UNIVERSITÄT FÜR BODENKULTUR WIEN
University of Natural Resources
and Life Sciences, Vienna

Masterarbeit

Bestandesvorratsermittlung mittels Airborne Laserscanning Daten

verfasst von

Bernhard Martin GILL, BSc.

im Rahmen des Masterstudiums

Forstwissenschaften

zur Erlangung des akademischen Grades

Diplom-Ingenieur

Litschau, November 2022

Betreut von:

Univ. Prof. Dr.rer.nat. Clement Atzberger

Mitbetreut von:

DI Dr. Markus Immitzer, MSc.
Dr. Markus Engel, MSc.

Department für Raum, Landschaft und Infrastruktur
Institut für Geomatik

Mitbetreut von:

DI Dr. Markus Immitzer, MSc.
Universität für Bodenkultur Wien
Institut für Geomatik
Peter-Jordan Straße 82
1190 Wien Austria

Dr. Markus Engel, MSc.
Landesanstalt für Wald und Forstwirtschaft
Abteilung 2 Boden und Klima
Hans-Carl-von-Carlowitz-Platz 1
85354 Freising Germany

Eidesstattliche Erklärung

Ich versichere an Eides statt, dass ich diese Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Gedanken, die im Wortlaut oder in grundlegenden Inhalten aus unveröffentlichten Texten oder aus veröffentlichter Literatur übernommen wurden, sind ordnungsgemäß gekennzeichnet, zitiert und mit genauer Quellenangabe versehen.

Die vorliegende Arbeit wurde bisher weder ganz noch teilweise in gleicher oder ähnlicher Form an einer Bildungseinrichtung als Voraussetzung für den Erwerb eines akademischen Grades eingereicht. Sie entspricht vollumfänglich den Leitlinien der wissenschaftlichen Integrität und den Richtlinien der guten wissenschaftlichen Praxis.

Litschau, 11. November 2022

Bernhard Martin GILL, BSc.

Danksagung

Mein Dank gilt Univ. Prof. Dr. Clement Atzberger, der mir die Verfassung meiner Masterarbeit am Institut für Geomatik ermöglichte und die Betreuung übernahm.

Bei Dr. Markus Immitzer möchte ich mich für die kollegiale und humorvolle Betreuung bedanken. Er stand mir mit Rat und Tat zur Seite, brachte mich mit seinen Ideen auf den richtigen Weg und war stets ein motivierender Diskussionspartner.
Ein großes Dankeschön gilt Dr. Markus Engel für seine inspirierenden Worte, wenn Schwierigkeiten bei der Arbeit aufgetaucht sind.

Weiters bedanke ich mich bei allen Kollegen des Instituts für Geomatik für ihre Unterstützung und Hilfestellung bei Problemen.

Ein großes Dankeschön gilt meiner Großmutter, Frau Grabenhofer, für das Lektorat an dieser Arbeit!

Zu guter Letzt ist es mir ein besonderes Anliegen, mich bei meiner Familie, allen voran bei meinen Eltern herzlich zu bedanken. Sie waren vor und während meines Studiums immer eine wesentliche Stütze und ein verlässlicher Rückhalt in meinem Leben.

Ihnen möchte ich diese Arbeit widmen.

Inhaltsverzeichnis

Eidesstattliche Erklärung ... iii
Danksagung ... iv
Inhaltsverzeichnis ... v
Kurzfassung ... vi
Abstract ... vii
1. Einleitung und Ziel der Arbeit ... 1
2. Methodik ... 3
2.1. Untersuchungsgebiet ... 3
2.2. Daten und verwendetes Material ... 4
2.3. Verarbeitung der ALS- Daten ... 7
2.3.1. Metriken aus dem nDSM- Raster ... 7
2.3.2. Metriken aus der ALS- Punktwolke ... 7
2.3.3. Metriken aus Sentinel – 2 Szene ... 12
2.4. Modellierung des Holzvorrates ... 13
2.4.1. Stratifizierung ... 13
2.4.2. Random Forest Modellierung ... 13
2.4.3. Moving- Window Approach ... 15
2.4.4. Waldmaske ... 15
2.4.5. Unterscheidung Laub- zu Nadelholz ... 15
3. Ergebnisse der Holzvorratsmodellierung ... 16
3.1. Gesamtmodelle ... 16
3.2. Nadelholzmodelle ... 21
3.3. Laubholzmodelle ... 26
3.4. Mischwaldmodelle ... 31
3.5. Validierung der extrahierten Metriken ... 35
3.6. Erstellung der Holzvorratskarte ... 36
3.7. Berechnung des Bestandesvorrats ... 38
4. Diskussion ... 39
4.1. Daten und Material ... 39
4.2. Diskussion der Ergebnisse ... 41
5. Schlussfolgerungen und Ausblick ... 44
6. Abbildungsverzeichnis ... 46
7. Tabellenverzeichnis ... 48
8. Literaturverzeichnis ... 49

Kurzfassung

Das Ziel dieser Arbeit ist die Erstellung einer Holzvorratskarte. Einerseits werden ausschließlich Höhenmetriken aus einem normalisierten Oberflächenmodell auf unterschiedlichen Kreisradien gerechnet. Der andere Weg ist die Ableitung von Höhen-, Intensitätsmetriken und die Berechnung der Stammzahl aus der Punktwolke. Für die Modellierung wurde der Random Forest Regression Algorithmus angewendet. Herausfordernd für die Modellierung war die zeitliche Abweichung zwischen Erhebung der Inventur (2014) zum Zeitpunkt der Befliegung (2018). Die Inventur wurde mit einer Winkelzählprobe (WZP) erhoben. Eine WZP hat den Nachteil, dass kein Flächenbezug vorhanden ist. Für die Anwendung eines flächenbezogenen Ansatzes ist dieser aber essenziell. Kreise mit Radien von 5 bis 27.5 m in 2.5 m Abstufungen wurden auf die WZP- Punkte gelegt und im Anschluss Metriken auf diesen Flächen berechnet. Um die Modellgüte zu verbessern, wurde die Inventur im Vorfeld in Straten unterteilt. Bei dem Vergleich der Modelle aus dem nDSM und der Punktwolke ist eindeutig ersichtlich, dass die zusätzlich gewonnene Information aus der Punktwolke einen starken Einfluss auf die Modellgüte hat. Auch die Verwendung der Kreiskombinationen waren den Einzelkreisradien überlegen. Die berechnete Stammzahl auf den Plots sowie die Information aus der Sentinel-2 Szene (4 Bänder und NDVI) brachten keine Verbesserung. Die besten Modelle aus der Punktwolke für Nadelholz erzielten ein R^2 von 0.67 und einen $RMSE_{rel}$ von 32.26 %. Das beste Modell für die Berechnung des Holzvorrats für Laubholz war am 15 m Einzelkreis mit Metriken abgeleitet aus der Punktwolke (R^2 = 0.64 bei $RMSE_{rel}$ = 34.49 %). Im Vergleich dazu schnitt der Radius des Einzelkreises des nDSM (bestes Modell Radius = 20 m) mit einem R^2 von 0.48 und $RMSE_{rel}$ von 41.76 % schlechter ab. In dieser Arbeit ist ersichtlich, dass die zusätzliche Information gewonnen aus der Punktwolke im Vergleich zu Höhenmetriken aus dem nDSM eine Modellverbesserung bewirkt.

Abstract

The aim of this thesis is the estimation of growing stock. On the one hand, only height metrics are calculated from a normalized digital surface model (nDSM) on different radii of circles. The other way was to derive height metrics, intensity metrics and the stand density from the point cloud. The Random Forest regression algorithm was used. Challenging for the modelling was the temporal discrepancy between survey of the inventory (2014) to the time of the aerial survey (2018). The inventory was collected using an angle count sample (ACS). An ACS has the disadvantage of not having an area reference. However, this is essential for the application of an area-based approach. The ACS points were visually corrected in their position. Circles with radii from 5 to 27.5 m in 2.5 m increments were placed on the ACS points and metrics were calculated on these areas. To improve model quality, the inventory was divided into strata (coniferous, deciduous, mixed). When comparing the models from the nDSM and the point cloud, it is clearly evident that the additional information obtained from the point cloud has a strong influence on the model quality. The use of the circle combinations was also superior to the single circle radii. The calculated stem density on the plots as well as the information from the Sentinel- 2 scene (4 bands and NDVI) was not chosen by the feature selection. The best models from the point cloud for softwood achieved an R^2 of 0.67 and an $RMSE_{rel}$ of 32.26 %. The best model for calculating the growing stock for hardwoods was at the 15 m single circle with metrics derived from the point cloud ($R^2 = 0.64$ with $RMSE_{rel} = 34.49$ %). In comparison, the radius of the single circle of the nDSM (best model radius = 20 m) performed worse with an R^2 of 0.48 and $RMSE_{rel}$ of 41.76 %. This clearly shows that the additional information from the point cloud causes an improvement of the models.

1. Einleitung und Ziel der Arbeit

Für die strategische und operative Planung von Forstbetrieben jeglicher Bewirtschaftungsgröße ist der Holzvorrat wie auch die Zusammensetzung der Baumarten (Laub- und Nadelholz) eine unverzichtbare Kennzahl. Derzeit wird in den meisten Betrieben der Holzvorrat mittels Inventur, detaillierter mit einer Stichprobeninventur, ermittelt. Dabei werden forstliche Parameter terrestrisch und punktuell mittels Winkelzählprobe (WZP) nach Bitterlich erhoben und anschließend auf die Fläche mit statistischen Ansätzen transferiert (Kramer & Akca, 1982). Der Vorteil dieser Methode liegt in der robusten Aussagekraft auf Betriebsebene. Der große Nachteil ist, dass es nicht möglich ist, die kostspielig gewonnene Information der Forstinventur auf die Bestandesebene (-fläche) zu skalieren. Um kostengünstige Rückschlüsse von der Betriebs- auf die Bestandesebene zu ziehen, wurden Erweiterte Forstliche Inventuren (EFI; Enhanced Forest Inventories) entwickelt (Næsset 1997, Lindberg und Hollaus 2012, White et al. 2013, White et al. 2017). Das Ziel von EFI ist die Unterstützung der strategischen wie auch operativen Planung. Es wird der Informationsfluss erhöht, da nun auch die Holzvorräte auf der Bestandesebene bekannt sind. Die EFI ist das Ergebnis eines flächenbezogenen Ansatzes (area- based approach; ABA) bei dem Holzvorräte mithilfe verschiedener Metriken entweder aus einer Light Detection and Ranging (LiDAR) Punktwolke oder aus einem normalisierten digitalen Oberflächenmodell (nDSM) errechnet und Bestandeskennzahlen geschätzt werden. nDSM sind in der Regel gewonnene Produkte aus einer LiDAR Punktwolke. In der Praxis wird diese Methode an großen Waldflächen (> 50000 ha) durchgeführt wobei Daten mittels Airborne Laser Scanning (ALS, Luftbefliegung mit LiDAR Sensor) oder photogrammetrischen Methoden generiert werden. In weiterer Folge werden aus der ALS- Punktewolke Metriken abgeleitet und es wird versucht, Schätzer für Bestandeskennwerte anhand dieser Metriken abzuleiten. Die Modelle werden rasterweise an der Gesamtfläche angewandt. Die Pixelgröße, an der Metriken abgeleitet werden ist essenziell- so muss die Flächengröße der Pixel der Validierungsflächen gleich groß sein wie die Flächen, bei denen die Werte geschätzt werden. In weiterer Folge ist jeder Pixel mit einem modellierten forstlichen Kennwert ausgestattet. Die Pixelwerte in einer definierten Fläche (Betrieb; Revier; […] Bestand) ergeben die Kennzahl, welche von Interesse ist, wieder (White et al. 2017). Die Methodik der EFI wird in der nördlichen Hemisphäre in homogenen Wäldern angewendet, um den Holzvorrat in großen Gebieten bestimmen zu können. Dafür werden in einem Raster (4 km x 4 km) Probekreise mit bis zu 625 m² und größer angelegt (White et al. 2017, White et al. 2013). Die Herausforderung in dieser Arbeit war unter anderem, den benötigten Flächenbezug aus Winkelzählproben herzustellen. Dafür wurde auf die Methodik von Immitzer et al. (2016) zurückgegriffen. Es wurden unterschiedliche Kreisradien und eine Kombination der Radien auf den WZP- Punkten angewendet

Eine weitere Methodik zur Ermittlung des Holzvorrates neben des ABA ist der Ansatz der Individual Tree Crown Delineation (individuelle Einzelbaumerkennung, ITCD). Dabei werden aus der ALS- Punktwolke Einzelbäume bzw. deren Kronen erkannt und mit einer eindeutigen ID versehen. Bei jedem Individuum wird seine gemessene Höhe aus der Punktwolke zugewiesen. In weiterer Folge wird mittels baumartenspezifischer Höhenfunktion bei jedem Individuum der BHD berechnet und mit der Formzahlfunktion wird am Ende das Volumen des Einzelbaumes ermittelt. Diese Methode wird im mitteleuropäischen Raum eher selten in der Praxis umgesetzt (Kaartinen et al. 2012). ITCD hat im Vergleich zum ABA den Nachteil, dass Bäume, welche sich unter dem Kronendach befinden, nicht erkannt werden. Zusätzlich steht das Problem im Raum, dass die Algorithmen zur eindeutigen Baumerkennung noch nicht das erforderliche Maß an Genauigkeit aufweisen, um präzise Baumarten identifizieren zu können. Dies ist ein sehr wichtiger Faktor, da Höhenkurven wie auch Formzahlfunktionen baumartenspezifisch sind. Als Vorteil zum ABA kann dennoch genannt werden, dass weniger Referenzdaten erhoben werden müssen (Kaartinen et al. 2012).

Der ABA unterstützt kostenschonend moderne Forstbetriebe bei der Bewirtschaftung der Bestände. Es werden keine hochauflösenden LiDAR Punktwolken benötigt, welche in der Anschaffung wie auch Speicherung höhere Kosten aufweisen (Kaartinen et al. 2012). Die bestehenden Inventurdaten werden als Trainingsdaten herangezogen – es muss nach Metriken in der Punktwolke gesucht werden, die mit dem Holzvorrat korrelieren.

Studien über die Holzvorratsermittlung reichen zurück bis zu Næsset (1997). In Norwegen, Finnland und Kanada wird die Nationale Forstinventur seit 20 Jahren mit ALS-Daten unterstützt. Dies ist kostengünstiger da der Stichprobenraster größer wird und die Anzahl der zu erhebenden Flächen sinkt. Vor 10 Jahren wurde in Finnland die Ersparnis der Forstinventur durch zusätzliche Verwendung von ALS- Daten auf rund 20 Mio. € beziffert (Hyyppä et al. 2012). Vergleichbare Arbeiten zeigen, dass dieses Forschungsthema auch nach über 20- jährigem Bestehen noch immer aktuell ist. Dies liegt einerseits an der preiswerten Verfügbarkeit der LiDAR Produkte und andererseits an dem Interesse seitens der Forschung und Privatwirtschaft (White et al. 2017).

Photogrammetrische Punktwolken werden hingegen aus mehreren Bildern vom selben Objekt aber aus unterschiedlichen Perspektiven generiert. Aus photogrammetrischen Punktwolken lassen sich ebenso nDSM berechnen. Immitzer et al. (2016) nutzten nDSM abgeleitet aus Stereo-Satellitendaten und zeigten, dass die Zuhilfenahme von Spektralinformationen zu den Höhendaten die Genauigkeit der Holzvorratsmodelle verbessern können ($R^2 = 0.60$, $RMSE_{rel} = 29.5$ %, nur Höhenvariablen: $R^2 = 0.53$, $RMSE_{rel} = 31.7$ %). Weiters wurde in der Arbeit gezeigt, dass der Multikreisansatz dem Einzelkreisansatz bei WZP-Daten überlegen ist. Kirchhoefer et al.(2019) hat Metriken aus Oberflächenmodellen abgeleitet und mithilfe linearer Modelle forstliche Kennwerte geschätzt. Die Bestimmtheitsmaße und die quadratische mittlere Abweichung unterscheiden sich in den Untersuchungsgebieten (R^2 von 0.55, 0.50, 0.57; $RMSE_{rel}$ = 38.6 %, 42.1 %, 37.9 %). In der Arbeit wurde anstelle eines Multikreisansatzes ein Ansatz mit fixem Kreisradius über die WZP- Punkte gelegt. Hill et al. (2018) hingegen hat Höhenmetriken mit einer Variable für die Baumart kombiniert und schreibt über das Problem der hohen Heterogenität, welche in großen Waldgebieten durch die unterschiedlichen Wuchsklassen auftritt. In der Arbeit von Socha et al. (2020) wurde ein area-based approach auf Weißkiefernbeständen in Polen durchgeführt. Das beste Modell hatte ein R^2 von 0.63 und einen $RMSE_{rel}$ von 22.8 %. Stepper et al. (2015) erzielte bei der Auswertung der OOB Daten der Random Forest Modelle am Plot $RMSE_{rel}$ von 30.92 % bis 36.02 %.
Hawryło und Wężyk (2018) schätzten den Holzvorrat mit Metriken aus einer photogrammetrischen Punktwolke aus Luftbildern mit respektablen Ergebnissen ($R^2 = 0.81$; $RMSE_{rel}$ =17.26 %). Vergleiche mit dieser Arbeit sind schwierig, da das Untersuchungsgebiet großteils mit einem homogenen Weißkiefernbestand bestockt ist.

Das Ziel dieser Arbeit ist, eine Holzvorratskarte unter Berücksichtigung der vorkommenden Nadel- und Laubholzarten für einen Forstbetrieb in Österreich zu erstellen. Es wird der ABA Ansatz wie in White et al. (2017) angewandt und es werden zusätzlich Spektraldaten nach Immitzer et al. (2016) für die Verbesserung der Modelle herangezogen.

Folgende Schritte werden dabei getestet:

- Unterschied der Modelle mit Höhenmetriken aus nDSM zu Höhenmetriken aus der ALS- Punktwolke mit zusätzlichen Intensitätsmetriken, Durchdringungsraten, Kronenvolumen, Variationskoeffizient
- Berechnung der Stammzahl nach ITC- Ansatz aus der LiDAR Punktwolke
- Verbesserung der Modelle mittels errechneter Stammzahl als Zusatzinformation
- Zusätzliche Information aus Sentinel- 2 Szene (4 Bänder + NDVI) für die Verbesserung der nDSM und ALS- Modelle

2. Methodik

2.1. Untersuchungsgebiet

Das Untersuchungsgebiet befindet sich in Niederösterreich, nördlich der Donau (Abbildung 1). Es erstreckt sich von Persenbeug bis nach Arbesbach und hat eine Ausdehnung von rund 12.500 ha und steht im Besitz der Familie Habsburg- Lothringen. Das Waldgebiet liegt im Wuchsgebiet 8.2 nach Mayer (Außeralpines Fichten- Tannen- Buchenwaldgebiet, Marschall (1975)) und verteilt sich über eine Seehöhe von 214 m bis 1061 m über Adria.

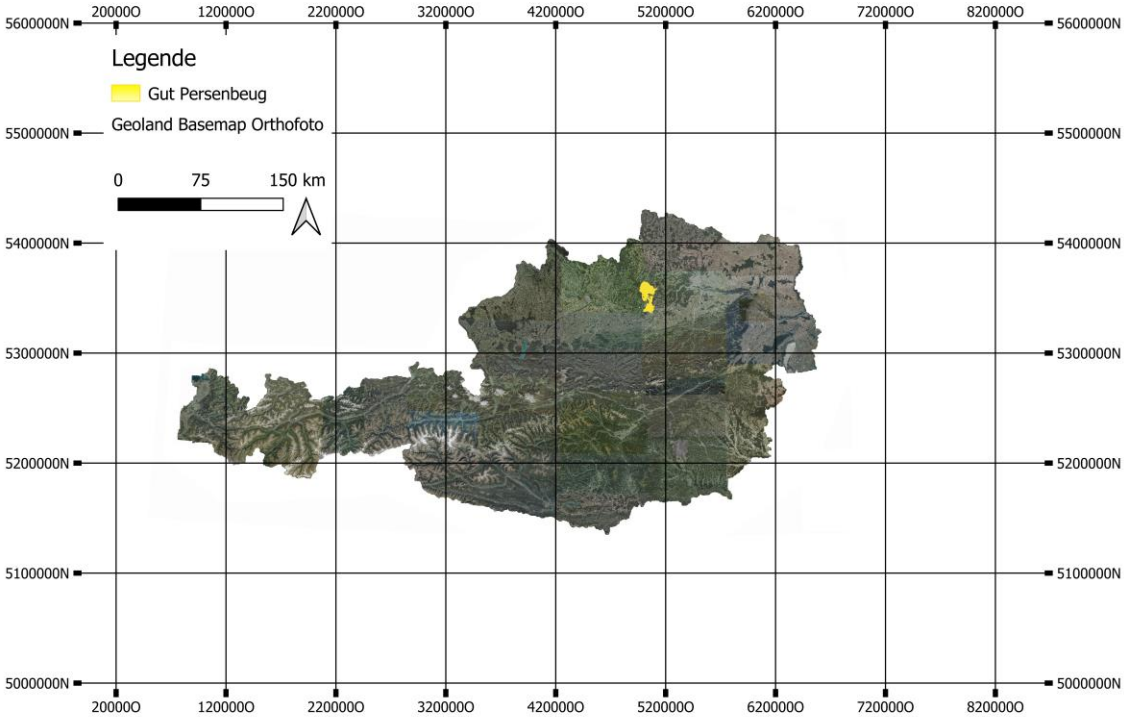

Abbildung 1: Das Untersuchungsgebiet (gelb) befindet sich im nördlichen Teil Österreichs, an der Grenze von Mühl- und Waldviertel.

