüglich der Genauigkeit.

Tabelle 4.1: Eigenschaften der sieben Referenzgebiete

Gebiet	Waldzusammensetzung	Geländehöhe	Neigung in Grad	Gebietsnummer
Gebhardsberg	Fichtenwald, Laubbäume am Waldrand	571 m ü.A.	10°-35°	1
Kennelbach	Laubmischwald	418 m ü.A.	0°-36°	2
Wolfurt	Nadelwald, hauptsächlich Tannen	470 m ü.A.	1°-32°	3
Dünserberg	Mischwald	960 m ü.A.	20°-55°	4
Latschau	Mischwald	987 m ü.A.	11°-40°	5
Langen b. B.	Tannenwald	547 m ü.A.	1°-45°	6
Bregenz	Mischwald/Auwald	413 m ü.A.	0°-7°	7

Es gibt verschiedene Methoden, um die Genauigkeit der Detektion zu bewerten. Hier wurde ein Matching Prozess durchgeführt, jedoch ist es auch möglich, eine Aussage auf Gebietebene anstatt auf Einzelbaumebene zu treffen, indem nur die Anzahl der detektierten Bäume mit der wahren Anzahl an Bäumen in dem jeweiligen Gebiet verglichen wird. Der Score bezüglich der detektierten Anzahl der Bäume liegt für die Testgebiete zwischen 60% und 94%. Im Zuge der Parameterwahl der Einzelbaumdetektion wurde auch eine Überdetektion vermieden. In allen der sieben Testgebiete wurde eine geringere Anzahl an Bäumen detektiert als tatsächlich in den

Gebieten vorhanden ist. Dies ist vor allem damit zu begründen, dass nah beieinanderstehende oder abgeschirmte Bäume in dem Oberflächenmodell nicht detektiert werden können. In Tabelle 4.2 sind die empirischen Genauigkeitsangaben des Validierungsprozesses zusammengefasst. Auf den folgenden Seiten wird näher auf die Detektionsergebnisse und Schwachstellen der Detektion in den einzelnen Gebieten eingegangen.

Tabelle 4.2: Empirische Ergebnisse des Matching Prozesses der detektierten Bäume mit den Referenzdaten

Gebiet	Anz. Referenzbäume	Anz. detektierte Bäume	Anz. Matches	Users Accuracy	Producers Accuracy	F1-Score	Score-Anz. Bäume
1	46	28	23	82%	50%	62%	61%
2	83	53	29	50%	35%	41%	64%
3	40	24	21	88%	53%	66%	60%
4	71	48	35	73%	49%	59%	68%
5	67	50	43	86%	64%	74%	75%
6	87	81	63	78%	72%	75%	93%
7	50	47	28	60%	56%	58%	94%

In den folgenden Abbildungen werden die Detektionsergebnisse der Einzelbaumdetektion in den Testgebieten dargestellt. Es werden die Positionen der true positives, false positives, false negatives und der Referenzbäume dargestellt. Darunter liegt ein digitales Oberflächenmodell des jeweiligen Gebietes. Die Rasterdatensätze enthalten absolute Höhen in Metern über Adria.

Das in Vorarlberg am nördlichsten liegende Testgebiet (6) befindet sich im Bregenzerwald auf einer mittleren Höhe von 547 m. Der in Abbildung 4.1 sichtbare nordwestliche Teil des Gebietes ist sehr steil mit einer Steigung von bis zu 45°, im restlichen Gebietsteil gibt es keine markanten Geländesprünge. Es konnte mit dem F1-Score eine Genauigkeit von 75% erreicht werden.

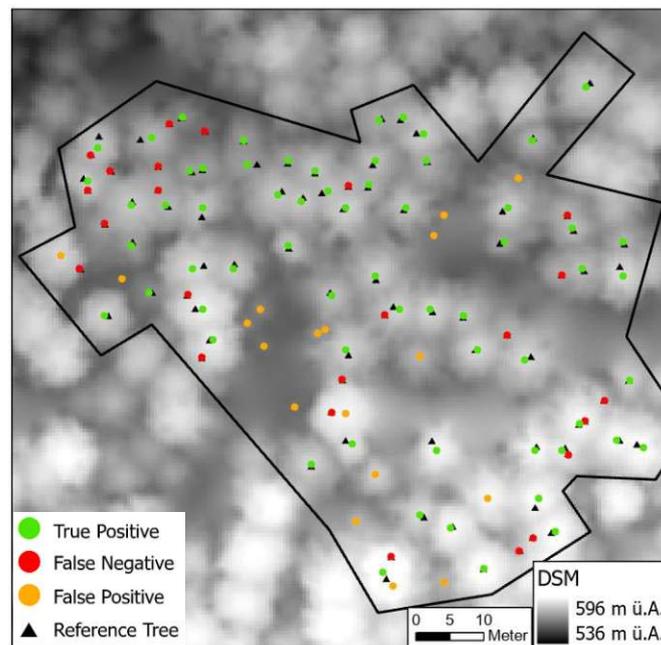


Abbildung 4.1: Ergebnisse des Matching Prozesses für Gebiet 6 mit einem F1-Score von 80%

In drei Bereichen sind Paare von false positives und false negatives zu sehen, die von dem Matching Algorithmus nicht als Paare erkannt wurden, jedoch durch visuelle Beurteilung als Matches anzunehmen sind, dies verbessert den F1-Score auf 80%. Die in orange dargestellten false positives können in einzelnen Fällen bei der Vermessung vergessen worden sein, in anderen Fällen kann es sich um nach oben wachsende Äste der Tannen handeln. Der Großteil dieser detektierten Bäume hat eine niedrige Baumhöhe, hier ist es möglich, dass die Bäume aufgrund des geringen BHD nicht eingemessen wurden. Andersrum handelt es sich auch bei den false negatives hauptsächlich um sehr dünne Bäume, diese stehen teilweise nahe bei hohen Bäumen, somit sind beim Übergang der Punktwolke zum Höhenmodell die Baumkronen nicht mehr voneinander zu trennen.

Im Gegensatz zu dieser sehr guten Detektionsrate stehen die Validierungen der Gebiete 2 und 7, welche beide einen sehr hohen Laubbaumanteil haben. Wie in Abbildung 4.2 und 10 ersichtlich, ist schon beim Betrachten der Oberflächenmodelle erkennbar, dass die Lokale Maxima Detektion aufgrund der verrauschten Kronenoberflächen schwieriger ist.

