



Flächendeckende Fernerkundungsbasierte Forstliche Strukturdaten

Verbundvorhaben „Entwicklung von Methoden und
Verfahren zur flächendeckenden und homogenen
Generierung von Waldparametern auf der Basis digitaler
Oberflächenmodelle aus Luftbilddaten (F³)“

Nordwestdeutsche Forstliche Versuchsanstalt (NW-FVA)
Grätzelstr. 2, 37079 Göttingen

Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg (FVA)
Wonnhaldestr. 4, 79100 Freiburg

*Gefördert durch das Bundesministerium für Ernährung und Landwirtschaft
(BMEL). Förderkennzeichen: 22025014 (FVA), 22024816 (NW-FVA)*

Dokumentation zu Meilenstein 8

„Verfahren zur Modellierung von Holzvorrat und
oberirdischer Biomasse“

Meilenstein 8

Verfahren zur Modellierung von Holzvorrat und oberirdischer Biomasse

A	Einleitung.....	1
B	Referenzdaten	2
C	Aufbereitung der Rasterdaten	3
C.1	Strukturmetriken aus Oberflächenmodellen	3
C.2	Topographiemetriken aus dem digitalen Geländemodell (DGM1).....	4
C.3	Klimadaten (Niederschlag und Temperatur).....	5
C.4	Bodendaten	6
C.5	Baumarten.....	7
D	Modellierung / Erstellen von Holzvorrats- und Biomassekarten	8
E	Modellierung von Holzvorrat und Biomasse in den Testgebieten des F ³ -Projekts.....	9
E.1	Nordwestdeutsche Testgebiete	9
E.2	Südwestdeutsche Testgebiete	13

A Einleitung

Für eine nachhaltige Waldbewirtschaftung sind Daten über die räumliche Verteilung von Holzvorräten eine wichtige Basis. Insbesondere in Zeiten häufig auftretender Kalamitäten durch Sturm und Insekten (z. B. Borkenkäfer) stellen sie eine wertvolle Grundlage für die Planung von Maßnahmen der Holzernte, der Verjüngung, des Waldschutzes und des Naturschutzes dar. Daten über die Verteilung und Höhe von Biomassevorräten spielen ebenfalls eine große Rolle, insbesondere im Rahmen der Erfassung der Kohlenstoffspeicherleistung der Wälder. Um flächendeckende Daten bezüglich Holzvorrat und Biomasse abzuleiten, stellen wir ein Verfahren zur Verfügung, welches basierend auf Oberflächenmodellen, Baumartenkarten und Standortdaten Vorräte von Holz und oberirdischer Biomasse modelliert.

Das Verfahren wurde in einem R-Skript implementiert, welches Daten aus verschiedenen Quellen integriert und für die Modellierung verwendet. Die wichtigsten Eingangsdaten sind Daten zur Struktur der Waldbestände, die aus normalisierten Oberflächenmodellen abgeleitet werden. Weitere Datensätze, welche die Baumartenzusammensetzung, die Bodeneigenschaften, die Topographie und das Klima beschreiben, haben das Potential die Vorhersage zu verbessern, wenn sie in die Modellierung eingebunden werden. Im Folgenden zeigen wir anhand von Datensätzen zu Boden, Klima, Topographie und Baumartenzusammensetzung, wie die Modellierung durch diese Datensätze unterstützt werden kann. Die verwendeten Daten liegen deutschlandweit als Rasterdaten vor und sind zum größten Teil frei zugänglich. Daten zur Baumartenzusammensetzung können mit frei verfügbaren Sentinel-2-Satellitendaten generiert werden (siehe Meilenstein M5¹). Daten zur Topographie können bei den Landesvermessungsinstitutionen bezogen werden, sind jedoch nicht immer kostenfrei. Kostenfreie Klimadaten werden z. B. vom Deutschen Wetterdienst (DWD) mit einer Auflösung von 1 x 1 km bereitgestellt. Bodendaten können aus der globalen Datenbank soilgrids.org mit einer Auflösung von 250 x 250 m bezogen werden. Die hier verwendeten Datensätze dienen als Beispiele, um zu illustrieren, wie ihre Einbindung in das Verfahren erfolgt. Die deutschlandweite Verfügbarkeit und der freie Zugriff auf die Daten spielten bei der Auswahl eine wichtige Rolle. Sollten den Nutzenden höher aufgelöste bzw. qualitativ höherwertigere Daten zur Verfügung stehen oder sollten andere bzw. zusätzliche Parameter für wichtig erachtet werden, können auch diese Datensätze in das Verfahren integriert und für die Modellierung verwendet werden. Es ist darauf zu achten, dass alle Datensätze in einer gemeinsamen Projektion und gleicher horizontalen Auflösung (20 x 20 m) vorliegen. Aufgrund der Verwendung eines Random Forest-Modells, das problemlos weitere Datensätze bei der Modellierung berücksichtigen kann, ist das Verfahren flexibel erweiterbar bzw. bei nicht Vorhandensein einzelner Datensätze reduzierbar.

Für die Modellierung unabdingbar sind Referenzdaten, z. B. aus forstlichen Inventuren, aus denen Informationen über Holzvorrat und oberirdische Biomasse pro Stichprobeneinheit abgeleitet und einem bestimmten Geländeausschnitt zugeordnet werden können. Solche Daten werden benötigt um die Modelle zur Vorhersage von Holzvorräten und Biomasse zu trainieren und zu validieren.

Im Folgenden erläutern wir das Verfahren zur Modellierung von Holzvorrat und Biomasse unter Verwendung der oben angesprochenen Datensätze. Die Aufbereitung der Eingangsdaten, die

¹ M5 – „Verfahren zur Ableitung von Baumarteninformationen aus Fernerkundungsdaten“:
https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021]

Modellierung und die Validierung der Ergebniskarten werden exemplarisch dargestellt. Wir möchten jedoch darauf hinweisen, dass die Qualität der Modellierung mitunter stark von der Qualität der verwendeten Daten abhängig ist. Die von den Referenzdaten erfasste Variabilität der Waldstrukturen, die Anzahl und Qualität der Referenzdatenpunkte sowie die Genauigkeit der Oberflächenmodelle, Baumartenkarten und Standortdaten spielen eine nicht zu unterschätzende Rolle. Da die Referenzdaten im Rahmen einer Kreuzvalidierung ebenfalls zur Validierung genutzt werden, ist darauf hinzuweisen, dass die daraus abgeleiteten Maßzahlen zur Einschätzung der Modellgenauigkeit sich auch nur auf die hierin abgebildeten Waldzustände beziehen. Wenn z. B. der Referenzdatensatz keine Daten aus Auwäldern enthält, werden die für diese Wälder modellierten Holz- bzw. Biomassevorräte hinsichtlich ihrer Güte nicht abgebildet. Wenn der Referenzdatensatz nur Inventurpunkte beinhaltet, die in Beständen mit einem Alter > 50 Jahre aufgenommen wurden, wird es nicht möglich sein, die Güte der Modellierung für Bestände jüngeren Alters zu quantifizieren. Die Abhängigkeit der Modellierungs- und der Validierungsergebnisse von den verwendeten Eingangsdaten sollte der Nutzer bei der Interpretation der Karten bzw. der Qualitätsangaben bedenken.

B Referenzdaten

Referenzdaten für den Holzvorrat und die Biomasse stammen üblicherweise aus forstlichen Stichprobeninventuren, wie z. B. der Betriebsinventur der Niedersächsischen Landesforsten oder auch der Bundeswaldinventur (BWI). Diese Daten enthalten Informationen zu Bruthöhendurchmesser (BHD) und Baumhöhe der an den Stichprobenpunkten ausgewählten Bäume. Aus diesen Einzelbaummessungen kann für jeden Baum mittels geeigneter Modelle das Holzvolumen bzw. die Biomasse modelliert werden. Für die Modellierung dieser Einzelbaumwerte sollten regional angepasste Modelle verwendet werden, um möglichst akkurat Holzvorrat und oberirdische Biomasse pro Baum bestimmen zu können. Sollten keine regional angepassten Modelle zur Verfügung stehen, kann z. B. auf Biomassemodelle zurückgegriffen werden, die für die dritte Bundeswaldinventur (BWI 3) entwickelt wurden² und im R-Paket *rbdatpro*³ implementiert sind. Die ermittelten Werte für Holzvorrat und Biomasse der Einzelbäume müssen anschließend unter Berücksichtigung der Größe der Probestfläche pro Stichprobenpunkt aggregiert und auf Schätzungen pro Hektar (ha) hochgerechnet werden. Pro Stichprobenpunkt muss außerdem die Höhe des höchsten Baumes bestimmt und in einer Spalte der Attributtabelle gespeichert werden. Die Inventurdaten sollten als Punkte in einer Shape-Datei abgespeichert werden, deren Attributtabelle für jeden Stichprobenpunkt die folgenden Angaben enthält:

1. Holzvorrat (m³/ha)
2. Oberirdische Biomasse (t/ha)
3. Höhe des höchsten Baumes

² Kändler, G., Bösch, B., 2012. Methodenentwicklung für die 3. Bundeswaldinventur: Modul 3 Überprüfung und Neukonzeption einer Biomassefunktion. Abschlussbericht Version 2b. Online: https://gitlab.com/vochr/rbdatpro/-/blob/809449728a9af21aa9f3a850645dd2dda4908d64/inst/doc/BDAT_References/KaendlerBoesch2012_BiomassFunction_german.pdf [letzter Zugriff 01.12.2020].