2.2. Daten und verwendetes Material

Forstinventur

Die Forstinventur wurde 2014 erhoben und ausgewertet. Die 874 Aufnahmepunkte verteilen sich über das ganze Untersuchungsgebiet und haben einen Rasterabstand von 400 m x 200 m und 600 m x 300 m. Die Probepunkte wurden im Wald mittels GNSS (Global Navigation Satellite System) eingemessen. Rund drei Viertel der Waldfläche im Untersuchungsgebiet ist mit Fichtenwald bestockt.

Folgende Daten wurden pro Winkelzählprobe- Punkt erhoben:

- Koordinaten des Probepunktes, Gauß- Krüger M34 & WGS84
- Alter der Bäume
- Abstand der Bäume zum WZP- Mittelpunkt
- Baumnummer
- Azimutwinkel von Probepunkt zu Baum
- Distanz von Probepunkt zu Baum
- Baumart
- Brusthöhendurchmesser in 1.3 m Höhe (BHD)
- Höhe der Oberhöhenbäume

<u>Die Winkelzählprobe</u>

Die Winkelzählprobe (WZP) wurde von BITTERLICH 1948 veröffentlicht. Messgeräte wie das Spiegel-Relaskop wurden erst 1955 entwickelt. Die WZP ist eine optische Stichprobe die durch Zählung von Bäumen die Bestandesgrundfläche pro Hektar [m²/ha] herleitet. Ob ein Baum in die Stichprobe fällt und gezählt wird, hängt einzig von dessen Durchmesser in Brusthöhe (BHD) und die Distanz zum Stichprobenmittelpunkt ab. Es werden die Zählbreiten (= k) 1 m², 2 m² und 4 m² angewandt. Jeder Baum, der in der Stichprobe gezählt wird, repräsentiert automatisch die Grundfläche mit dem Wert der Zählbreite (Kramer und Akça 2008).

Für die Ermittlung des Vorrats pro Hektar muss nun die Höhe gefunden werden. Dies kann auf mehrere Arten, je nach gewünschter Genauigkeit durchgeführt werden. Die Bestandesoberhöhe wird vom zweitstärksten Baum, auf den Durchmesser bezogen, repräsentiert. Dies ist pro Baumart anzuwenden: Findet sich nur ein Baum einer Baumart, so wird dessen Höhe gemessen. Somit steht pro Baumart und pro Winkelzählprobe eine Höhe zur Verfügung. Dies ist in der statistischen Verteilung aller Höhen zu diesem Alter die Bestandesoberhöhe (Kramer und Akça 2008).

Der Durchmesser in 1.3 m Höhe, auch Brusthöhendurchmesser genannt, spiegelt den Querschnitt (Kreisfläche) eines idealisierten Zylinders wider. Bei der Winkelzählprobe wird die Kreisfläche in m² im Verhältnis zur Zählbreite in m² gesetzt. Dieser Wert ist die repräsentive Stammzahl pro Hektar.

Die Baumart zu bestimmen ist essenziell, denn jede Baumart weist andere Wuchseigenschaften auf. Für die Hauptbaumarten gibt es je nach Höhenlage oder Wuchsgebiet angepasste Höhenfächer (Wuchsreihen der Baumhöhe versus Alter). Edellaubholzarten werden als sonstiges Laubholz angeführt. Inventuren konzentrieren sich auf Hauptbaumarten die bestandesbildend sind. Wenn Edellaubholzarten in Reinbeständen vorkommen, werden diese gesondert aufgenommen. Kommt dies in Betrieben oft vor, ist eine geschichtete Stichprobe von Vorteil, wobei Straten mit Baumarten empfohlen werden (Kramer und Akça 2008).

Laserscan- Befliegung

Die Befliegung erfolgte Mitte April 2018 und lieferte eine mittlere Punktdichte von 16 Punkte pro m². Das Flugzeug war mit einem Riegl VQ580 Scanner ausgestattet und flog in Streifen mit einer Querabdeckung von 50 %. Die Flughöhe betrug im Mittel 500 m über Adria Die erhobenen Daten wurden in Boden und nicht Boden vorklassifiziert.

Erstellung des nDSM

Die Ableitung des nDSM erfolgte aus den ALS Daten. Zuerst wurde ein digitales Geländemodell (DTM) und im Weiteren ein digitales Oberflächenmodell (DSM) kalkuliert, beide hatten eine Pixelgröße von 0.5 m x 0.5 m. Durch Subtraktion (DSM – DTM) erhält man das normalisierte Oberflächenmodell, in dem die Rasterzellenwerte die Höhe der Objekte vor Ort repräsentieren.

Co- Registrierung der Inventurpunkte

Die im Wald mit Überschirmung erhobenen Inventurpunkte sind durch die Einmessung mittels GNSS nicht lagerichtig. Dies ist in Abbildung 2 sichtbar gemacht. Um diesen Lagefehler zu korrigieren, wurden alle WZP- Punkte visuell kontrolliert und auf Grund der Struktur der Bäume (grüne Punkte) in der WZP an das nDSM angepasst.

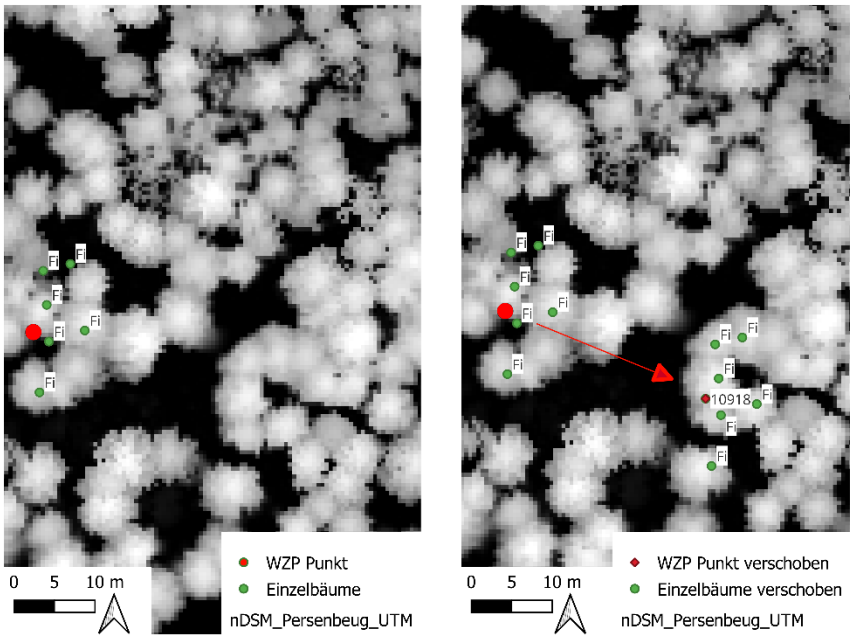

Abbildung 2: Korrektur der Lage des WZP- Punktes 10918: vor (links) und rechts (nach) der Verschiebung.

Durch den Abstand von vier Jahren zwischen Inventur und Befliegung muss von Veränderungen wie Zuwachs oder Nutzung ausgegangen werden. WZP Punkte mit einer Höhenveränderung von mehr als 25 % können als Ausreißer gewertet und ausgeschlossen werden (Kirchhoefer et al. 2019).

Dies kann nach Kirchhoefer et al. (2019) wie folgt durchgeführt werden:

$$devH_{max} = \left|100 - \frac{100}{\max(H_{FI}) * \max(H_{CHM\&ALS})}\right| \qquad (1)$$

Wobei $devH_{max}$ der Unterschied der Höhe in Prozent, $\max(H_{FI})$ die größte gemessene Höhe am WZP Punkt in der Natur in m und $\max(H_{CHM\&ALS})$ die größte gemessene Höhe am WZP Punkt am nDSM bzw. ALS Punktwolke ist.

2.3. Verarbeitung der ALS- Daten

Die ursprüngliche, rohe Punktwolke wurde mit LAStools (rapidlasso GmbH 2022) und lidR (Roussel J. R., 2022) bearbeitet. Um eine schnellere Bearbeitung der Daten zu gewährleisten, wurden die ALS Daten indexiert (LAStools; lasindex) und klassifiziert (LAStools; lasground, lasclassify). Im Folgenden werden die beiden Arten der Weiterverarbeitung beschrieben.

2.3.1. Metriken aus dem nDSM- Raster

Um die Metriken aus den unterschiedlich großen Kreisradien zu erhalten wurden die Höhenmetriken mit dem R- Paketes *terra* (Hijmans 2022) extrahiert. Die Kreisradien beginnen bei 5 m Radius und erweitern jeweils in 2.5 m Stufen den Radius auf 27.5 m. Für jeden Radius wurde von der Höhe das Minimum, das Maximum, der Mittelwert, die Standardabweichung und verschiedene Quantilen (5 %, 25 %, 50 %, 75 %, 95 %) berechnet.

Auffallend bei allen Kreisradii war, dass mit zunehmender Baumhöhe die Datenpunkte eine extreme Streuung über alle Metriken aufweisen. Diese Beobachtung wurde auch bei den Metriken generiert aus der ALS Punktwolke gemacht. Ähnliche Erfahrungen hat auch Hill et al. (2018) dokumentiert. Dies entsteht unter anderem durch die großen Unterschiede hinsichtlich der Bonität (Wuchsleistung) aber auch durch die Bewirtschaftung. Dichtere Bestände investieren mehr in das Höhenwachstum bzw. wird mehr oder weniger bei 30 m bis 40 m die maximale Höhe erreicht und nur noch in das Dickenwachstum investiert.

2.3.2. Metriken aus der ALS- Punktwolke

Um aus der gesamten Punktwolke Metriken ableiten zu können, musste diese als erstes normalisiert werden. Das R- Paket lidR (Roussel, Auty, & De Boissieu, 2022) bietet dafür optimale open- source Möglichkeiten an. Das Paket kann aus der gesamten Punktewolke, die eine Größe von 300 GB hat, kleine, vordefinierte Punktwolken (chunks) in den Arbeitsspeicher laden und prozessieren. Die Größe der chunks sind frei wählbar, für die Prozessierung wurde eine Größe von 250 m x 250 m verwendet. Mithilfe eines Buffers (gewählt wurden 30 m) können Randprobleme minimiert werden. Für die Normalisierung der Punktwolke wurde der Befehl *normalize_height()* benützt. Aus der bereits klassifizierte Punktewolke wurde on- the- fly für jeden chunk ein Geländemodell (DTM) errechnet und von der Punktwolke subtrahiert. Das DTM wird mittels k- nearest- neighbour (knn) und inverser Distanzgewichtung (IDW) berechnet. Die Idee der inversen Distanzgewichtung ist, dass räumliche Nachbarn vom Wert ähnlich sind (Riedl, 2022). Die normalisierte Punktwolke wurde für die weitere Prozessierung verkleinert. Dabei wurde ein Puffer von 50 m um die lagerichtigen WZP- Punkte gelegt und diese Bereiche ausgeschnitten. Für die Berechnung der Metriken wurden folgende Kreisradii um die WZP- Punkte gewählt: 5 bis 27.5 m in 2.5 m Schritten. Das lidR Paket (Roussel, Auty, & De Boissieu, 2022) bringt mit der Funktion *plot_metrics()* eine Möglichkeit, verschiedene Metriken aus den ALS Daten zu berechnen. Zusätzlich wurde auch die Stammzahl für verschiedene Kreisradii abgeleitet und für die Modellierung herangezogen. Für die Höhen- und Intensitätsmetriken wurde eine Vegetationsmaske angewandt (Immitzer, 2016). Um Ausreißer aus der Punktwolke zu eliminieren, bietet das lidR Paket an, einen definierten z- Bereich einzulesen. Das Minimum der Höhe betrug 2 m und das Maximum wurde auf 52 m eingestellt, um Höhenausreißer zu eliminieren. Der Wertebereich unter 2 m ist nicht eindeutig

als Baum zu identifizieren, und um Ausreißer im negativen Wertebereich wie auch im oberen (> 52 m) zu eliminieren wurde dieser Filter angewandt. 52 m wurde auch deshalb als Maximum gewählt, da die größte gemessene Höhe laut Inventur 2014 46.4 m betrug.

2.3.2.1. Berechnung der Höhenmetriken

Das R- Paket *terra* von Roussel et al. (2020) bietet die Möglichkeit an, vordefinierte Höhenmetriken aus der Punktwolke zu berechnen. Die Liste der bestehenden Möglichkeiten ist beachtlich, und durch die open- source Verfügbarkeit des Paketes ist es sehr wahrscheinlich, dass die Autorengemeinschaft die Liste stetig erweitern wird. Folgend ist der aktuelle Stand der Metriken angeführt, welche mit dem Paket berechnet werden können.

- Maximum (zmax)
- Mittelwert (zmean)
- Standardabweichung (zsd) der Höhenverteilung
- Skewness (zskew) der Höhenverteilung
    - Die Skewness (Schiefe) gibt an, wie viel eine Häufigkeitsverteilung von der Normalverteilung abweicht. Null bedeutet keine Abweichung, ist der Wert größer als Null liegt eine Rechtsschiefe vor und bei kleiner Null eine Linksschiefe
- Kurtosis (zkurt) der Höhenverteilung
    - Bei einem Kurtosis Wert von Null folgen die Daten einer Normalverteilung. Bei positiven Werten der Kurtosis übergipfeln die Daten in der Mitte die Normalverteilung, bei negativen Kurtosis Werten ergibt sich ein ausgeprägter Randbereich.
- Streuungsmaß (zentropy) der Höhenverteilung
- Prozent der Returns (Punkte) über dem Mittelwert (pzabovemean)
- Quantilen der Höhenverteilung
    - Prozent: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95
- Kumulierter Prozentsatz der Returns (nach Woods et al. (2008))

2.3.2.2. Berechnung der Intensitätsmetriken

Das *terra* Paket von Roussel et al. (2020) bietet ebenso die Berechnung von vordefinierte Intensitätsmetriken an. Diese sind folgende:

- Summe der Intensität für jeden Return (itot)
- Maximale Intensität (imax)
- Mittelwert (imean)
- Standardabweichung (isd)
- Skewness (iskew)
- Kurtosis (ikurt)
- Prozentsatz Intensität von Bodenpunkten (ipground)
- Prozentsaz der Intensität unter der Quantile der Höhe (ipcumzq)
    - Quantilen: 10%, 30%, 50%, 70%, 90%

2.3.2.3. Berechnung der Punktanzahl (Roussel J. R., 2022)

Die Punktanzahl wurde für unterschiedliche Bereiche auf der z- Achse berechnet (bis 2 m, 2 m bis 52 m und 0 m bis 52 m), um in weiterer Folge die Durchdringungsrate (PR; penetration rate) und das Kronenvolumen (VC) berechnen zu können. Die Punktanzahl in der Punktwolke wurde ebenfalls für alle Kreisradien berechnet.

Durchdringungsrate nach Yu et al. (2015)

$$PR = \frac{N_{h \leq 2}}{N_{total}} \qquad (2)$$

PR = Durchdringungsrate (penetration rate)

$N_{h<=2}$ = Punktanzahl von 0 m bis 2 m

N_{total} = Punktanzahl total (0 m bis 52 m)

Variationskoeffizient (H_{CV}) nach Yu et al. (2015)

$$H_{CV} = \frac{zsd}{zmean} \qquad (3)$$

Wobei zsd die Standardabweichung der Höhenverteilung und zmean der Mittelwert der Höhe am Plot ist.

Kronenvolumen (VC) nach Yu et al. (2015)

$$VC = zmean * (1 - PR) \qquad (4)$$

Wobei $zmean$ der Mittelwert der Höhe und PR die Durchdringungsrate am Plot ist.

2.3.2.4. Berechnung der Stammzahl

Die Stammzahl innerhalb des jeweiligen Kreisradius wurde mit dem R Paket *lidR* (Roussel et al. 2020) prozessiert. Zuerst wurden für jeden WZP Punkt Regionen mit einem Radius von 50 m aus der gesamten ALS Punktwolke ausgeschnitten (*clip_circle(),* Abbildung 3). In weiterer Folge wurden die Koordinaten der Baumhöhen mit einem lokalen Maximumfilter (lmf) und einem fixen Suchfenster berechnet. Das Suchfenster wurde mit 3 m und mit 5 m Durchmesser berechnet und visuell am nDSM abgeglichen. Das Suchfenster von 5 m errechnet eine bessere Erfassung der Stammzahl. Die Berechnung der Anzahl der Bäume im jeweiligen Probekreis erfolgte mit QGIS (Punkte im Polygon zählen). In Abbildung 4 ist ersichtlich, dass die Stammzahl am Probekreis nicht linear, sondern quadratisch mit dem Vorrat korreliert. Dies ist insofern eine interessante Beobachtung, da die Annahme bestand, dass in juvenilen Beständen die Stammzahl hoch aber dafür der Vorrat gering ist. Bei alten Beständen hingegen die Stammzahlhaltung gering aber der Vorrat hoch.

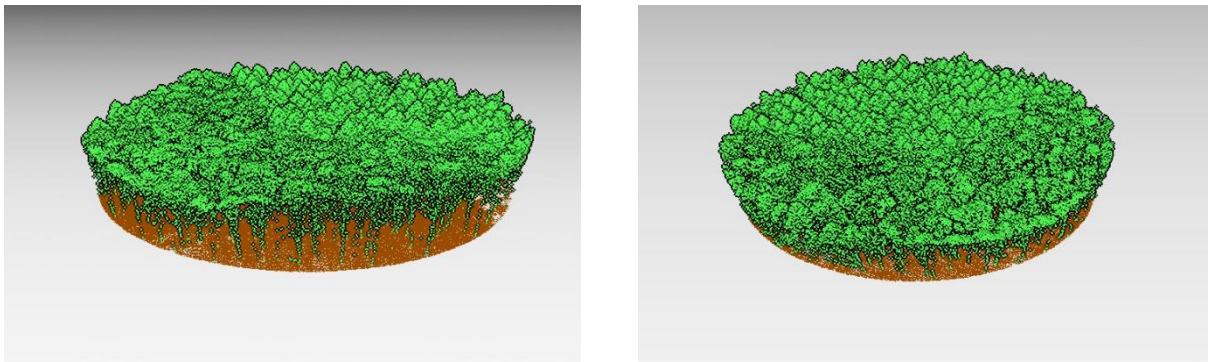

Abbildung 3: Darstellung der Punktwolke mit 50 m Radius zur Sichtbarmachung der ableitbaren Baumhöhen. Die beiden Abbildungen unterscheiden sich in der Perspektive.

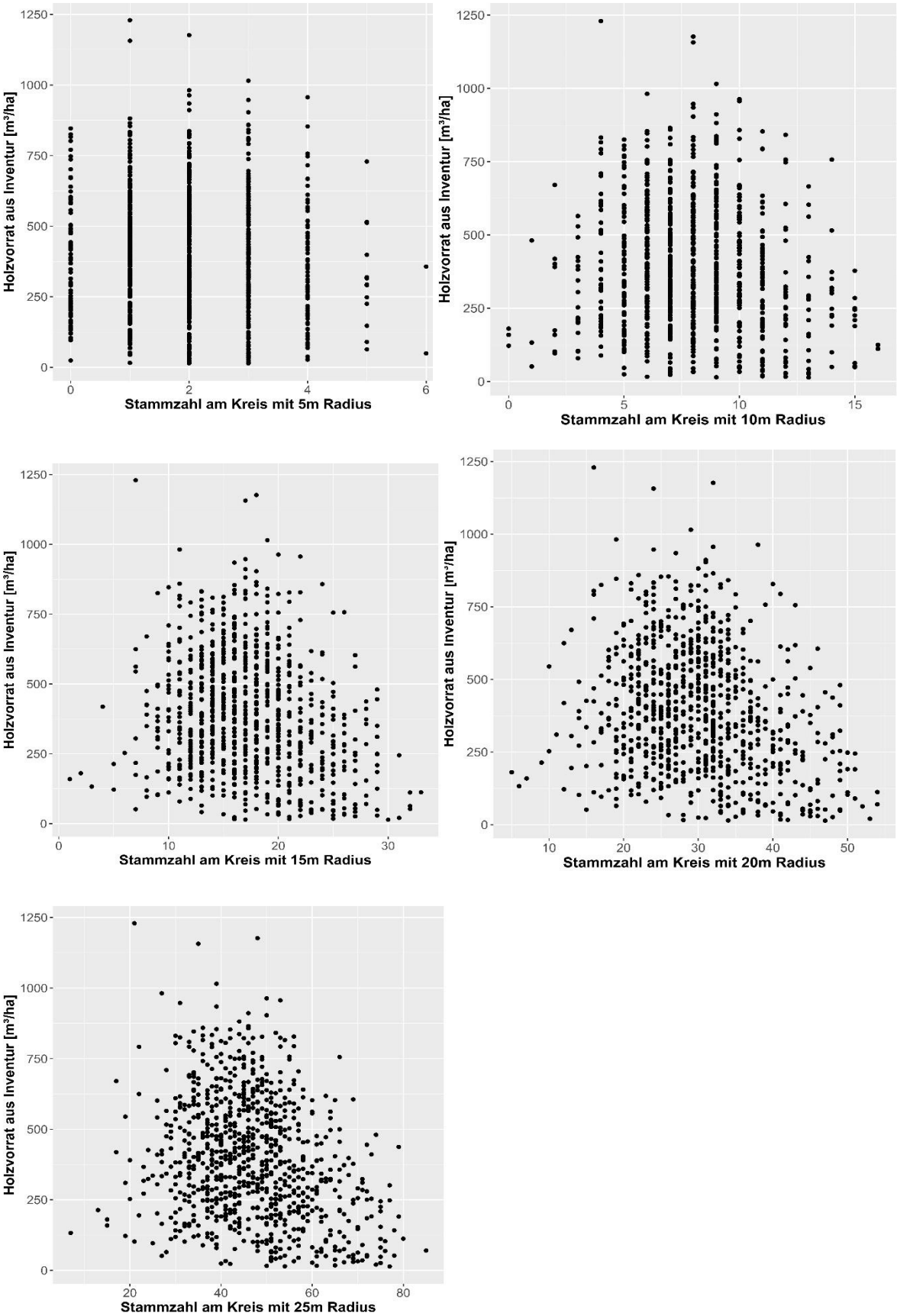

Abbildung 4: Nach der Berechnung der Stammzahl im Probekreis wurde diese dem Holzvorrat für die verschiedenen Kreisradien gegenübergestellt. Abgebildet sind die Kreisradien 5,10,15,20,25 m.