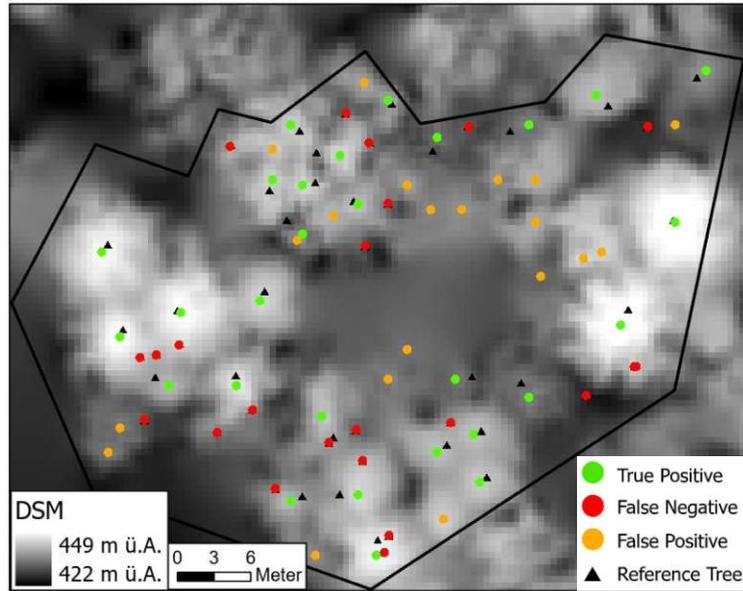


Abbildung 4.2: Gebiet 7 zeigt die Schwierigkeiten der Detektion in Laubwäldern

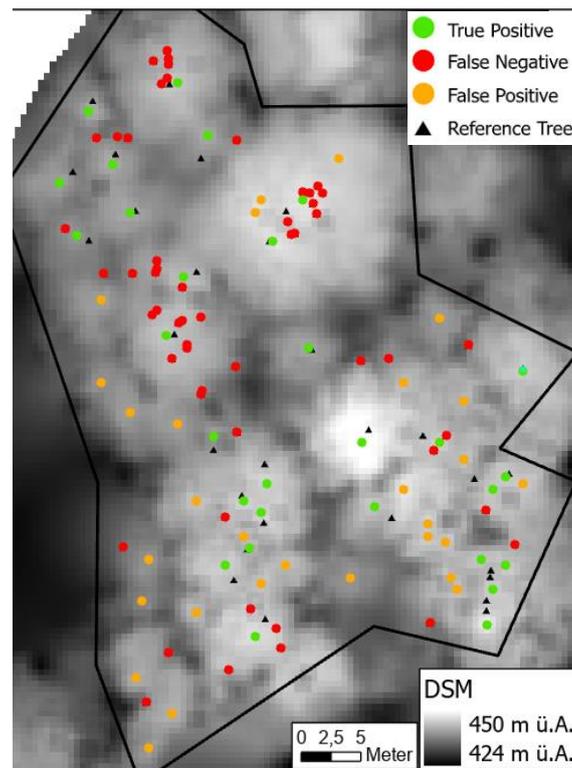


Abbildung 4.3: In Gebiet 2 kommen sehr viele enge Gruppierungen und stark geneigte Bäume vor

Es sind auch hier einzelne Paare zu finden, die nur aufgrund des Distanzkriteriums von zwei Metern im Matching Prozess nicht als Match klassifiziert wurden. Vorallem die Producers accuracy, welche angibt wie viele der Referenzbäume detektiert werden können, ist mit 35% und 56% niedrig. Bei Gebiet 2 sind im nördlichen Teil zwei Gruppierungen von Laubbäumen zu sehen, die in einer großen Baumkrone enden. Schon aufgrund der Auflösung des Oberflächenmodells sind nah beieinander stehende Bäume nicht mehr zu trennen. Mehr als 10 der als false negative klassifizierten Bäume sind laut Messprotokoll überschattet, außerdem wurde bei mehreren Bäumen eine starke Neigung dokumentiert. Die im westlichen Bildrand (Abbildung 4.3) sichtbaren false positives in orange stellen bei näherer Betrachtung die nordöstlich liegenden Bäume dar, die sich stark Richtung Gebietsgrenze neigen. Auch in Gebiet 7 aus Abbildung 4.2 sind einige der false negatives sehr dünne Bäume, die möglicherweise aufgrund der geringen Höhe nicht detektiert wurden, und wiederum die false positives hauptsächlich niedrige Bäume, die teilweise aufgrund des zu geringen BHDs nicht eingemessen wurden.

Im Gegensatz zu diesem Laubwald, steht das Gebiet 1, in Abbildung 4.4 zu sehen. Hier wurden hauptsächlich Fichten eingemessen. Am westlichen Gebietsrand, gleichzeitig

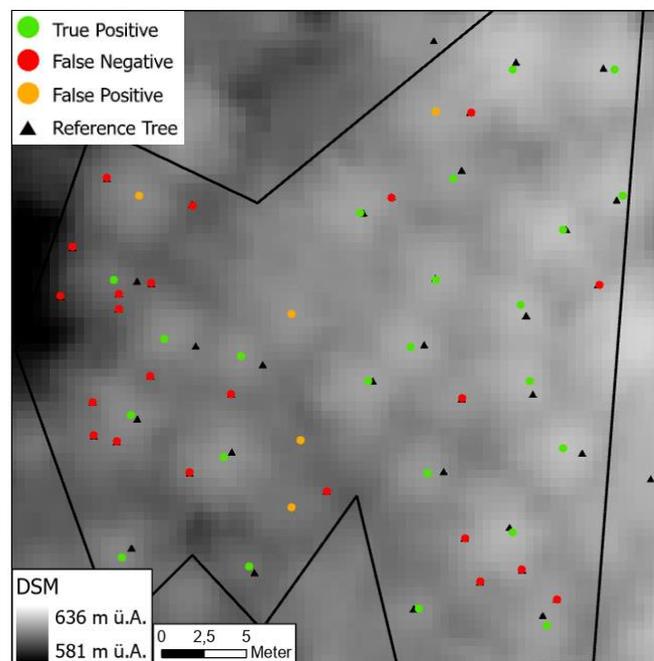


Abbildung 4.4: Gebiet 1, hauptsächlich Fichten auf einem Gelände mit einer Neigung von durchschnittlich 25-30°

auch der Waldrand, gibt es Gruppierungen von jungen Laubbäumen, die nicht detektiert werden konnten. Auch in diesem Gebiet kann der F1-Score aufgrund drei nicht automatisch gemachter Paare von false positives und false negatives auf 71% aufgebessert werden.

In Gebiet 3 aus Abbildung 4.5 wurden nur 40 Referenzbäume aufgenommen. Das arithmetische Mittel zwischen der Users und Producers accuracy liegt nach manueller Bearbeitung bei 75%. Vor allem Tannen und Laubbäume mit einem geringen Brusthöhendurchmesser zwischen 14 cm und 25 cm konnten nicht detektiert werden.

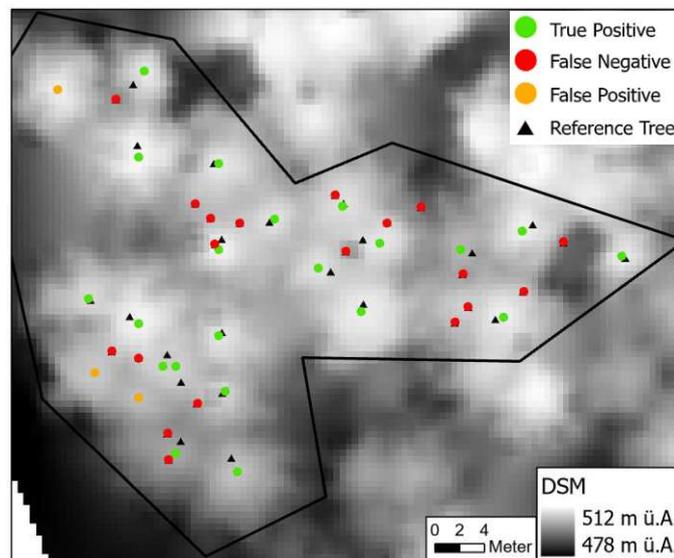


Abbildung 4.5: Gebiet 3, Bäume mit einem Brusthöhendurchmesser unter 40 cm und vor allem unter 25 cm sind schwer zu detektieren

Das Gebiet 4 ist mit einer Geländeneigung von bis zu 45° das steilste innerhalb der vorhandenen Testgebiete. Diese hohe Neigung spielt auch bei den Ergebnissen der Detektion eine Rolle. Sehr steiles Gelände kann dazu führen, dass die Bäume eine starke Neigung haben, obwohl diese vor Ort visuell nicht erkennbar ist. Dies führt dazu, dass aufgrund der geringen zulässigen Distanz des Matching Prozesses wenige Paare gefunden wurden. Bei dreien der nicht detektierten Bäume in Abbildung 4.6 wurde auch vor Ort eine starke Neigung festgestellt, einige niedrige Bäume sind überschattet.

