



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 5

„Verfahren zur Ableitung von Baumarteninformationen aus
Fernerkundungsdaten“



Meilenstein 5

Verfahren zur Ableitung von Baumarteninformationen aus Fernerkundungsdaten

A	Einleitung.....	1
B	Vorprozessierung von Sentinel-2-Satellitendaten.....	1
B.1	Daten	1
B.2	Benötigte Software.....	4
B.3	Prozessierung	5
B.3.1	Erstellen einer Wolkenmaske (Fmask 4.0)	5
B.3.2	Atmosphärische Korrektur	6
B.3.2.1	Alternative 1: SNAP – Sen2Cor Plugin	6
B.3.2.2	Alternative 2: Sen2Cor als eigenständige Software	8
B.3.3	Resampling der 20 m Bänder auf 10 m Auflösung (DSen2)	9
B.3.4	Topographische Korrektur und Berechnung von Vegetationsindizes (R)	10
C	Ableitung von Baumarteninformationen aus Sentinel-2-Satellitendaten	15
C.1	Referenzdaten	15
C.2	Baumartenerfassung	16
C.3	Validierungsergebnisse aus dem F ³ -Projekt	18
ANHANG	22



A Einleitung

Das Ziel der Baumartenklassifizierung im F³-Projekt ist die Herleitung von Zusatzinformationen für die Modellierung von Holzvorrat und oberirdischer Biomasse (siehe hierzu die Dokumentation zu Meilenstein M8 „Modellierung von Holzvorrat und oberirdischer Biomasse“¹). Die Baumart beeinflusst Holzvorrat und Biomasse. Als Zusatzinformation in einem Modellierungsprozess sollte eine Baumartenkarte deshalb helfen, die Genauigkeit der Vorhersage zu verbessern. Die mit den nachfolgend beschriebenen Methoden erzeugte Baumartenkarte erreicht eine Genauigkeit, die für den vorgesehenen Zweck im Projekt ausreicht, nicht aber in anderen Bereichen mit höheren Genauigkeitsansprüchen, etwa bei der forstbetrieblichen Inventur. Es sollte immer zuerst geprüft werden ob die entsprechenden Genauigkeitsanforderungen erfüllt werden.

Die Methodenbeschreibung gliedert sich in zwei Teile. In Abschnitt B ist die Vorprozessierung der Sentinel-2-Satellitendaten beschrieben. In Abschnitt C ist die Klassifizierung beschrieben sowie das Ergebnis der Kartenvalidierung dargestellt.

B Vorprozessierung von Sentinel-2-Satellitendaten

Ziel der hier beschriebenen Vorprozessierung von Sentinel-2-Satellitendaten ist die Gewinnung einer Datengrundlage, die zur Erstellung von Baumartenkarten genutzt werden kann. Die entwickelten Schritte der Prozessierung wurden daher für die Analyse von Waldflächen optimiert.

Die im Folgenden vorgestellte Prozessierung umfasst das Maskieren von Wolken, Wolkenschatten und Schnee, eine Korrektur atmosphärischer Einflüsse, das Resampling der 20 m-Bänder der Sentinel-2-Satellitendaten auf 10 m Auflösung, eine topographische Korrektur und die Berechnung von Vegetationsindizes.

B.1 Daten

Für das Verfahren werden Sentinel-2 **Level-1C²**-Satellitendaten benötigt. Das heißt, orthorektifizierte aber noch nicht atmosphärisch und topografisch korrigierte Daten. Die Daten können z. B. über den Copernicus Open Access Hub³ kostenfrei bezogen werden. Zum Auffinden der Level-1C-Daten auf dieser Plattform muss in den Suchkriterien der „Mission: Sentinel-2“ im Feld „Product Type“ der Eintrag „S2MSI1C“ ausgewählt werden (Abbildung 1).

¹ M8 – „Verfahren zur Modellierung von Holzvorrat und oberirdischer Biomasse“:

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

² Sentinel-2 Level-1C: <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/product-formatting> [letzter Zugriff 28.11.2020].

³ Copernicus Open Access Hub: <https://scihub.copernicus.eu/dhus/#/home> [letzter Zugriff 28.11.2020].

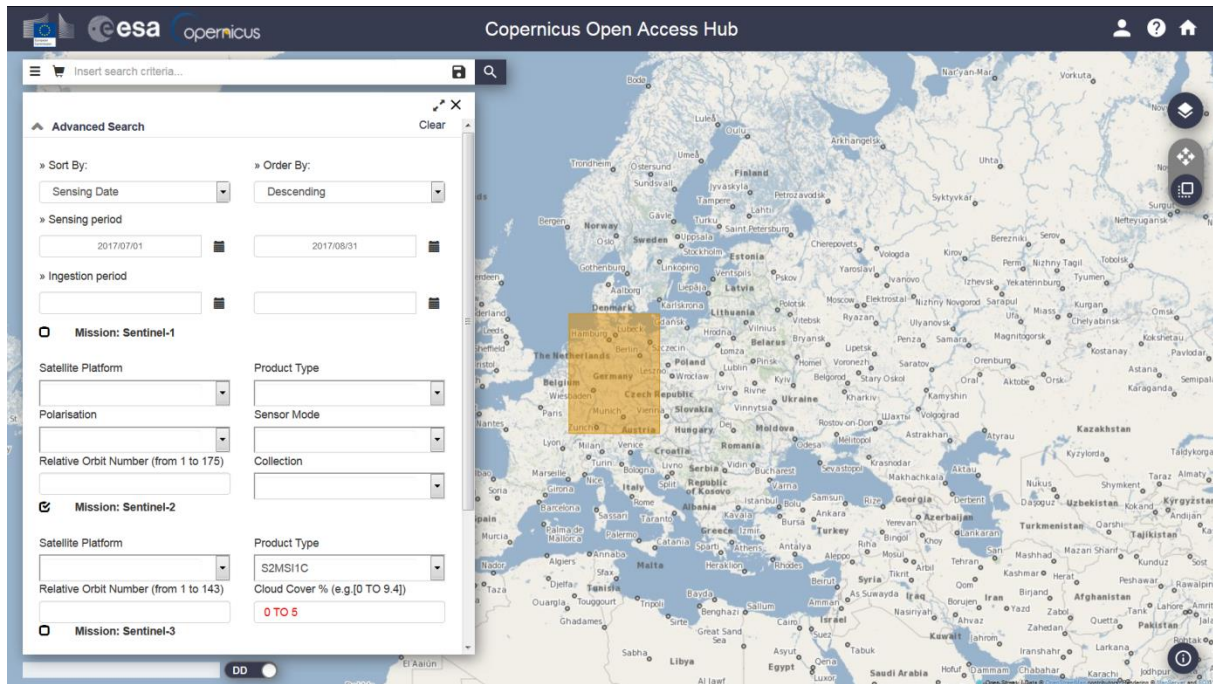


Abbildung 1: Suchmaske für Sentinel-Satellitendaten im Copernicus Open Access Hub. Einstellungen für Sentinel-2 Level-1C (Product Type: S2MS1C) mit Wolkenanteil von 0-5% aufgenommen zwischen 01. Juli und 31. August 2017 im Bereich des auf der Karte selektierten Gebietes.