2.3.3. Metriken aus Sentinel–2 Szene

Die Sentinel–2 Szene wurde am 29. August 2018 aufgenommen. Um weitere Metriken zu erhalten, wurde nach Immitzer et al. (2016) der normalisierte differenzierte Vegetationsindex mit einer Rastergröße von 10 m x 10 m berechnet (Abbildung 5). Der NDVI wird aus den Reflexionswerten im nahen Infrarotbereich und des roten sichtbaren Bereichs des Lichtes berechnet (Rouse et al. 1973). Das Ergebnis ist ein Wertebereich von -1 bis +1, wobei negative Werte Wasserflächen repräsentieren. Zusätzlich zum NDVI wurden von den Bändern Infrarot, Rot, Blau und Grün folgende Metriken extrahiert: Mittelwert, Standardabweichung und die Quantilen für 5 %, 25 %, 50 %, 75 % und 95 % wie die Minimum- und Maximumwerte.

Normalisierter differenzierter Vegetationsindex

$$NDVI = \frac{NIR - Rot}{NIR + Rot} \qquad (5)$$

NIR = naher Infrarotbereich, Band 8

Rot = Roter Bereich, Band 4

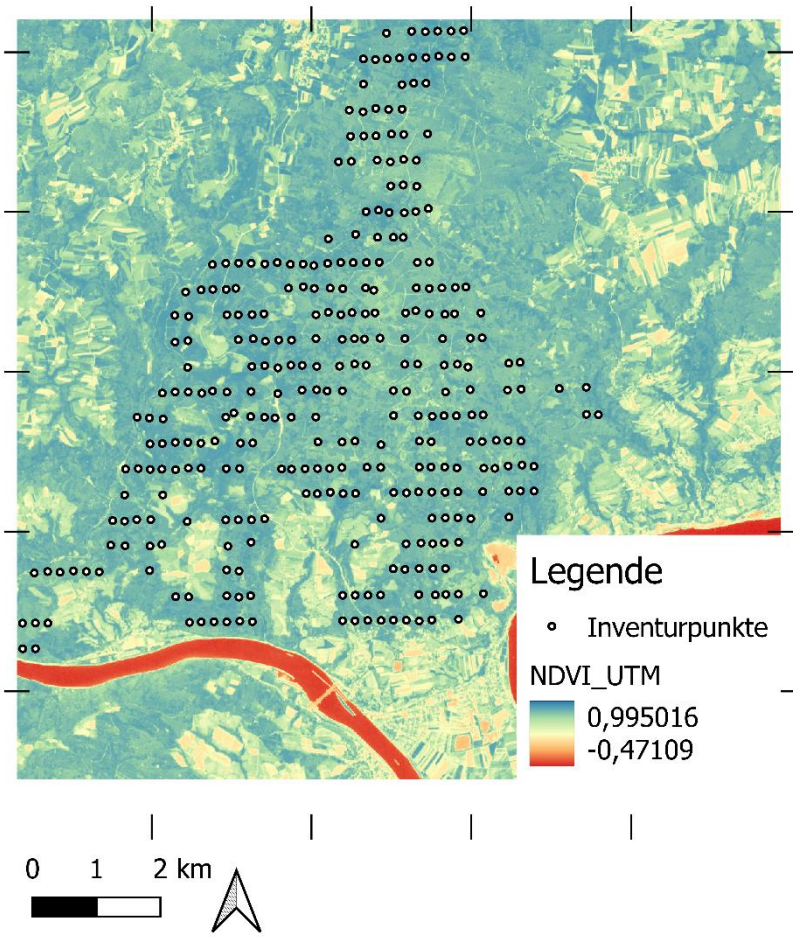

Abbildung 5: Karte mit der Information des NDVI errechnet aus der Sentinel-2 Szene

2.4. Modellierung des Holzvorrates

2.4.1. Stratifizierung

Bei einer Stratifizierung werden aus einer großen Grundgesamtheit Unterstichproben gezogen, die Einfluss auf die Genauigkeit haben (Kramer und Akça 2008). Da bei der WZP auch die Baumarten erhoben wurden, können drei Straten bestimmt werden. Ein WZP- Punkt muss am ermittelten Vorrat gleich oder größer als 80 % Nadel- oder Laubholzanteil haben, um als Nadel- oder Laubwald klassifiziert zu werden. Treffen beide Bedingungen nicht zu, dann wird der Punkt als Mischwald ausgewiesen.

2.4.2. Random Forest Modellierung

Für die Modellierung des Bestandesvorrats wurde ein Random Forest Regression Algorithmus (Breiman 2001) mit dem R- Paket *caret* von Kuhn (2021) berechnet. Um die Genauigkeit zu optimieren, wurde mittels rückwärtiger Feature Selection (Immitzer et al. 2016) der mittlere quadratische Fehler (mean square error, MSE) niedrig gehalten. Dabei bestimmt die Anzahl der erklärenden Variablen die Iterationsschritte. Bei jeder Iteration wird ein neuer Random Forest erstellt es wird die Variable mit dem geringsten Importance- Wert eliminiert (Immitzer et al. 2016). In weiterer Folge wird jener Random Forest gewählt, bei dem die Güte (MSE gering) am höchsten ist. Wenn der MSE gering ist, hat das zu bedeuten, dass die Verzerrung und Varianz zum Schätzer gering sind. Der MSE ist ein Maß der Güte einer Schätzfunktion (Regression) bei der ein unbekannter Parameter (beispielsweise Bestandesvorrat) ermittelt wird. Je näher der Wert des MSE bei 0 liegt, desto geringer ist die Abweichung. Der MSE ist folgend definiert:

$$MSE = \sum_{i=1}^{n} \frac{(Y_i - \hat{Y})^2}{n-2} \qquad (6)$$

Wobei Y_i der wahre und $\hat{Y}$ der geschätzte Wert aus der Regressionsgeraden des Random Forest ist. Random Forest basiert auf Entscheidungsbäumen, die nicht miteinander korrelieren. Jeder Baum „wächst“ individuell und stellt eine Regression dar. Für jedes Modell wurden 1000 Regressionsbäume „gepflanzt“ und das Modell (Entscheidungsbaum im Wald) mit dem geringsten MSE bzw. der höchsten Güte wurde ausgewählt. Random Forest Algorithmen haben folgende Vorteile (Immitzer et al. 2016; Breiman 2001):

- Zielvariablen müssen nicht einer Normalverteilung unterliegen
- Mehrdimensionale Datensätze und stark korrelierende Schätzer können sehr effizient verarbeitet werden
- Überanpassung (overfitting[1]) wird nahezu verhindert
- Güte des Random Forest kann mit den out–of-bag (OOB[2]) Daten ($Y_i, \hat{Y}$) bestimmt werden
- Informationen über die Wichtigkeit, also den Einfluss von jeder Eingangsvariable auf den MSE

Der Nachteil bzw. die Schwäche des Random Forest Algorithmus liegt darin, dass die zu schätzenden Variablen nur in der Spannweite des trainierten Modells liegen können. Eine Extrapolation, wie beispielsweise mit linearen Modellen realisierbar, ist nicht möglich (Breiman 2001). Dass niedrige überschätzt und hohe Werte unterschätzt werden, gehört ebenfalls zu

[1] Überanpassung (overfitting) bedeutet, dass zu viele erklärende Variablen das Modell beschreiben. Es wird immer das Modell gesucht, welches mit der geringsten Anzahl an Variablen die höchste Güte (Genauigkeit) aufweist

[2] Out-of-bag (OOB) Daten sind die im Random Forest enthaltenen geschätzten Ergebnisse.

den Nachteilen das Random Forest. Dies liegt daran, dass bei den errechneten Bäumen der Wert gemittelt wird (Baccini et al. 2004).

In der vorliegenden Arbeit wurde aus dem nDSM wie auch aus der Punktwolke die gleiche Anzahl von Modellen berechnet. Dies ist in Tabelle 1 ersichtlich.

Tabelle 1: Übersicht über die Anzahl der erstellten Random Forest Feature Selection Modelle. EK = Einzelkreis, MK = Multikreis

	nDSM Höhenmetriken	Metriken aus Punktwolke
Gesamtmodell	10 EK + 1 MK = 11 Modelle	10 EK + 1 MK = 11 Modelle
Nadelholzmodell	10 EK + 1 MK = 11 Modelle	10 EK + 1 MK = 11 Modelle
Laubholzmodell	10 EK + 1 MK = 11 Modelle	10 EK + 1 MK = 11 Modelle
Mischwaldmodell	10 EK + 1 MK = 11 Modelle	10 EK + 1 MK = 11 Modelle
Summe	44 Modelle erstellt	44 Modelle erstellt

Statistische Auswertung der Modelle

Für die Auswertung der statistischen Kenngrößen der Modelle wurde mit der Statistik Software R (R Core Team 2022) in RStudio (RStudio team 2019) mit dem Paket *caret* von Kuhn (2021) Berechnungen durchgeführt. Es wurden der absolute und relative *root mean square error (RMSE, Wurzel des mittleren Standardfehlers)*, der absolute und relative *bias* sowie das Bestimmtheitsmaß (R^2) berechnet.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \qquad (7)$$

$$RMSE_{rel} = \frac{RMSE}{\bar{y}} \times 100 \qquad (8)$$

$$bias = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{\bar{y}} \qquad (9)$$

$$bias_{rel} = \frac{bias}{\bar{y}} \times 100 \qquad (10)$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (11)$$

Wobei y_i die Holzvorräte laut Inventur sind. Die Variable $\hat{y}_i$ ersetzt in der Funktion die vorhergesagten Holzvorräte, also die out- of- bag (OOB) Daten der Random Forest Modelle. Zuletzt beschreibt $\bar{y}$ den Mittelwert der Holzvorräte laut Inventur. n ist die Anzahl der untersuchten Variablen, also die Anzahl an Inventurpunkten.

2.4.3. Moving-Window Approach

Der Moving-Window Approach ist ein bewegliches Fenster, dass sich um seine Größe, respektive Durchmesser, über den Raster bewegt (Immitzer et al. 2016) . Mit dem Befehl *focal statistics* in ArcMap (10.7.1) werden die benötigten Zielvariablen von den Random Forest Modellen auf dem nDSM berechnet. Wichtig ist hierbei, dass die Fläche (Durchmesser) annähernd ident mit dem der Random Forest Modelle ist (White et al. 2017) . Bei der Prozessierung der ALS- Punktwolke kann man nicht wie bei *focal statistics* in ArcMap einen Kreis als Moving Window auswählen. In lidR (Roussel J. R., 2022) wird mit dem Befehl *pixel_metrics ()* eine Seitenlänge eines Quadrates verlangt (res). Somit muss hierbei darauf geachtet werden, dass die Flächengröße annähernd ident zu den Kreisradiusn des Random Forests Modelles ist (White et al. 2017).

2.4.4. Waldmaske

Die Modelle wurden mit Pixelwerte aus nDSM auf Waldflächen angewendet. Der Forstbetrieb hat in seiner Bestandeskarte in der Attributtabelle angegeben, welche Flächen Wald und nicht Wald sind. Aus dieser Information wurde ein Raster erstellt (ArcGIS *„rasterize“*) und im Anschluss den nicht Wald- Flächen „NO DATA“ zugewiesen (ArcGIS *„set null“*). Als Vergleich wurde die Waldkarte („Waldmaske“) vom Bundesforschungszentrum für Wald (BFW) angefordert. Diese Karte wurde über mehrere Jahre aus der laufenden Österreichischen Waldinventur (ÖWI) erstellt. Diese Karte ist kostenlos verfügbar (Bauerhansl, 2022). Die Waldmaske wurde im Koordinatensystem Gauß-Krüger M34 geliefert und musste in das UTM (EPSG: 32633) reprojiziert werden. Im Anschluss wurde die Auflösung von rund 1 m x 1 m auf 0.5 m x 0.5 m erhöht. Für die finale Erstellung der Holzvorratskarte wurde die Waldmaske des Forstbetriebes verwendet da dort auch kleine Flächen wie Wildwiesen, Holzplätze, etc. als Nicht- Wald gekennzeichnet wurden. Forststraßen sind i.S.d. ForstGes 1975 Wald. Da auf diesen Flächen für die Modellierung aber kein Holz stockt, und der Höhenwert der Straße 0 m beträgt, so manipuliert eine Forststraße beispielsweise die Mittelwertberechnung in einem definierten Kreis. Aus dieser Überlegung heraus werden auch Forststraßen als Nicht- Wald geführt und diesen Pixeln wurde ebenfalls „NO DATA“ zugewiesen.

2.4.5. Unterscheidung Laub- zu Nadelholz

Um die entwickelten Algorithmen, die für Laub- und Nadelholz getrennt ermittelt wurden, auf die Fläche zu bekommen muss diese in Laub- und Nadelholz unterschieden werden. Für diesen und viele weitere Zwecke wurde das Projekt Land Information System Austria (LISA) ins Leben gerufen. Der LISA- Datensatz ist kostenlos verfügbar (GeoVille, 2022). Es wurden im speziellen auf Daten der „LandCoverMap“ zugegriffen, also Ergebnisse der Landbedeckungskarte aus dem Jahr 2016. Die Daten wurden aus Satellitendaten, Orthofotos, Höhenkarten und ALS- Befliegungen erstellt (GeoVille, 2022). Dieser Datensatz wurde sortiert und eine Laubholzkarte für das Untersuchungsgebiet erstellt. Es wurde unterstellt, dass alle Flächen, die nicht von der Laubholzkarte überdeckt werden als Nadelholzflächen zu behandeln sind. Die errechneten Vorräte mit dem Nadelholzalgorithmus wurden invers mit den Laubholzkarten maskiert. Das bedeutet, dass nur Pixel der Nadelholzflächen mit Daten versehen sind. Flächen die als Laubholz klassifiziert wurden werden nun bei den Laubholzvorratskarten maskiert. Diese beiden Karten (Laubholzvorratskarte exklusive Nadelholz und Nadelholzvorratskarte exklusive Laubholz) werden im Anschluss als Mosaik zusammengeführt (R- Paket *terra „mosaic“*).

3. Ergebnisse der Holzvorratsmodellierung

Random Forest Modelle

Für beide Varianten (Metriken aus nDSM und aus der ALS- Punktwolke) wurden Random Forest Modelle erstellt. Es wurden verschiedene Kreisradii, Kreiskombinationen und in Baumarten unterteilte Modelle berechnet. Die statistischen Kenngrößen geben Aufschluss über die Genauigkeit der Modelle.

Die Stammzahl am Plot und auch die extrahierten Werte aus den Sentinel- 2 Aufnahmen und die NDVI-Werte wurden in keinem Modell gewählt. Das bedeutet, dass diese Metriken keine Verbesserung der Vorhersagegenauigkeit liefern.

3.1. Gesamtmodelle

Bei den Gesamtmodellen wurde nicht zwischen Laub-, Nadel- oder Mischwald unterschieden und alle WZP- Punkte wurden für die Modellierung herangezogen (Anzahl = 874). Die besten Ergebnisse konnte hier das Gesamtmodell aus der Punktwolke mit kombinierten Kreisen erzielen (R^2 = 0.60, $RMSE_{rel}$ = 34.05 %. Im Vergleich dazu erzielte die Kreiskombination mit Metriken abgeleitet aus dem nDSM ein R^2 von 0.57 und einen $RMSE_{rel}$ von 35.75 %. Auffallend in Tabelle 2 ist, dass die Kreiskombination in jedem Szenario (nDSM & Punktwolke) besser als das beste Einzelkreismodell ist. Vergleicht man die Gesamtmodelle anhand der von der Feature Selection gewählten Metriken, ist auffallend, dass die zusätzliche Verwendung von Intensitätsmetriken eine Verbesserung der Modelle bewirkt. Dies bezieht sich auf Einzelkreise sowie Kreiskombinationen. Die zusätzliche Information der Sentinel Daten plus NDVI hat hingegen keine Verbesserung gebracht. Bei den Einzelkreisen brachte das beste Ergebnis laut Tabelle 2 der 12.5 m Kreisradius mit Metriken abgeleitet aus der Punktwolke hervor. Das R^2 liegt bei 0.58 und der $RMSE_{rel}$ bei 35.19 %. Der R^2 des 12.5 m Kreisradius des nDSM liegt bei 0.55 und der $RMSE_{rel}$ bei 36.58 %. Dass die Feature Selection bei beiden Varianten (nDSM & Punktwolke) denselben Einzelkreisradius gewählt hat ist Zufall. In der grafischen Auswertung (Variable Importance Plot) der unterschiedlichen Modelle (Kreiskombination & Einzelkreis) sind die von der Feature Selection ausgewählten Variablen ersichtlich. Abbildung 6 zeigt die Modellvariablen der Kreiskombination des Gesamtmodells auf dem nDSM. In Abbildung 7 sind die Variablen der Kreiskombination des Gesamtmodells aus der Punktwolke mit deren Einfluss auf den MSE abgebildet. Abbildung 8 zeigt den Variable Importance Plot am 12.5 m Kreis der Punktwolke. In Abbildung 9 ist der Variable Importance Plot am 12.5 m Kreis des nDSM abgebildet. Die grafischen Vergleiche der unterschiedlichen Modelle findet sich in Abbildung 10.

Interessant ist, dass bei dem Gesamtmodell mit Kreiskombination und Einzelkreisradius als wichtigste Variable (%IncMSE) die Kurtosis der Intensität (k12_ikurt) des 12.5 m Kreises gewählt wurde. Die grafische Darstellung der Out- Of- Bag (OOB) Daten (Scatterplots) in Abbildung 10 macht mithilfe der Ausgleichsgeraden (rot) ersichtlich, dass alle erstellten Modelle niedrige Holzvorräte überschätzen und hohe Holzvorräte unterschätzen.

Tabelle 2: Auflistung der statistischen Auswertung der Gesamtmodelle mit abgeleiteten Metriken aus nDSM (oben) und ALS Punktwolke (unten).

Gesamtmodelle nDSM

Kreisradius [m]	MSE	RMSE [m³/ha]	$RMSE_{rel}$ [%]	bias	$bias_{rel}$ [%]	R^2
5.0	23472.89	153.21	38.74	-1.76	-0.44	0.50
7.5	22260.44	149.20	37.73	-2.04	-0.51	0.52
10.0	21275.97	145.86	36.88	-1.49	-0.38	0.54
12.5	20926.21	144.66	**36.58**	-1.13	-0.29	**0.55**
15.0	21291.94	145.92	36.90	-2.06	-0.52	0.54
17.5	21427.77	146.38	37.01	-1.54	-0.39	0.54
20.0	21229.47	145.70	36.84	-2.02	-0.51	0.54
22.5	21556.00	146.82	37.12	-1.18	-0.30	0.53
25.0	22098.49	148.66	37.59	-1.25	-0.32	0.52
27.5	22343.80	149.48	37.80	-1.71	-0.43	0.52
Kombiniert	19987.77	141.38	**35.75**	-3.01	-0.76	**0.57**

Gesamtmodelle Punktwolke

Kreisradius [m]	MSE	RMSE [m³/ha]	$RMSE_{rel}$ [%]	bias	$bias_{rel}$ [%]	R^2
5.0	22046.77	148.48	37.40	-2.73	-0.69	0.52
7.5	20842.76	144.37	36.37	-2.66	-0.67	0.55
10.0	20202.98	142.14	35.81	-2.88	-0.73	0.56
12.5	19508.54	139.67	**35.19**	-1.92	-0.48	**0.58**
15.0	19651.66	140.18	35.31	-1.64	-0.41	0.57
17.5	20824.76	144.31	36.35	-1.72	-0.43	0.55
20.0	20383.86	142.77	35.97	-2.14	-0.54	0.56
22.5	21442.48	146.43	36.89	-1.94	-0.49	0.54
25.0	21768.89	147.54	37.17	-2.41	-0.61	0.53
27.5	21881.39	147.92	37.26	-2.12	-0.53	0.53
Kombiniert	18268.21	135.16	**34.05**	-3.30	-0.83	**0.60**

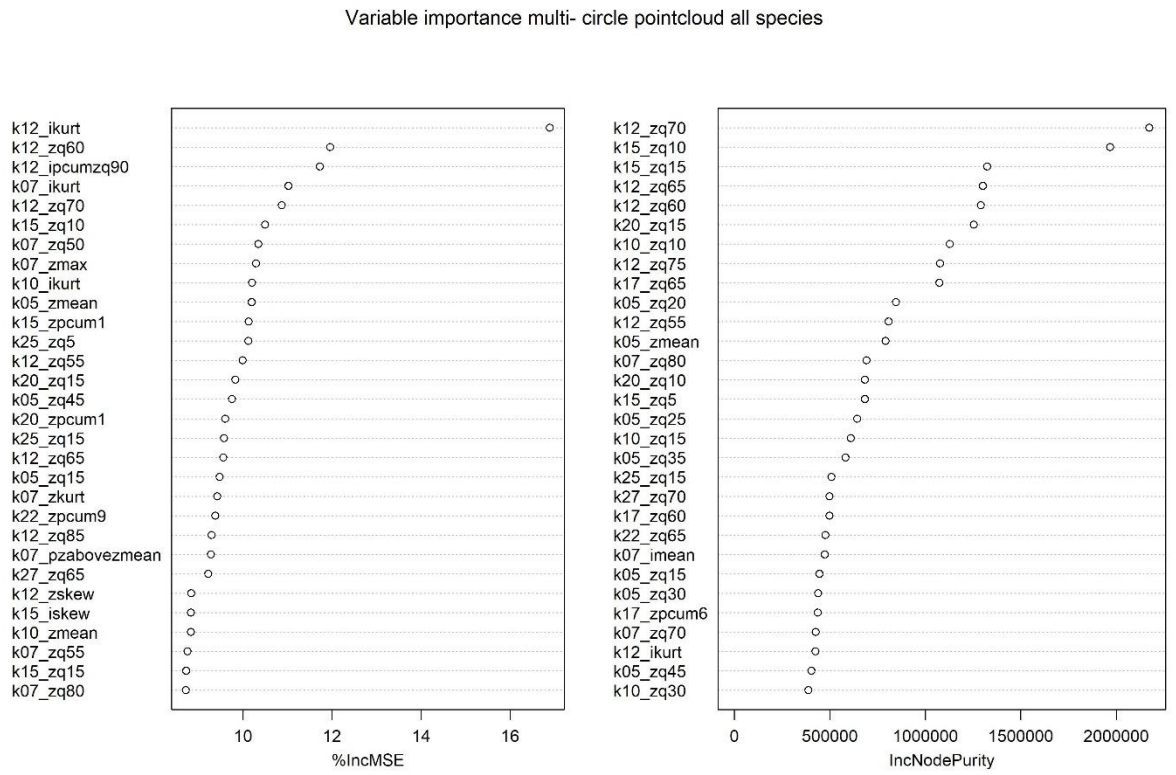

Abbildung 6: Variable Importance Plot des Gesamtmodells mit Metriken abgleitet aus der Punkwolke. „k“ steht für Kreis, „i“ = Intensitätsmetriken und „z“ = Höhenmetriken.