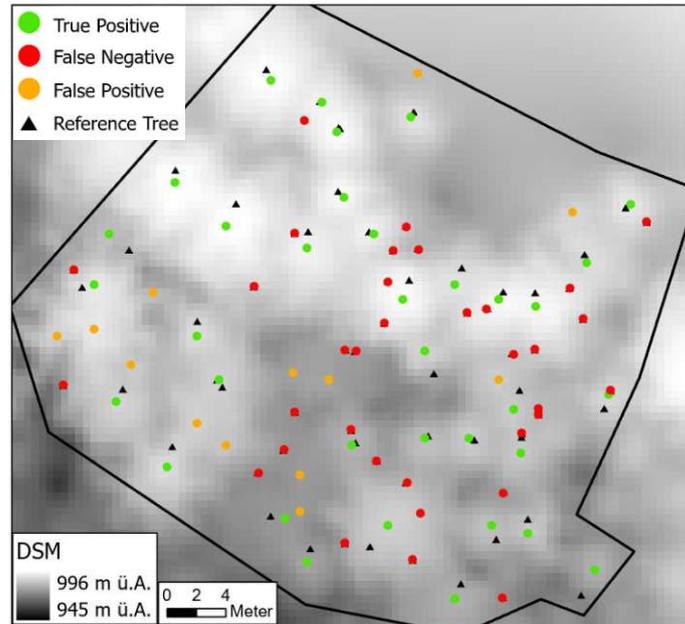


Abbildung 4.6: Referenzgebiet (4) auf 960 m.ü.A. mit sehr starker Geländeneigung

Die im automatisierten Matching Prozess gefundenen Paare wurden anschließend noch bearbeitet um von dem Algorithmus nicht gefundene Matches als solche zu klassifizieren. Dazu zählen Paare, die aufgrund einer Distanz von mehr als zwei Metern nicht erkannt wurden. Die maximal zulässige Distanz wurde mit zwei Metern sehr gering gewählt, um falsche Zuordnungen zu vermeiden, jedoch ist zu beachten, dass hier die Positionen der eingemessenen Baumstämme am Boden mit den Positionen der detektierten Baumspitzen aus der Punktwolke verglichen werden. Eine Differenz von zwei Metern entsteht bereits bei einer Neigung von 2.9 Grad, wenn eine Baumhöhe von 40 Metern angenommen wird. Insgesamt wurden in den sieben Gebieten, die zur Validierung der Einzelbaumdetektion genutzt wurden, 444 Bäume terrestrisch eingemessen. Davon wurden in denselben Gebieten 331 Bäume, oder 75% detektiert, davon wiederum waren nach visueller Beurteilung 282 korrekt detektierte Bäume und die übrigen 49 fälschlicherweise detektierte Bäume. Von den 444 Referenzbäumen konnten 159 nicht detektiert werden. Aus diesen Parametern wurden die in Tabelle 4.3 zusammengefassten empirischen Werte berechnet. Es ergab sich ein F1-Score der Detektion in allen Gebieten von 73%. Mit 85% ist die Users accuracy, das Verhältnis der korrekt detektierten Bäume zur Anzahl aller detektierten Bäume, deutlich höher, als die Producers accuracy von 64%, die die Zahl der korrekt detektierten Bäume mit der Anzahl der Referenzbäume ins Verhältnis stellt.

Tabelle 4.3: Zusammenfassung der empirischen Parameter aller Referenzgebiete und gesamt berechnet. Die Werte der true positives (TP), false negatives (FN) und false positives (FP) sind die Resultate des Matching Prozesses, daraus wurden die Parameter Users accuracy (UA), Producers accuracy (PA), F1-Score und die Detektionsrate berechnet

Gebiet	Ref. Bäume	Det. Bäume	TP	FN	FP	UA	PA	F1-Score	Det.-Rate
1	46	28	26	19	2	0.93	0.58	0.71	0.61
2	83	53	46	35	11	0.81	0.57	0.67	0.64
3	40	24	24	16	0	1.00	0.60	0.75	0.60
4	71	48	42	29	6	0.88	0.59	0.71	0.68
5	67	50	45	22	5	0.90	0.67	0.77	0.75
6	87	81	66	21	13	0.84	0.76	0.80	0.93
7	50	47	33	17	12	0.73	0.66	0.69	0.94
Gesamt	444	331	282	159	49	0.85	0.64	0.73	0.75

4.2 Baumartenklassifizierung

4.2.1 Validierung

Die Auswahl der Merkmale wurde mithilfe von Funktionen des python-Moduls scikit-learn durchgeführt. So konnte zusätzlich zur visuellen Beurteilung der Pairplots von jeweils zwei Merkmalen auch über die „Feature Importance“ beurteilt werden, wieviel das Merkmal zur Unterscheidung der Hauptbaumartengruppen beiträgt. Um anschließend unter den wichtigsten Baummerkmalen auch die richtige Anzahl an Merkmalen zur Klassifizierung zu bestimmen, um ein Overfitting des Modells zu vermeiden, wurden die Test- und Trainingsgenauigkeiten miteinander verglichen (Abbildung 4.7). Auch in der Grafik ist zu erkennen, dass eine Anzahl von fünf Merkmalen geeignet ist.

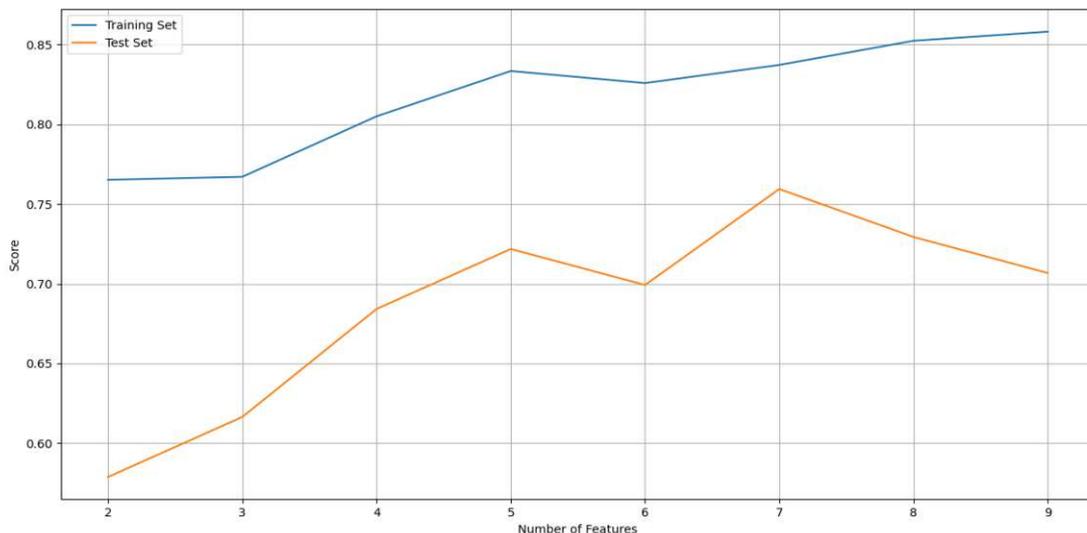


Abbildung 4.7: Vergleich der Trainings- und Test-Scores des Klassifizierungsmodells nach unterschiedlicher Anzahl an Merkmalen

Folgend werden die fünf Merkmale beschrieben, welche für die Baumartenklassifizierung eingesetzt und aus den Punktwolken der einzelnen Bäume berechnet wurden.