³ R-Paket *rbdatpro*: <https://gitlab.com/vochr/rbdatpro> [letzter Zugriff 30.11.2020]

C Aufbereitung der Rasterdaten

C.1 Strukturmetriken aus Oberflächenmodellen

Python Skript: *Metriken_Holzvolumen_F3.py*

*Bevor das Python Skript *Metriken_Holzvolumen_F3.py* gestartet werden kann, muss in **Zeile 196** der Pfad zum Eingabeordner, der die normalisierten Oberflächenmodelle im LAZ-Format enthält, spezifiziert werden. In **Zeile 198** der Pfad zum Ausgabeordner, der die generierten Metriken enthalten soll, angegeben werden. Der Ausgabeordner muss zuvor manuell angelegt werden. In **Zeile 202** muss das Koordinatenreferenzsystem (UTM Zone 32N oder Gauß-Krüger Zone 3) spezifiziert werden, in dem die Daten vorliegen. Dies geschieht über die Angabe des entsprechenden EPSG-Codes: 25832 (UTM Zone 32 N), 25833 (UTM Zone 33 N) oder 31466 / 31467 / 31468 (GK 2 / 3 / 4). Außerdem muss in **Zeile 205** der Pfad zum „bin“-Ordner der LAStools spezifiziert werden.*

Die Berechnung der Strukturmetriken basiert auf normalisierten Oberflächenmodellen (nDOM) im LAZ-Format, wie sie durch das in Meilenstein M3⁴ beschriebene Verfahren bereitgestellt werden. Mittels des Python-Skripts *Metriken_Holzvolumen_F3.py*⁵ werden unter Verwendung der Software LAStools⁶ aus den nDOM-Dateien Statistiken abgeleitet, die die Struktur des Waldes pro 20 x 20 m-Flächen beschreiben. Die Auswahl der mittels des Python-Skripts berechneten Metriken orientiert sich maßgeblich an den von Schumacher et al. (2019)⁷ für die Modellierung des Holzvorrats in Baden-Württemberg verwendeten Metriken. Das Skript berechnet kachelweise die folgende Metriken und legt sie in Unterordnern ab:

- Mittlere Höhe des nDOM pro 20 x 20 m-Pixel
 - o LAStools „lasgrid“
 - o Ordner: mean_20
- Maximale Höhe des nDOM pro 20 x 20 m-Pixel
 - o LAStools „lasgrid“
 - o Ordner: max_20
- Standardabweichung der Höhenwerte des nDOM pro 20 x 20 m-Pixel
 - o LAStools „lasgrid“
 - o Ordner: std_20

⁴ M3 – „Standardisiertes Verfahren zur Ableitung normalisierter Oberflächenmodelle (nDOM) aus Luftbildern“: https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021]

⁵ Python-Skript *Metriken_Holzvolumen_F3.py*: https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021]

⁶ LAStools: <https://rapidlasso.com/lastools/> [letzter Zugriff 30.11.2020]

⁷ Schumacher, J., Rattay, M., Kirchhöfer, M., Adler, P., Kändler, G. 2019. Combination of Multi-Temporal Sentinel 2 Images and Aerial Image Based Canopy Height Models for Timber Volume Modelling. *Forests* 2019, 10, 746.

- 95. Perzentil der Höhenwerte des nDOM pro 20 x 20 m-Pixel
 - LAStools „lascanopy“
 - Ordner: perc95_20
- 75. Perzentil der Höhenwerte des nDOM pro 20 x 20 m-Pixel
 - LAStools „lascanopy“
 - Ordner: perc75_20
- Prozentualer Anteil der Fläche (20 x 20 m), die mit Höhenwerten > 6 m überschirmt ist
 - LAStools „lascanopy“
 - Ordner: cctot6_20
- Prozentualer Anteil der Fläche (20 x 20 m), die mit Höhenwerten > 20 m überschirmt ist
 - LAStools „lascanopy“
 - Ordner: cctot20_20

Nachdem die Metriken berechnet wurden, müssen die Kacheln zu Mosaiken zusammengefügt werden. Hierzu kann die Software ArcGIS (Tool: Mosaic To New Raster) oder QGIS (Tool: Verschmelzen) oder das R-Skript *„mosaic_tiles_F3.R“*⁸ verwendet werden. Bei Verwendung des R-Skripts muss lediglich in Zeile 24 der Pfad zum Verzeichnis, das die Ordner mit den zuvor erzeugten Strukturmetriken enthält, angepasst werden (dieser entspricht dem im Python-Skript *„Metriken_Holzvolumen_F3.py“*⁵ spezifizierten Ausgabeordner). Das R-Skript erstellt in diesem Verzeichnis ein neues Unterverzeichnis mit dem Namen „Mosaik“, innerhalb dessen die zu einem Mosaik zusammengefügt Dateien abgelegt werden.

C.2 Topographiemetriken aus dem digitalen Geländemodell (DGM1)

Zur Ableitung topographischer Parameter wird das digitale Geländemodell DGM1 im LAZ-Format verwendet. Um die Auflösung des DGM von 1 m an die benötigte Auflösung von 20 m anzupassen und um die Parameter mittlere Geländehöhe und Standardabweichung (Indikator für die Steigung des Geländes) zu berechnen, wird das LAStools Modul lasgrid verwendet. Das Geländemodell ist im Normalfall keinen Änderungen unterworfen, sodass die Ableitung der genannten Parameter nur einmalig durchgeführt werden muss. Daher wird hierfür kein Software-Skript zur Verfügung gestellt. Die Berechnung kann über das GUI der Software LAStools durch die Wahl der folgenden Einstellungen erfolgen:

Mittlere Geländehöhe:

1. Auswahl der zu prozessierenden DGM1.laz-Dateien
2. Auswahl der Option „merge files into one“
3. Spezifikation des Ausgabeordners
4. Spezifikation des Dateinamens
5. Spezifikation „pixel/step size“ = 20
6. Spezifikation „item“ = elevation
7. Spezifikation „op“ = average (Berechnung der mittleren Geländehöhe)
8. Spezifikation „fill n pixels“ = 3

⁸ R-Skript *„mosaic_tiles_F3.R“*: https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021]

9. Spezifikation „format“ = tif

Standardabweichung der Höhenwerte

Zur Berechnung der Standardabweichung sind dieselben Einstellungen zu wählen, wie zur Berechnung der mittleren Geländehöhe (siehe oben), einzig die unter Punkt 7 genannte Einstellung für „op“ muss angepasst werden. Hier muss **„op“ = stddev** gewählt werden.

C.3 Klimadaten (Niederschlag und Temperatur)

Flächendeckende Daten zur Beschreibung der klimatischen Bedingungen liegen deutschlandweit in einer Auflösung von 1 x 1 km vor und können kostenfrei vom DWD bezogen werden. Folgende Datensätze wurden im Rahmen des F³-Projektes für die Modellierung verwendet:

- Das 30-jährige Mittel (1981 – 2010) der Niederschlagshöhe
 - o Abgeleitet unter Berücksichtigung der langjährigen Klimatologie aus Daten der DWD-Stationen und qualitativ gleichgestellten Partnernetzstationen in Deutschland.
 - o https://opendata.dwd.de/climate_environment/CDC/grids_germany/multi_annual/precipitation [letzter Zugriff 30.11.2020]
 - Datei: grids_germany_multi_annual_precipitation_1981-2010_17.asc.gz
- Das 30-jährige Mittel (1981 – 2010) des monatlich gemittelten täglichen Lufttemperaturmaximums (in 2 m Höhe)
 - o Abgeleitet unter Berücksichtigung der Höhenabhängigkeiten aus Daten der DWD-Stationen und qualitativ gleichgestellten Partnernetzstationen in Deutschland.
 - o https://opendata.dwd.de/climate_environment/CDC/grids_germany/multi_annual/air_temperature_max [letzter Zugriff 30.11.2020]
 - Datei: grids_germany_multi_annual_air_temp_max_1981-2010_17.asc.gz

Nach dem Entpacken der ZIP-Archive liegen die Datensätze in einer Gauß-Krüger-Projektion (EPSG: 31467) im ASC-Dateiformat vor. Um sie für die weitere Verarbeitung aufzubereiten, muss ihre Auflösung auf 20 m geändert werden sowie die Projektion gegebenenfalls an die der übrigen Rasterdaten angepasst werden. Außerdem müssen die Dateien im TIF-Format abgespeichert werden. Alle nötigen Prozessierungen lassen sich in einem Schritt mit dem ArcGIS-Tool „Project Raster“ durchführen. Nach dem Aufrufen des Tools müssen das zu prozessierende Raster ausgewählt und die Projektion des Rasters spezifiziert werden (z. B. EPSG: 31467). Bei der Spezifikation der Ausgabedatei muss darauf geachtet werden, dass die Dateiendung „.tif“ verwendet wird. Die gewünschte Projektion der Ausgabedatei sowie die geographische Transformation müssen ebenfalls angegeben werden, z. B. Projektion = EPSG: 25832 und Transformation = DHDN_To_ETRS_1989_8_NTv2. Als „Resampling Technique“ sollte die Option „BILINEAR“ gewählt werden und als gewünschte Pixelgröße X = 20 und Y = 20. Um sicherzustellen, dass die Lage der Pixel der erzeugten Datei exakt der Lage der Pixel der anderen Daten entspricht, muss am unteren Rand des Dialogfensters die Schaltfläche „Environments...“ angeklickt werden. Im sich daraufhin öffnenden Dialogfenster „Environment Settings“ muss der Eintrag „Processing Extent“ geöffnet und ein „Snap Raster“ ausgewählt werden. Als „Snap Raster“ sollte eines der Strukturparameter-Raster (z. B. die Datei mean_20_mosaic.tif) aus Abschnitt C.1 ausgewählt werden. Schließen Sie nun das Fenster „Environment Settings“ durch klicken auf „OK“ und klicken Sie im Fenster „Project Raster“ ebenfalls „OK“, um die Prozessierung zu starten.