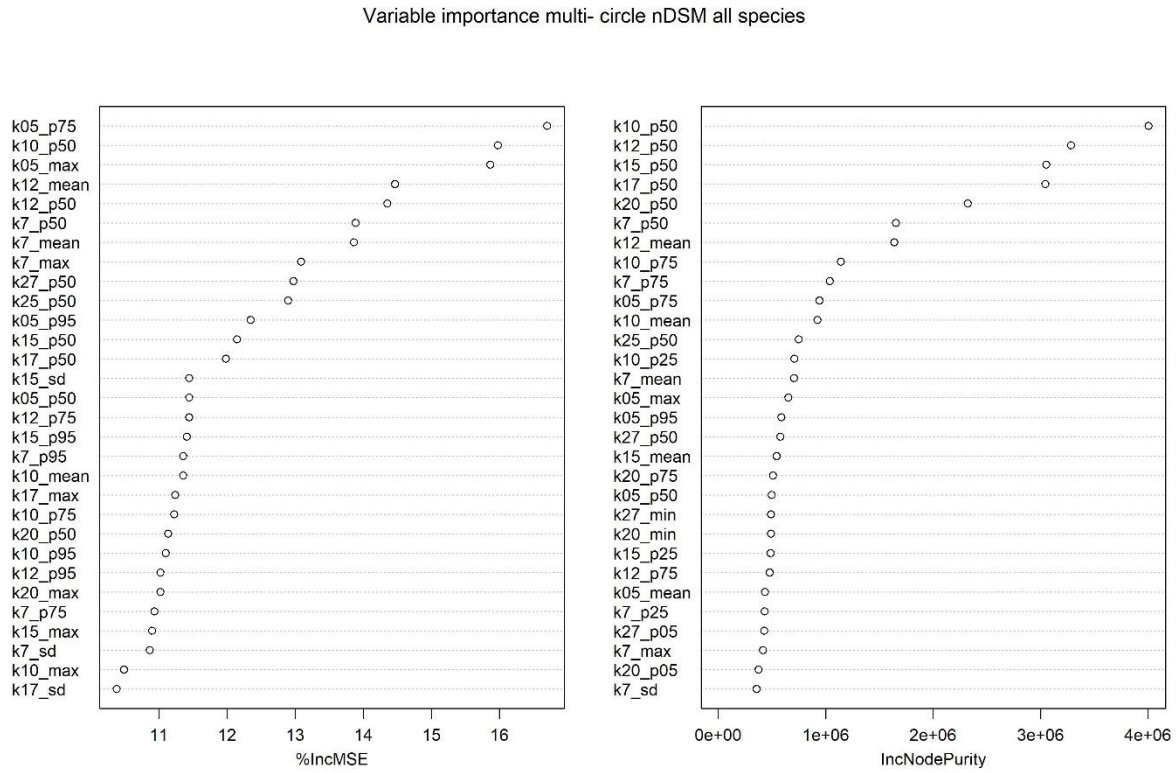

Abbildung 7: Variable Importance Plot der Kreiskombination mit Höhenmetriken abgeleitet aus dem nDSM. "k" = Kreis, die Zahl "05" für den Kreisradius in m. "p" sind Prozent der Quantil.

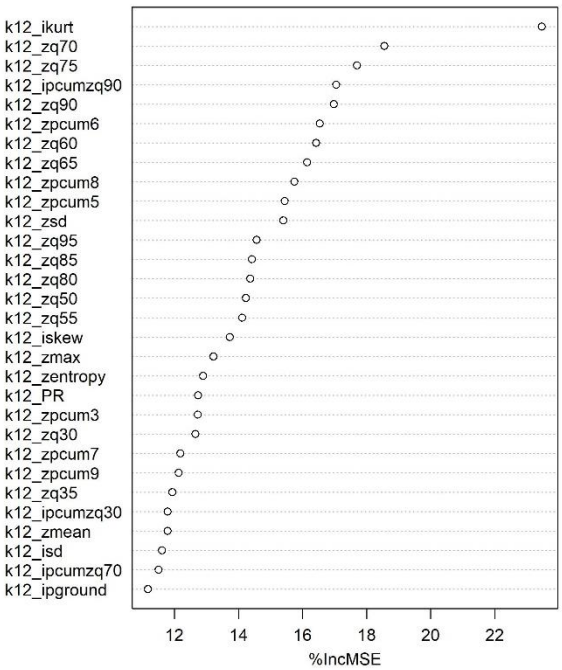

Abbildung 8: Variable Importance Plot nach der Feature Selection am 12.5 m Kreisradius der Punktwolke. "k" = Kreis, "12" = Kreisradius in m (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken

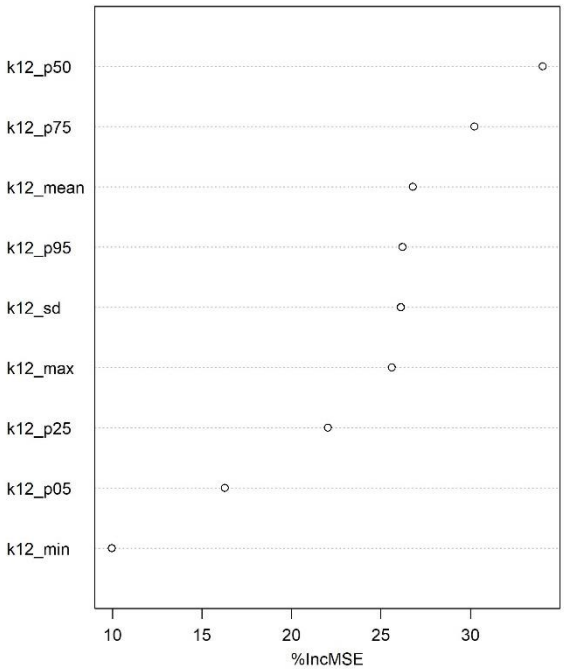

Abbildung 9: Variable Importance Plot nach der Feature Selection am 12.5 m Kreisradius des nDSM. "k" = Kreis, "12" = Kreisradius in m (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken

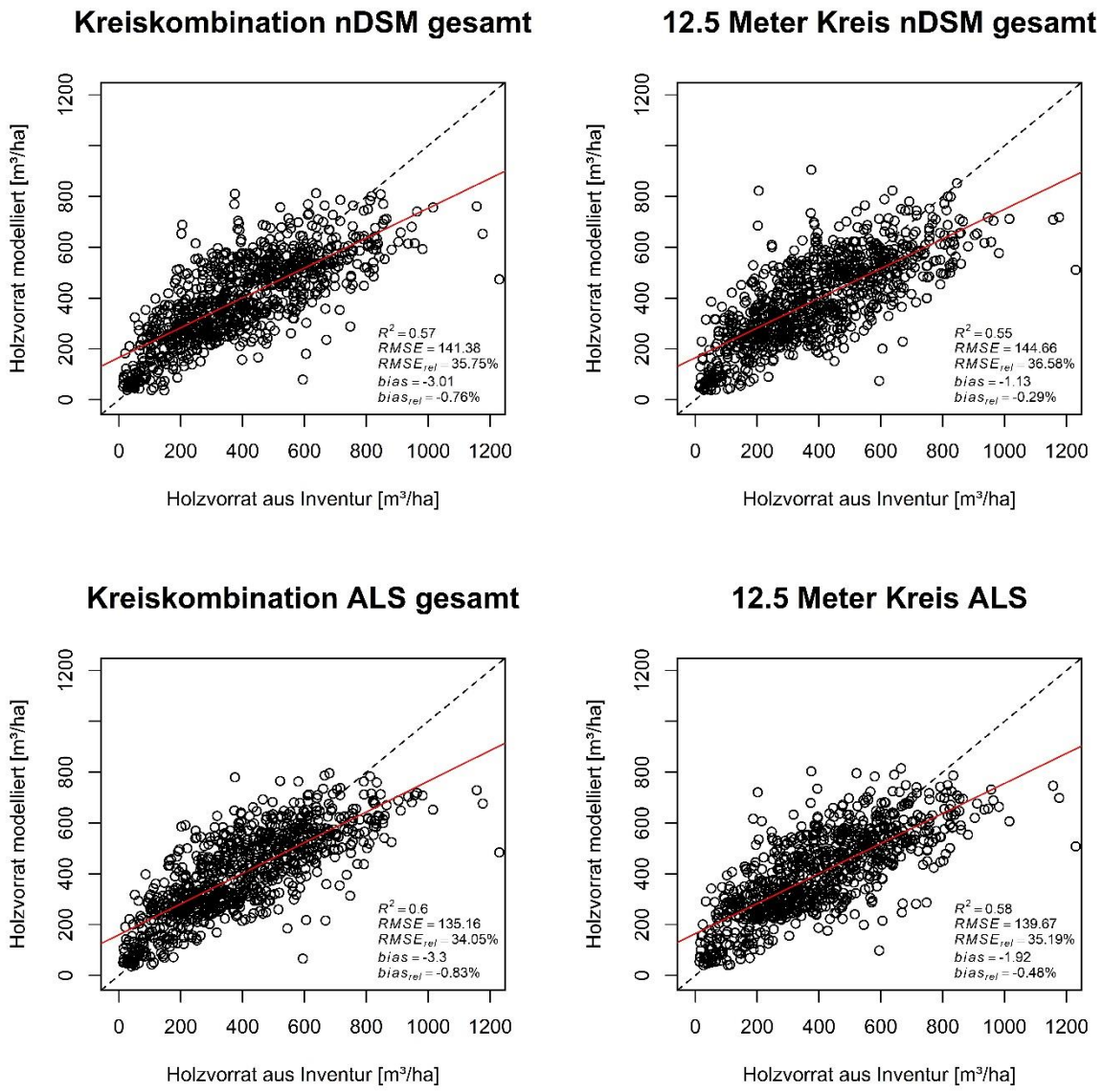

Abbildung 10: Scatterplots mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts) der Gesamtmodelle. Punktierte Linie = 1:1 Gerade, rote durchgezogene Linie = Ausgleichsgerade.

3.2. Nadelholzmodelle

Die Kombination der Kreisradius erzielen in beiden Varianten (Metriken aus nDSM & ALS) die besten Ergebnisse (Tabelle 3). Es zeigt sich, dass die Stratifizierung ebenso einen Vorteil gegenüber dem Gesamtmodell hat. Die statistische Auswertung hat bei der Kreiskombination mit den ALS- Metriken ein R^2 von 0.67 und einen $RMSE_{rel}$ von 32.28 %. Im Vergleich dazu hat das Random Forest Modell mit den Höhenmetriken aus dem nDSM ein R^2 von 0.66 bei einem $RMSE_{rel}$ von 32.17 %. Somit ist bei den Nadelholzmodellen ersichtlich, dass die Zusatzinformation, die aus der Punktwolke gewonnen werden, sich nicht gravierend auf die Modellgüte auswirken. Bei den Einzelkreisen erzielen die beiden Varianten ebenfalls ähnliche Genauigkeiten, es wurden aber bei der Variante der Kreisradius von 10 m (R^2 = 0.63, $RMSE_{rel}$ = 34.50 %) und bei den abgeleiteten Metriken aus der Punktwolke der Radius von 12.5 m (R^2 =0.65, $RMSE_{rel}$ = 33.26 %) als beste Einzelkreisvariante errechnet. Bei den Modellen ist ebenso auffallend, dass niedrige Holzvorräte systematisch überschätzt und hohe Vorräte unterschätzt werden. Die Variable Importance Plots von den Kreiskombinationen aus der Punktwolke (Abbildung 11) und dem nDSM (Abbildung 12) stellen visuell den Einfluss der von der Feature Selection gewählten Variablen am MSE dar.

Beim Nadelholzmodell Kreiskombination ALS ist zu beachten, dass die Feature Selection die Kurtosis der Intensität und auch das Kronenvolumen als wichtige Variablen zur Modellgüte gewählt hat (Abbildung 11). Auch beim ALS Einzelkreismodell (Abbildung 13) ist die Kurtosis der Intensität am Variable Importance Plot angeführt. Diese Variablen können beim nDSM nicht aufscheinen, da Metriken zur Intensität nur bei aus der Punktwolke ableitbar sind. In Abbildung 14 ist der Variable Importance Plot mit den gewählten Metriken der Feature Selection von Kreisradius 12.5 m abgebildet. Die grafischen Vergleiche der unterschiedlichen Modelle findet sich in Abbildung 15.

Tabelle 3: Statistische Auswertung der erstellten Nadelholzmodelle. Oben sind die Einzelkreise und Kombination der aus dem nDSM abgeleiteten Modelle ersichtlich. Unten sind die Modelle (Einzelkreis & Kombination) der Metriken aus der Punktwolke angeführt.

		Nadelholzmodelle nDSM				
Kreisradius [m]	MSE	RMSE [m³/ha]	$RMSE_{rel}$ [%]	bias	$bias_{rel}$ [%]	R^2
5.0	20315.52	142.53	36.88	0.66	0.17	0.58
7.5	18591.86	136.35	35.28	-0.18	-0.05	0.61
10.0	17774.19	133.32	**34.50**	-0.92	-0.24	**0.63**
12.5	17814.32	133.47	34.54	-0.02	0.00	0.63
15.0	18170.17	134.80	34.88	-1.63	-0.42	0.62
17.5	18722.41	136.83	35.41	-1.49	-0.38	0.61
20.0	18169.91	134.80	34.88	-1.73	-0.45	0.62
22.5	18554.36	136.21	35.25	-1.04	-0.27	0.61
25.0	19556.69	139.85	36.19	0.17	0.04	0.59
27.5	19383.49	139.22	36.03	0.27	0.07	0.59
Kombiniert	16427.11	128.17	**33.17**	-1.03	-0.27	**0.66**

		Nadelholzmodelle Punktwolke				
Kreisradius [m]	MSE	RMSE [m³/ha]	$RMSE_{rel}$ [%]	bias	$bias_{rel}$ [%]	R^2
5.0	18877.52	137.40	35.47	-1.75	-0.45	0.60
7.5	18308.17	135.31	34.93	-1.27	-0.33	0.62
10.0	17920.19	133.87	34.56	-2.84	-0.73	0.62
12.5	16592.75	128.81	**33.26**	-0.10	-0.02	**0.65**
15.0	17390.05	131.87	34.05	-1.22	-0.32	0.63
17.5	17719.06	133.11	34.37	-1.32	-0.34	0.63
20.0	19338.82	139.06	35.90	-1.04	-0.27	0.60
22.5	18634.41	136.51	35.24	-0.09	-0.02	0.61
25.0	19446.81	139.45	36.00	0.22	0.06	0.59
27.5	19703.44	140.37	36.24	0.20	0.05	0.59
Kombiniert	15630.76	125.02	**32.28**	-0.87	-0.22	**0.67**

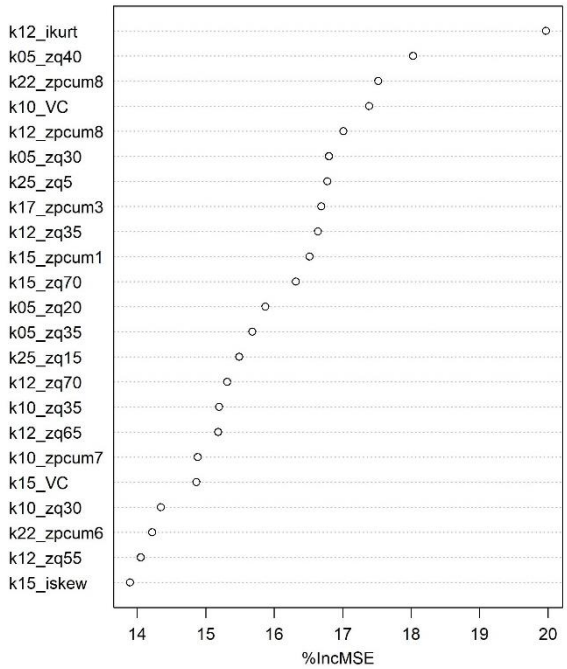

Abbildung 11: Variable Importance Plot der Kreiskombination an der Punktwolke. "k" = Kreis, "12" = Kreisradius in m (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken, „VC „= Kronenvolumen

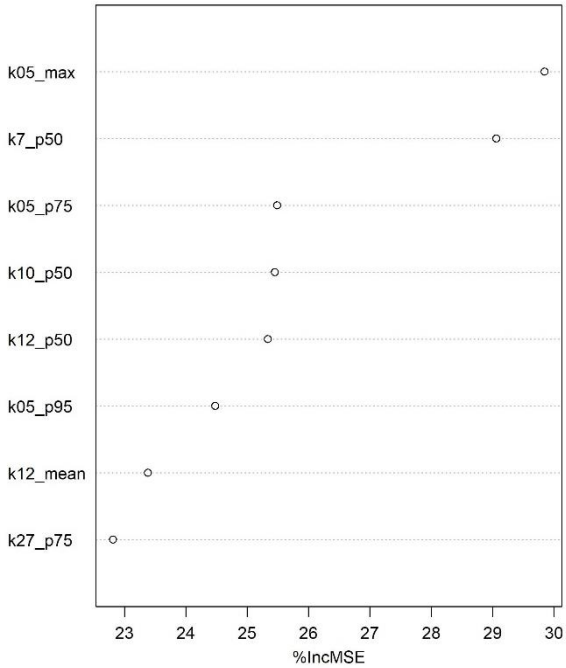

Abbildung 12: Variable Importance Plot der Kreiskombination am nDSM abgeleiteten Höhenmetriken. "k" = Kreis, "05" = Kreisradius [m], "p75" = 75 % Quantil

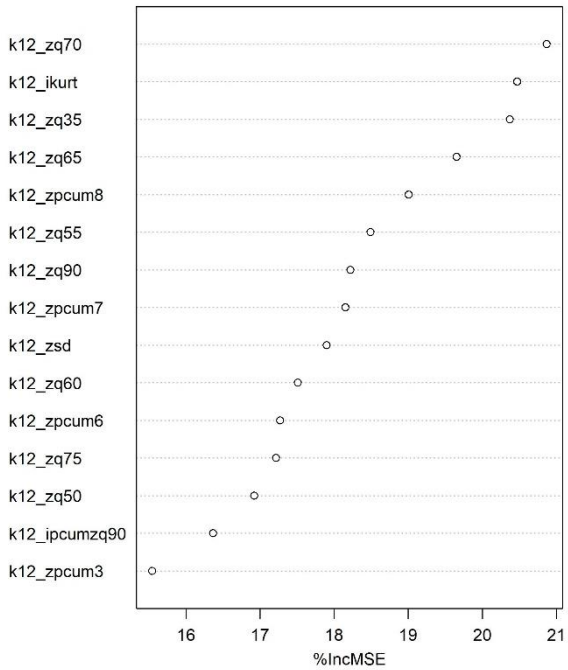

Abbildung 13: Variable Importance Plot nach der Feature Selection am 12.5 m Kreisradius der Punktwolke. "k" = Kreis, "12" = Kreisradius in m (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken

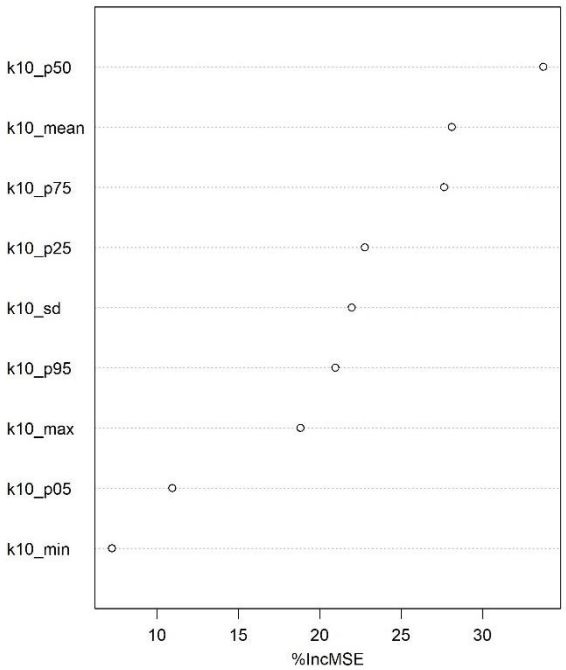

Abbildung 14: Variable Importance Plot am Kreisradius 10 m Nadelholz des nDSM. "k" = Kreis, "10" = Kreisradius [m], "p50" = 50 % Quantil (Median)