Für die Merkmalsberechnung wurden die Punkte des gesamten Baumsegments nach verschiedenen Teilmengen gefiltert. Im Folgenden stehen die Bezeichnungen H, I und P für Höhe, Intensität und Punkte. Für die Teilmengen werden die Abkürzungen Veg und G für Punkte der Klasse hohe Vegetation bzw. für Bodenpunkte eingesetzt. Wird keine der beiden Abkürzungen angegeben, ist die gesamte Punktwolke gemeint. Für zwei der

Merkmale ist der Abstand der Punkte zur Baumoberfläche relevant, deswegen wurden die Punktwolken nicht nur mittels dem Geländemodell normalisiert, sondern zusätzlich auch mit dem Oberflächenmodell. So konnten Merkmale berechnet werden, welche die Beziehung zwischen den Punkten bis zu 3 m unter der Waldoberfläche und den Punkten unterhalb und somit innerhalb des Waldes beschreiben. I_{K_M} beschreibt die Intensitäten der Bodenpunkte, indem das Verhältnis zwischen dem Median der Intensitäten aller Punkte bis zu 3 m unterhalb der Oberfläche und dem Median der Intensitäten der darunter liegenden Punkte berechnet wird. Ein ähnliches Merkmal stellt I_{K_all} dar, welches das Verhältnis der Summe der Intensitäten in den oberen 3 m und der Summe der Intensitäten von allen Punkten des Baumsegments darstellt. Das Merkmal I_{H_9} beschreibt die Eindringung in die Baumkrone. Dafür wird das Baumsegment der Punkte hoher Vegetation in zehn gleich große Höhenschichten unterteilt und der Median der Intensität der zweitobersten Schicht dargestellt. Auch für das Merkmal Z_{90} werden nur die Punkte hoher Vegetation herangezogen, um das 90%-Perzentil der Punktwolke der tatsächlichen Baumhöhe gegenüberzustellen. Für das fünfte Merkmal, mit I_{Veg_G} bezeichnet, wurde das Verhältnis des Medians der Intensitäten der Punkte hoher Vegetation und des Medians der Intensitäten der Bodenpunkte berechnet. Die Formeln (4.1) bis (4.5) zeigen die Berechnungen der Features.

$$I_{K_M} = \frac{\text{Median}(\text{Intensität}[H > DSM - 3])}{\text{Median}(\text{Intensität}[H < DSM - 3])} \quad (4.1)$$

$$I_{K_all} = \frac{\text{Summe}(\text{Intensität}[H > DSM - 3])}{\text{Summe}(\text{Intensität})} \quad (4.2)$$

$$I_{H_9} = \text{Median}(\text{Intensität}[8 * H/10 \leq H \leq 9 * H/10]) \quad (4.3)$$

$$Z_{90} = \frac{H_{0,90}[\mathit{Veg.}]}{\mathit{Max}[H]} \quad (4.4)$$

$$I_{\mathit{Veg}_G} = \frac{\mathit{Median}(\mathit{Intensität}[\mathit{Veg.}])}{\mathit{Median}(\mathit{Intensität}[\mathit{Boden}])} \quad (4.5)$$

Diese oben beschriebenen Merkmale der Punktwolke eines Baumes wurden in dem finalen Machine-Learning Modell eingesetzt. Durch Hinzunahme der Orthophotos konnte für jedes Baumsegment auch der NDVI (Normalized Difference Vegetation Index) berechnet werden. Durch die Erstellung unterschiedlicher Klassifizierungsmodelle wurde getestet, ob zusätzliche Daten wie Orthophotos oder True Orthophotos die Klassifizierungsergebnisse verbessern. Der Modellvergleich zeigt, dass die Hinzunahme weiterer Daten keine Verbesserung bringt. Somit wurde eine Baumartenklassifizierung basierend allein auf Laserscandaten für den gesamten Forstbestand der Illwerke vkw durchgeführt. Mit dem Modell wurde eine Genauigkeit von 71% erreicht.

Zusätzlich wurde auch ein Modell zur alleinigen Unterscheidung zwischen Laub- und Nadelbäumen erstellt. Hierbei ergab sich eine Genauigkeit von 89%.

Wie schon in Kapitel 3 erwähnt, wurden die Modellparameter mittels Kreuzvalidierung getestet. Der gesamte Datensatz wurde in 80% Trainingsdaten und 20% Validierungsdaten geteilt. Somit gab es fünf verschiedene Validierungsdatensätze, mit jeweils unterschiedlichen Trainingsdatensätzen. Die unterschiedlichen Stichproben zeigten bei gleichen Modellparametern sehr ähnliche Werte der statistischen Gütekriterien. Die folgende Tabelle 4.4 zeigt die statistischen Genauigkeitsparameter für einen der Validierungsdatensätze.

Tabelle 4.4: Ergebnisse der statistischen Gütekriterien eines Validierungsdatensatzes nach Anwendung des Klassifizierungsmodells.

	Precision	Recall	F1-Score	Support
Fichte	0,70	0,84	0,77	64
Laub	0,68	0,71	0,69	24
Nadel	0,00	0,00	0,00	5
Tanne	0,74	0,57	0,65	40
Accuracy			0,71	133
Weighted Avg	0,68	0,71	0,69	133

Aufgrund von nicht genügend vorhandener Datensätze der Klasse „andere Nadelbäume“ können diese mit dem Modell nicht von Tannen oder Fichten unterschieden werden. Die resultierenden Hauptbaumartengruppen des Validierungsdatensatzes nach der Klassifizierung können auch der Konfusionsmatrix in Tabelle 4.5 entnommen werden. Die Diagonale zeigt die Anzahl der korrekt klassifizierten Datensätze, hier 94 von 133. Es ist zu sehen, dass Tannen zwar sehr gut von Laubbäumen zu unterscheiden sind, jedoch wurden bei diesem Datensatz 39% fälschlicherweise als Fichten klassifiziert.

Tabelle 4.5: Konfusionsmatrix der Klassifizierung des Validierungsdatensatzes.

		Klassifizierung			
		Fichte	Laub	Tanne	Nadel
Referenz	Fichte	54	5	0	5
	Laub	6	17	0	1
	Nadel	2	1	0	2
	Tanne	15	2	0	23

Nach Anwendung des finalen Klassifizierungsmodells auf alle Referenzdaten ergab sich eine Genauigkeit von ebenfalls 71%. Aufgrund der geringen Anzahl an Referenzdaten wurde hier kein unabhängiger Validierungsdatensatz erstellt, sondern nur jener zum Testen des Modells. Für eine finale Validierung wäre ein neuer Referenzdatensatz außerhalb der Referenzgebiete sinnvoll.

4.3 Holzvorrat

Aufgrund des exponentiellen Zusammenhanges zwischen Baumhöhe und dem resultierenden Holzvolumen des Baumstammes wurde eine Exponentialfunktion zur Beschreibung des Holzvorrates gewählt. Formel (4.6) zeigt diesen Zusammenhang, hierbei steht HV für den Holzvorrat und $h(B)$ für die Baumhöhe. Die Funktionsparameter a und b wurden mithilfe der Referenzdaten bestimmt. Dafür wurde jene Exponentialkurve bestimmt, bei welcher die Summe der Abstandsquadrate der Datenpunkte zur Funktion minimal ist.

$$HV[m^3] = a * e^{h(B)*b} \quad (4.6)$$

Die resultierende Kurve als Funktion zwischen Holzvorrat und Höhe für Nadelbäume ist in Abbildung 4.8 dargestellt. Die eingesetzten Referenzbäume waren zwischen 5 und 38 m hoch mit einem Holzvorrat zwischen 0 - 7 m³. Mit dieser Funktion sollten nur Aussagen über Bäume mit einer Baumhöhe innerhalb dieses Bereiches getroffen werden.