C.4 Bodendaten

Um Bodeneigenschaften in der Modellierung von Holzvorrat und Biomasse berücksichtigen zu können, kann auf Daten von ISRIC – World Soil Information⁹ („Soilgrids-Daten“) zurückgegriffen werden. Hierbei ist jedoch zu beachten, dass es sich bei diesen Daten um einen globalen Datensatz handelt, der nach einem global konsistenten Verfahren basierend auf globalen Kovariaten und global angepassten Modellen erzeugt wurde. Es wird daher empfohlen, für nationale bzw. regionale Anwendungen Datensätze zu verwenden, die besser an die lokalen Gegebenheiten angepasst sind. Da unseres Wissens kein deutschlandweit nach einem einheitlichen Verfahren generierter Datensatz bezüglich der Bodeneigenschaften verfügbar ist, ist im Folgenden die Verarbeitung von Soilgrids-Daten erläutert. Es sei aber darauf hingewiesen, dass bei Verfügbarkeit von akkurateren Daten die Verwendung selbiger in Betracht gezogen werden sollte.

Soilgrids-Daten können über das Webportal www.soilgrids.org kostenfrei heruntergeladen werden. Die verfügbaren Datenkacheln haben eine Ausdehnung von 200 x 200 km und eine Auflösung von 250 m. Alternativ kann auf die Daten auch über einen Web Coverage Service (WCS)¹⁰ zugegriffen werden. Um die potentielle Kationenaustauschkapazität, den Stickstoffvorrat (Indikatoren für die Nährstoffversorgung/Bodenfruchtbarkeit) sowie die Lagerungsdichte als Indikator für die Durchwurzelbarkeit aus den Soilgrids-Daten berechnen zu können, müssen folgende Daten-Layer für alle 6 verfügbaren Bodentiefen (0 – 5 cm, 5 – 15 cm, 15 – 30 cm, 30 – 60 cm, 60 – 100 cm, 100 – 200 cm) über das Portal heruntergeladen werden:

- Bulk Density (cg/cm³)
- Coarse Fragments (cm³/dm³)
- Cation Exchange Capacity (at pH7) (mmol(c)/kg)
- Nitrogen (cg/kg)

Die Dateien müssen nach dem Herunterladen umbenannt werden. Hierbei ist darauf zu achten, dass die Dateien entsprechend ihrer Bodentiefe einen Indexwert erhalten. Dieser Indexwert muss so vergeben werden, dass der Bodentiefe 0 – 5 cm der Wert 1 zugewiesen wird, der Bodentiefe 5 – 15 cm der Wert 2 usw. bis zur Bodentiefe 100 – 200 cm, der der Wert 6 zugewiesen wird.

- Bulk Density → `bulkdensity_1.tif` bis `bulkdensity_6.tif`
- Coarse Fragments → `coarse_fragments_1.tif` bis `coarse_fragments_6.tif`
- Cation Exchange Capacity (at pH7) → `cation_exchange_capacity_1.tif` bis `cation_exchange_capacity_6.tif`
- Nitrogen → `nitrogen_1.tif` bis `nitrogen_6.tif`

Nachdem die Datenlayer heruntergeladen und umbenannt wurden, kann das R-Skript `soilgrids_F3.R`¹¹ verwendet werden um die Daten zu prozessieren. Das R-Skript berechnet aus den Datenlayern:

1. Maximale Bodendichte des Hauptwurzelraums (kg/dm³)
 - a. Dateiname: `bulkdensity_max.tif` (Hauptwurzelraum 0 – 100 cm)

⁹ International Soil Reference and Information Centre (ISRIC): <https://www.isric.org/> [letzter Zugriff 30.11.2020]

¹⁰ soilgrids WCS URLs: <https://maps.isric.org/> [letzter Zugriff 23.11.2020]

¹¹ R-Skript `soilgrids_F3.R`: https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021]

2. Stickstoffvorrat (t/ha)
 - a. Dateiname: nitrogen1.tif (Oberboden 0 – 30 cm)
 - b. Dateiname: nitrogen2.tif (Hauptwurzelraum 0 – 100 cm)
 - c. Dateiname: nitrogen3.tif (Gesamtes Profil 0 – 200 cm)
3. Potentielle Kationenaustauschkapazität (kmol(c)/ha)
 - a. Dateiname: cec1.tif (Oberboden 0 – 30 cm)
 - b. Dateiname: cec2.tif (Hauptwurzelraum 0 – 100 cm)
 - c. Dateiname: cec3.tif (Gesamtes Profil 0 – 200 cm)

R-Skript: „soilgrids_F3.R“

*Bevor das R-Skript „soilgrids_F3.R“ gestartet werden kann, muss in **Zeile 55** der Pfad zum Eingabeordner, der die heruntergeladenen Soilgrids-Daten enthält, spezifiziert werden. In **Zeile 57** muss der Pfad zum Ausgabeordner, der die prozessierten Daten enthalten soll, angegeben werden. Der Ausgabeordner muss zuvor manuell angelegt werden.*

Um die Daten für die weitere Verarbeitung aufzubereiten, muss ihre Auflösung auf 20 m erhöht sowie die Projektion der Soilgrids-Daten (Homolosine WGS84¹²) an die Projektion der übrigen Daten angepasst werden. Hierzu kann abermals das ArcGIS-Tool „Project Raster“ Verwendung finden. Nach dem Aufrufen des Tools muss das zu prozessierende Raster ausgewählt und die Ausgabedatei (Dateiendung „.tif“) spezifiziert werden. Die gewünschte Projektion der Ausgabedatei ist ebenfalls anzugeben. Als „Resampling Technique“ sollte die Option „BILINEAR“ gewählt werden und als gewünschte Pixelgröße X = 20 und Y = 20. Um sicherzustellen, dass die Lage der Pixel der erzeugten Datei exakt der Lage der Pixel in den anderen Dateien entspricht, muss am unteren Rand des Dialogfensters die Schaltfläche „Environments...“ gedrückt werden. Im sich daraufhin öffnenden Dialogfenster „Environment Settings“ muss der Eintrag „Processing Extent“ geöffnet und anschließend ein „Snap Raster“ ausgewählt werden. Als „Snap Raster“ sollte eines der Strukturparameter-Raster (z. B. die Datei mean_20_mosaic.tif) ausgewählt werden. Schließen Sie nun das Fenster „Environment Settings“ durch klicken auf „OK“ und klicken Sie im Fenster „Project Raster“ ebenfalls „OK“ um die Prozessierung zu starten. Wiederholen Sie diesen Vorgang für alle sieben Daten-Layer.

C.5 Baumarten

Sollte die Ihnen zur Verfügung stehende Baumartenkarte (z. B. erzeugt nach dem in Meilenstein M5¹³ beschriebenen Verfahren) nicht in derselben Projektion und Auflösung wie die übrigen für die Modellierung verwendeten Datensätze vorliegen, muss auch sie mit dem ArcGIS-Tool „Project Raster“ prozessiert werden. Nach dem Aufrufen des Tools muss das zu prozessierende Raster ausgewählt und die Ausgabedatei (Dateiendung „.tif“) spezifiziert werden. Die gewünschte Projektion der Ausgabedatei ist ebenfalls anzugeben. Als „Resampling Technique“ sollte die Option „NEAREST“ gewählt werden und als gewünschte Pixelgröße X = 20 und Y = 20. Um sicherzustellen, dass die Lage der Pixel der erzeugten Datei exakt der Lage der Pixel in den anderen Dateien entspricht, muss am

¹² Homolosine WGS84 Koordinatensystem: <https://www.isric.org/explore/soilgrids/faq-soilgrids#How can I use the Homolosine projection> [letzter Zugriff 23.11.2020]

¹³ M5 – „Verfahren zur Ableitung von Baumarteninformationen aus Fernerkundungsdaten“: https://www.waldwissen.net/assets/FVA/Technik und Planung/F3-Projekt/F3_Dokumentation und Skripte 01.zip [letzter Zugriff 11.01.2021]

unteren Rand des Dialogfensters die Schaltfläche „Environments...“ gedrückt werden. Im sich daraufhin öffnenden Dialogfenster „Environment Settings“ muss der Eintrag „Processing Extent“ geöffnet und anschließend ein „Snap Raster“ ausgewählt werden. Als „Snap Raster“ sollte eines der Strukturparameter-Raster (z. B. die Datei mean_20_mosaic.tif) ausgewählt werden. Schließen Sie nun das Fenster „Environment Settings“ durch klicken auf „OK“ und klicken Sie im Fenster „Project Raster“ ebenfalls „OK“ um die Prozessierung zu starten.