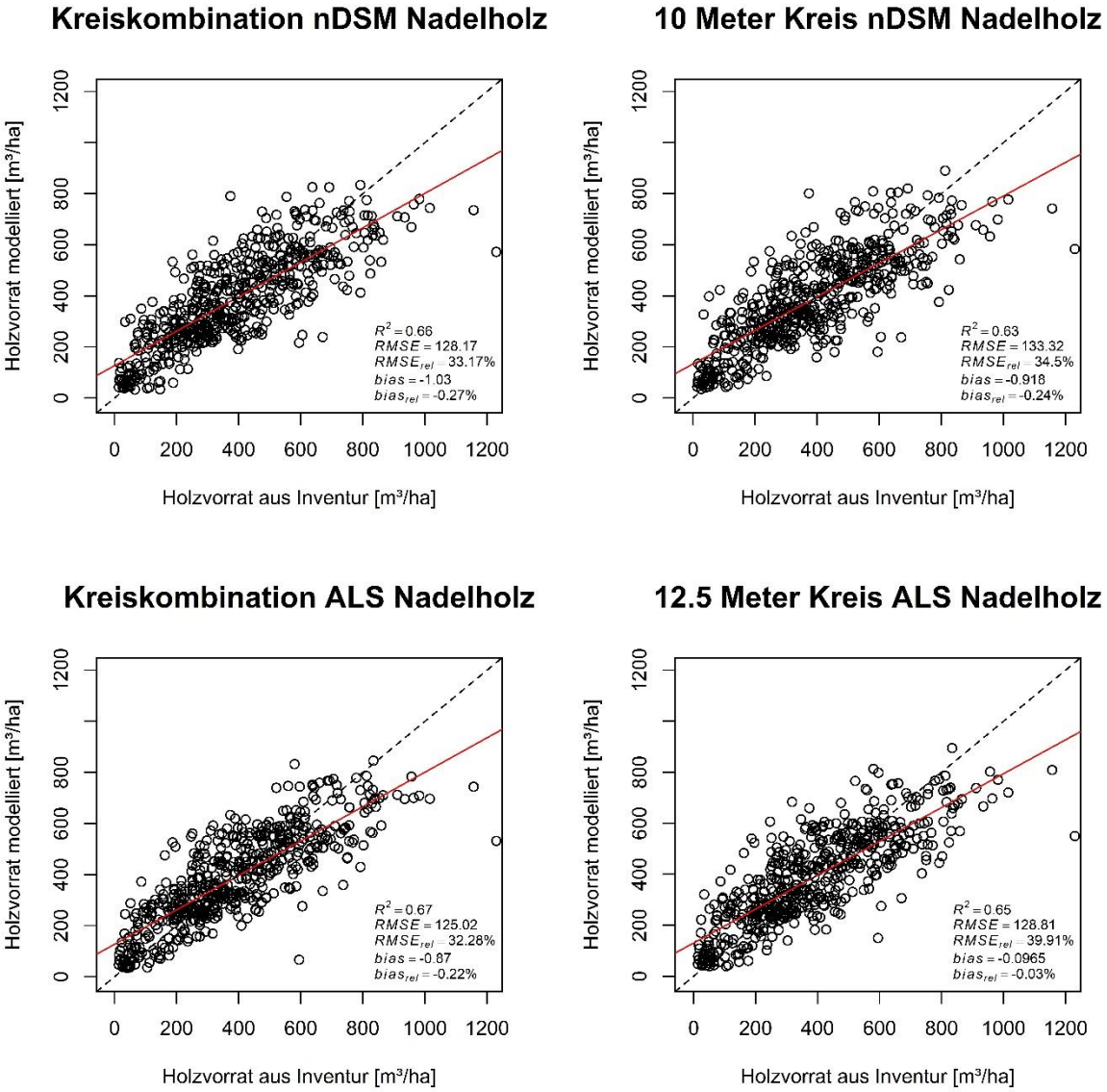

Abbildung 15: Scatterplots der Nadelholzmodelle mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts). Punktierte Linie = Diagonale im Plot, rote durchgezogene Linie = Ausgleichsgerade. Der Einzelkreisradius vom besten ALS-Modell (unten rechts) beträgt 12.5 m.

3.3. Laubholzmodelle

Bei der Stratifizierung der Inventur wurden 129 WZP- Punkte mit einem Laubholzanteil von größer oder gleich 80 % der Klasse Laubholz zugewiesen. Die Ergebnisse haben hier überrascht, da das errechnete Modell aus der ALS Punktwolke mit einem Kreisradius von 15 m deutlich die anderen Modelle übertroffen hat (R^2 = 0.64 bei $RMSE_{rel}$ = 34.49 %). Im Vergleich dazu schnitt das Modell des Einzelkreises des nDSM (bestes Modell Radius = 20 m) mit einem R^2 von 0.48 und $RMSE_{rel}$ von 41.76 % schlechter ab (Tabelle 4). Auch das Kreiskombinationsmodell Laubholz aus dem nDSM konnte mit einem R^2 von 0.47 und $RMSE_{rel}$ von 42.30 % nicht die Vorhersagegenauigkeit wie das Einzelmodell aus der Punktwolke erzielen. Das Laubholzmodell mit Kreiskombinationen aus den ALS- Daten kommt auf ein R^2 von 0.61 bei einem $RMSE_{rel}$ von 36.31 %. Die gewählten Variablen des Kombinationsmodelles errechnet mit Metriken aus der Punktwolke ist in Abbildung 16 ersichtlich. Bei dem Modell *„Kreiskombination Punktwolke Laubholz“* ist die Punktanzahl der Kreisradien 10 m und 12.5 m in die Auswahl der Feature Selection gekommen. Die statistische Auswertung aller Laubholzmodelle ist Tabelle 4 zu entnehmen. Der Variable Importance Plot für die Kreiskombination am nDSM ist in Abbildung 17 abgebildet. Für das beste Einzelkreismodell Laubholz wurde von der Feature Selection der Kreisradius von 15 m aus der Punktwolke ausgewählt, das zugehörige Variable Importance Plot ist in Abbildung 18. Der Kreisradius mit der höchsten Güte am nDSM hat 20 m, die gewählten Variablen des Modelles sind in Abbildung 19 abgebildet. Der Scatterplot in Abbildung 20 zeigt, dass die Laubholzmodelle niedrige Vorräte überschätzen und hohe Vorräte unterschätzen.

Nach der Auswertung der Feature Selection wurde nach Abbildung 18 die Intensitätsmetrik „ipcumzq70“ als größter Einflussfaktor auf die Genauigkeit (%IncMSE) gewählt. Diese Metrik ist der Prozentsatz der reflektierten Intensität unterhalb des 70. Perzentils der Höhe. Diese Intensitätsmetrik hat auch Einzug in die Modellauswahl der Kreiskombination der Punktwolke gefunden. Dem Feature Selection Plot aus Abbildung 16 ist zu entnehmen, dass diese Intensitätsmetrik nicht denselben Einfluss auf die Genauigkeit hat.

Tabelle 4: Statistische Auswertung der erstellten Laubholzmodelle. Oben sind die Einzelkreise und Kombination der aus dem nDSM abgeleiteten Modelle ersichtlich. Unten sind die Modelle (Einzelkreis & Kombination) der Metriken aus der Punktwolke angeführt.

			Laubholzmodell nDSM			
Kreisradius [m]	MSE	RMSE [m³/ha]	RMSE rel [%]	bias	bias rel [%]	R²
5.0	26562.99	162.98	48.30	0.10	0.03	0.32
7.5	23758.03	154.14	45.68	-3.45	-1.02	0.39
10.0	22664.51	150.55	44.61	-3.30	-0.98	0.41
12.5	21442.36	146.43	43.39	-2.40	-0.71	0.44
15.0	21694.03	147.29	43.65	0.71	0.21	0.44
17.5	20610.35	143.56	42.54	-1.74	-0.52	0.47
20.0	19858.95	140.92	**41.76**	3.02	0.89	**0.48**
22.5	21028.86	145.01	42.97	-1.02	-0.30	0.45
25.0	21368.25	146.18	43.32	-2.37	-0.70	0.45
27.5	21979.37	148.25	43.93	-0.41	-0.12	0.43
Kombiniert	20377.00	142.75	**42.30**	-1.04	-0.31	**0.47**

			Laubholzmodell Punktwolke			
Kreisradius [m]	MSE	RMSE [m³/ha]	RMSE rel [%]	bias	bias rel [%]	R²
5.0	20765.71	144.10	42.50	-0.18	-0.05	0.46
7.5	18620.21	136.46	40.25	1.90	0.56	0.52
10.0	15465.58	124.36	36.68	-1.36	-0.40	0.60
12.5	16862.90	129.86	38.30	-0.71	-0.21	0.56
15.0	13671.38	116.92	**34.49**	-3.13	-0.92	**0.64**
17.5	19497.97	139.64	41.18	-0.32	-0.09	0.49
20.0	18431.42	135.76	40.04	-3.67	-1.08	0.52
22.5	18456.32	135.85	40.07	2.82	0.83	0.52
25.0	16846.19	129.79	38.28	-3.55	-1.05	0.56
27.5	17990.42	134.13	39.56	0.23	0.07	0.53
Kombiniert	15158.89	123.12	**36.31**	2.17	0.64	**0.61**

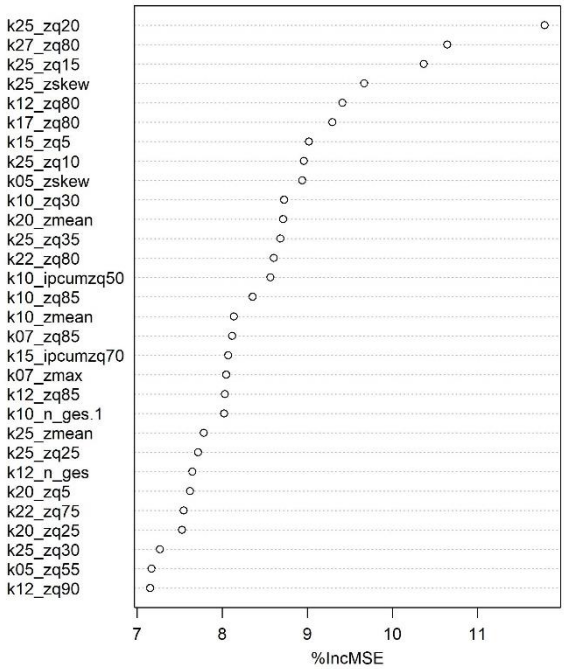

Abbildung 16: Variable Importance Plot der Kreiskombination Laubholzmodell der Punktwolke. "k" = Kreis, "12" = Kreisradius [m], "i" = Intensitätsmetriken, "z" = Höhenmetriken, „n_ges“ = Punktanzahl

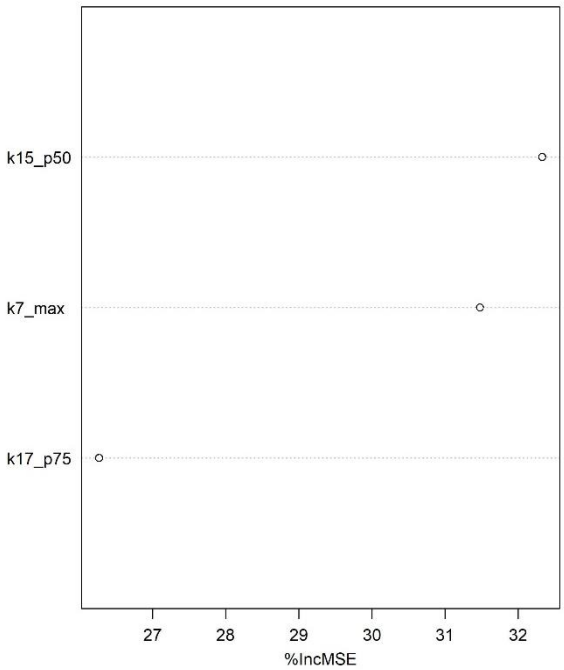

Abbildung 17: Variable Importance Plot Kreiskombination nDSM Laubholz. "k" = Kreis, "15" = Kreisradius in m,"p50" = 50 % Quantil der Höhe.

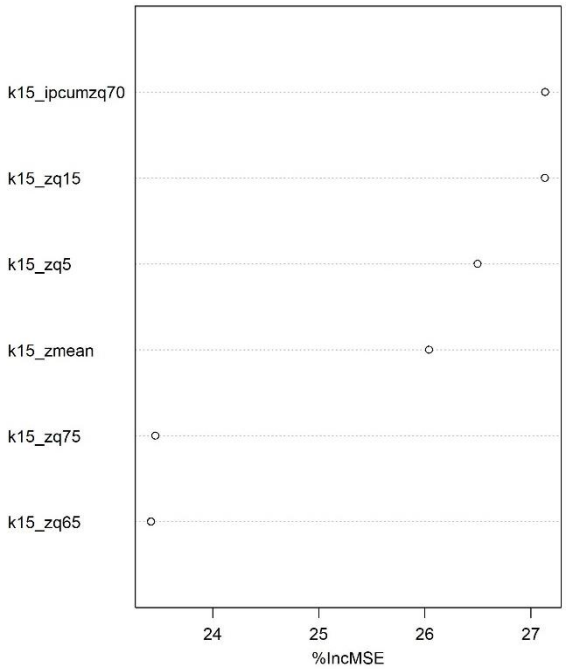

Abbildung 18: Variable Importance Plot am Kreisradius 15 m Laubholz der Punktwolke. "k" = Kreis, "15" = Kreisradius [m], „z“ = Höhenmetrik, „mean“ = Mittelwert, „ipcumzq70“ = Prozentsatz der zurückgegebenen Intensität unterhalb des 70. Perzentils der Höhe

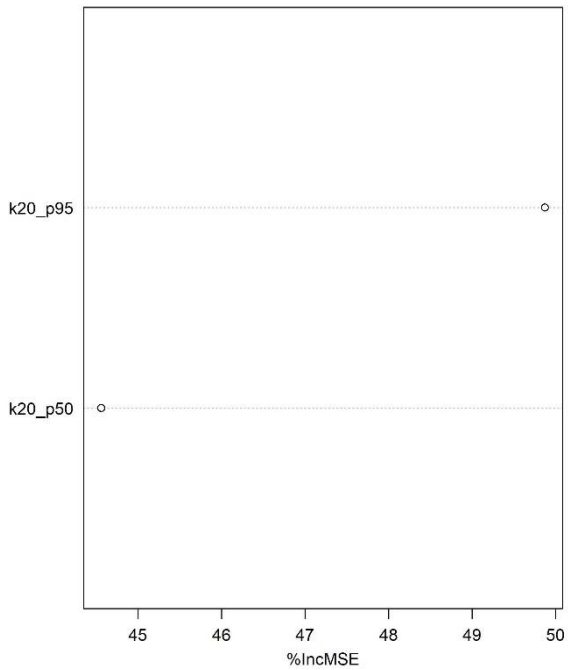

Abbildung 19: Variable Importance Plot 20 m Radius Einzelkreis Laubholz mit Höhenmetriken aus dem nDSM abgeleitet. "k" = Kreis, "20" = 20 m Kreisradius, "p95" = 95 % Quantil der Höhe.

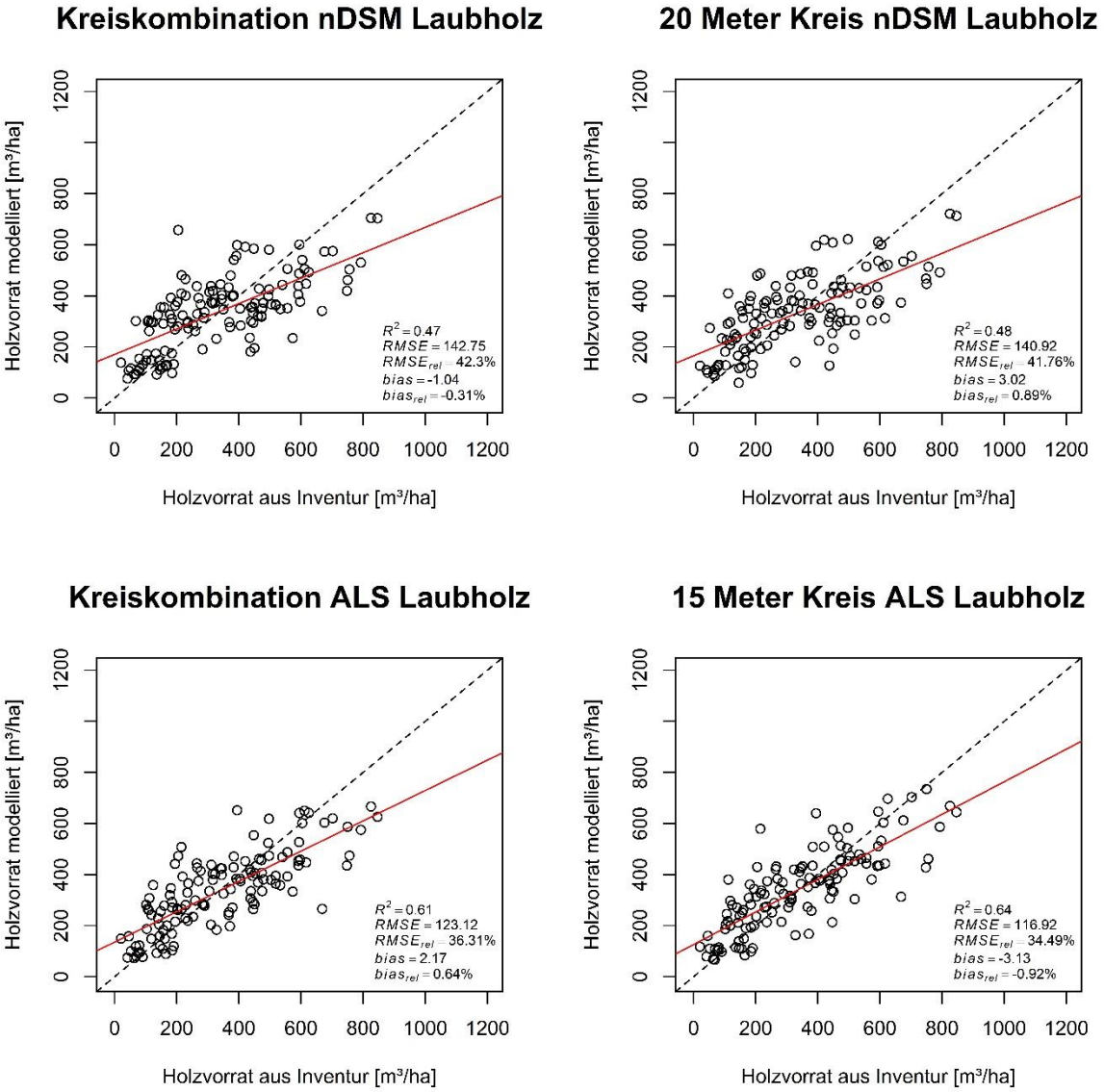

Abbildung 20: Scatterplots der Laubholzmodelle mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts). Punktierte Linie = Diagonale im Plot, rote durchgezogene Linie = Ausgleichsgerade.

3.4. Mischwaldmodelle

WZP-Punkte, welche weder dem Nadel- noch Laubholzstratum zugeordnet werden konnten, wurden als Mischwald klassifiziert. Dieses Stratum war mit 190 Punkten größer als das Laubholzstratum (129 WZP-Punkte). Die Modelle mit Kreiskombinationen konnten sich auch hier durch ihre höhere statistische Güte von den Einzelkreismodellen abheben. Dies ist in Der Scatterplot in Abbildung 25 visualisiert, dass die Mischwaldmodelle niedrige Vorräte überschätzen und hohe Vorräte unterschätzen.
Tabelle 5 ersichtlich. Das Kreiskombinationsmodell mit Metriken aus der Punktwolke erzielte ein R^2 von 0.41 und einen $RMSE_{rel}$ von 33.23 %, die gewählten Metriken sind im Variable Importance Plot in Abbildung 21 ersichtlich. Das Kreiskombinationsmodell mit Höhenmetriken abgeleitet aus dem nDSM weist ein R^2 von 0.41 und einen $RMSE_{rel}$ von 33.72 % auf, die Variablen des Modells sind im Variable Importance Plot in Abbildung 22 ersichtlich. Bei den Einzelkreismodellen weist das Mischwaldmodell mit Metriken aus der Punktwolke eine höhere Güte (r = 10 m, R^2 = 0.40, $RMSE_{rel}$ = 33.43 %) als das Modell mit Höhenmetriken aus dem nDSM (r = 15 m, R^2 = 0.37, $RMSE_{rel}$ =34.76 %). Der Variable Importance Plot des Einzelkreismodells mit Metriken aus der Punktwolke ist in Abbildung 23 abgebildet. Die ausgewählten Variablen des 15 m Mischwaldmodells des nDSM, ersichtlich im Variable Importance Plot, findet sich in Abbildung 24 wider. Der Scatterplot in Abbildung 25 visualisiert, dass die Mischwaldmodelle niedrige Vorräte überschätzen und hohe Vorräte unterschätzen.