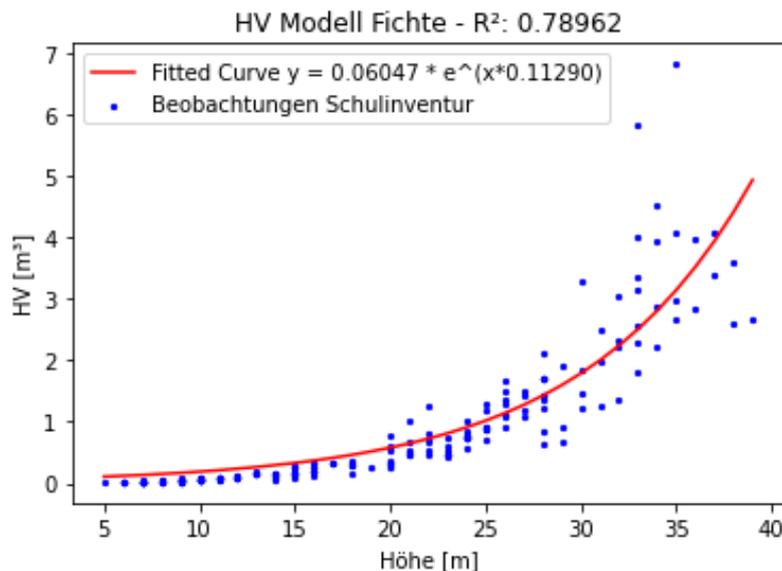


Abbildung 4.8: Funktion zur Beschreibung der Beziehung zwischen Holzvorrat und Baumhöhe für Nadelbäume

Der Determinationskoeffizient R^2 , als Maß zur Bestimmung der Anpassungsgüte, wurde wie folgt berechnet.

$$R^2 = 1 - \frac{SQR}{SQT} = \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - y(\text{mean}))^2} \quad (4.7)$$

Die Definition von R^2 in Formel (4.7) zeigt, dass es sich hierbei um das Verhältnis der Fehlerquadratsumme zur Gesamtvarianz der Variable handelt. Für das Holzvorratsmodell der Fichten ergab sich ein Determinationskoeffizient von ca. 79%. Für die beiden Modelle für Laubbäume und für Tannen ergaben sich Werte des R^2 Bestimmtheitsmaß von 80% und 85%. Zusätzlich wurde ein Modell aller Nadelbäume abgeleitet, hier ergab sich ein R^2 von 84%. Abbildung 4.9 zeigt den Vergleich der vier angepassten Funktionen. Obwohl für das Modell der Nadelbäume 207 Fichten, 167 Tannen und nur 3 weitere Nadelbäume zur Funktionsanpassung eingesetzt wurden, zeigt sich, dass die hohen HV Werte der Tannen im Bereich hoher Bäume großen Einfluss auf die resultierende Funktion haben. Dies liegt vor allem daran, dass es für Baumhöhen im Bereich zwischen 38 m und 45 m keine Datensätze für Fichten, sondern nur für Tannen und andere Nadelbäume gibt. Für das Modell der Laubbäume wurden 96 Referenzbäume verwendet.

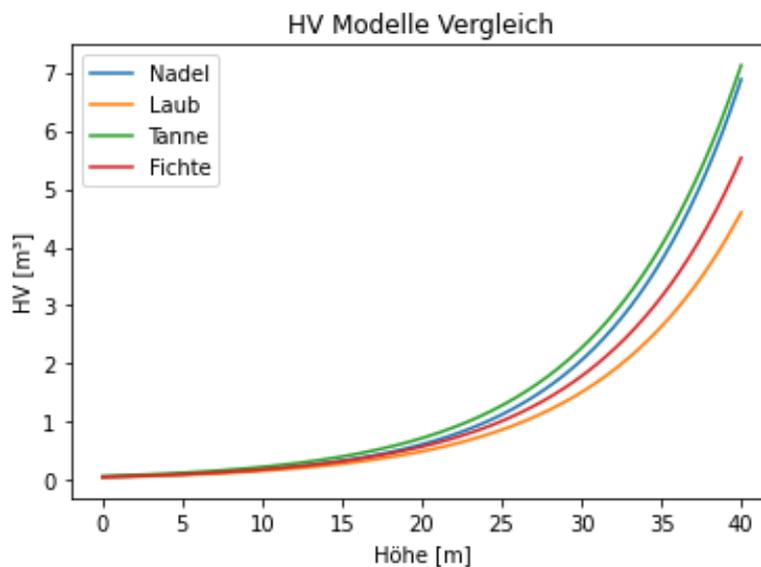


Abbildung 4.9: Vergleich der vier Holzvorratsmodelle

Dieses Modell kann für Gebiete mit Waldstrukturen ähnlich jener der Referenzgebiete zur Abschätzung des Holzvorrates eingesetzt werden. Für großflächige Holzvorrats-schätzungen kann das Modell durch neue Referenzdaten in den jeweiligen Gebieten verbessert werden. Außerdem ist es sinnvoll, durch Stichproben die Abschätzung im Nachhinein zu validieren.

4.4 Diskussion der Ergebnisse

Die Einzelbaumdetektion in einem Oberflächenmodell anstatt direkt in der Punktwolke bringt immer einen Informationsverlust mit sich. Ob durch eine Detektion direkt in der Punktwolke tatsächlich bessere Ergebnisse erreicht werden können, ist jedoch von der Punktdichte der ALS Daten und der Waldzusammensetzung abhängig. Bei einer Rasterisierung der Daten gehen Informationen der unteren Schichten des Waldes verloren, wodurch überschattete Bäume nichtmehr detektiert werden können. Dasselbe gilt für Gruppierungen von Bäumen. Abbildung 4.10 zeigt einen Ausschnitt der Punktwolke eines Referenzgebietes, in welchem sich drei Doppelbäume und fünf weitere Bäume befinden. Die Baumstämme verschwinden alle in einer gemeinsamen Baumkrone, hier konnten in der Analyse fünf Baumspitzen detektiert werden.

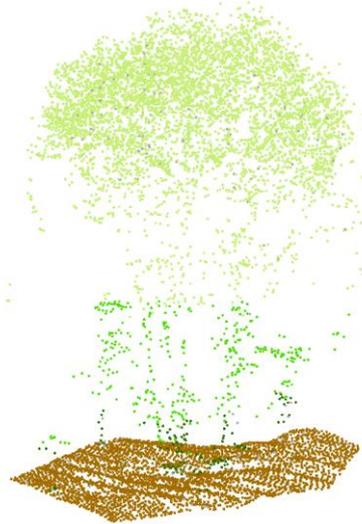


Abbildung 4.10: Punktwolke einer Gruppierung von mehreren Bäumen mit einer gemeinsamen Baumkrone

Bei solchen Gruppierungen werden auch Algorithmen zur Detektion in der Punktwolke an ihre Grenzen stoßen. Auch Bäume mit einer starken Neigung können Fehlerquellen sein. Vor allem in steilem Gelände können dadurch große Abweichungen zwischen den abgeleiteten Höhen und den tatsächlichen Baumhöhen entstehen. Da für die in dieser Analyse eingesetzten Validierungsdatensätze keine Höhen der Referenzbäume zur Verfügung standen, konnte keine Aussage über die Höhengenaugigkeit getroffen werden.

In einem der sieben Referenzgebiete wurde zusätzlich ein Vergleich der Einzelbaumdetektion mit einer punktwolkenbasierten Auswertung durchgeführt. Dafür

wurde an der TU Wien ein Algorithmus eingesetzt, welcher abhängig von der Baumhöhe unterschiedliche Suchradien nutzt und zusätzlich die Punktdichte untersucht um aus der Baumkrone herausragende Äste zu erkennen und als Maxima zu entfernen. Abbildung 4.11 zeigt den Vergleich der beiden Detektion.