Die Sentinel-2 Level-1C-Satellitendaten werden in 100 x 100 km Bildkacheln („Granules“) im UTM/WGS84-Koordinatensystem bereitgestellt. Die Granules sind um 5 km gepuffert, sodass bei Auswertungen Randeffekte vermieden werden können. Die Daten werden im Format „SAFE“ (Standard Archive Format for Europe) zur Verfügung gestellt. Diesem Format liegt eine feste Ordnerstruktur zugrunde, in der neben den Satellitendaten auch Metadaten gespeichert sind. Der Dateiname folgt einer einheitlichen Namenskonvention⁴:

MMM_MSIXXX_YYYYMMDDHHMMSS_Nxxyy_ROOO_Txxxxx_<Product Discriminator>.SAFE

- MMM: Missionsidentifikation (S2A/S2B)
- MSIXXX: MSIL1C steht für ein Level-1C-Produkt (MSIL2A stünde für ein Level-2A-Produkt, das hier nicht verwendet wird)
- YYYYMMDDHHMMSS: Datum und Startzeit der Datenaufnahme
- Nxxyy: PDGS Processing Baseline Nummer (z. B. N0204)
- ROOO: Relative Orbitnummer (R001 - R143)
- Txxxxx: Kachelnummer (z. B. T32UMU)
- Product Discriminator: Datums- und Zeitangabe mit 15 Stellen um zwischen verschiedenen Endprodukten der gleichen Datenaufnahme zu unterscheiden (z. B. 20170105T013443)
- SAFE: Produktformat (Standard Archive Format for Europe)

Die Kachelnummer stellt den Bezug zur geografischen Abdeckung der Kachel her. Die weltweite Kacheleinteilung (KML-Format) kann unter folgendem Link heruntergeladen werden: https://sentinel.esa.int/documents/247904/1955685/S2A_OPER_GIP_TILPAR_MPC_20151209T095117_V20150622T000000_21000101T000000_B00.kml [letzter Zugriff 28.11.2020]. Ein auf das

⁴ Sentinel-2 Namenskonvention: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/naming-convention> [letzter Zugriff 02.12.2020].

deutsche Bundesgebiet reduzierter Datensatz wird vom Julius Kühn Institut als „Web Feature Service“ (WFS) angeboten⁵. Dieser Datensatz wird in Abbildung 2 dargestellt.

Die Sentinel-2-Satelliten nehmen Daten in 13 verschiedenen spektralen Bändern mit unterschiedlichen räumlichen Auflösungen (10, 20 und 60 m) auf. Eine Übersicht der Sentinel-2-Bänder und ihrer räumlichen Auflösung ist in Tabelle 1 und Abbildung 3 zu finden.

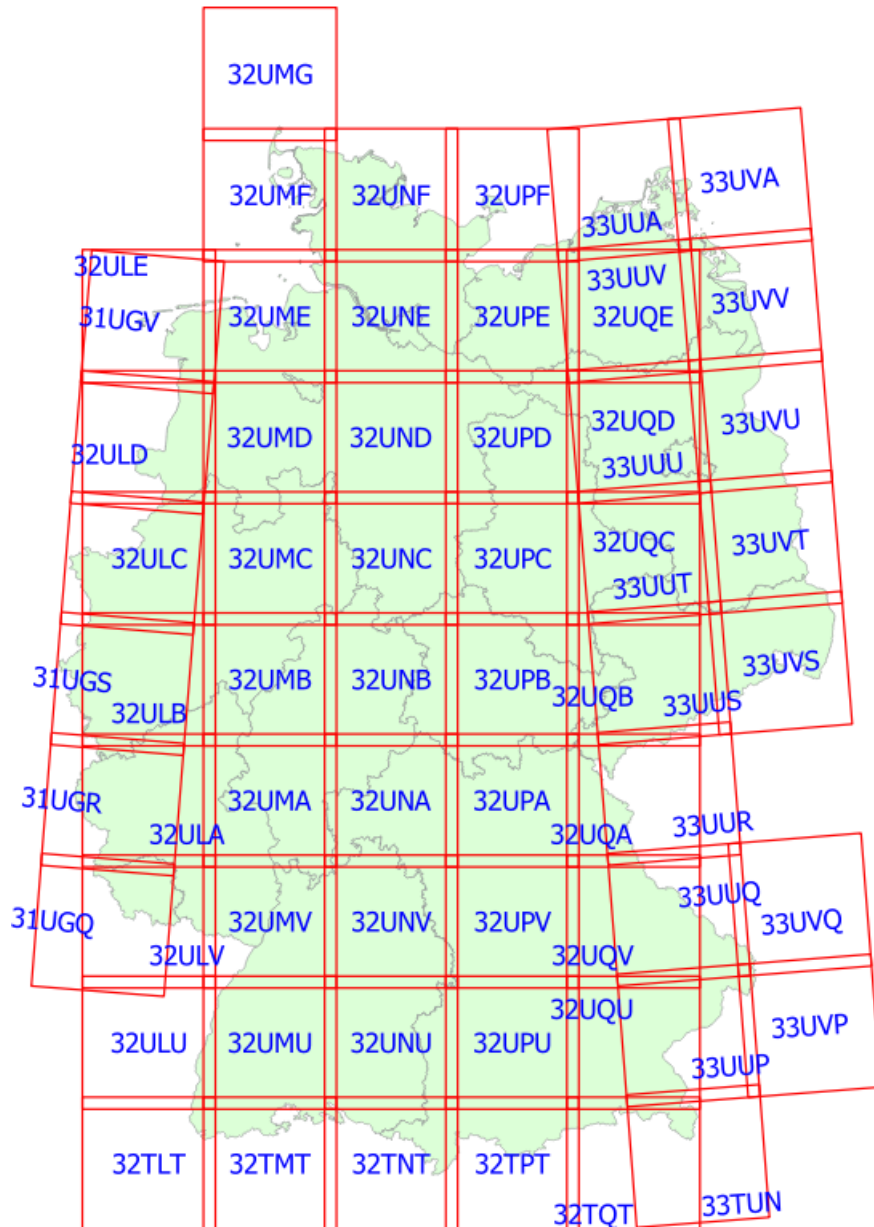


Abbildung 2: Sentinel-2-Kacheln für das Gebiet der BRD (Quelle: <https://ffj.julius-kuehn.de/webdienste/webdienste-des-ffj/sentinel-2-kacheln.html>).