D Modellierung / Erstellen von Holzvorrats- und Biomassekarten

Wenn alle für die Modellierung benötigten Datensätze (Referenzdaten als Shape-Datei, Prädiktoren als Rasterdateien) entsprechend der oben beschriebenen Vorgehensweise aufbereitet wurden, kann die Erstellung der Holzvorrats- und Biomassekarten mit Hilfe des R-Skripts *modellierung_volumen_biomasse_F3.R*¹⁴ durchgeführt werden.

R-Skript: *modellierung_volumen_biomasse_F3.R*

Bevor das R-Skript modellierung_volumen_biomasse_F3.R gestartet werden kann, müssen folgende Anpassungen vorgenommen werden:

Zeile 51 Definition des Arbeitsverzeichnisses (Working Directory)

Zeilen 55 & 56 Pfad und Dateiname unter dem die Ausgabedateien (Holzvolumen- und Biomassekarte) gespeichert werden sollen

Zeile 61 Pfad zum Referenzdatensatz (Shape-Datei)

Zeile 65 Name der Spalte der Referenzdaten-Attributtabelle, die das Holzvolumen enthält

Zeile 66 Name der Spalte der Referenzdaten-Attributtabelle, die die Biomasse enthält

Zeile 67 Name der Spalte der Referenzdaten-Attributtabelle, die die Höhe des höchsten Baumes pro Stichprobenpunkt enthält

Zeilen 72 – 78 Pfade zu den aus dem Oberflächenmodell abgeleiteten Strukturparametern: cctot6, cctot20, mean20, perc75, perc95, std20 und max20

Zeile 82 Dateiname des Rasters „max20“ ohne Dateiendung .tif

Zeilen 89 & 90 Pfade zu den aus dem DGM abgeleiteten Topographiemetriken

Zeile 95 Pfad zur Baumartenkarte

Zeile 97 Dateiname der Baumartenkarte

Zeilen 100 – 106 Pfade zu den Bodendaten: bulk density, cation exchange capacity 1 bis 3, nitrogen 1 bis 3

Zeilen 111 & 112 Pfade zu den Klimadaten: Temperatur und Niederschlag.

¹⁴ R-Skript *modellierung_volumen_biomasse_F3.R*:

https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip

Das Skript liest die aufbereiteten Prädiktoren (Rasterdateien) sowie Referenzdaten (Shape-Datei) ein und führt die folgenden Schritte durch:

- 1.) Maskieren aller Pixel mit einer Kronenüberschirmung < 25%.
- 2.) Zuschneiden aller Rasterdaten auf eine gemeinsame Ausdehnung.
- 3.) Auslesen der Pixelwerte der Rasterdaten an den Koordinaten der Referenzdatenpunkte (Stichprobenpunkte).
- 4.) Detektion und Löschen von Ausreißern: Weicht die höchste Baumhöhe eines Stichprobenpunktes der Referenzdaten um 25% von der höchsten Höhe des Oberflächenmodells an diesem Punkt ab, gilt dieser Datenpunkt als Ausreißer und wird aus dem Datensatz entfernt.
- 5.) Trainieren je eines Random Forest-Modells zur Modellierung von Holzvorrat und Biomasse.
- 6.) Anwendung der Modelle auf die Rasterdaten zur Erzeugung von Holzvorrats- und Biomassekarten.
- 7.) Berechnung der Vorhersagefehler (totaler und relativer RMSE) der Modelle basierend auf den Out-Of-Bag-Samples des Random Forest-Modells. Graphische Ausgabe eines Scatterplots (vorhergesagte vs. beobachtete Werte) inklusive der Vorhersagefehlerwerte.

Anders als bei der Erstellung der Waldstrukturparameter (siehe Meilenstein M4¹⁵), wird den erstellten Holzvorrats- und Biomassekarten kein Farbschema automatisch zugewiesen. Soll ein bestimmtes Farbschema, z. B. das in Abschnitt E dargestellte, Verwendung finden, muss dies durch die Nutzenden zugewiesen werden, z. B. in ArcGIS oder QGIS.

E Modellierung von Holzvorrat und Biomasse in den Testgebieten des F³-Projekts

E.1 Nordwestdeutsche Testgebiete

Für die Testgebiete des F³-Projekts in Nordwestdeutschland in Solling und Harz wurden unter Verwendung der oben beschriebenen Daten und des dargestellten Verfahrens Holzvorrats- und Biomassekarten erstellt. Für das Testgebiet in der Heide wurde keine Modellierung durchgeführt, da hier keine geeigneten Oberflächenmodelle zur Verfügung standen. Dies liegt daran, dass die Luftbildbefliegung des Gebiets Heide zu einem Zeitpunkt im Frühjahr (5. Mai 2016) stattfand, zu dem die Laubbäume noch nicht ausgetrieben hatten. Aufgrund der fehlenden Belaubung ist in den Luftbildern der Waldboden durch die Äste der Laubbäume zu sehen, was während des Image-Matchings dazu führt, dass die abgeleiteten Höhen der Höhe des Waldbodens, also der Geländehöhe, entsprechen und nicht der Vegetationshöhe.

Als Referenzdaten wurden für die Modellierung in den Testgebieten in Solling und Harz Stichprobendaten der Betriebsinventur der Niedersächsischen Landesforsten (NLF) aus den Jahren 2011 und 2012 verwendet. An den Stichprobenpunkten ($n = 2.133$) wurden auf konzentrischen Probekreisen mit Radien von 6 bzw. 13 m alle Bäume mit einem BHD von 7 – 30 cm resp. BHD > 30 cm aufgenommen. Die in den Jahren 2011 und 2012 gemessenen Baumhöhen und BHD wurden mit dem

¹⁵ M4 – „Methodenbeschreibung zur Ableitung von Waldstrukturparametern aus Oberflächenmodellen“:
https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021]

Waldwachstumssimulator ForestSimulator (Hansen und Nagel 2014¹⁶) auf das Jahr der Luftbildbefliegung 2016 fortgeschrieben. Die Modellierung des Derbholzvorrats jedes Baumes erfolgte ebenfalls mit dem ForestSimulator. Die Modellierung der oberirdischen Biomasse jedes Baumes erfolgte mit Hilfe des R-Paket *,rbdapro³*. Auf Grundlage dieser Daten wurde für jeden Stichprobenpunkt eine Schätzung des Derbholzvorrats pro ha und der oberirdischen Biomasse pro ha ermittelt.

Unter Verwendung des R-Skripts *modellierung_volumen_biomasse_F3.R* wurden die Referenzdaten mit den Rasterdaten verschnitten und ein Random Forest-Modell trainiert, das zur flächigen Modellierung von Derbholzvorrat bzw. Biomasse auf die Rasterdaten angewendet wurde. Die Daten der Testgebiete in Solling und Harz wurden zu einem Datensatz zusammengefügt und es wurde ein globales Modell erstellt, das für beide Testgebiete anwendbar ist.

Die Kreuzvalidierung des Random Forest-Modells ergab, dass das Modell in der Lage ist den Derbholzvorrat mit einem Vorhersagefehler von 117,2 m³/ha (RMSE) bzw. 30,9% (rel. RMSE) zu modellieren (vgl. Abbildung 1). Die mittels des Modells kartierten Derbholzvorräte der Testgebiete in Solling und Harz sind in Abbildung 2 und Abbildung 3 dargestellt. Das Modell für die Vorhersage der oberirdischen Baumbiomasse (BHD > 7 cm) weist einen Vorhersagefehler von 69,4 t/ha (RSME) bzw. 31,2% (rel. RMSE) auf (vgl. Abbildung 1). Die mittels des Modells kartierten Vorräte oberirdischer Biomasse der Testgebiete in Solling und Harz sind in Abbildung 4 und Abbildung 5 dargestellt.