Tabelle 5: Statistische Auswertung der erstellten Mischwaldmodelle. Oben sind die Einzelkreise und Kombination der aus dem nDSM abgeleiteten Modelle ersichtlich. Unten sind die Modelle (Einzelkreis & Kombination) der Metriken aus der Punktwolke angeführt.

			Mischmodell nDSM			
Kreisradius [m]	MSE	RMSE [m³/ha]	$RMSE_{rel}$ [%]	bias	$bias_{rel}$ [%]	R^2
5.0	27048.64	164.46	35.65	-2.79	-0.60	0.34
7.5	27486.99	165.79	35.94	0.08	0.02	0.34
10.0	27729.83	166.52	36.10	-0.67	-0.15	0.33
12.5	25757.63	160.49	34.79	1.14	0.25	0.37
15.0	25707.07	160.33	**34.76**	-1.18	-0.26	**0.37**
17.5	27786.66	166.69	36.14	-0.94	-0.20	0.33
20,0	28302.62	168.23	36.47	3.40	0.74	0.32
22.5	27573.33	166.05	36.00	0.70	0.15	0.33
25.0	26131.92	161.65	35.04	-1.51	-0.33	0.36
27.5	26493.55	162.77	35.29	-2.11	-0.46	0.35
Kombiniert	24199.23	155.56	**33.72**	0.33	0.07	**0.41**
			Mischmodell Punktwolke			
Kreisradius [m]	MSE	RMSE [m³/ha]	$RMSE_{rel}$ [%]	bias	$bias_{rel}$ [%]	R^2
5.0	24624.87	156.92	33.77	-3.81	-0.82	0.39
7.5	25172.03	158.66	34.14	0.53	0.11	0.38
10.0	24137.42	155.36	**33.43**	-3.17	-0.68	**0.40**
12.5	25428.14	159.46	34.32	-1.91	-0.41	0.37
15.0	24741.98	157.30	33.85	-2.18	-0.47	0.38
17.5	28078.38	167.57	36.06	-3.24	-0.70	0.31
20.0	27356.39	165.40	35.59	-2.07	-0.45	0.32
22.5	28166.36	167.83	36.12	-3.29	-0.71	0.31
25.0	27689.84	166.40	35.81	0.19	0.04	0.31
27.5	30342.24	174.19	37.49	-2.78	-0.60	0.26
Kombiniert	23841.42	154.41	**33.23**	-3.02	-0.65	**0.41**

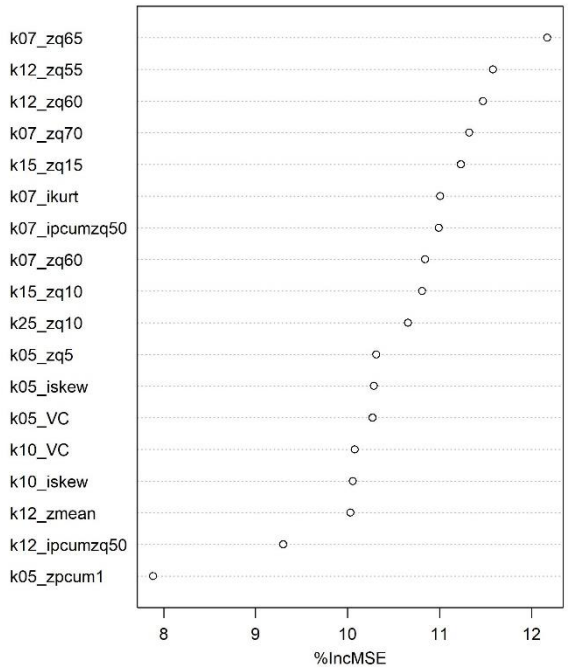

Abbildung 21: Variable Importance Plot der Kreiskombination Mischwaldmodell der Punktwolke. "k" = Kreis, "12" = Kreisradius [m], "i" = Intensitätsmetriken, "z" = Höhenmetriken, „n_ges“ = Punktanzahl

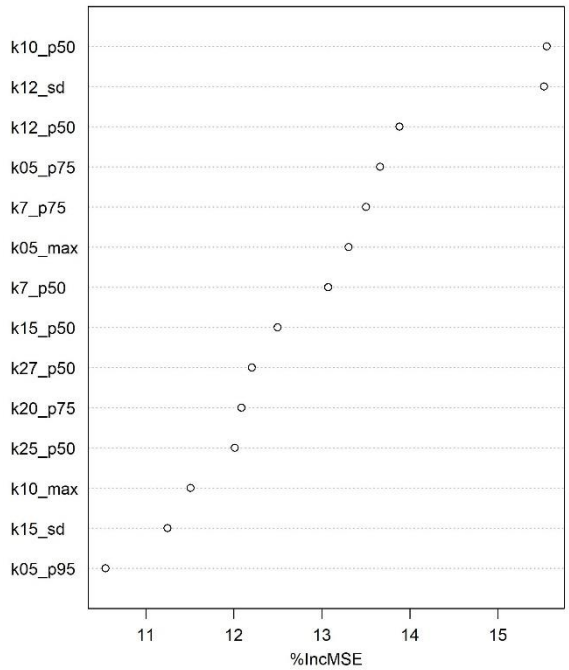

Abbildung 22: Variable Importance Plot der Kreiskombination Mischwaldmodell des nDSM. "k" = Kreis, "12" = Kreisradius [m], "i" = Intensitätsmetriken, "z" = Höhenmetriken, „n_ges“ = Punktanzahl

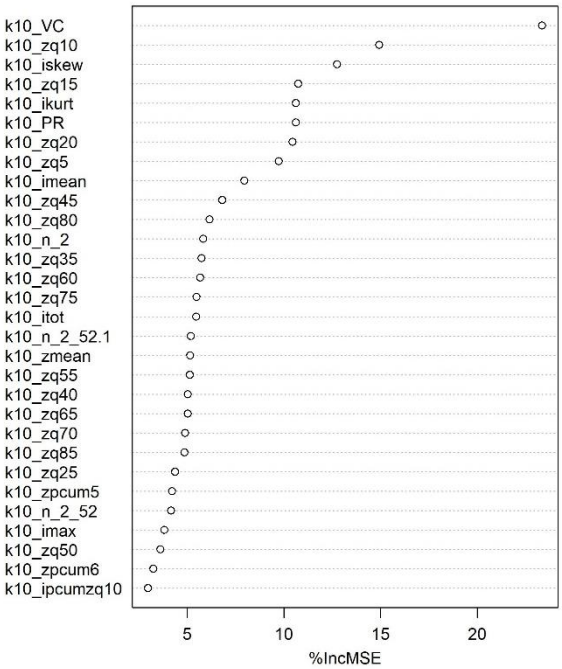

Abbildung 23: Variable Importance Plot 10 m Radius Einzelkreis Mischwald mit Metriken aus der Punktwolke abgeleitet. "k" = Kreis, "10" = 10 m Kreisradius, „z“ = Höhenmetriken, „i“ = Intensitätsmetriken, „VC“ = Kronenvolumen, „PR“ = Durchdringungsrate, „n“ = Punktanzahl

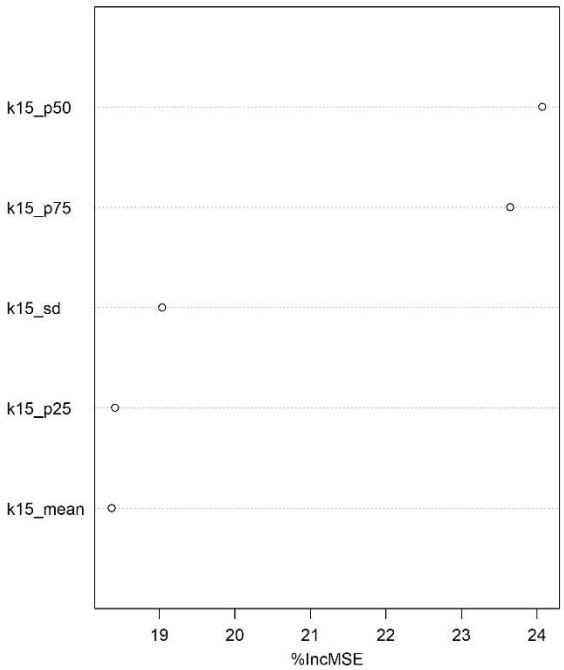

Abbildung 24: Variable Importance Plot 15 m Radius Einzelkreis Mischwald mit Höhenmetriken aus dem nDSM abgeleitet. "k" = Kreis, "20" = 20 m Kreisradius, "p95" = 95 % Quantil der Höhe.

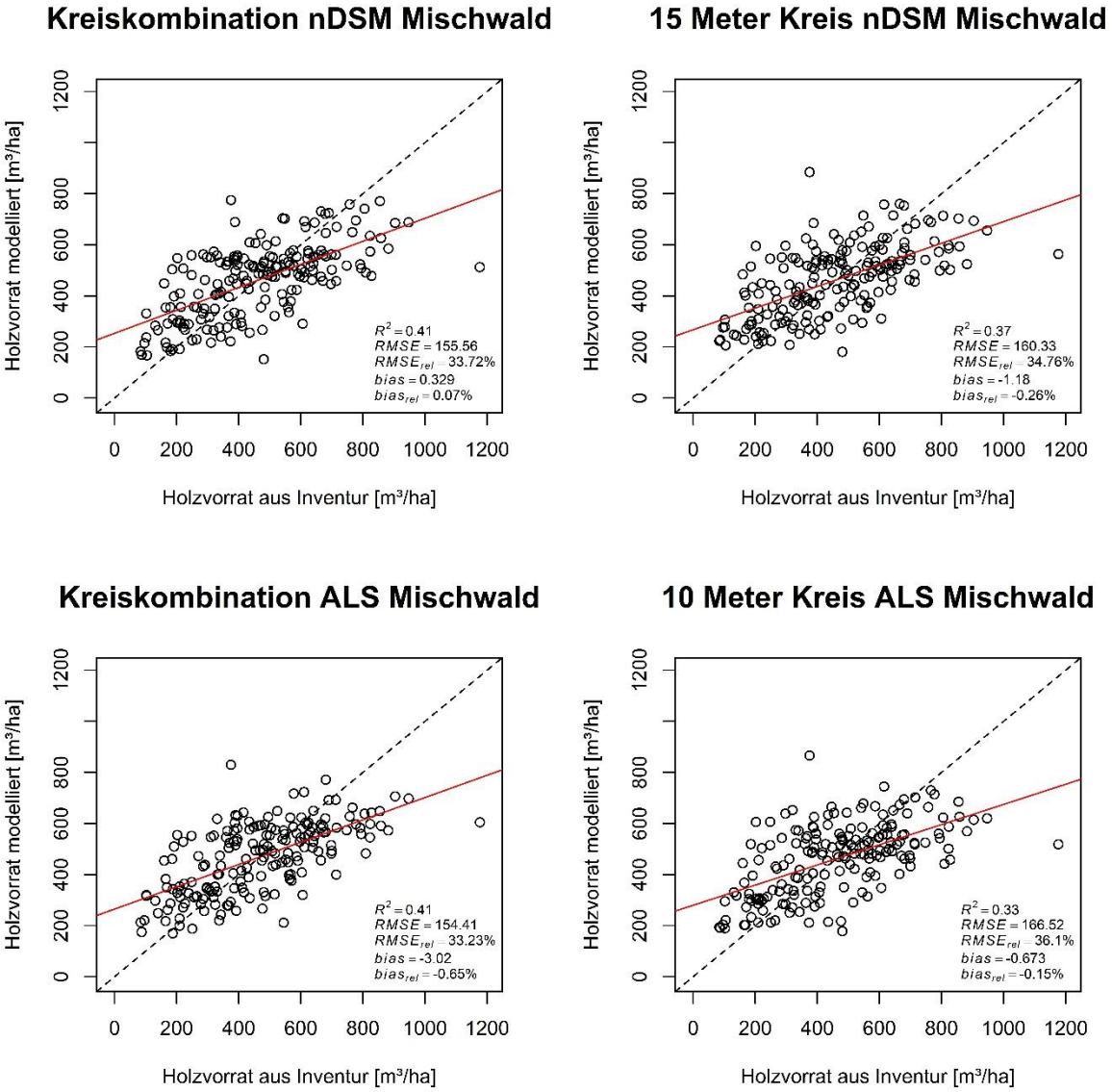

Abbildung 25: Scatterplots der Mischwaldmodelle mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts). Punktierte Linie = Diagonale im Plot, rote durchgezogene Linie = Ausgleichsgerade.

3.5. Validierung der extrahierten Metriken

Für die Erstellung der Holzvorratskarten wurde ein wall-to-wall mapping Ansatz von Immitzer et al. (2016) angewendet. Dabei wurden die Metriken der Kreiskombinationen Nadel- und Laubholzmodelle aus dem nDSM abgeleitet. Die Random Forest Modelle aus dem nDSM haben eine ähnliche Genauigkeit wie die Modelle aus der Punktwolke, sind aber wesentlich einfacher umzusetzen. Für das Nadelholzmodell wurden mit *focal statistics* (ArcMap, Esri 2021) acht Metriken berechnet. Für das Laubholzmodell drei Metriken. Als Output gibt es einen Raster für jede Metrik mit der Größe und der Auflösung des Eingaberasters (Eingaberaster = nDSM; 0.5 m x 0.5 m). Um zu testen, ob der wall-to-wall Ansatz funktioniert, wurden die Ergebnisse der Extracts („extract“, terra) mit den auf der Fläche angewendeten focal statistics an den Inventurpunkten verglichen.

3.6. Erstellung der Holzvorratskarte

Um die finale Holzvorratskarte (Shapefile) zu erstellen wurde aus der Vorratskarte (Raster) die Pixelwerte aus den Beständen extrahiert (*terra „extract“*). Es liegt nun eine Liste auf, in der jeder Pixelwert (Holzvorrat) im Bestand gespeichert ist. Für jeden Bestand wurde der Mittelwert der Holzvorräte berechnet. Die Random Forest Modelle schätzen den Holzvorrat pro Hektar. Um den Bestandesvorrat zu ermitteln, muss nun mit der Flächengröße in Hektar multipliziert werden. Die Flächengröße wurde aus der Größe der Polygone im Shapefile berechnet.

Im nächsten Schritt wurden die 8 (Nadelholz) und 3 (Laubholz) Raster zu einem großen Raster zusammengeführt (*stack*, package terra, R). Wichtig hierbei ist, dass die Rasterwerte denselben Namen wie die Metriken aufweisen, ansonsten findet das Random Forest Modell die Variablen nicht um den Holzvorrat schätzen zu können. Auch die Ausmaße sowie die Pixelgröße muss ident sein. Wenn diese Bedingungen erfüllt sind, lässt sich mit *predict* (terra, R) der Vorrat schätzen. Dies wurde auf der ganzen Fläche mit dem Nadel- sowie Laubholzmodell durchgeführt. Als Output erhält man zwei Holzvorratskarten, die nur auf Waldflächen den Holzvorrat bestimmt haben. Im nächsten Schritt werden beide Holzvorratskarten mit den LISA- Laubholzmasken maskiert (*terra „mask“*). Der Befehl *„mask“* hat in den Einstellungen die Option, die Maske mit *inverse = TRUE* zu setzen. Dabei werden die Pixel unter der Maske mit „NO DATA“ überschrieben und die Flächen außerhalb behalten ihre Pixelwerte bei. Dies wurde bei der Nadelholzvorratskarte angewandt, um Laubholzflächen ausschließen zu können. Bei der Laubholzvorratskarte wurde Nadelholz maskiert, also mit „NO DATA“ versehen. Im Weiteren Verarbeitungsprozess werden die beiden maskierten Karten zusammengeführt (*terra „mosaik“*). Das Ergebnis, die finale Holzvorratskarte, ist in Abbildung 26 ersichtlich, bei der jeder Pixelwert den Holzvorrat für Nadel- oder Laubholz beinhaltet.

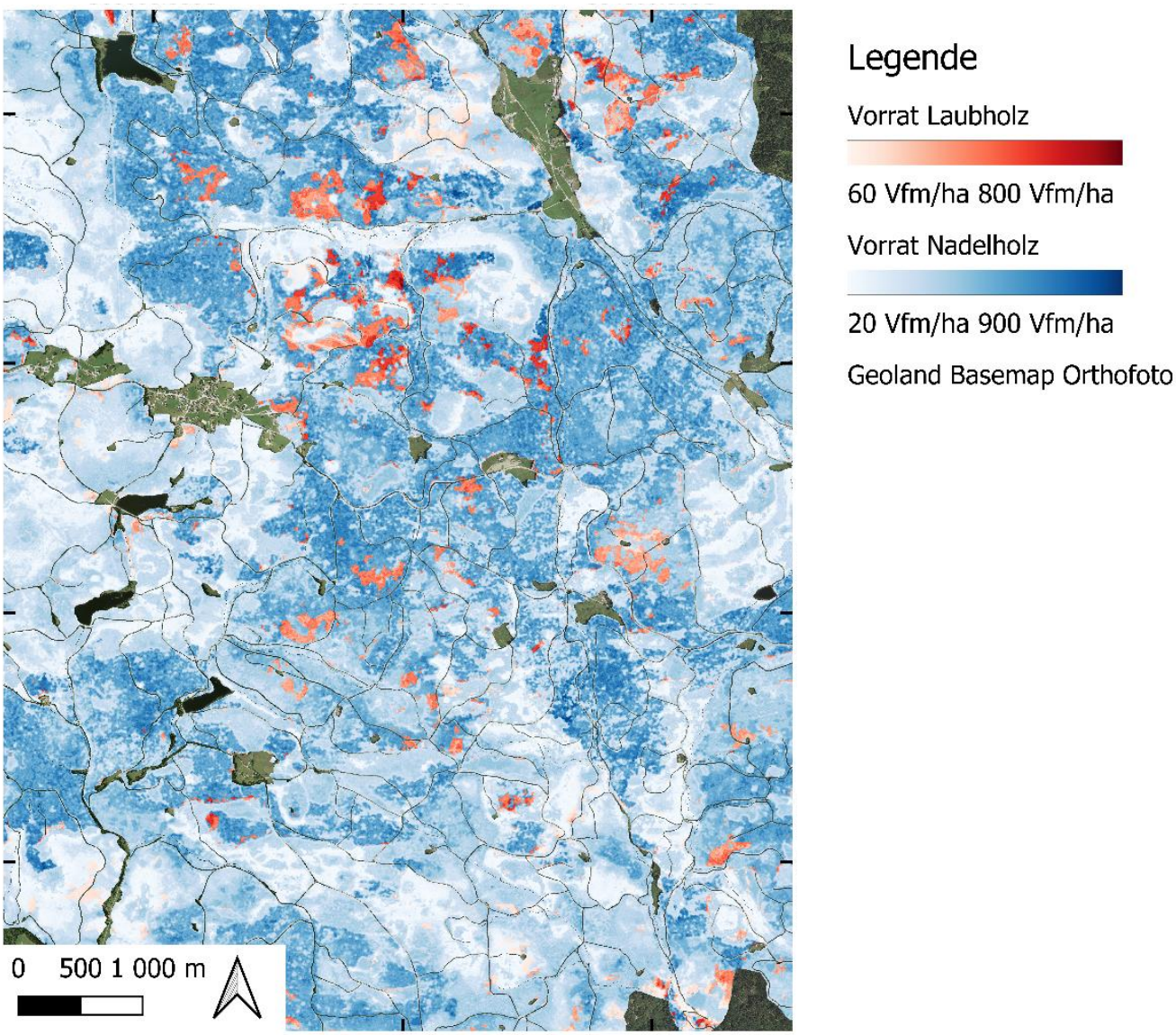

Abbildung 26: Holzvorratskarte zusammengestellt aus Nadel- und Laubholzvorrat für das Untersuchungsgebiet. Rot eingefärbte Pixel sind Laubholzbestände. Nadelholzbestände wurden blau eingefärbt. Nur Waldflächen wurden eingefärbt. Als Hintergrund dient das Orthophoto von Geoland Basemap.

3.7. Berechnung des Bestandesvorrats

Für die Berechnung des Bestandesvorrats wurde die Bestandeskarte des Betriebes herangezogen. Im Bearbeitungsvorgang werden aus der Holzvorratskarte aus Abbildung 26 die Pixelwerte extrahiert (R- Paket *terra „extract“*) und in Listen abgespeichert. Nun liegt die Anzahl der Pixelwerte pro Bestandeseinheit vor. Wichtig ist zu erwähnen, dass ein Pixelwert den Holzvorrat pro Hektar und nicht am Ort repräsentiert. Mit den Holzvorräten in der Bestandeseinheit (= Polygon) wurde der Mittelwert berechnet (White et al. 2017). Es liegt nun die Information auf, wie hoch die Vorratshaltung im Bestand pro Hektar ist (Vorratsfestmeter pro Hektar). Diese Information wird auch bei der klassischen Inventur mit der WZP erhoben und kann nun verglichen werden. Da das Interesse auch dem Bestandesvorrat gilt, werden die Vorräte pro Hektar mit den Bestandesflächen in Hektar multipliziert. Die neuen Informationen (Vorrat pro Hektar und Bestandesvorrat) werden nun der Attributtabelle der Bestandeskarte angehängt und kann wie in Abbildung 27 ersichtlich, mit QGIS und anderen Programmen visualisiert werden.