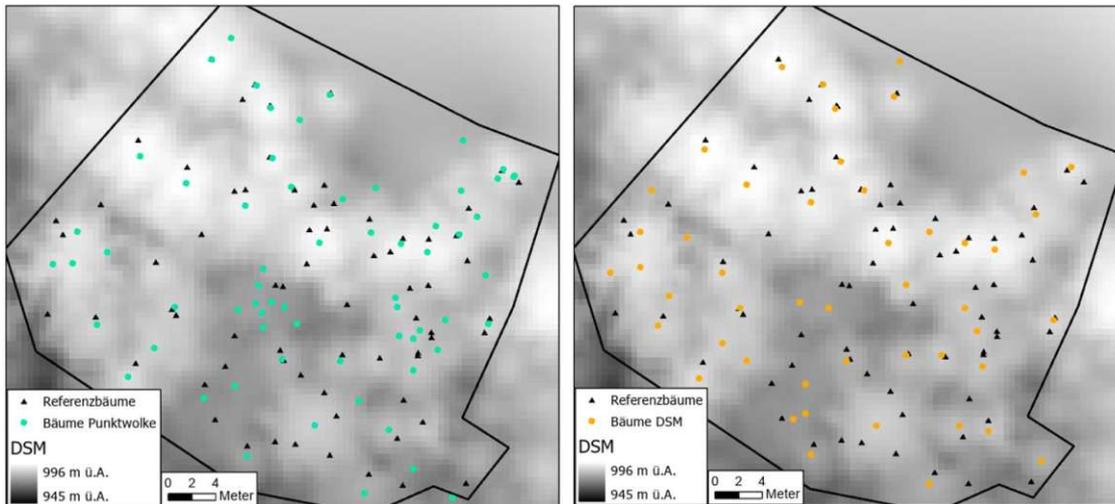


Abbildung 4.11: Vergleich der Einzelbaudetektion der Analyse (rechts) mit einer Detektion direkt in der Punktwolke (links)

In einzelnen Bereichen des Referenzgebietes konnten mit beiden Analysen dieselben Bäume detektiert werden, es zeigt sich jedoch, dass mittels der Maxima Suche direkt in der Punktwolke mehr Bäume detektiert wurden. Das hier gewählte Testgebiet ist aufgrund der Zusammensetzung des Waldes und dem sehr steilen Gelände für eine Rasteranalyse schwierig.

Um die Performance der Einzelbaumdetektion im Vergleich zu anderen entwickelten Detektionsmethoden zu beurteilen, wurde ein Vergleich mit den Ergebnissen aus Eysn et al. (2015) durchgeführt. In der Benchmark wurden verschiedene Methoden miteinander verglichen, darunter sieben rasterbasierte Detektionen und eine punktwolkenbasierte Baumdetektion. Auch in diesem Methodenvergleich zeigte sich, dass die Detektion in der Punktwolke, in den unteren Waldschichten bessere Ergebnisse bringt. Insgesamt brachte diese Methode jedoch nicht die besten Ergebnisse. Das beste Verhältnis zwischen einer hohen Matching Rate und einer niedrigen Commission Rate wurde mit einer Maxima Suche auf einem Oberflächenmodell mit einem variablen Suchfenster erreicht. Die in der Studie berechneten Parameter der Omission Rate von 59% und der Commission Rate von 9% können zur Vergleichbarkeit umgerechnet werden in eine Users accuracy von 91% und

einer Producers accuracy von 59%. Daraus ergibt sich ein F1-Score von 71%. In der in dieser Arbeit umgesetzten Analyse ergaben sich diese Werte für alle Testgebiete gesamt zu 85%, 64% und 73%. Somit zeigt sich, dass die erzielten Ergebnisse mit den Genauigkeiten anderer Methoden vergleichbar sind.

Die digitale Forstanalyse wurde innerhalb der Illwerke vkw bereits für Analysen betreffend dem Gefahrenpotential eingesetzt. Ein Beispiel ist eine Auswertung in der Nähe einer Seilbahn. Hierfür wurde für die detektierten Bäume auf Basis der Baumhöhe und deren dreidimensionalem Abstand zur Seillinie das Gefahrenpotential abgeleitet. Die Information des Holzvorrates der zu fällenden Bäume führte zu einer Vereinfachung in der Planung. Für die aufgrund des Gefahrenpotentials zu fällenden Bäume wurde im Zuge der Analyse ein Holzvorrat von 123 m³ berechnet. Bei der anschließend durchgeführten Holzung ergaben sich 118 m³ Sägerundholz. Es ist zu beachten, dass sich der berechnete Holzvorrat aufgrund der verwendeten ALS Daten auf die Baumhöhen von 2017 bezieht. Auch der Verlust durch die Holzung müsste für eine direkte Vergleichbarkeit noch einberechnet werden. Trotz der limitierenden Faktoren zur Vergleichbarkeit zeigte sich, dass die Holzvorratsabschätzung in diesem Gebiet gute Ergebnisse brachte.

Für eine Weiterführung der Analyse wäre eine Zuwachsabschätzung zwischen dem Aufnahmezeitpunkt der Daten und dem Analysezeitpunkt erstrebenswert.

Für eine Weiterverarbeitung der digitalen Forstanalyse, wäre es auch interessant zu testen, ob durch eine Klassifizierung mittels Punktwolkenmerkmalen auch weitere Baumarten voneinander unterschieden werden können. Dafür ist eine große Anzahl weiterer Referenzbäume nötig. Mit den hier vorhandenen Referenzdaten von 665 Einzelbäumen, war das Testen von weiteren Unterscheidungen nicht möglich.

Eine weitere Optimierungsmöglichkeit besteht bezüglich der Laufzeit der Analyse. Der Fokus wurde in dieser Arbeit auf die Analyseergebnisse und nicht auf die Optimierung der Datenverarbeitung gelegt. Vor allem die Extraktion der Punktwolken der Einzelbäume und die anschließende Merkmalsberechnung daraus führten zu einer langen Programmlaufzeit. Durch zukünftige neue Verarbeitungsmöglichkeiten von Punktwolken in den eingesetzten Softwarepaketen ist hier eine Optimierung denkbar.

5 ZUSAMMENFASSUNG

In der vorliegenden Arbeit wurde eine digitale Forstanalyse auf Einzelbaumniveau erstellt. Es wurden die Parameter Baumposition, Höhe, Hauptbaumartengruppe und Holzvorrat der Bäume bestimmt. Zur Umsetzung der Analyse wurden auch eigene Datensätze in sieben Referenzgebieten aufgenommen. Datenbasis der Analyse sind Airborne Laser Scanning Punktwolken. In daraus abgeleiteten Höhenmodellen werden mittels Lokaler Maximum Suche die Baumspitzen detektiert. Für die Einzelbaumdetektion wird ein variables Suchfenster eingesetzt, dessen Größe abhängig von der Geländehöhe und der Baumhöhe ist. Zur Parameterwahl und auch für die anschließende Validierung wird ein Validierungsalgorithmus, bestehend aus Matching Prozess und Kennwert Berechnungen, eingesetzt. Mittels eines abgeleiteten normalisierten Oberflächenmodells werden die Baumhöhen der detektierten Baumpositionen bestimmt. Auf Basis der Punktwolke wird anschließend eine Klassifizierung zwischen den Hauptbaumartengruppen Fichte, Tanne, Laubbaum und anderen Nadelbäumen durchgeführt. Dafür werden aus den ALS Punkten der Einzelbäume Merkmale berechnet. Für die Abgrenzung der Punktwolken der einzelnen Bäume werden Zylinder mit einem Radius von drei Metern mit der Baumposition als Mittelpunkt eingesetzt. Für die Berechnungen werden Attribute wie die Koordinaten, Intensität, Klassifizierung, Nummer der Echos und Anzahl der Echos verwendet, um die Geometrie- und Intensitätsverteilungen zu beschreiben. Auf Basis dieser Merkmale und den Referenzdaten werden durch den Einsatz von maschinellem Lernen die Hauptbaumartengruppen unterschieden. Dafür wird eine Klassifizierung mittels Random Forest Classifier in Python durchgeführt. Mit den Ergebnissen der Einzelbaumdetektion und der Baumartenklassifizierung wird anschließend das Holzvolumen des Baumstammes abgeleitet. Hierfür dienen Referenzdaten eines Schulprojektes, bei dem zusätzlich zu der Baumart und der Höhe, auch das Holzvolumen erfasst wurde. Es wird mittels Kurvenanpassung eine Exponentialfunktion für den Zusammenhang zwischen Baumhöhe und Holzvolumen abgeleitet.