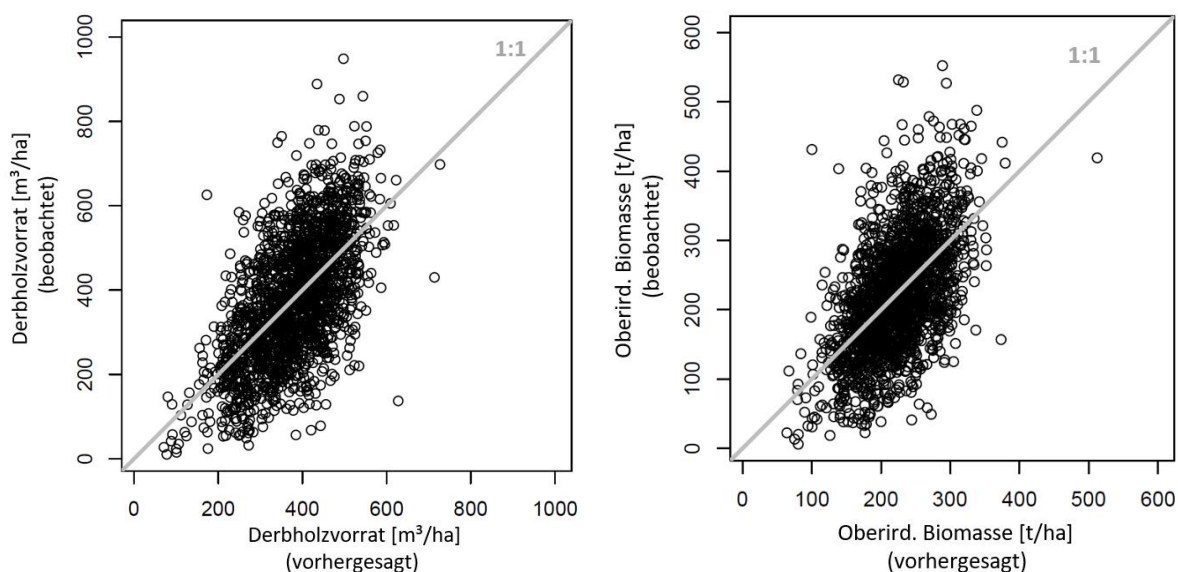


Abbildung 1: Darstellung der in den Testgebieten in Solling und Harz beobachteten Derbholz- bzw. Biomassevorräte, abgetragen gegen die mit dem Random Forest-Modell vorhergesagten Werte.

¹⁶ Hansen, J., Nagel, J., 2014. Waldwachstumskundliche Softwaresysteme auf Basis von TreeGroSS – Anwendung und theoretische Grundlagen. Beiträge aus der Nordwestdeutschen Forstlichen Versuchsanstalt Band 11. Göttingen: Universitätsverlag. DOI: 10.17875/gup2014-757

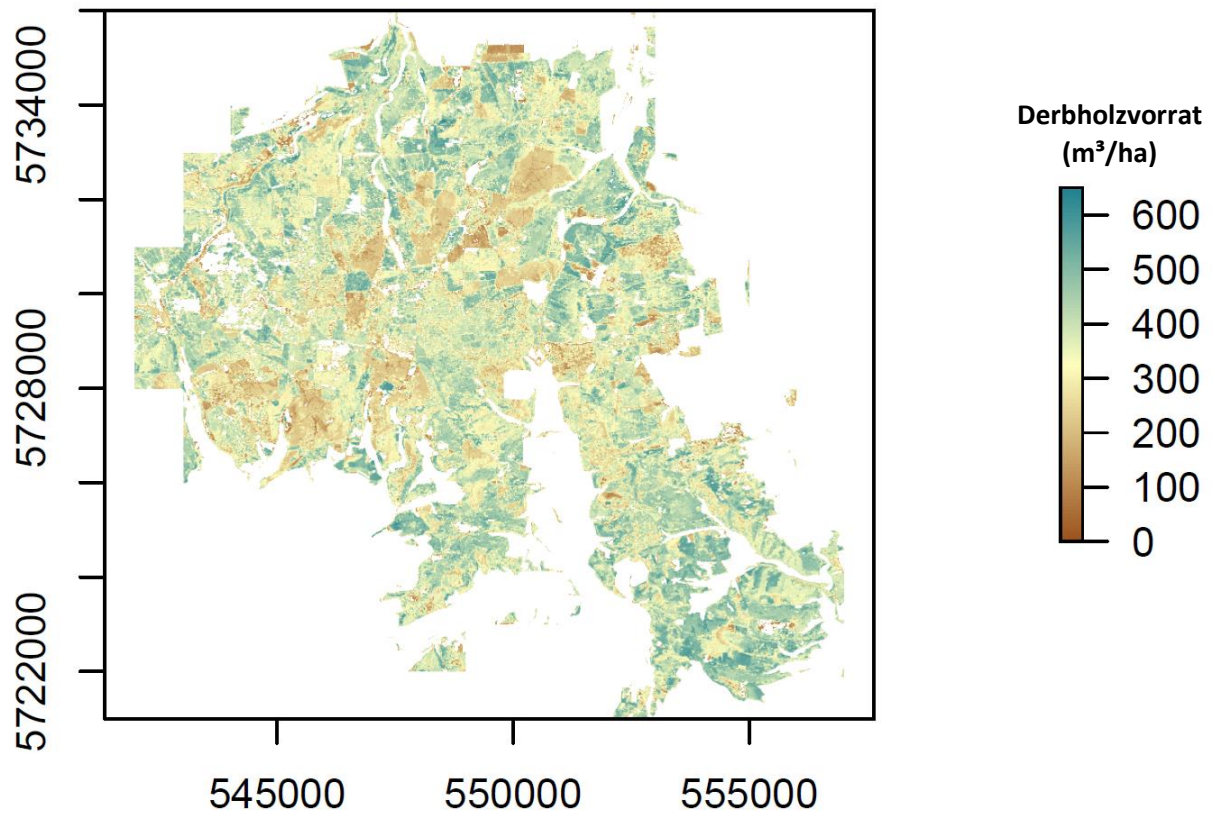


Abbildung 2: Verteilung der Derbholzvorräte (m³/ha) im Testgebiet Solling.

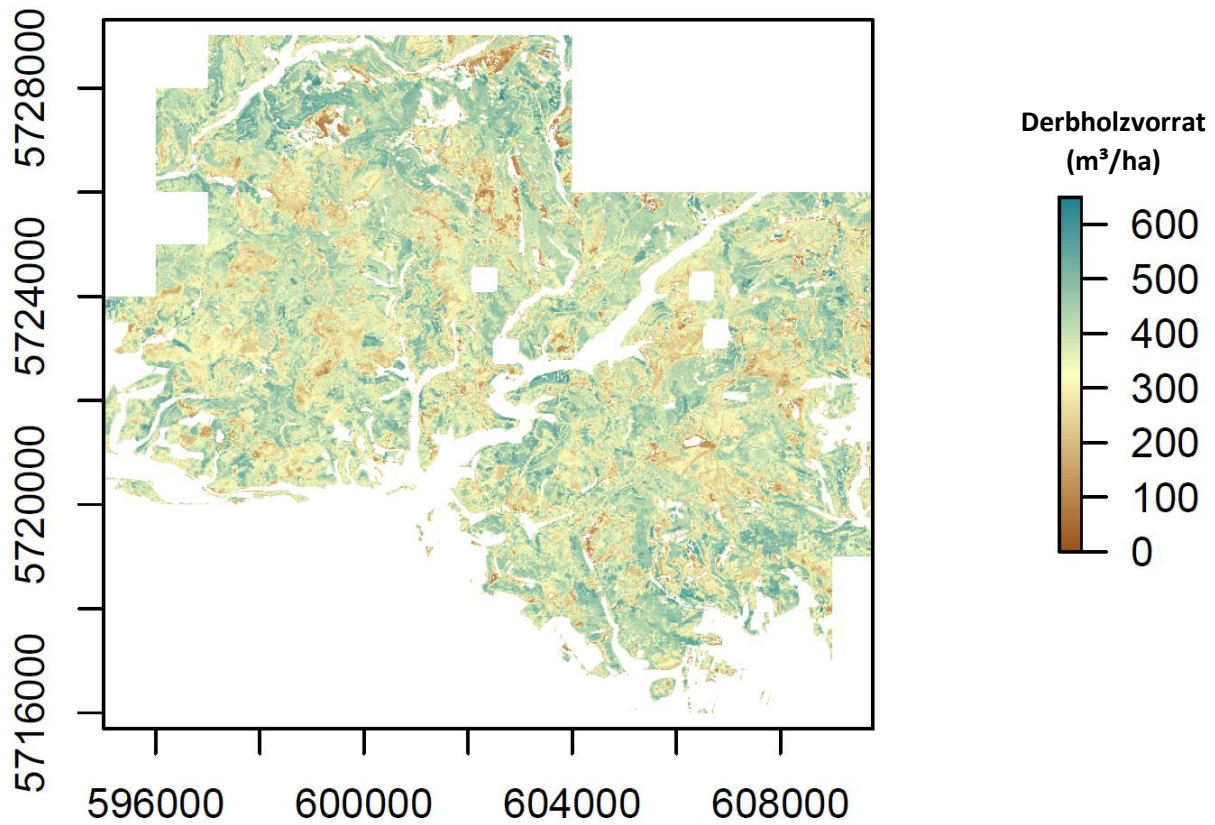


Abbildung 3: Verteilung der Derbholzvorräte (m³/ha) im Testgebiet Harz.