⁵ Sentinel-2-Granule BRD: <https://ffj.julius-kuehn.de/webdienste/webdienste-des-ffj/sentinel-2-kacheln.html> [letzter Zugriff 28.11.2020].

Tabelle 1: Eigenschaften der spektralen Bänder der Sentinel-2-Satellitendaten.

Sentinel-2-Bänder	Mittlere Wellenlänge	Räumliche Auflösung
1 Aerosole	442 nm	60 m
2 Blau	492 nm	10 m
3 Grün	559 nm	10 m
4 Rot	665 nm	10 m
5 Red Edge	704 nm	20 m
6 Red Edge	739 nm	20 m
7 Red Edge	780 nm	20 m
8 NIR	833 nm	10 m
8A Narrow NIR	864 nm	20 m
9 Wasserdampf	943 nm	60 m
10 SWIR–Cirruswolken	1377 nm	60 m
11 SWIR	1610 nm	20 m
12 SWIR	2186 nm	20 m

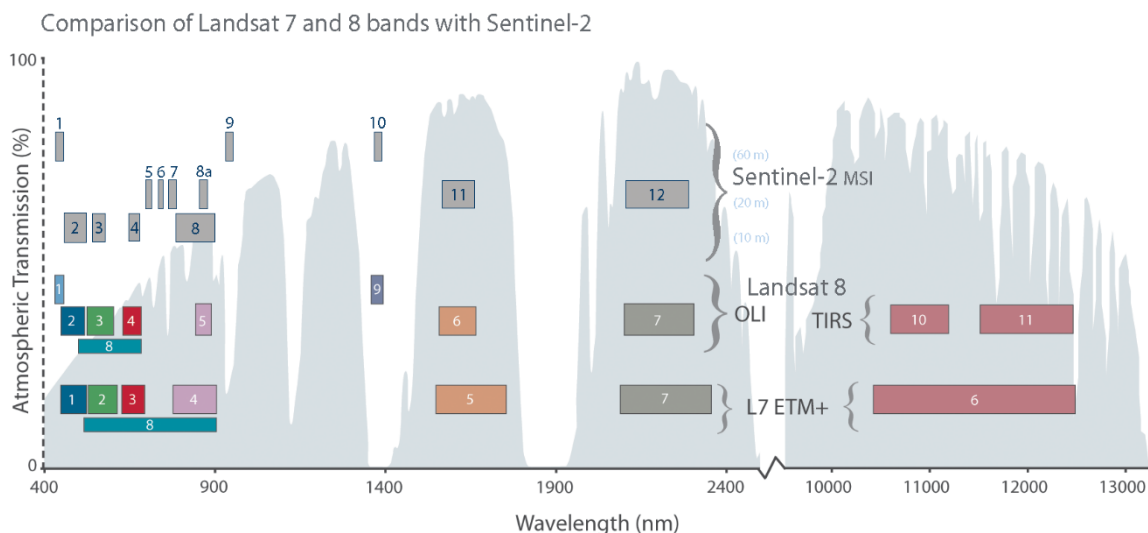


Abbildung 3: Darstellung der spektralen Bänder der Sentinel-2-Satellitendaten im Vergleich zu Landsat-8- und Landsat-7-Satellitendaten (Quelle: <https://landsat.gsfc.nasa.gov/article/sentinel-2a-launches-our-compliments-our-complements> [letzter Zugriff 11.01.2021]).

Neben den Sentinel-2-Satellitendaten wird ein digitales Geländemodell (DGM) mit einer Auflösung von 10 m oder höher benötigt, um eine topographische Korrektur der Satellitendaten zu ermöglichen (siehe Abschnitt B.3.4).

B.2 Benötigte Software

Für die Atmosphärenkorrektur von Sentinel-2 Level-1C-Satellitendaten wird die Sentinel Toolbox der Software SNAP sowie das Plugin Sen2Cor verwendet. Alternativ kann Sen2Cor auch als eigenständige Software genutzt werden. Zur Erstellung einer Wolkenmaske wird die Software Fmask 4.0 benötigt. Das Resampling der 20 m-Bänder erfolgt mit dem Python-Skript DSen2. Die topographische Korrektur sowie die Berechnung der Vegetationsindizes werden mit der Statistiksoftware R durchgeführt.



Alle benötigten Programme sind kostenfrei verfügbar und stehen unter folgenden Internetadressen zum Download bereit.

- SNAP (Sentinels Application Platform)
 - <http://step.esa.int/main/download/snap-download/> [letzter Zugriff 11.01.2021]
- Sen2Cor (SNAP Plugin und als eigenständige Software)
 - http://step.esa.int/main/snap-supported-plugins/sen2cor/sen2cor_v2-8/ [letzter Zugriff 11.01.2021]
- Fmask 4.0
 - <https://github.com/gersl/fmask> [letzter Zugriff 11.01.2021]
- R
 - <https://www.r-project.org/> [letzter Zugriff 11.01.2021]
- DSen2
 - <https://github.com/lanha/DSen2> [letzter Zugriff 11.01.2021]

B.3 Prozessierung

B.3.1 Erstellen einer Wolkenmaske (Fmask 4.0)

Zur Erstellung einer Sentinel-2-Wolkenmaske wird das Programm Fmask 4.0 verwendet. Nach dem Start des Programms müssen lediglich wenige Einstellungen vorgenommen werden (siehe Abbildung 4). Die im F³-Projekt ermittelten Einstellungen resultierten für die meisten Bilder nach visueller Überprüfung in zufriedenstellenden Wolkenmasken. Die Einstellungen wurden so gewählt, dass der Wolkenanteil eher überschätzt wird, also tendenziell mehr Wolken detektiert werden als im Bild vorhanden sind. Dadurch wird verhindert, dass Gebiete mit Wolkeneinfluss später bei der Klassifizierung der Baumarten verwendet werden und Fehlklassifikationen verursachen. Es wird empfohlen, die generierte Wolkenmaske mit der tatsächlichen Verteilung von Wolken im Bild visuell zu vergleichen und die Einstellungen des Programms Fmask anzupassen, sollten die Wolken nicht ausreichend genau erfasst worden sein. Den größten Einfluss auf die Menge erfasster Wolken hat der Wert „Cloud Probability Threshold“. Werden hier höhere Werte eingetragen, detektiert das Programm weniger Wolken, bei niedrigeren Werten steigt die Menge detektierter Wolken an. Auch der Menüpunkt „Result Dilations“ beeinflusst die Größe der ausgewiesenen Wolken- und Wolkenschatten stark. Dieser Wert definiert, wie groß der Buffer sein soll, mit dem die detektierten Wolken umgeben werden. Eine Anpassung dieser Werte kann, je nach Wolkenart, angeraten sein. Wolken, die zu den Rändern hin ausfransen und dünne, semi-transparente Schleier bilden, werden an den Rändern häufig nur unzureichend genau erfasst. Durch die Wahl eines höheren Wertes für den Buffer können die Wolkenränder dennoch in der Wolkenmaske berücksichtigt werden. Dicke und kompakte Wolken, die sich eindeutig abgrenzen lassen, benötigen hingegen nur einen Buffer von geringer Größe. Nachfolgend ist der Arbeitsablauf mit der Software Fmask beschrieben:

- 1.) Öffnen des Programms Fmask 4.0 und Anpassen der folgenden Einstellungen.
- 2.) *Image Path*: Navigieren zum Ordner, der die Sentinel-2 Level-1C-Daten enthält, für die eine Wolkenmaske erstellt werden soll. Das Programm prozessiert alle Sentinel-2-Satellitendaten, die in diesem Ordner enthalten sind.
- 3.) *Cloud Probability Threshold (%)*: Durch Klicken auf den Button „Custom“ wird die Eingabe eines Wertes ermöglicht. Mit dem Schieberegler den Wert auf 5.0 setzen, alternativ kann der Wert mittels Tastatur in das Feld rechts des Reglers eingetragen werden.

- 4.) *Result Dilations (Pixels)*: Die hier eingetragenen Werte geben die Größe des Puffers an, der um detektierte Wolken, Wolkenshatten und Schnee gezeichnet werden soll. Werte für Cloud auf 100, Cloud Shadow auf 100 und Snow auf 3 setzen.
- 5.) *Run*: Nachdem die Einstellungen angepasst wurden, wird mit dem Button „Run“ das Programm gestartet. Die Prozessierungsdauer beträgt in der Regel 5 bis 10 Minuten pro Sentinel-2-Kachel.
- 6.) *Output*: Die generierte Wolkenmaske wird innerhalb der Ordnerstruktur der Sentinel-2-Satellitendaten im Ordner „FMASK_DATA“ als TIFF-Datei gespeichert.

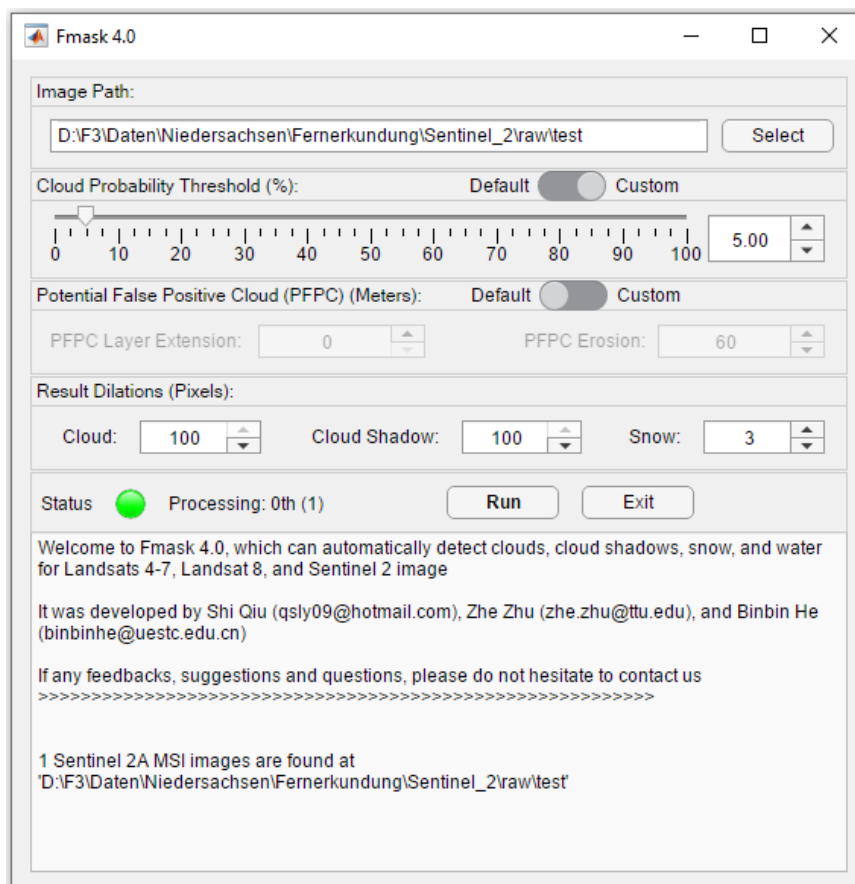


Abbildung 4: GUI des Programms Fmask 4.0 mit den im F³-Projekt genutzten Einstellungen zur Erzeugung einer Sentinel-2-Wolkenmaske.

B.3.2 Atmosphärische Korrektur

Die Software Sen2Cor wird verwendet, um eine Atmosphären- sowie Cirruskorrektur der Sentinel-2-Satellitendaten durchzuführen. Die Korrektur reduziert den Einfluss atmosphärischer Störungen auf die Pixelwerte der Satellitendaten. Sie kann entweder als Plugin über die Software SNAP genutzt oder als eigenständiges Programm über die Kommandozeile ausgeführt werden.

B.3.2.1 Alternative 1: SNAP – Sen2Cor-Plugin

Das Sen2Cor-Plugin für die Software SNAP stellt eine grafische Benutzeroberfläche bereit, in der alle Einstellungen getätigt werden können. Der Vorteil besteht in der dadurch leichteren Handhabung der Software.

- 1.) Öffnen der Sentinel-2 Level-1C-Satellitendaten in der Software SNAP.
- 2.) Öffnen des Plugins Sen2Cor über die Menüpunkte „Optical“ → „Thematic Land Processing“ → „Sen2Cor Processor“ → „Sen2Cor280“ (Abbildung 5).