Bei der Einzelbaumdetektion wurden in den unterschiedlichen Referenzgebieten Genauigkeiten zwischen 67% und 80% erreicht. Es zeigten sich deutliche Unterschiede aufgrund der Waldzusammensetzung. Mit der anschließenden Unterteilung der Einzelbäume in die Kategorien Fichte, Tanne, Laubbaum und andere Nadelbäume konnte

eine Klassifizierungsgenauigkeit von 71% erreicht werden. Zusätzlich wurde ein Modell zur alleinigen Unterscheidung zwischen Laub- und Nadelbäumen erstellt, wobei eine Genauigkeit von 89% erreicht wurde. Die Ergebnisse der Abschätzung des Holzvolumens werden mittels dem R^2 Bestimmtheitsmaß beurteilt. Der Determinationskoeffizient R^2 beschreibt die Anpassungsgüte der Referenzdaten an das Modell. Für die Holzvorratsmodelle der Fichten, Tannen und Laubbäume wurden R^2 Werte von 79%, 80% und 85% erreicht.

Die Analyse dieser Arbeit dient der flächendeckenden Abschätzung des Forstbestandes und wird auch für die Identifizierung des Gefahrenpotentials in der Nähe von Anlagen der kritischen Infrastruktur und anderen technischen Anlagen eingesetzt. Die beschriebene Analyse wurde mittels Python und dem Einsatz der ArcGIS API for Python automatisiert um eine wiederverwendbare Analyse zu erzeugen. In dieser Arbeit wurden ALS Daten aus dem Jahr 2017 eingesetzt. Die Ergebnisse der durchgeführten Analyse können nach einer erneuten ALS Befliegung, welche in Vorarlberg alle sechs Jahre durchgeführt wird, mit Ergebnissen einer neuen Analyse verglichen werden.

Wünschenswert ist auch ein Vergleich der Ergebnisse mit Daten der herkömmlichen Forstinventur. Bei der Einzelbaumdetektion zeigte sich ganz klar der Nachteil der durchgeführten Einzelbaumdetektion für die Detektion von Bäumen in unteren Schichten des Waldes. Ein Vorteil gegenüber herkömmlichen Forstinventurdaten ist jedoch die flächendeckende Datenverfügbarkeit auf Einzelbaumniveau. Für die Erstellung einer digitalen Forstinventur sind immer manuelle Daten für die Kalibrierung notwendig, weshalb sie die herkömmliche Forstinventur nicht ersetzen sondern ergänzen soll.

Auch die hier erstellten Modelle für die Baumartenklassifizierung und die Holzvorratsschätzung könnten durch zusätzliche Daten noch verbessert werden.

Durch die hier vorgestellte automatische und wiederverwendbare Analyse wird eine Digitalisierung des Forstbestandes der Illwerke vkw AG geschaffen und somit kann das Management und die Erhaltung der Wälder verbessert werden.

6 ABBILDUNGSVERZEICHNIS

Abbildung 2.1: Aufbau eines Lidar-Systems. Abbildung aus Höfle und Rutzinger (2011)	4
Abbildung 2.2: Die häufigsten Schlüsselwörter in Artikeln bezogen auf Technologien zur Datengewinnung für die Kartierung von Wäldern im Zeitraum 2010 bis 2020. Abbildung aus Nitoslowski et al. (2021)	14
Abbildung 2.3: Datenstruktur des Entscheidungsbaumes zur Klassifikation aus Frochte (2021)	19
Abbildung 3.1: Vergleich von Orthophoto (links) und True Orthophoto (rechts), mit den detektierten Baumspitzen in Rot dargestellt.....	23
Abbildung 3.2: Aufnahmegebiete der Schulinventur in Grün. Orthophoto Land Vorarlberg.....	24
Abbildung 3.3: In den orangemarkierten Gebieten wurden eigenständig Referenzdaten aufgenommen. Orthophoto Land Vorarlberg	25
Abbildung 3.4: Beispiele von Kreis- und Rechtecknachbarschaften für die lokale Maxima Suche (Funktionsweise von „Focal Statistics“—ArcGIS Pro Dokumentation, o. D.) .	27
Abbildung 3.5: Punktwolken der drei Baumarten Fichte, Tanne, Laubbaum (von links nach rechts).....	30
Abbildung 4.1: Ergebnisse des Matching Prozesses für Gebiet 6 mit einem F1-Score von 80%.....	37
Abbildung 4.2: Gebiet 7 zeigt die Schwierigkeiten der Detektion in Laubwäldern.....	38
Abbildung 4.3: In Gebiet 2 kommen sehr viele enge Gruppierungen und stark geneigte Bäume vor	38
Abbildung 4.4: Gebiet 1, hauptsächlich Fichten auf einem Gelände mit einer Neigung von durchschnittlich 25-30°	39
Abbildung 4.5: Gebiet 3, Bäume mit einem Brusthöhendurchmesser unter 40 cm und vor allem unter 25 cm sind schwer zu detektieren.....	40
Abbildung 4.6: Referenzgebiet (4) auf 960 m.ü.A. mit sehr starker Geländeneigung...	41
Abbildung 4.7: Vergleich der Trainings- und Test-Scores des Klassifizierungsmodells nach unterschiedlicher Anzahl an Merkmalen	43
Abbildung 4.8: Funktion zur Beschreibung der Beziehung zwischen Holzvorrat und Baumhöhe für Nadelbäume.....	47

Abbildung 4.9: Vergleich der vier Holzvorratsmodelle	48
Abbildung 4.10: Punktwolke einer Gruppierung von mehreren Bäumen mit einer gemeinsamen Baumkrone.....	49
Abbildung 4.11: Vergleich der Einzelbaudetektion der Analyse (rechts) mit einer Detektion direkt in der Punktwolke (links)	50

7 TABELLENVERZEICHNIS

Tabelle 3.1: Parameter und daraus berechnete Kenngrößen des ALS Fluges, aus Würländer (2019b)	22
Tabelle 3.2: Kernelgrößen der Einzelbaumdetektion abhängig von der Geländehöhe (Zellenwert im DTM) und der Baumhöhe (Zellenwert im nDSM). Ein $n \times n$ Kernel, bedeutet eine Maxima Suche innerhalb einer Rechteknachbarschaft, bei einer Kreisnachbarschaft wird der Radius angegeben, beides wird jeweils in Zelleneinheiten angegeben	28
Tabelle 3.3: Werte der maximalen Distanz und der maximalen Höhendifferenz zwischen detektierten Bäumen und Referenzbäumen im Matching Prozess, mit der Baumhöhe $h(B)$	33
Tabelle 4.1: Eigenschaften der sieben Referenzgebiete	35
Tabelle 4.2: Empirische Ergebnisse des Matching Prozesses der detektierten Bäume mit den Referenzdaten	36
Tabelle 4.3: Zusammenfassung der empirischen Parameter aller Referenzgebiete und gesamt berechnet. Die Werte der true positives (TP), false negatives (FN) und false positives (FP) sind die Resultate des Matching Prozesses, daraus wurden die Parameter Users accuracy (UA), Producers accuracy (PA), F1-Score und die Detektionsrate berechnet.....	42
Tabelle 4.4: Ergebnisse der statistischen Gütekriterien eines Validierungsdatensatzes nach Anwendung des Klassifizierungsmodells.	46
Tabelle 4.5: Konfusionsmatrix der Klassifizierung des Validierungsdatensatzes.	46