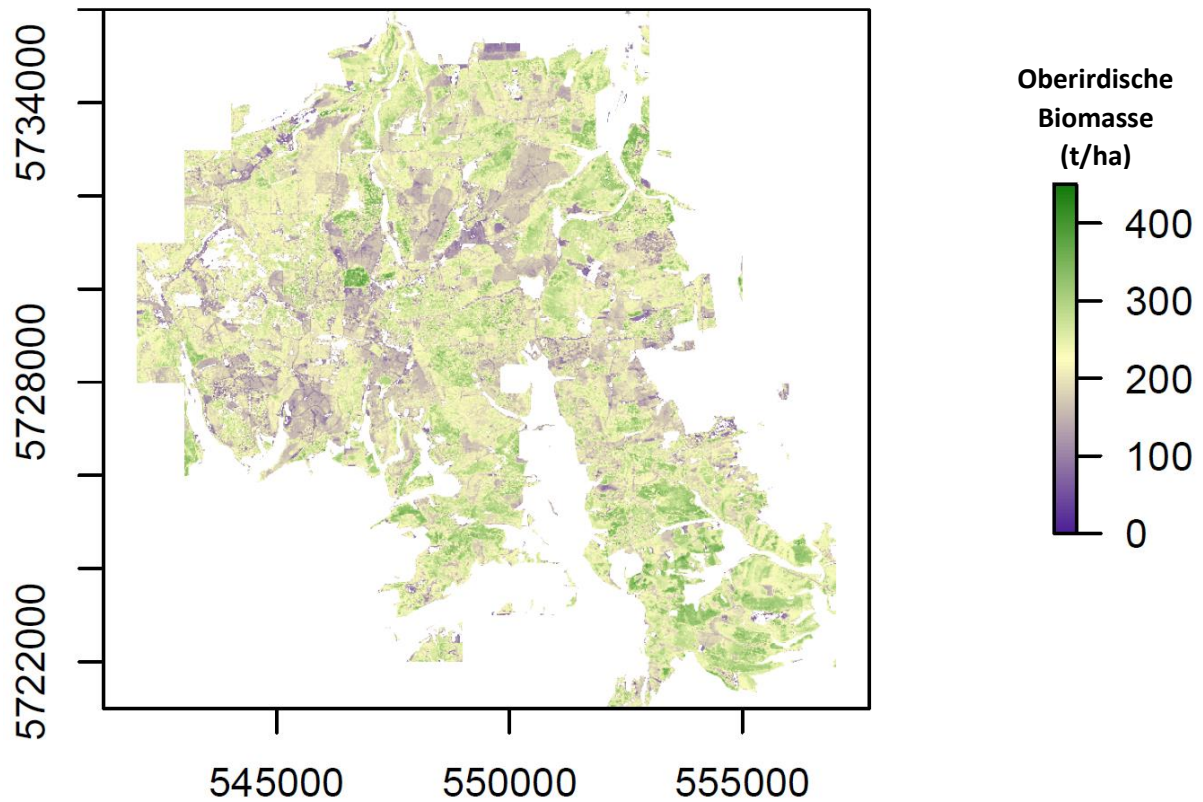


Abbildung 4: Verteilung der oberirdischen Biomasse (t/ha) im Testgebiet Solling.

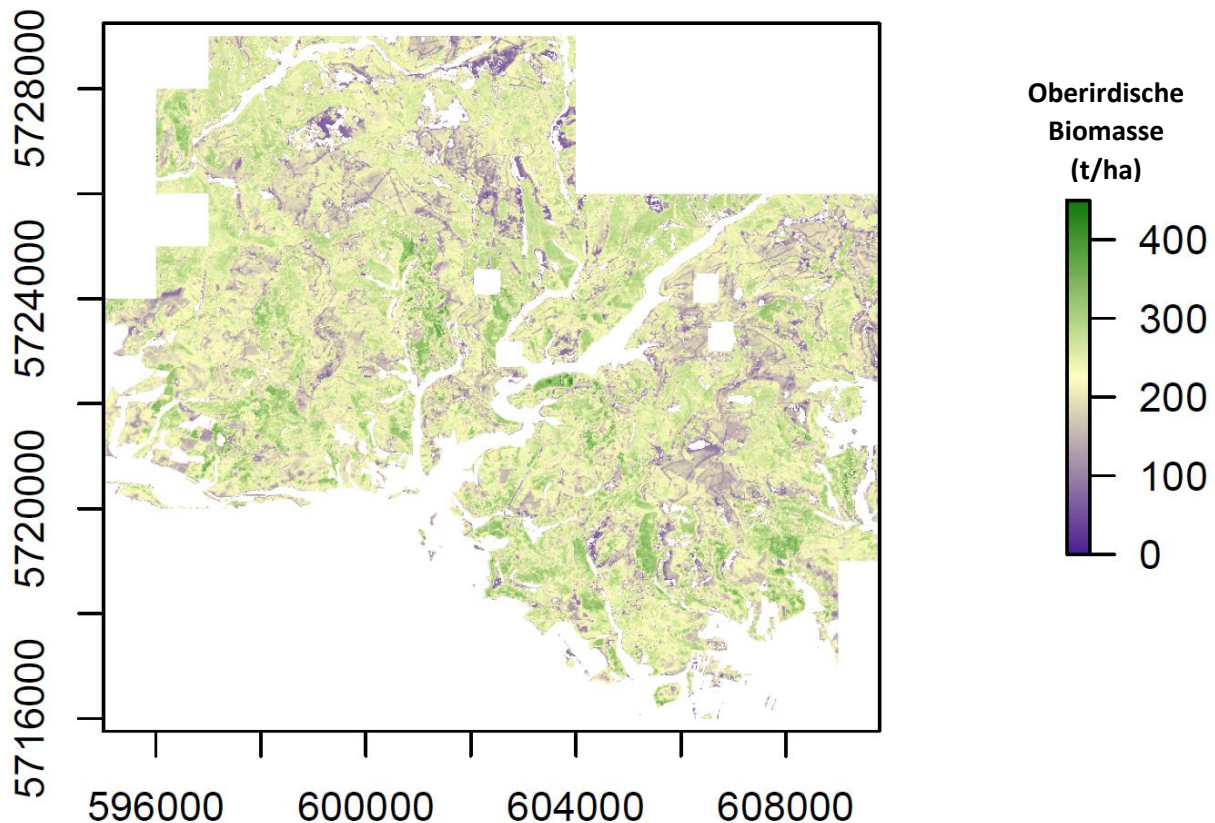


Abbildung 5: Verteilung der oberirdischen Biomasse (t/ha) im Testgebiet Harz.

E.2 Südwestdeutsche Testgebiete

Die F³-Testgebiete in Südwestdeutschland liegen in den Bereichen Südschwarzwald und südliche Oberrheinebene, Schwäbische Alb sowie Südwestdeutsches Alpenvorland. Für alle drei Testgebiete wurden mit dem oben beschriebenen Verfahren Holzvorrats- und Biomassekarten erstellt. Als Referenzdaten wurden hierfür Daten der BWI 3 genutzt, die von 2011 bis 2012 durchgeführt wurde. Die Auswahl der Probestämme bei der BWI 3 wurde mit der Winkelzählprobe (Zählfaktor 4) durchgeführt¹⁷. Für die Ableitung der Oberflächenmodelle wurden Luftbilddaten der regelmäßigen Befliegungen aus den Jahren 2012 und 2013 verwendet. Es wird davon ausgegangen, dass sich die Baumartenzusammensetzung innerhalb weniger Jahre nicht wesentlich ändert, weshalb es vertretbar ist, eine auf Sentinel-2-Satellitendaten aus dem Jahr 2017 basierende Baumartenkarte in der Modellierung zu verwenden. Gleiches gilt für die Bodendaten.

Der genutzte BWI 3-Datensatz beinhaltet Werte für Holzvorrat und oberirdische Biomasse pro Probestamm, sodass diese nicht gesondert modelliert werden mussten. Unter Nutzung der Information „Stammzahl pro ha“ wurden Vorrat und oberirdische Biomasse für jeden Probestamm auf einen Hektar hochgerechnet und anschließend pro Traktecke der BWI Stichprobenpunkte aggregiert. Im Rahmen dieses Prozesses wurde auch die maximale Höhe der Probestämme pro Traktecke bestimmt. Mit den Vorrats- bzw. Biomassedaten der $n = 2.164$ Traktecken des BWI-Datensatzes sowie den dort extrahierten Prädiktoren wurde je ein Random Forest-Modell trainiert.

¹⁷ Die dritte Bundeswaldinventur BWI 2012. Inventur- und Aufnahmemethoden: https://www.bundeswaldinventur.de/fileadmin/SITE_MASTER/content/Downloads/BWI_Methodenband_web.pdf [letzter Zugriff 30.11.2020]

Die Ergebniskarten der Holzvorrats- und Biomassemodellierung der südwestdeutschen Testgebiete sind in Abbildung 6 bis Abbildung 11 dargestellt. Mit Random Forest wurde der Derbholzvorrat in den Südwestdeutschen Testgebieten mit einem Vorhersagefehler von 198,7 m³/ha (RMSE) bzw. 43,7% (rel. RMSE) modelliert. Das Modell für die Vorhersage der oberirdischen Baumbiomasse (BHD > 7 cm) weist einen Vorhersagefehler von 103,3 t/ha (RMSE) bzw. 41,7% (rel. RMSE) auf (Vgl. Abbildung 12).