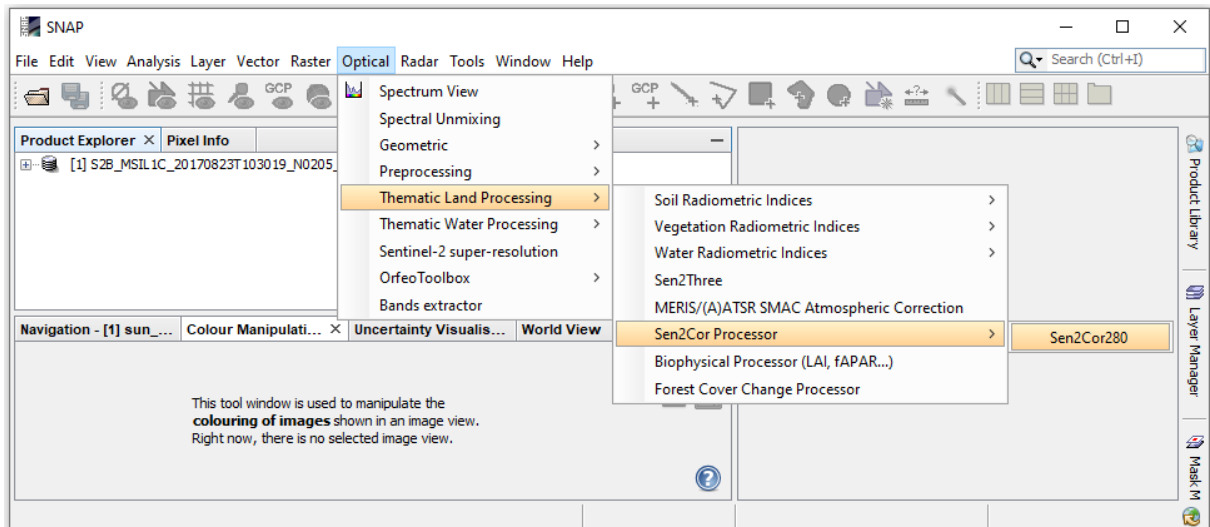


Abbildung 5: Öffnen des Sen2Cor 2.8 Plugins in der Software SNAP.

- 3.) Im geöffneten GUI des Sen2Cor-Plugins die Registerkarte „Processing Parameter“ wählen und die folgenden Einstellungen vornehmen (Abbildung 6):
 - a. Resolution: ALL
 - b. *Nb threads*: Anzahl der Prozessoren des PCs, die für die Prozessierung genutzt werden sollen.
 - c. Cirrus correction: TRUE
 - d. Alle anderen Einstellungen bleiben unverändert.
 - e. Auf „Run“ klicken, um den Prozess zu starten.
 - f. Output: Nach Beendigung der Berechnung befindet sich im Ordner, der auch das verwendete Level-1C-Bild enthält, ein neuer Ordner. Dieser enthält das korrigierte Bild und ist an dem Namensbestandteil „MSIL2A“ zu erkennen.

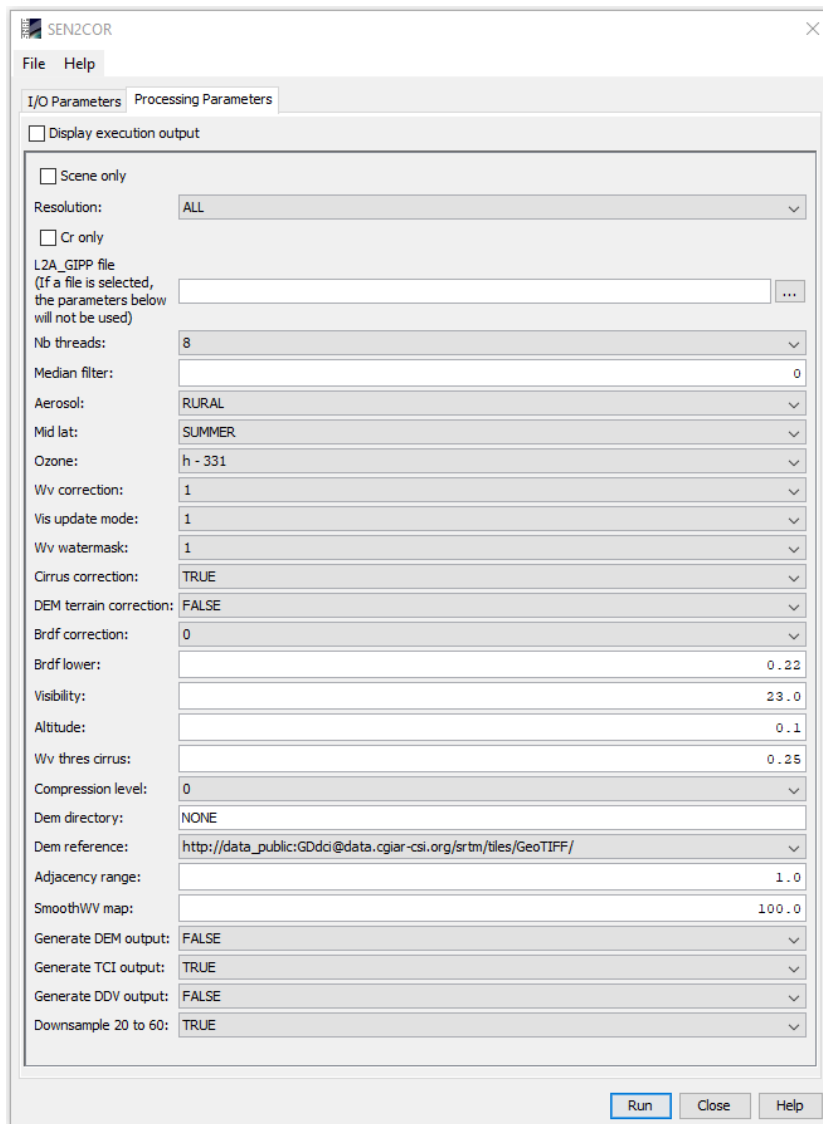


Abbildung 6: GUI des Sen2Cor-Plugins mit den für die atmosphärische Korrektur empfohlenen Einstellungen.

B.3.2.2 Alternative 2: Sen2Cor als eigenständige Software

Bei der Nutzung von Sen2Cor als eigenständige Software muss die Software SNAP nicht installiert werden. Zudem werden alle Einstellungen im XML-Format gespeichert, was eine weitestgehend automatisierte Prozessierung von Sentinel-2-Satellitendaten ermöglicht. Nachfolgend wird auf die Nutzung von Sen2Cor V2.8 über die Kommandozeile eingegangen:

- 1.) Sen2Cor V2.8 über http://step.esa.int/main/snap-supported-plugins/sen2cor/sen2cor_v2-8 [letzter Zugriff 11.01.2021] herunterladen und installieren. Bei der Installation wird das Verzeichnis „C:\Users\nutzer.name\Documents\sen2cor\2.8\“ angelegt. Dies ist das \$HOME-Verzeichnis.
- 2.) Im \$HOME-Verzeichnis befindet sich im Ordner „cfg“ die Datei „L2A_GIPP.xml“ mit der durch Anpassung der Inhalte bestimmter XML-Tags verschiedene Nutzungseinstellungen vorgenommen werden können. Die Namen der XML-Tags haben Bezug zu den Parameternamen in Abbildung 6 und die Einstellungen können entsprechend aus dieser Abbildung übernommen werden.