8 LITERATURVERZEICHNIS

- Amiri, Nina; Krzystek, Peter; Heurich, Marco; Skidmore, Andrew (2019):
Classification Of Tree Species As Well As Standing Dead Trees Using Triple
Wavelength Als In A Temperate Forest. In: *Remote Sensing* 11 (22), S. 2614. Doi:
10.3390/Rs11222614.
- Apostol, Bogdan; Petrilă, Marius; Lorent, Adrian; Ciceu, Albert; Gancz, Vladimir;
Badea, Ovidiu (2020): Species discrimination and individual tree detection for predicting
main dendrometric characteristics in mixed temperate forests by use of airborne laser
scanning and ultra-high-resolution imagery. In: *The Science of the total environment* 698,
S. 134074. DOI: 10.1016/j.scitotenv.2019.134074.
- Avdeeva, I.; Belolipezkaj, A.; Golovina, T.; Polyaniin, A. (2020): Digitalization as a new
factor of sustainable development of the Russian forest industry complex. In: *IOP Conf.
Ser.: Earth Environ. Sci.* 595 (1), S. 12051. DOI: 10.1088/1755-1315/595/1/012051.
- Bishop, Christopher M. (2006): Pattern recognition and machine learning. New York:
Springer (Information science and statistics).
- Dostálová, Alena; Lang, Mait; Ivanovs, Janis; Waser, Lars T.; Wagner, Wolfgang (2021):
European Wide Forest Classification Based on Sentinel-1 Data. In: *Remote Sensing* 13
(3), S. 337. DOI: 10.3390/rs13030337.
- Dubayah, Ralph O.; Drake, Jason B. (2000): Lidar Remote Sensing for Forestry
Applications.
- Eysn, Lothar; Hollaus, Markus; Lindberg, Eva; Berger, Frédéric; Monnet, Jean-Matthieu;
Dalponte, Michele et al. (2015): A Benchmark of Lidar-Based Single Tree Detection
Methods Using Heterogeneous Forest Data from the Alpine Space. In: *Forests* 6 (12), S.
1721–1747. DOI: 10.3390/f6051721.
- Freer-Smith, P. H.; Broadmeadow, M. S. J.; Lynch, J. M. (2007): Forestry and climate
change. Wallingford: CABI.
- Frochte, Jörg (2021): Maschinelles Lernen. Grundlagen und Algorithmen in Python. 3.,
überarbeitete und erweiterte Auflage. München: Hanser (Plus.Hanser-Fachbuch).

Höfle, Bernhard; Rutzinger, Martin (2011): Topographic airborne LiDAR in geomorphology: A technological perspective. In: *Zeit fur Geo Supp* 55 (2), S. 1–29. DOI: 10.1127/0372-8854/2011/0055S2-0043.

Hollaus, M.; Dorigo, W.; Wagner, W.; Schadauer, K.; Höfle, B.; Maier, B. (2009a): Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data. In: *International Journal of Remote Sensing* 30 (19), S. 5159–5175. DOI: 10.1080/01431160903022894.

Hollaus, M.; Mücke, W.; Höfle, B.; Dorigo, W.; Pfeifer, N.; Wagner, W. (2009b): Tree species classification based on full-waveform airborne laser scanning data.

Hollaus, M.; Wagner, W.; Eberhöfer, C.; Karel, W. (2006): Accuracy of large-scale canopy heights derived from LiDAR data under operational constraints in a complex alpine environment. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 60 (5), S. 323–338. DOI: 10.1016/j.isprsjprs.2006.05.002.

Hollaus, M.; Wagner, W.; Schadauer, K.; Maier, B.; Gabler, K. (2009c): Growing stock estimation for alpine forests in Austria: a robust lidar-based approach.

Julian Evens (2001): An overview of forest science. Oxford: Blackwell Science (The forests handbook / ed. by Julian Evans, Vol. 1).

Kaushik, Kavita; Kumar, Deepak (2020): Forest Disturbance Mapping with Microwave Remote Sensing. In: *Journal of Landscape Ecology* 13 (2), S. 95–128. DOI: 10.2478/jlecol-2020-0012.

Kislov, Dmitry E.; Korznikov, Kirill A. (2020): Automatic Windthrow Detection Using Very-High-Resolution Satellite Imagery and Deep Learning. In: *Remote Sensing* 12 (7), S. 1145. DOI: 10.3390/rs12071145.

Lindberg, Eva; Eysn, Lothar; Hollaus, Markus; Holmgren, Johan; Pfeifer, Norbert (2014): Delineation of Tree Crowns and Tree Species Classification From Full-Waveform Airborne Laser Scanning Data Using 3-D Ellipsoidal Clustering. In: *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 7 (7), S. 3174–3181. DOI: 10.1109/JSTARS.2014.2331276.

Markus Melin; Aurélie C. Shapiro; Paul Glover-Kapfer (2017): Remote Sensing: LiDAR. In: *WWF Conservation Technology* (Series 1(3)).

Mustafic, Sead; Schardt, Matthias (2019): Deep Learning-basierte Baumartenklassifizierung auf Basis von ALS-Daten.

Nevalainen, Olli; Honkavaara, Eija; Tuominen, Sakari; Viljanen, Niko; Hakala, Teemu; Yu, Xiaowei et al. (2017): Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. In: *Remote Sensing* 9 (3), S. 185. DOI: 10.3390/rs9030185.

Nitoslawski, S. A.; Wong-Stevens, K.; Steenberg, J. W. N.; Witherspoon, K.; Nesbitt, L.; van den Konijnendijk Bosch, C. C. (2021): The Digital Forest: Mapping a Decade of Knowledge on Technological Applications for Forest Ecosystems. In: *Earth's Future* 9 (8). DOI: 10.1029/2021EF002123.

Piermattei, Livia; Hollaus, Markus; Milenković, Milutin; Pfeifer, Norbert; Quast, Raphael; Chen, Yuwei et al. (2017): An Analysis of Ku-Band Profiling Radar Observations of Boreal Forest. In: *Remote Sensing* 9 (12), S. 1252. DOI: 10.3390/rs9121252.

Pitkänen, Timo P.; Sirro, Laura; Häme, Lauri; Häme, Tuomas; Törmä, Markus; Kangas, Annika (2020): Errors related to the automatized satellite-based change detection of boreal forests in Finland. In: *International Journal of Applied Earth Observation and Geoinformation* 86, S. 102011. DOI: 10.1016/j.jag.2019.102011.

Pollanschütz, J. (1974): Formzahlenfunktionen der Hauptbaumarten. In: *Österreichische Allgemeine Forstzeitung* 85 (12).

Pourghasemi, Hamid Reza; Gayen, Amiya; Lasaponara, Rosa; Tiefenbacher, John P. (2020): Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. In: *Environmental research* 184, S. 109321. DOI: 10.1016/j.envres.2020.109321.

Shardt, Matthias; Ziegler, Michaela; Wimmer, Andreas; Wack, Roland; Hyyppä, Juha (2002): Assessment of Forest Parameters by Means of Laserscanning.

Windrim, Lloyd; Bryson, Mitch (2020): Detection, Segmentation, and Model Fitting of Individual Tree Stems from Airborne Laser Scanning of Forests Using Deep Learning. In: *Remote Sensing* 12 (9), S. 1469. DOI: 10.3390/rs12091469.

Würländer, Dipl. Ing. Roland (2019a): Airborne Laserscanner Daten Vorarlberg 2017 - Projektabschlussbericht.

Würländer, Roland (2019b): Airborne Laserscanner Daten Vorarlberg 2017 Projektabschlussbericht.

Yan, Wanqian; Guan, Haiyan; Cao, Lin; Yu, Yongtao; Li, Cheng; Lu, JianYong (2020): A Self-Adaptive Mean Shift Tree-Segmentation Method Using UAV LiDAR Data. In: *Remote Sensing* 12 (3), S. 515. DOI: 10.3390/rs12030515.

Bundeforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft (BFW). (2022, 11. Januar). Österreichische Waldinventur. BFW. Abgerufen am 24. September 2022, von <https://geo.bfw.ac.at/oewi/#/>

Bundeforschungszentrum für Wald (BFW). BFW. Abgerufen am 27.02.2022, von https://bfw.ac.at/rz/bfwcms.web_print?dok=7229

Funktionsweise von „Focal Statistics“—ArcGIS Pro | Dokumentation. (o. D.). Abgerufen am 9. Oktober 2022, von <https://pro.arcgis.com/de/pro-app/latest/tool-reference/spatial-analyst/how-focal-statistics-works.htm>