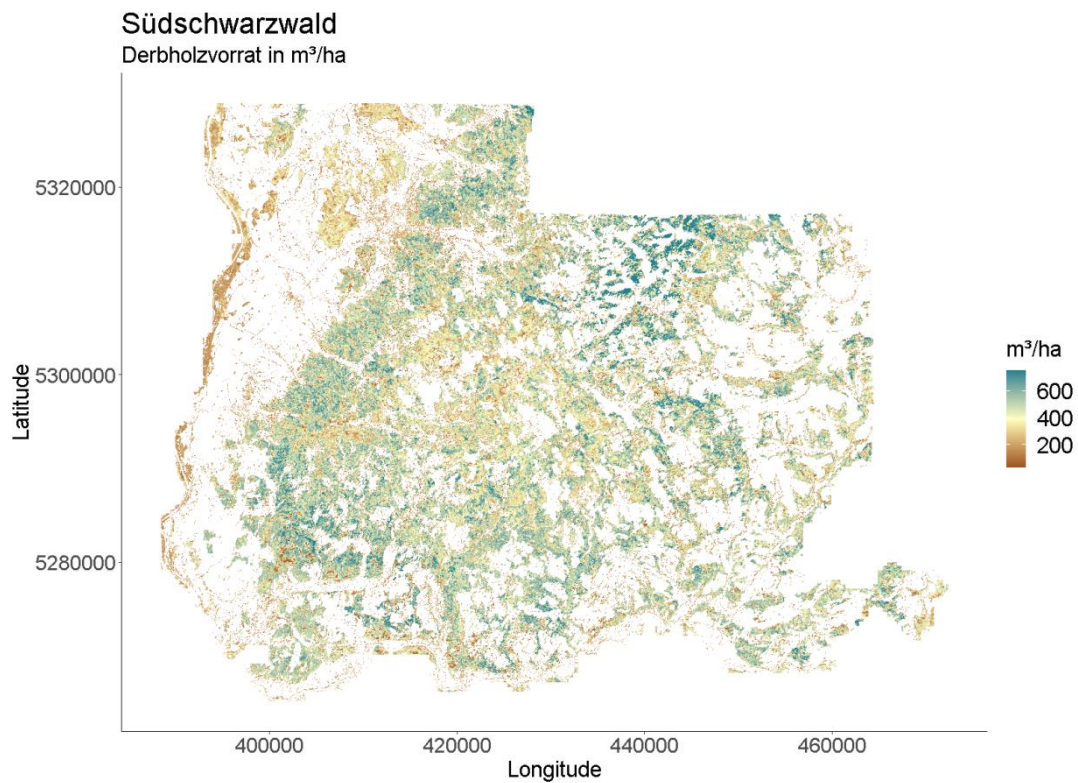


Abbildung 6: Verteilung der Derbholzvorräte (m³/ha) im Testgebiet „Südschwarzwald und südliche Oberrheinebene“.

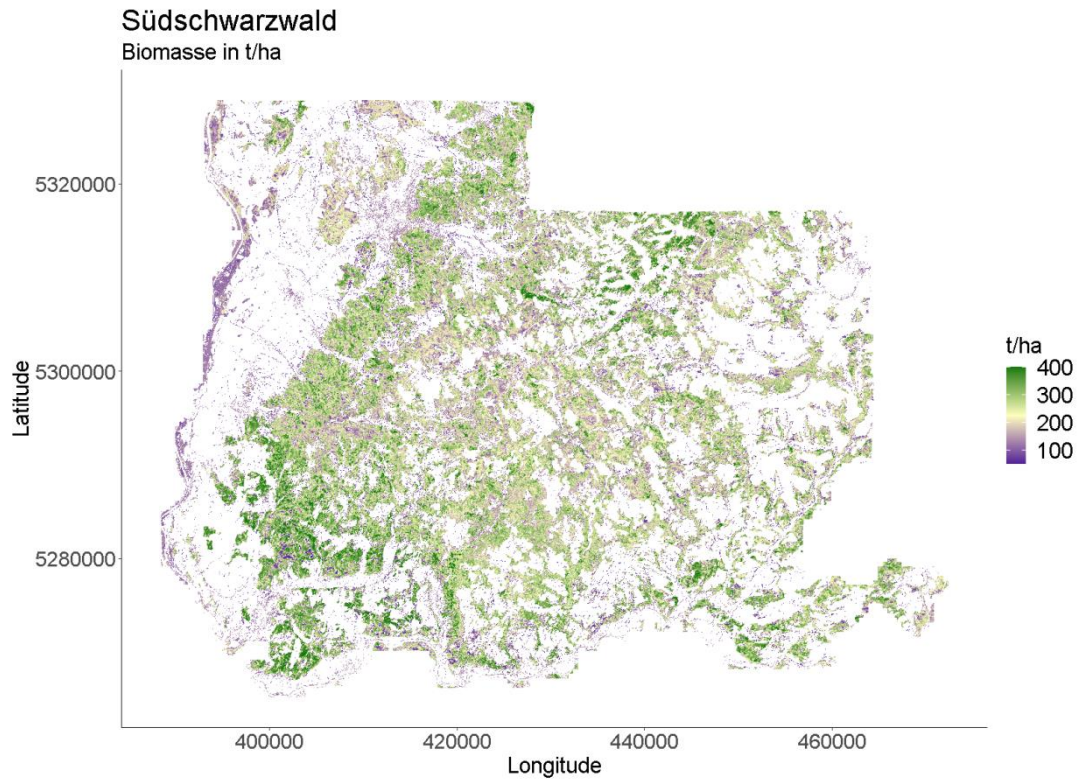


Abbildung 7: Verteilung der oberirdischen Biomasse (t/ha) im Testgebiet „Südschwarzwald und südliche Oberrheinebene“.

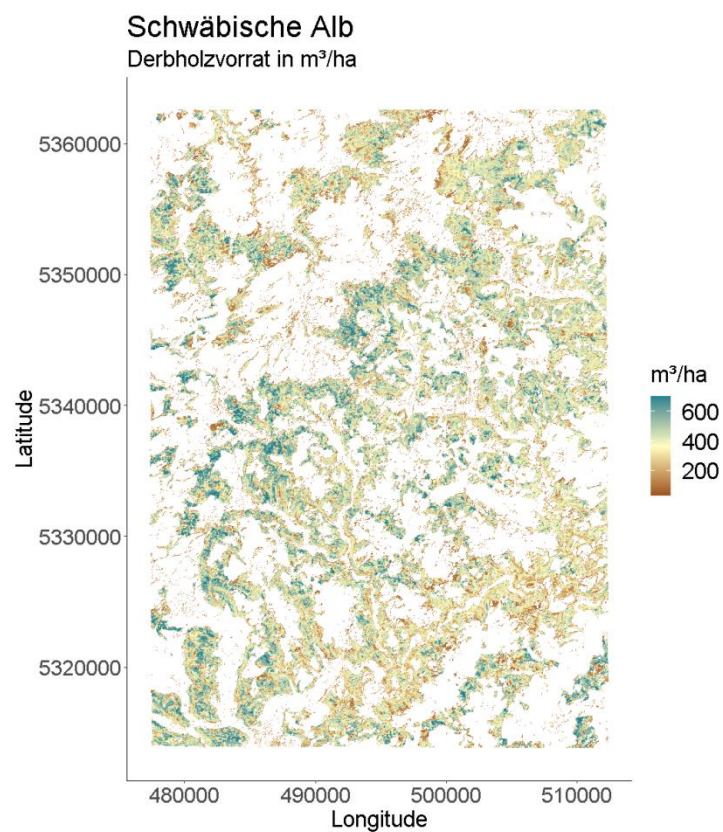


Abbildung 8: Verteilung der Derbholzvorräte (m³/ha) im Testgebiet „Schwäbische Alb“.