Zum Starten des Prozesses muss in der Kommandozeile zum Installationsverzeichnis von Sen2Cor (hier C:\Sentinel\Sen2Cor-02.08.00-win64) navigiert werden. Dort wird folgende Eingabe gemacht und anschließend die Prozessierung gestartet:

```
L2Aprocess.bat
S:\Pfad\Eingabe\S2A_MSIL1C_20190619T103031_N0207_R108_T32UMU_20190
619T141311.SAFE --resolution 10 --output_dir S:\Pfad\Ausgabeordner
```

Wenn die Auflösung („resolution“), wie hier, auf 10 m gesetzt ist, werden auch die 20 m-Bänder automatisch mitprozessiert. Andere Einstellmöglichkeiten für „resolution“ sind „20“ und „60“. Die GIPP-Datei, wird automatisch gefunden, weshalb es unbedingt nötig ist, dass diese im vorgegebenen Ordner unter C:\Users\nutzer.name\Documents\sen2cor\2.8\cfg\ (siehe oben) verbleibt. Auch der Name darf nicht geändert werden.

Die Ergebnisse aus Sen2Cor werden in einer festgelegten Dateistruktur gespeichert. Dabei wird jedes einzelne Band in einer separaten Datei (*.jp2) abgelegt. Die ursprünglichen 10 m-Bänder (B02 (Blau), B03 (Grün), B04 (Rot), B08 (nahes Infrarot)) werden nach der Korrektur in einer Auflösung von 10 m sowie 20 m abgelegt. Die 20 m-Bänder (B05, B06, B07, B8A, B11, B12) behalten ihre Auflösung bei.

B.3.3 Resampling der 20 m-Bänder auf 10 m-Auflösung (DSen2)

Sentinel-2-Satellitendaten beinhalten 13 Spektralbänder, die in unterschiedlichen Auflösungen aufgezeichnet wurden. Mit Hilfe des DSen2 Python-Skripts werden die in 20 m-Auflösung vorliegenden Bänder auf eine Auflösung von 10 m hochskaliert. Hierbei werden die 20 m-Pixel mittels eines *Deep Neural Networks* unter Zuhilfenahme der 10 m-Bänder spektral entmischt. Details der angewendeten Methode können hier nachgelesen werden:

<https://arxiv.org/pdf/1803.04271.pdf>⁶ [letzter Zugriff 11.01.2021]

<https://github.com/lanha/DSen2> [letzter Zugriff 11.01.2021]

- 1.) Das Skript ist derzeit nur für die Prozessierung von Sentinel-2-Satellitendaten verfügbar, die in der Ordnerstruktur der Level-1C vorliegen. Aus diesem Grunde müssen die im vorherigen Schritt erstellten Level-2A-Daten in den ursprünglichen Level-1C-Ordner kopiert werden. Hierzu wird in den „IMG_DATA“ Ordner der Level-1C-Daten navigiert und es werden alle dort vorhandenen Dateien gelöscht (Level-1C-Ordnerpfad: z. B. D:\F3\S2B_MSIL1C_20170823T103019_N0205_R108_T32UND_20170823T103018.SAFE\GRANULE\L1C_T32UND_A002422_20170823T103018\IMG_DATA)
- 2.) Im Ordner „IMG_DATA“ der Level-2A-Daten befinden sich drei Ordner mit den Bezeichnungen „R10m“, „R20m“ und „R60m“.
 - a. Den Ordner „R10m“ öffnen und die .jp2-Dateien der Bänder B02, B03, B04 und B08 in den Ordner „IMG_DATA“ des Level-1C-Produkts kopieren.
 - b. Den Ordner „R20m“ öffnen und die .jp2-Dateien der Bänder B05, B06, B07, B08A, B11 und B12 in den Ordner „IMG_DATA“ des Level-1C-Produkts kopieren.
 - c. Aus allen Dateinamen im Level-1C-Ordner die Zeichenfolgen „_10m“ bzw. „_20m“ entfernen, sodass die Dateinamen wieder der ursprünglichen Level-1C-Namensgebung entsprechen, z. B. T32UNC_20170823T103019_B02_10m.jp2 → T32UNC_20170823T103019_B02.jp2

⁶ Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltasvias, E., Schindler, K., 2018. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 146, pp. 305–319.



- 3.) Das Python-Skript DSen2 wird über die Kommandozeile des Rechners mit folgendem Befehl gestartet:

```
python.exe /pfad/zu/s2_tiles_supres.py /pfad/zu/input_file.SAFE  
/MTD_MSIL1C.xml /pfad/zu/output_file.tif --copy_original_bands
```

Beispiel:

```
python.exe "C:/Users/pbeckschaefer/Documents/DSen2-  
master/testing/s2_tiles_supres.py"  
"D:/F3/S2B_MSIL1C_20170823T103019_N0205_R108_T32UNC_20170823T103018.S  
AFE/MTD_MSIL1C.xml"  
"D:/F3/S2B_MSIL1C_20170823T103019_N0205_R108_T32UNC_20170823T103018_D  
Sen2.tif" --copy_original_bands
```

B.3.4 Topographische Korrektur und Berechnung von Vegetationsindizes (R)

Um Beleuchtungsunterschiede innerhalb einer Sentinel-2-Szene auszugleichen, die durch die Topographie des Geländes bedingt sind, wird eine topographische Korrektur durchgeführt. Hierzu wird das R-Skript „**S2_topocorr.R**“ verwendet, das auf die R-Pakete *raster*⁷, *RStoolbox*⁸ und *plotKML*⁹ zugreift. Im Anschluss an die topographische Korrektur berechnet das Skript 16 Vegetationsindizes.

Das Skript ist in drei Teile gegliedert: (1) Vorbereitung des digitalen Geländemodells, (2) topographische Korrektur und (3) Berechnung von Vegetationsindizes.

Bevor das Skript gestartet werden kann, müssen im Skript folgende Angaben gemacht werden:

- (a) **Zeile 38:** Pfad zum Arbeitsverzeichnis (Working Directory)
- (b) **Zeile 41:** Pfad zum digitalen Geländemodell (DGM5 oder DGM10)
- (c) **Zeile 46:** Pfad zur Sentinel-2-Szene. Achtung: Der Pfad ist ohne die Dateierweiterung .tif anzugeben.
- (d) **Zeile 49:** Pfad zur Wolkenmaske (erzeugt mit Fmask 4.0).
- (e) **Zeile 52 und 53:** *Sun Zenith* und *Sun Azimuth* der Sentinel-2-Szene. Die Werte können in der Software SNAP abgelesen werden (siehe unten).
- (f) **Zeile 59:** Sollen Vegetationsindizes berechnet werden?
 - i. Ja: setzen Sie den Wert auf TRUE (Default).
 - ii. Nein: setzen Sie den Wert auf FALSE.
- (g) **Zeile 64:** Soll eine NDVI-basierte Vegetationsmaske erstellt werden, um nicht-Vegetationsflächen zu maskieren und von der Prozessierung auszuschließen?
 - i. Ja: setzen sie den Wert auf TRUE (Default).
 - ii. Nein: setzen Sie den Wert auf FALSE.
- (h) **Zeile 67:** Sind die Pixelwerte der Sentinel-2-Szene zwischen 0 und 1 skaliert?
 - i. Ja: setzen sie den Wert auf TRUE.
 - ii. Nein: setzen Sie den Wert auf FALSE (Default).
- (i) **Zeile 75:** Wurde die Sentinel-2-Szene mit der Software DSen2 prozessiert?
 - i. Ja: setzen sie den Wert auf TRUE (Default).
 - ii. Nein: setzen Sie den Wert auf FALSE.