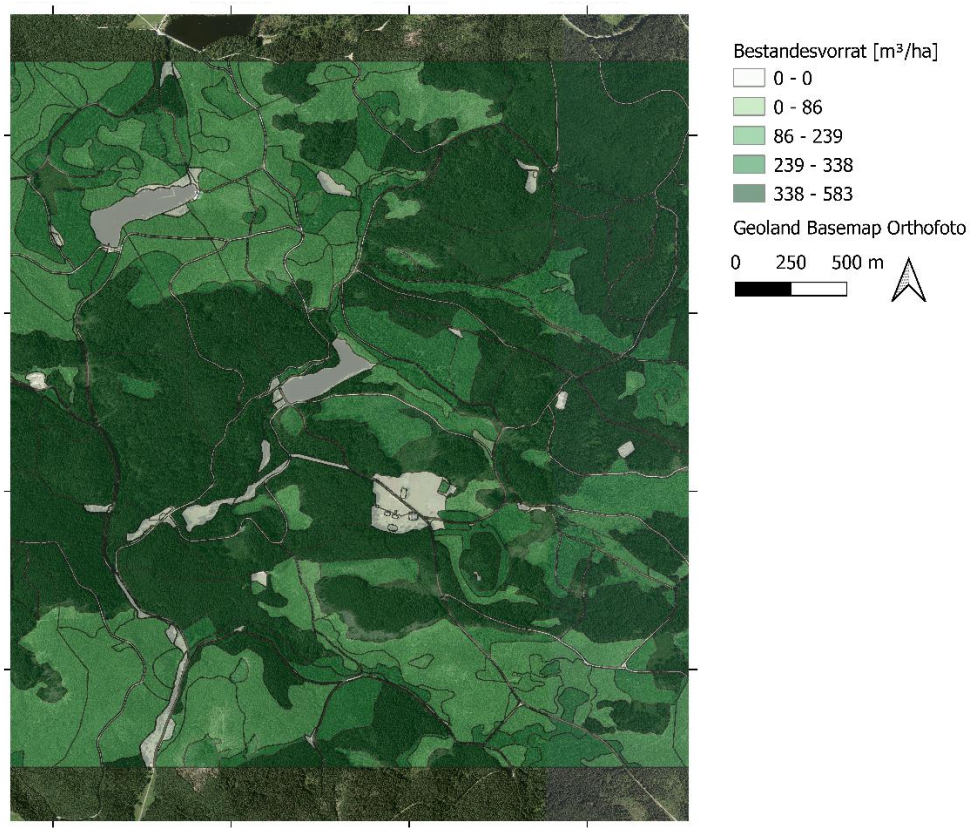

Abbildung 27: Bestandeskarte mit Information des Holzvorrats. In weiß eingefärbte Flächen entsprechen keinem Wald wie beispielsweise Wiesen, Wege oder Teiche. Als Hintergrund dient das Orthophoto von Geoland Basemap.

4. Diskussion

4.1. Daten und Material

Die vorliegenden ALS Daten hatten eine sehr hohe Auflösung (von 16 Punkten pro m²). Mit dieser hohen Punktdichte lassen sich sehr genaue Oberflächenmodelle sowie Geländemodelle erstellen. Zusätzlich zur Höheninformation können aus der Punktwolke auch Informationen zur Intensität, Punktdichte, Punktanzahl und Punktverteilung gewonnen werden. Diese Informationen sind entscheidend, um weitere, detailliertere Analysen über Untersuchungsgebiete zu treffen. Die Analysen reichen von der Unterscheidung der Baumarten mit ALS Daten (Briechle et al. 2021) bis zur Ermittlung des Stammdurchmessers (Amiri et al. 2017). Bei der Erkennung von Einzelstämmen ist jedoch eine noch höhere Punktdichte (300 Punkte pro m²) essenziell. Die Möglichkeiten zur Analyse von ALS-Punktwolken sind sehr vielseitig, in der Forstwirtschaft werden sie immer öfter eingesetzt, um schneller qualitative Informationen von der Waldfläche zu generieren. Gerade nach Sturm- oder Schneebruchereignissen muss auf eine schnelle und koordinierte Holzabfuhr gesetzt werden, und dafür ist es wichtig zu wissen, wie hoch die Schadholzmenge ist. Für diese Art der Anwendung wurden Modelle von Nothdurft et al. (2021) entwickelt.

Für die Erstellung der vorliegenden Arbeit war einer der ersten Arbeitsschritte die Inventurpunkte (Winkelzählprobe) von 2014 an den richtigen Ort zu verschieben. Das normalisierte digitale Oberflächenmodell (nDSM) stammt aus dem Jahr 2018. In diesem Zeitraum gab es Holznutzungen und Nutzungen aus Kalamitäten. Bäume sind zusätzlich eingewachsen. Dieser Umstand kann nicht garantieren, dass alle Punkte lagerichtig verschoben wurden, da Bäume oder Bestände fehlten. Für eine präzisen Ersteinrichtung ist es essenziell, dass der Zeitraum zwischen Erhebung der Daten und Befliegung so kurz wie möglich gehalten wird (White et al. 2017). Dies ist bei der Ersterstellung der Modelle zu beachten, und kann bei Folgeanwendung der Modelle auf der Fläche vernachlässigt werden. Die Verschiebung der WZP- Punkte ist nur bei der Ersteinrichtung einmalig nötig. Bei der Folgeanwendung kann darauf zurückgegriffen werden.

Das Stichprobendesign spielt eine große Rolle für die Erstellung des area- based- approach (flächenbezogener Ansatz, ABA) – es muss ein Flächenbezug hergestellt werden. Dies ist insofern schwierig, da die Winkelzählprobe keinen eindeutigen Flächenbezug hat. Eine Methode die sich auch in dieser Arbeit bewährt hat, ist die Kombination von mehreren Einzelkreisen mit unterschiedlichen Radien (Immitzer et al. 2016). In Kirchhoefer et al. (2019) wurde der Median von den Stichprobenmittelpunkten zu den Bäumen errechnet und als Radius festgelegt. Die Ergebnisse in den Modellen „Gesamtmodell“, „Nadelholz“ und „Mischwald“ der vorliegenden Arbeit haben gemeinsam, dass die Kreiskombination besser ist als das jeweilige Einzelkreismodell. Außer im Modell „Laubholz“, dort übertrifft das Einzelkreismodell abgeleitet aus dem nDSM und der Punktwolke das Kreiskombinationsmodell geringfügig.

Bei Levick et al. (2016) wurde getestet wie sich die Größe der Probekreise auf die Modellgenauigkeit auswirkt. Trainingsflächen mit der Größe von 1 ha brachten die besten Modelle hervor (R^2 = 0.92; RMSE = 50.57 m³). Modelle, die mit Flächen von 0.05 ha (500 m², Radius von 12.61 m) kamen über eine Genauigkeit von R^2 = 0.68 und RMSE = 101.01 m³ nicht hinaus. Des Weiteren wurde festgestellt, dass bei den kleinen Versuchsflächen von 0.05 ha die Güte (R^2, RMSE) der Modelle nicht mehr zunimmt ab einer Stichprobengröße von 350. Diese Werte sind mit dieser Arbeit insofern vergleichbar, da die Auswahl des besten Einzelkreises im Gesamt- und Nadelholzmodell in beiden Varianten (nDSM, Punktwolke) auf den Kreis mit einem Radius von 12.5 m fiel. Das beste Modell war das Nadelholzmodell der Punktwolke mit einem R^2 von 0.65 und einem RMSE von 128.81 m³ ($RMSE_{rel}$ = 33.26 %)

Die Stratifizierung der Daten hat eine Verbesserung der Modellgenauigkeit gebracht. Eine Unterteilung in Laub-, Nadel- und Mischwald macht aber von Beginn an nur dann einen Sinn, wenn es die Möglichkeit zur Klassifizierung der Punktwolke oder des Untersuchungsgebietes in diese Straten gibt. Die LISA Daten (Kozak et al. 2017) sind sehr gut für eine Unterteilung geeignet. Die Unterteilung der Waldtypen in Baumarten würde eine weitere Steigerung der Genauigkeit der Modelle erzwecken (Hill et al. 2018). Dies ist auch ein wichtiger Schritt für die Zukunft, anhand von ALS – Daten Baumarten gezielt zu unterscheiden.

Beim Moving- Window Approach ist es wichtig, dass die Größe und die Form der Trainingsfläche (Stichprobenfläche; Probekreis) gleich bzw. annähernd gleich sein muss wie die der zu schätzenden Fläche (White et al. 2017). Probekreise sind einfacher als Quadrate oder Rechtecke in der Natur zu erheben. Besser geeignet als Winkelzählproben sind definitiv Probekreise.

Eine Waldmaske zu verwenden ist ein wichtiger Schritt für die Erstellung einer Holzvorratskarte. Die Modelle sollen den Holzvorrat nur auf Waldflächen schätzen, und sollen nicht von Höhenwerten von Forststraßen oder Wiesen verzerrt werden. Die getestete Waldmaske vom BFW ist für ganz Österreich frei verfügbar und wird mit der Österreichischen Wald- Inventur (ÖWI) trainiert (Bauerhansl, 2022). Bei der Anwendung der Waldmaske fiel positiv auf, dass der Detailgrad sehr hoch ist. Da Forststraßen i.S.d. ForstGes 1975 Waldflächen sind, sind Forststraßen in dieser Maske als Wald mitinbegriffen. Dies sollte unbedingt bei einer Anwendung der Waldkarte des BFW beachtet werden, denn dies kann zu Verzerrungen in den Ergebnissen kommen. Die Erstellung einer eigenen Waldmaske aus der Bestandeskarte des Forstbetriebes war ein probates Mittel. Diese Karte ist sehr detailliert geführt und bietet eine erstklassige Möglichkeit zur Erstellung einer Waldmaske.

4.2. Diskussion der Ergebnisse

Die Ergebnisse aus Tabelle 6 veranschaulichen, dass es eine Verbesserung der Vorhersagegenauigkeit gibt, wenn zusätzlich zu den Höhenmetriken ALS Metriken hinzugezogen werden. Gerade bei Nadelholz und dem Gesamtwaldmodell ist die Kurtosis der Intensität des 12.5 m Kreisradius die ausschlaggebende Variable für die Genauigkeit (MSE) der Modelle. Diesen Zusammenhang hat die Feature Selection des Random Forest ausgewählt. Alle Modelle, die mit Metriken aus der ALS Punktwolke trainiert wurden, übertreffen die Modelle mit Höhenmetriken aus dem normalisierten digitalen Oberflächenmodell. Random Forest ist ein nichtlinearer Regressionsansatz. Das bedeutet, dass zwischen Zielvariable und erklärender Variable nicht zwingend Zusammenhänge bestehen müssen.

Tabelle 6: Überblick über die Ergebnisse der besten Random Forest Modelle. NDSM = normalisiertes digitales Oberflächenmodell

MODELL	KREISRADIUS	$RMSE_{REL}$ [%]	R^2
GESAMTMODELL PUNKTWOLKE	Kombiniert	34.05	0.60
GESAMTMODELL NDSM	Kombiniert	35.75	0.57
NADELHOLZMODELL PUNKTWOLKE	Kombiniert	32.28	0.67
NADELHOLZMODELL NDSM	Kombiniert	33.17	0.66
LAUBHOLZMODELL PUNKTWOLKE	15 m	36.31	0.61
LAUBHOLZMODELL NDSM	Kombiniert	42.30	0.47

Beim Vergleich der Ergebnisse mit anderen Studien sollten sich die Waldgebiete bzw. die Zusammensetzung der Baumarten der Untersuchungsgebiete ähnlich sein. Dies ist beispielsweise bei Immitzer et al. (2016), Hill et al. (2018) , Stepper et al. (2015) oder auch bei Kirchhoefer et al. (2019) der Fall. In der vorliegenden Arbeit wird im besten Modell ein R^2 von 0.48 und ein $RMSE_{rel}$ von 46.69 % erzielt.

Immitzer et al.(2016) steigerte die Genauigkeit seiner Modelle mit der Zuhilfenahme von Spektraldaten zu den Höhendaten (R^2 = 0.60, $RMSE_{rel}$ = 29.5 %, nur Höhenvariablen: R^2 = 0.53, $RMSE_{rel}$ = 31.7 %). Immitzer et al. (2016) hat ebenfalls Modelle mit einem Multikreisansatz mit der Ausgangslage auf WZP- Punkten trainiert. Kirchhoefer et al.(2019) hat Metriken aus Oberflächenmodellen abgeleitet und mithilfe linearer Modelle forstliche Kennwerte geschätzt. Die Bestimmtheitsmaße und die quadratische mittlere Abweichung unterscheiden sich in den Untersuchungsgebieten (R^2 von 0.55, 0.50, 0.57; $RMSE_{rel}$ = 38.6 %, 42.1 %, 37.9 %). In der Arbeit wurde anstelle eines Multikreisansatzes ein Ansatz mit fixem Kreisradius über die WZP- Punkte gelegt. Hill et al. (2018) hingegen hat Höhenmetriken mit einer Variable für die Baumart kombiniert und schreibt über das Problem der hohen Heterogenität, welche in großen Waldgebieten durch die unterschiedlichen Wuchsklassen auftritt. In der Arbeit von Socha et al. (2020) wurde ein area- based approach auf Weißkiefernbeständen in Polen durchgeführt. Das beste Modell hatte ein R^2 von 0.63 und

einen $RMSE_{rel}$ von 22.8 %. Stepper et al. (2015) erzielte bei der Auswertung der OOB Daten der Random Forest Modelle am Plot $RMSE_{rel}$ von 30.92 %bis 36.02 %.

Hawryło und Wężyk (2018) schätzten den Holzvorrat mit Metriken aus einer photogrammetrischen Punktwolke aus Luftbildern mit respektablen Ergebnissen (R^2 = 0.81; $RMSE_{rel}$ =17.26 %). Vergleiche mit dieser Arbeit sind schwierig, da das Untersuchungsgebiet großteils mit einem homogenen Weißkiefernbestand bestockt ist.

Dass die Aufteilung (Stratifizierung) in Baumarten eine Modellverbesserung bringen kann, wird in der Arbeit von Rahlf et al. (2021) sehr deutlich. Ein Ergebnis von ihnen war, dass die Nationale Forstinventur in Norwegen ausreicht, um Rückschlüsse auf den Holzvorrat auf Betriebsebene zu ziehen. In dieser Arbeit wurden Modelle für das Fichten- Stratum erstellt. Die besten Ergebnisse wurden mit einem linearen gemischten Modell für Fichte erzielt (R^2 = 0.82; $RMSE_{rel}$ = 24 %).

Jedes Modell wurde genau wie auch in dieser Arbeit auf das Untersuchungsgebiet angepasst und trainiert. Interessant ist dennoch, dass ähnlichen Ergebnisse mit unterschiedlicher Herangehensweise erzielt werden.

Die vorliegenden Ergebnisse liegen im Mittelfeld der oben angeführten Studien.

Hill et al. (2018) beschreibt die Schwierigkeiten der Vorratsmodellierung in heterogenen Waldgesellschaften. Auch wurde in dieser Arbeit die Verzerrung (*lag*) der Daten durch den zeitlichen Abstand von terrestrischer Erhebung der Daten bis zum Zeitpunkt der Befliegung. Aufgrund dieses Artikels wurde in dieser Arbeit versucht, zu testen, ob es sich beim Untersuchungsgebiet um ein hetero- oder homogenes Waldgebiet handelt. Bei der Erstellung der Inventur wurden die Österreichischen Ertragstafeln angewandt. Die Information der Bonitätsstufe lag nicht bei der Inventur bei. In der Inventur des Forstbetriebes gab es bei dem Großteil der Inventurpunkte die Vorräte aus der letzten Inventur. Die Idee war, die Vorräte aus 2014 (aktuell) mit den der letzten Inventur zu subtrahieren, um die Wüchsigkeit abbilden zu können. Negative Werte (Abgänge) oder zu hohe Werte wurden dabei ausgeschieden. In Abbildung 28 ist ersichtlich, dass WZP- Punkte mit derselben Höhe unterschiedliche Holzvorräte einnehmen können und die Kurve mit steigenden Werten auf der x- Achse wie ein Fächer (oder Trichter) sich öffnet. Dies liegt daran, da der BHD bei der Volumenberechnung mit dem Quadrat in die Formel einbezogen wird. Auf guten Standorten wachsen die Bäume auch schneller in die Breite, dadurch entsteht dieser Fächer. Abhilfe dagegen würde eine Stratifizierung in den Wuchsklassen bringen.

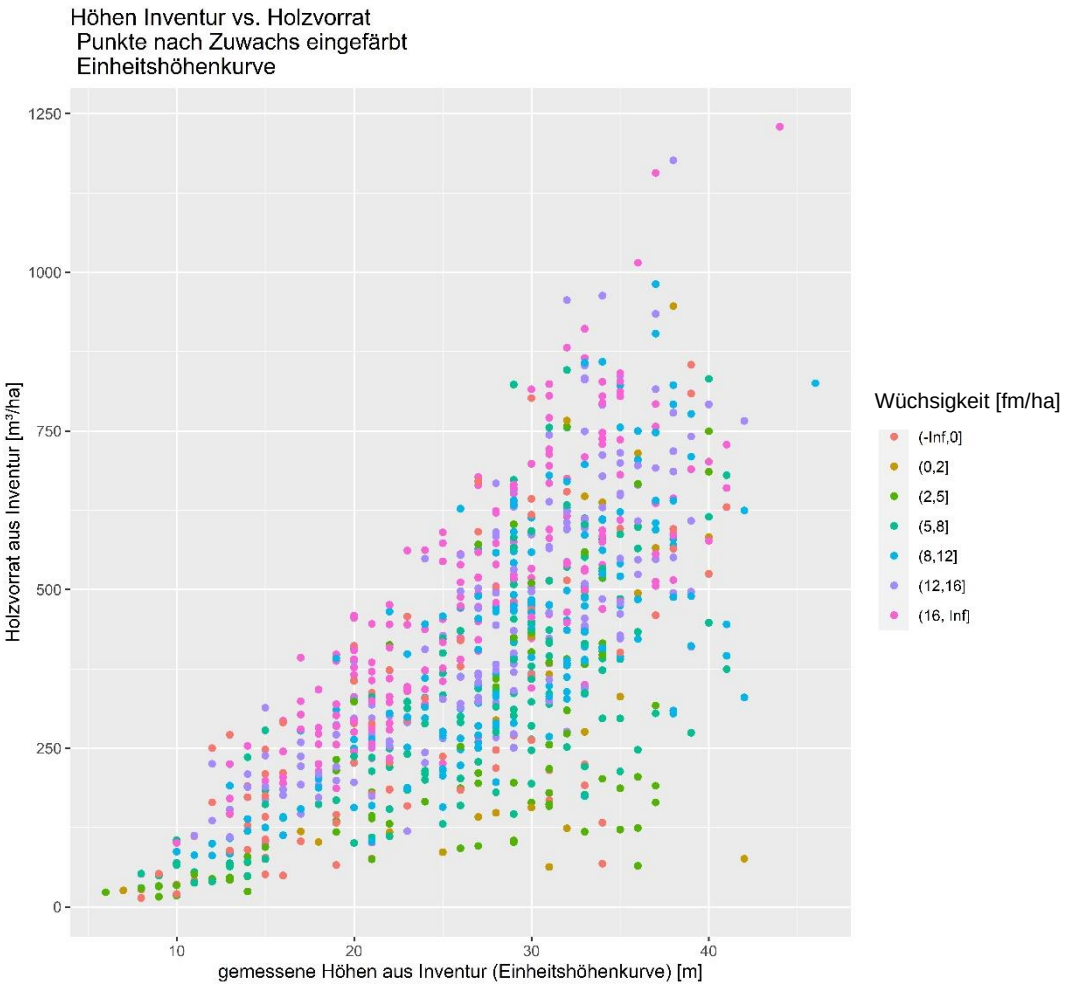

Abbildung 28: Holzvorrat der Inventur beschrieben mit den Höhen aus der Einheitshöhenkurve. Die Holzvorräte wurden anhand der Bonität eingefärbt. Ersichtlich ist, dass Bäume mit derselben Höhe unterschiedliche Volumina annehmen können.

5. Schlussfolgerungen und Ausblick

Die Bestandesvorratsschätzung mittels ALS-Daten bietet vielseitige Einsatzmöglichkeiten für die Unterstützung der strategischen und operativen Planung von Forstbetrieben wie auch in der Forschung. Der Vorteil liegt ebenso in der einfachen und kostengünstigen Vorratsschätzung in schlecht erschlossenen Gebieten wie auch der Abschätzung der anfallenden Schadholzmenge nach Kalamitäten wie Windwürfe oder Lawinenabgänge. Nach mehrmaligen Befliegungen lassen sich für die einzelnen Bäume und Baumarten Höhenfächer anlegen – das Wachstum auf unterschiedlichen Standorten kann so detailliert beobachtet werden. Für etliche Anwendungen bedarf es jedoch einer vorrausschauenden Planung und einer existierenden Bestandesvorratskarte. Durch die zukünftig preiswertere Verfügbarkeit von ALS-Daten wird es in Zukunft immer einfacher werden, Daten zu erhalten und auch auswerten zu können. Durch die Regionalisierung von ALS-Daten mithilfe von bestehender, terrestrisch erhobener Inventurpunkten, wurde mit dem ABA eine einfache wie auch robuste Methodik vorgestellt. Die Ergebnisse zeigen, dass zusätzlich Metriken aus ALS-Daten eine Verbesserung der Modellgüte im Vergleich zu Höhenmetriken aus nDSM bewirken. Es ist deutlich zu erwähnen, dass in der vorliegenden Arbeit das volle Potential von ALS- Daten nicht ausgeschöpft wurde. Dennoch konnte aufgezeigt werden, dass die Modelle mit zusätzliche Information gewonnen aus der Punktwolke die Modelle mit Höhenmetriken aus dem nDSM hinsichtlich Genauigkeit übertreffen. Dass die zusätzliche Information aus der Sentinel-2 Szene keinen Mehrwert gebracht hat, ist vermutlich der geringen Auflösung (10 m x 10 m) geschuldet. Für zukünftige Studien sollte die Information aus einer Befliegung oder hochauflösende Satellitenbilder zusätzlich getestet werden. Dass die Stammzahl auf den Probekreisen keinen Einzug in die Modelle gefunden hat war interessant zu beobachten. Hinsichtlich der Verschneidung des ABA- Ansatzes mit dem ITC-Ansatzes sollte ebenso in Zukunft eingegangen werden.