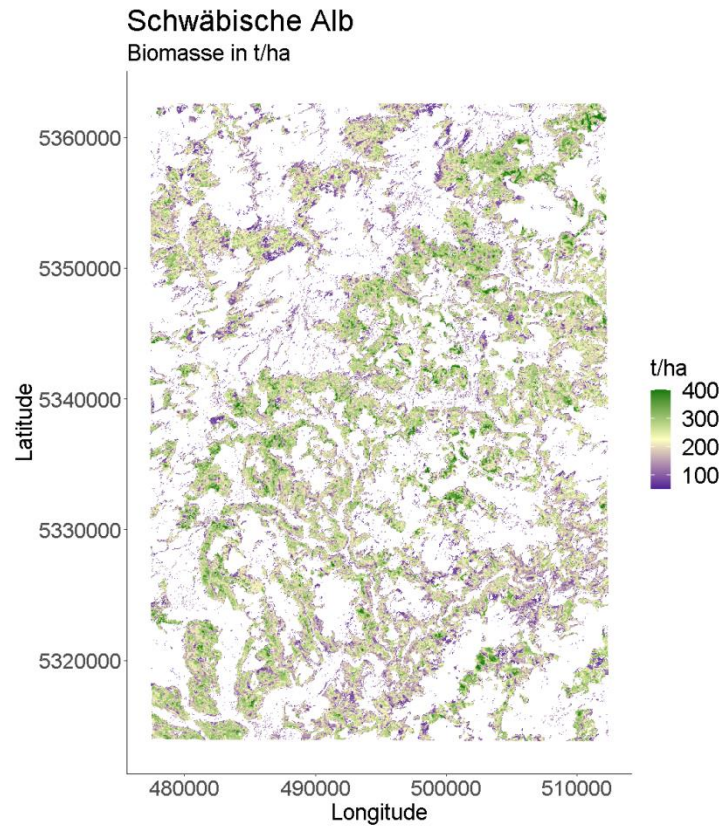


Abbildung 9: Verteilung der oberirdischen Biomasse (t/ha) im Testgebiet „Schwäbische Alb“.

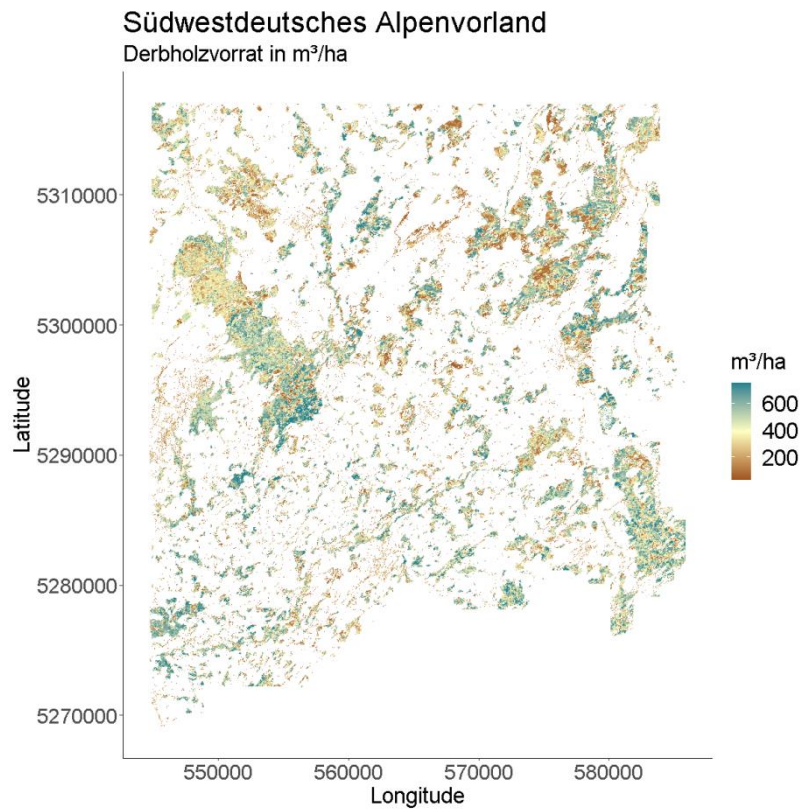


Abbildung 10: Verteilung der Derbholzvorräte (m³/ha) im Testgebiet „Südwestdeutsches Alpenvorland“.

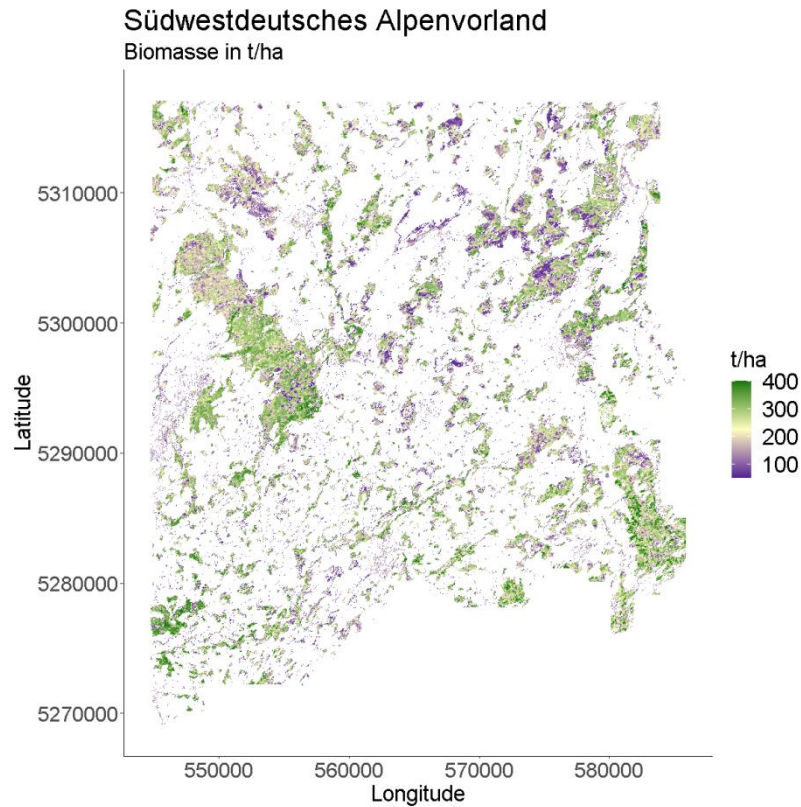


Abbildung 11: Verteilung der oberirdischen Biomasse (t/ha) im Testgebiet „Südwestdeutsches Alpenvorland“.

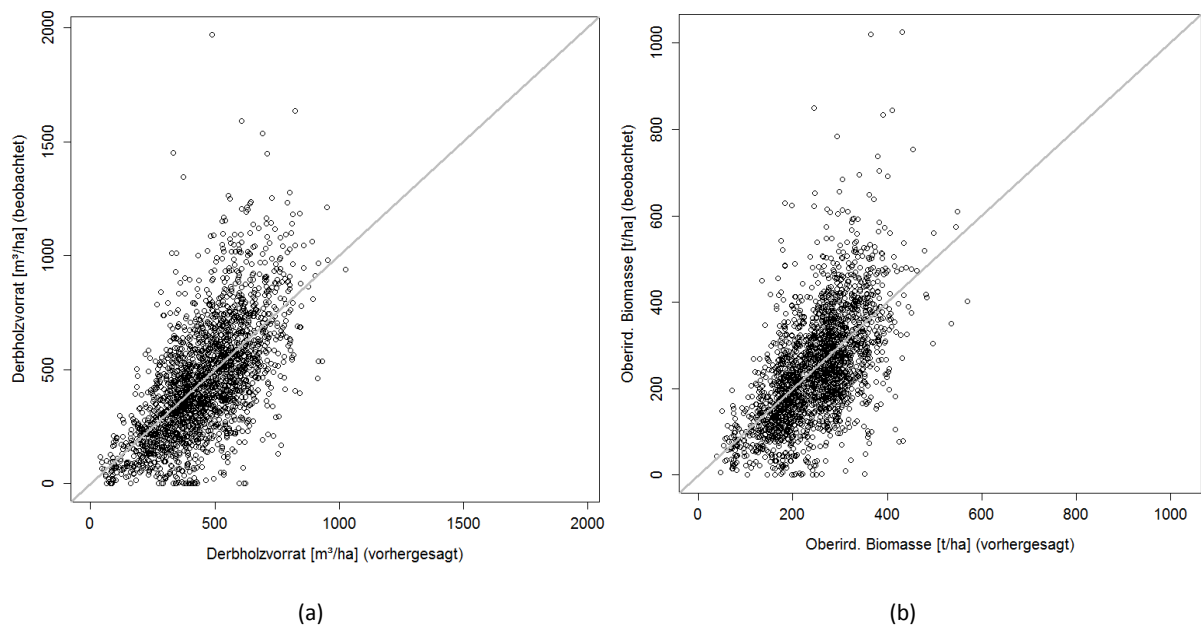


Abbildung 12: Darstellung der beobachteten (a) Derbholzvorrats- bzw. (b) Biomassewerte, abgetragen gegen die vorhergesagten Werte in den südwestdeutschen Testgebieten.

Um die Auswirkungen einer Erhöhung der Trainingsdatenmenge auf die erzielbare Genauigkeit zu untersuchen wurde der baden-württembergische Datensatz mit den entsprechenden Daten aus den



niedersächsischen Projektgebieten (Abschnitt E.1) kombiniert. Mit diesem kombinierten Datensatz wurden weitere Random Forest-Modelle trainiert. Im Vergleich zu den Modellen, die nur mit den baden-württembergischen Daten trainiert wurden, weisen die Modelle, die auf dem kombinierten Datensatz basieren, einen etwas geringeren Vorhersagefehler auf. So wurde beim Holzvorrat ein RMSE von 164,7 m³/ha (RMSE%: 39,3) und bei der oberirdischen Biomasse ein RMSE von 88,4 t/ha (RMSE%: 37,5) erreicht.