⁷ R-Paket *raster*: <https://cran.r-project.org/web/packages/raster/> [letzter Zugriff 02.12.2020].

⁸ R-Paket *RStoolbox*: <https://bleutner.github.io/RStoolbox/> [letzter Zugriff 02.12.2020].

⁹ R-Paket *plotKML*: <https://CRAN.R-project.org/package=plotKML> [letzter Zugriff 02.12.2020].



Nachdem diese Angaben gemacht wurden, kann das komplette Skript gestartet werden ohne weitere Änderungen am Code vorzunehmen.

Nach Abschluss der Prozessierung mit dem R-Skript liegen die spektralen Bänder der Sentinel-2-Satellitendaten in folgender Reihenfolge vor:

- (1) Blue (492.4 nm)
- (2) Green (559.8 nm)
- (3) Red (664.6 nm)
- (4) Red Edge 1 (704.1 nm)
- (5) Red Edge 2 (740.5 nm)
- (6) Red Edge 3 (782.8 nm)
- (7) Near Infrared (832.8 nm)
- (8) Narrow Near Infrared (864.7 nm)
- (9) Short Wave Infrared 1 (1613.7 nm)
- (10) Short Wave Infrared 2 (2202.4 nm)

1. Teil I – Vorbereitung des digitalen Geländemodells (DGM)

Teil I dient dazu, das DGM für die topographische Korrektur vorzubereiten. Hierfür wird das DGM auf die Ausdehnung der Sentinel-2-Kachel zugeschnitten und an die 10 m-Auflösung der Sentinel-2-Satellitendaten angepasst (Resampling). Es wird empfohlen, das so prozessierte DGM zu speichern und für künftige Prozessierungen desselben Geländeausschnitts vorzuhalten. Künftige Prozessierungen können dann unmittelbar mit Teil II des Skripts begonnen werden.

2. Teil II – Topographische Korrektur

Teil II des Skripts maskiert die Wolken- und Wolkenschattenflächen der Sentinel-2-Szene aus und führt die topographische Korrektur der Sentinel-2-Satellitendaten mit einer modifizierten Variante der Minnaert Funktion durch. Die Modifikation reduziert den Korrekturfaktor k der Minnaertgleichung (siehe Formel 5 in Riaño et. al 2003¹⁰) um 40%. Dies verhindert, dass es zu Überkorrekturereffekten kommt, die dazu führen, dass dunkle Bildbereiche zu stark aufgehellt werden. Die Korrektur wird stratifiziert nach Hangneigungsklassen durchgeführt. Um die Stratifikation des Bildes in Hangneigungsklassen durchzuführen, wird Code des R-Paket *RStoolbox*⁸ verwendet. Für jedes Stratum wird ein eigener Korrekturfaktor k berechnet.

Die korrekten Werte für Sun Zenith und Azimuth einer jeden Sentinel-2-Szene sind in den Sentinel-2-Satellitendaten als Bänder hinterlegt, die in SNAP geöffnet werden können. Hierzu öffnen Sie die entsprechende Sentinel-2-Szene in SNAP und navigieren in den Unterordner Bands → Sun. Durch Doppelklick auf den Eintrag sun_zenith öffnet sich im unteren linken Fenster des SNAP GUI ein Diagramm, auf dessen x-Achse Minimum, Mittelwert und Maximum der in der Szene ermittelten sun_zenith Werte angegeben sind (Abbildung 7). Lesen Sie den Mittelwert ab und übertragen Sie ihn in das Skript (Zeile 52). Den Mittelwert für den sun_azimuth ermitteln Sie anschließend auf die gleiche Weise und tragen ihn in Zeile 53 des Skripts ein. Alternativ können beide Werte auch aus der innerhalb des SAFE-Verzeichnisses abgelegten Datei „MTD_TL.xml“ (z. B. ..\S2A_MSIL1C_20170619T103021_N0205_R108_T32UNA_20170619T103021.SAFE\GRANULE\L1C_T

¹⁰ Riaño, D., Chuvieco, E., Salas, J., Aguado, I., 2003. Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. IEEE Transactions on Geoscience and Remote Sensing, 41(5): 1056 – 1061.



32UNA_A010401_20170619T103021(MTD_TL.xml) ausgelesen werden. Innerhalb des XML-Tags <Mean_Sun_Angle> geben die XML-Tags <ZENITH_ANGLE> und <AZIMUTH_ANGLE> die entsprechenden Werte an.

Ein Beispiel der Auswirkungen der topographischen Korrektur ist in Abbildung 8 zu sehen.