Punktwolken mit einer niedrigen Punktdichte (< 5 Punkte / m^2) sind hinsichtlich Beschreibung der vertikalen Bestandesstruktur Grenzen gesetzt. Bedingt durch die geringe Punktanzahl sinkt die Wahrscheinlichkeit, dass die Punkte durch das Kronendach in den Bestand eindringen können, um das innere des Bestandes beschreiben zu können. Abhilfe dagegen schaffen Produkte mit einer weit höheren Punktdichte, sogenannte High-Density Punktwolken, welche eine Punktdichte von über 300 Punkte / m^2 aufweisen. Studien zeigen, dass Größen wie der BHD und weitere Durchmesser am Stammverlauf messbar sind, wenn High- Density Punktwolken verwendet werden. Dies ist ebenso für die Berechnung des Einzelbaumvolumens nach ITC essenziell. Des Weiteren bieten solche qualitativ hochwertige Punktwolke in Zukunft die Möglichkeit, Bestandeskennzahlen hinsichtlich der Qualität der Bäume, aus den Befliegungsdaten ableiten zu können. Einige Gedankenansätze hierfür wäre die automatische Ermittlung der Astanzahl pro m Stamm (Astigkeit) oder auch die automatisierte Messung des Kronenansatzes. Auch die Kombination des ABA mit dem ITC bietet großes Potential für die Steigerung der Genauigkeit der Modelle. Der Schlüssel für die detaillierte und fast fehlerlose Beschreibung von Waldflächen liegt in der Baumartenerkennung.

In der vorliegenden Arbeit wurde die Unterscheidung zwischen Laub- und Nadelholz mittels der bestehenden und frei verfügbaren (Open-Source) Klassifikation der Landnutzung des Landinformationsservice Austria zugegriffen. Die Ergebnisse haben gezeigt, dass eine Stratifizierung in Laub- und Nadelholz die Modellgüte im Vergleich zu Modellen, in denen alle Stichprobenpunkte Einzug finden, verbessert. Die Schwierigkeit der Schätzung des Bestandesvorrats liegt in der Heterogenität hinsichtlich Wachstumsleistung der Bäume im Untersuchungsgebiet. Bäume können bei gleichem BHD unterschiedliche Höhen haben. Dies führt zu einer grundsätzlichen Streuung in der Datengrundlage. Dieses Problem kann insofern behoben werden, wenn die Grunddaten in Wuchsklassen und Baumarten stratifiziert werden würden. Des Weiteren würde sich für ein optimales Modell ein anderes Stichprobendesign zur

Erhebung der Trainingsdaten empfehlen. Die WZP als Stichprobendesign hat den gravierenden Nachteil, dass kein Flächenbezug gegeben ist – die vorliegende Arbeit zeigt, dass mit Kombination von unterschiedlichen Kreisradii Abhilfe geschaffen werden kann. Empfehlenswert sind dennoch fixe Probekreise bei der Erhebung, beispielsweise mit einer Grundfläche von 625 m².

In Zukunft wird die Verfügbarkeit der ALS-Daten steigen und deren Preise fallen. Bei derzeit steigender Popularität hinsichtlich Verwendung von Informationen aus der Luft ist es essenziell auch mit den akquirierten Daten arbeiten zu können. Dies beginnt bei der Speicherung der Daten mit Information (Fluginformationen, Informationen zum Sensor, etc.) und endet mit der Verarbeitung der Punktwolke. In dieser Arbeit wurde darauf Wert gelegt, so viel wie möglich mit Open-Source Software zu prozessieren. Festzuhalten ist, dass die Verwendung von ALS-Daten sehr vielseitig ist, und dass es de facto keine Begrenzung gibt. Seit QGIS-Version 3.18 und aufsteigend ist es erstmals möglich LiDAR Punktwolken grafisch darzustellen. Dies bietet die Möglichkeit, alle Visualisierungen in dem Open-Source Programm darstellen zu können. Für die Prozessierung der ALS-Daten hat sich als bestes Werkzeug das lidR Paket in R herausgefiltert. Ebenso ist das Programm lastools eine herausragende Software, mit dem Nachteil, dass etliche Befehle wiederum kostenpflichtig sind. Abhilfe dagegen schafft nur dass für die unterschiedlichen Bearbeitungsschritte auf unterschiedliche Software zugegriffen wird. Für die Prozessierung von Fernerkundungsdaten, respektive Geoinformationsdaten, hat sich als Nonplusultra das R-Paket „terra“ im Verlauf der Arbeit bewiesen. Mithilfe des Paktes lassen sich Shapefiles wie auch Rasterdaten spielend übereinanderlegen und miteinander verschneiden, wie auch Datensätze ohne räumlichen Index mit Shapefiles verknüpfen. Der Vorteil ist im Vergleich zu QGIS, dass die Daten nur dann in den Arbeitsspeicher geladen werden, wenn es für den Rechenschritt nötig ist. Die Open- Source Pakete basierend auf R bieten allesamt eine optimale Ausnutzung der Hardware – rechenintensive (bezogen auf die Dauer) Arbeitsschritte werden minimiert.

Die Verwendung von ALS-Daten für die Lösung und Unterstützung von forstrelevanten Problemstellungen ist hinsichtlich Forstwirtschaft 4.0 ein wichtiges Themengebiet. Die Verfügbarkeit der Daten wie auch die Bereitschaft für die Verwendung der gewonnen Information ist gegeben. Die Anwendungsbereiche wie auch das Potential der ALS-Daten sind noch weit nicht ausgeschöpft und die Forschung in diesem Themenbereich ist als neu anzusehen. In der Zukunft sind Strukturanalysen von Beständen und auch Einzelbäumen, die Weiterentwicklung der Baumartenerkennung und auch die Optimierung der Algorithmen für die flächendeckende Stammzahlermittlung für das ITC-Verfahren als Forschungsfelder zu identifizieren.

6. Abbildungsverzeichnis

Abbildung 1: Das Untersuchungsgebiet (gelb) befindet sich im nördlichen Teil Österreichs, an der Grenze von Mühl- und Waldviertel. 3
Abbildung 2: Korrektur der Lage des WZP- Punktes 10918: vor (links) und rechts (nach) der Verschiebung. 5
Abbildung 3: Darstellung der Punktwolke mit 50 m Radius zur Sichtbarmachung der ableitbaren Baumhöhen. Die beiden Abbildungen unterscheiden sich in der Perspektive. 10
Abbildung 4: Nach der Berechnung der Stammzahl im Probekreis wurde diese dem Holzvorrat für die verschiedenen Kreisradien gegenübergestellt. Abgebildet sind die Kreisradien 5,10,15,20,25 m. 11
Abbildung 5: Karte mit der Information des NDVI errechnet aus der Sentinel- 2 Szene 12
Abbildung 6: Variable Importance Plot des Gesamtmodells mit Metriken abgleitet aus der Punkwolke. „k“ steht für Kreis, „i“ = Intensitätsmetriken und „z“ = Höhenmetriken. 18
Abbildung 7: Variable Importance Plot der Kreiskombination mit Höhenmetriken abgeleitet aus dem nDSM. "k" = Kreis, die Zahl "05" für den Kreisradius in Meter. "p" sind Prozent der Quantil. 18
Abbildung 8: Variable Importance Plot nach der Feature Selection am 12.5 m Kreisradius der Punktwolke. "k" = Kreis, "12" = Kreisradius in Meter (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken 19
Abbildung 9: Variable Importance Plot nach der Feature Selection am 12.5 m Kreisradius des nDSM. "k" = Kreis, "12" = Kreisradius in Meter (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken 19
Abbildung 10: Scatterplots mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts) der Gesamtmodelle. Punktierte Linie = 1:1 Gerade, rote durchgezogene Linie = Ausgleichsgerade. 20
Abbildung 11: Variable Importance Plot der Kreiskombination an der Punktwolke. "k" = Kreis, "12" = Kreisradius in Meter (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken, „VC „= Kronenvolumen. 23
Abbildung 12: Variable Importance Plot der Kreiskombination am nDSM abgeleiteten Höhenmetriken. "k" = Kreis, "05" = Kreisradius [m], "p75" = 75 % Quantil 23
Abbildung 13: Variable Importance Plot nach der Feature Selection am 12.5 m Kreisradius der Punktwolke. "k" = Kreis, "12" = Kreisradius in Meter (12 = 12.5 m), "i" = Intensitätsmetriken, "z" = Höhenmetriken 24
Abbildung 14: Variable Importance Plot am Kreisradius 10 m Nadelholz des nDSM. "k" = Kreis, "10" = Kreisradius [m], "p50" = 50 % Quantil (Median) 24
Abbildung 15: Scatterplots der Nadelholzmodelle mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts). Punktierte Linie = Diagonale im Plot, rote durchgezogene Linie = Ausgleichsgerade. Der Einzelkreisradius vom besten ALS Modell (unten rechts) beträgt 12.5 m. 25
Abbildung 16: Variable Importance Plot der Kreiskombination Laubholzmodell der Punktwolke. "k" = Kreis, "12" = Kreisradius [m], "i" = Intensitätsmetriken, "z" = Höhenmetriken, „n_ges“ = Punktanzahl 28
Abbildung 17: Variable Importance Plot Kreiskombination nDSM Laubholz. "k" = Kreis, "15" = Kreisradius in Meter,"p50" = 50 % Quantil der Höhe. 28
Abbildung 18: Variable Importance Plot am Kreisradius 15 mLaubholz der Punktwolke. "k" = Kreis, "15" = Kreisradius [m], „z“ = Höhenmetrik, „mean“ = Mittelwert, „ipcumzq70“ = Prozentsatz der zurückgegebenen Intensität unterhalb des 70. Perzentils der Höhe 29
Abbildung 19: Variable Importance Plot 20 m Radius Einzelkreis Laubholz mit Höhenmetriken aus dem nDSM abgeleitet. "k" = Kreis, "20" = 20 Meter Kreisradius, "p95" = 95 % Quantil der Höhe. 29

Abbildung 20: Scatterplots der Laubholzmodelle mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts). Punktierte Linie = Diagonale im Plot, rote durchgezogene Linie = Ausgleichsgerade. 30
Abbildung 21: Variable Importance Plot der Kreiskombination Mischwaldmodell der Punktwolke. "k" = Kreis, "12" = Kreisradius [m], "i" = Intensitätsmetriken, "z" = Höhenmetriken, „n_ges“ = Punktanzahl 32
Abbildung 22: Variable Importance Plot der Kreiskombination Mischwaldmodell des nDSM. "k" = Kreis, "12" = Kreisradius [m], "i" = Intensitätsmetriken, "z" = Höhenmetriken, „n_ges“ = Punktanzahl 32
Abbildung 23: Variable Importance Plot 10 m Radius Einzelkreis Mischwald mit Metriken aus der Punktwolke abgeleitet. "k" = Kreis, "10" = 10 m Kreisradius, „z“ = Höhenmetriken, „i“ = Intensitätsmetriken, „VC“ = Kronenvolumen, „PR“ = Durchdringungsrate, „n“ = Punktanzahl 33
Abbildung 24: Variable Importance Plot 15 m Radius Einzelkreis Mischwald mit Höhenmetriken aus dem nDSM abgeleitet. "k" = Kreis, "20" = 20 m Kreisradius, "p95" = 95 % Quantil der Höhe. 33
Abbildung 25: Scatterplots der Mischwaldmodelle mit Gegenüberstellung Holzvorrat Inventur zu Holzvorrat modelliert mit Kreiskombination (links) zu bestem Einzelkreis (rechts). Punktierte Linie = Diagonale im Plot, rote durchgezogene Linie = Ausgleichsgerade. 34
Abbildung 26: Holzvorratskarte zusammengestellt aus Nadel- und Laubholzvorrat für das Untersuchungsgebiet. Rot eingefärbte Pixel sind Laubholzbestände. Nadelholzbestände wurden blau eingefärbt. Nur Waldflächen wurden eingefärbt. Als Hintergrund dient das Orthofoto von Geoland Basemap. 37
Abbildung 27: Bestandeskarte mit Information des Holzvorrats. In weiß eingefärbte Flächen entsprechen keinem Wald wie beispielsweise Wiesen, Wege oder Teiche. 38
Abbildung 28: Holzvorrat der Inventur beschrieben mit den Höhen aus der Einheitshöhenkurve. Die Holzvorräte wurden anhand der Bonität eingefärbt. Ersichtlich ist, dass Bäume mit derselben Höhe unterschiedliche Volumina annehmen können. 43

7. Tabellenverzeichnis

Tabelle 1: Übersicht über die Anzahl der erstellten Random Forest Feature Selection Modelle. EK = Einzelkreis, MK = Multikreis 14
Tabelle 2: Auflistung der statistischen Auswertung der Gesamtmodelle mit abgeleiteten Metriken aus nDSM (oben) und ALS Punktwolke (unten). 17
Tabelle 3: Statistische Auswertung der erstellten Nadelholzmodelle. Oben sind die Einzelkreise und Kombination der aus dem nDSM abgeleiteten Modelle ersichtlich. Unten sind die Modelle (Einzelkreis & Kombination) der Metriken aus der Punktwolke angeführt. 22
Tabelle 4: Statistische Auswertung der erstellten Laubholzmodelle. Oben sind die Einzelkreise und Kombination der aus dem nDSM abgeleiteten Modelle ersichtlich. Unten sind die Modelle (Einzelkreis & Kombination) der Metriken aus der Punktwolke angeführt. 27
Tabelle 5: Statistische Auswertung der erstellten Mischwaldmodelle. Oben sind die Einzelkreise und Kombination der aus dem nDSM abgeleiteten Modelle ersichtlich. Unten sind die Modelle (Einzelkreis & Kombination) der Metriken aus der Punktwolke angeführt. 31
Tabelle 6: Überblick über die Ergebnisse der besten Random Forest Modelle. NDSM = normalisiertes digitales Oberflächenmodell 41

8. Literaturverzeichnis

Amiri, N.; Polewski, P.; Yao, W.; Krzystek, P.; Skidmore, A. K. (2017): Detection of single tree stems in forested areas from hight density ALS point clouds using 3D shape descriptors. In: *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* IV-2/W4, S. 35–42. DOI: 10.5194/isprs-annals-IV-2-W4-35-2017.

Baccini, A.; Friedl, M. A.; Woodcock, C. E.; Warbington, R. (2004): Forest biomass estimation over regional scales using multisource data. In: *Geophys. Res. Lett.* 31 (10). DOI: 10.1029/2004GL019782.

Breiman, L. (2001): Random Forests. In: *Machine Learning* (45), S. 5–32.

Briechle, S.; Krzystek, P.; Vosselman, G. (2021): Silvi-Net – A dual-CNN approach for combined classification of tree species and standing dead trees from remote sensing data. In: *International Journal of Applied Earth Observation and Geoinformation* 98, S. 102292. DOI: 10.1016/j.jag.2020.102292.

Esri (2021): ArcMap Desktop. Version 10.7.1: Esri Deutschland GmbH.

Hawryło, P.; Wężyk, P. (2018): Predicting Growing Stock Volume of Scots Pine Stands Using Sentinel-2 Satellite Imagery and Airborne Image-Derived Point Clouds. In: *Forests* 9 (5), S. 274. DOI: 10.3390/f9050274.

Hijmans, J.-R. (2022): terra: Spatial Data Analysis. Version 1.5-17. Online verfügbar unter https://CRAN.R-project.org/package=terra.

Hill, A.; Buddenbaum, H.; Mandallaz, D. (2018): Combining canopy height and tree species map information for large-scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes. In: *Eur J Forest Res* 137 (4), S. 489–505. DOI: 10.1007/s10342-018-1118-z.

Hyyppä, J.; Yu, X.; Hyyppä, H.; Vastaranta, M.; Holopainen, M.; Kukko, A.; Kaartinen, H. (2012): Advances in Forest Inventory Using Airborne Laser Scanning. In: *Remote Sensing* 4 (5), S. 1190–1207. DOI: 10.3390/rs4051190.

Immitzer, M.; Stepper, C.; Böck, S.; Straub, C.; Atzberger, C. (2016): Use of WorldView-2 stereo imagery and National Forest Inventory data for wall-to-wall mapping of growing stock. In: *Forest Ecology and Management* 359, S. 232–246. DOI: 10.1016/j.foreco.2015.10.018.

Kaartinen, H.; Hyyppä, J.; Yu, X.; Vastaranta, M.; Hyyppä, H.; Kukko, A. (2012): An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. In: *Remote Sensing* 4.

Kirchhoefer, M.; Schumacher, J.; Adler, P. (2019): Potential of remote sensing-based forest attribute models for harmonising large-scale forest inventories on regional level: a case study in Southwest Germany. In: *Annals of Forest Science* 76 (2). DOI: 10.1007/s13595-019-0804-4.

Kozak, J.; Gimmi, U.; Houet, T.; Bolliger, J. (2017): Current practices and challenges for modelling past and future land use and land cover changes in mountainous regions. In: *Reg Environ Change* 17 (8), S. 2187–2191. DOI: 10.1007/s10113-017-1217-2.

Kramer, H.; Akça, A. (2008): Leitfaden zur Waldmesslehre. 5., überarb. Aufl.: Frankfurt a. M. : Sauerländer. Online verfügbar unter https://permalink.obvsg.at/bok/AC06831037.

Kuhn, M. (2021): caret: Classification and Regression Training. R package version 6.0-90. Online verfügbar unter https://CRAN.R-project.org/package=caret.

Levick, S. R.; Hessenmöller, D.; Schulze, E-D. (2016): Scaling wood volume estimates from inventory plots to landscapes with airborne LiDAR in temperate deciduous forest. In: *Carbon balance and management* 11 (1), S. 7. DOI: 10.1186/s13021-016-0048-7.

Lindberg, E.; Hollaus, M. (2012): Comparison of Methods for Estimation of Stem Volume, Stem Number and Basal Area from Airborne Laser Scanning Data in a Hemi-Boreal Forest. In: *Remote Sensing* 4 (4), S. 1004–1023. DOI: 10.3390/rs4041004.

Marschall, J. (1975): Hilfstafeln für die Forsteinrichtung. Zehnte Auflage. Wien: Österreichischer Forstverein. Online verfügbar unter https://permalink.obvsg.at/bok/AC15477207.

Næsset, E. (1997): Estimating timber volume of forest stands using airborne laser scanner data. In: *Remote Sensing of Environment* 61 (2), S. 246–253. DOI: 10.1016/S0034-4257(97)00041-2.

Nothdurft, A.; Gollob, C.; Kraßnitzer, R.; Erber, G.; Ritter, T.; Stampfer, K.; Finley, A. O. (2021): Estimating timber volume loss due to storm damage in Carinthia, Austria, using ALS/TLS and spatial regression models. In: *Forest Ecology and Management* 502, S. 119714. DOI: 10.1016/j.foreco.2021.119714.

R Core Team (2022): R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna. Online verfügbar unter https://www.R-project.org/.

Rahlf, J.; Hauglin, M.; Astrup, R.; Breidenbach, J. (2021): Timber volume estimation based on airborne laser scanning — comparing the use of national forest inventory and forest management inventory data. In: *Annals of Forest Science* 78 (2). DOI: 10.1007/s13595-021-01061-4.

rapidlasso GmbH (2022): LAStools: rapidlasso GmbH. Online verfügbar unter https://rapidlasso.com/lastools/.

Rouse, John; Haas, R.; Schell, J.; Deering, D. (1973): Monitoring vegetation systems in the great plains with ERTS. In: *Remote Sensing Center*.

Roussel, J.-R.; Auty, D.; Coops, N. C.; Tompalski, P.; Goodbody, T. R.H. (2020): lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. In: *Remote Sensing of Environment* 251, S. 112061. DOI: 10.1016/j.rse.2020.112061.

RStudio team (2019): RStudio: Integrated Development for R. Boston: RStudio, Inc. Online verfügbar unter http://www.rstudio.com/.

Socha, J.; Hawryło, P.; Pierzchalski, M.; Sterenczak, K.; Krok, G.; Wężyk, P.; Tyminska-Czabanska, L. (2020): An allometric area- based aproach - a cost- effective method for stand volume estimation based on ALS and NFI data. In: *Forestry: An International Journal of Forest Research* 93 (1), S. 344–358. DOI: 10.1093/forestry/cpz061.

Stepper, C.; Straub, C.; Pretzsch, H. (2015): Using semi-global matching point clouds to estimate growing stock at the plot and stand levels: application for a broadleaf-dominated forest in central Europe. In: *Can. J. For. Res.* 45 (1), S. 111–123. DOI: 10.1139/cjfr-2014-0297.

White, J.; Tompalski, P.; Vastaranta, M.; Wulder, M.; Stepper, C. (2017): A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach. In: *Canadian Forest Service*.

White, J. C.; Wulder, M. A.; Varhola, A.; Vastaranta, M.; Coops, N. C. (2013): A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. In: *The Forestry Chronicle* 89 (06), S. 722–723. DOI: 10.5558/tfc2013-132.

Woods, M.; Lim, K.; Treitz, P. (2008): Predicting forest stand variables from LiDAR data in the Great Lakes – St. Lawrence forest of Ontario. In: *The Forestry Chronicle* 84 (6), S. 827–839. DOI: 10.5558/tfc84827-6.

Yu, X.; Hyyppä, J.; Karjalainen, M.; Nurminen, K.; Karila, K.; Vastaranta, M. (2015): Comparison of Laser and Stereo Optical, SAR and InSAR Point Clouds from Air- and Space-Borne Sources in the Retrieval of Forest Inventory Attributes. In: *Remote Sensing* 7 (12), S. 15933–15954. DOI: 10.3390/rs71215809.