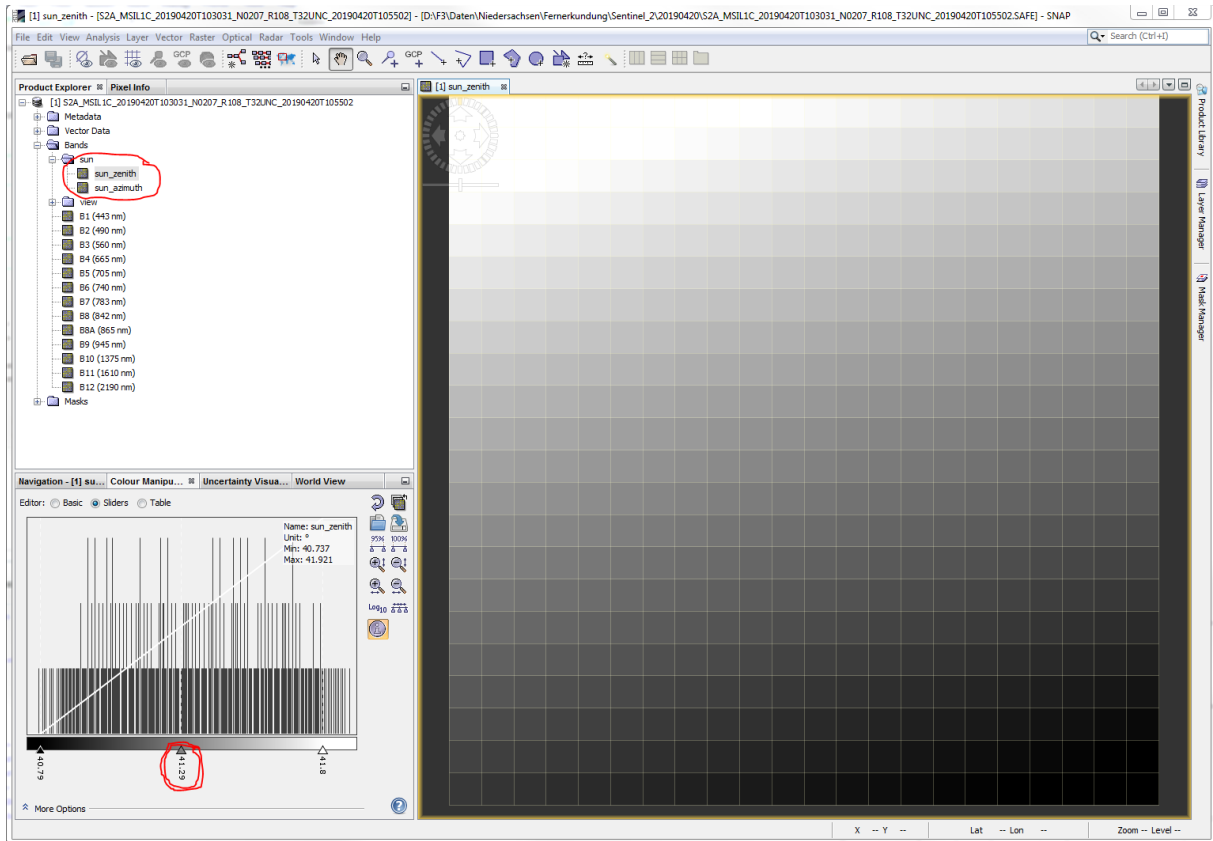


Abbildung 7: Screenshot des SNAP GUI. Rot umrandet ist der Ordner, der die Bänder sun_zenith und sun_azimuth enthält (oben links). Außerdem ist der Mittelwert des sun_zenith, der in das R-Skript übertragen werden muss, rot umrandet (unten links).

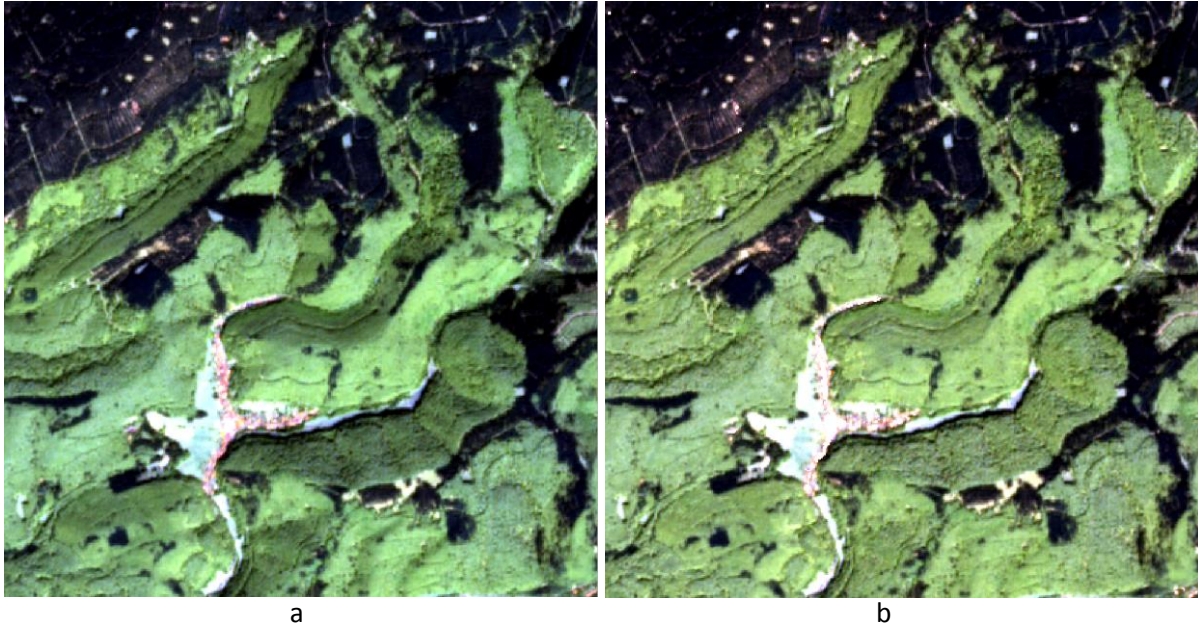


Abbildung 8: Ausschnitt aus Sentinel-2-Kompositbild (Bänder SWIR2, SWIR1, RedEdge1) vor (a) und nach (b) topographischer Korrektur.

3. Teil III – Berechnung von Vegetationsindizes

Im Anschluss an die topographische Korrektur berechnet das Skript 16 Vegetationsindizes und speichert diese in zwei TIFF-Dateien mit der Endung „VI1.tif“ bzw. „VI2.tif“.

Die Datei mit der Endung VI1.tif enthält Indizes, die ausschließlich unter Verwendung der Sentinel-2-Bänder, die ursprünglich mit einer Auflösung von 10 m aufgezeichnet wurden (Blue, Green, Red und Near Infrared (NIR)), berechnet wurden:

- NDVI (Normalized Difference Vegetation Index)
 - $(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}) * 10000$
- GLI (Green Leaf Index)
 - $((2 * \text{Green} - \text{Red} - \text{Blue}) / (2 * \text{Green} + \text{Red} + \text{Blue})) * 10000$
- PBI (Plant Biochemical Index)
 - $\text{NIR} / \text{Green} * 10000$
- NGRDI (Normalized Green Red Difference Index)
 - $(\text{Green} - \text{Red}) / (\text{Green} + \text{Red}) * 10000$
- CVI (Chlorophyll Vegetation Index)
 - $\text{NIR} * (\text{Red} / \text{Green}^2) * 10000$
- GNDVI (Green Normalized Difference Vegetation Index)
 - $(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green}) * 10000$
- BNDVI (Blue Normalized Difference Vegetation Index)
 - $(\text{NIR} - \text{Blue}) / (\text{NIR} + \text{Blue}) * 10000$



Die Datei mit der Endung VI2.tif enthält Indizes, für deren Berechnung auch Bänder verwendet wurden, die ursprünglich mit 20 m-Auflösung aufgezeichnet wurden.

- MCARI (Modified Chlorophyll Absorption Ratio Index)
 - $((\text{Red Edge 1} - \text{Red}) - (\text{Red Edge 1} - \text{Green})) * (\text{Red Edge 1}/\text{Red})$
- MNDWI (Modified Normalized Difference Water Index)
 - $(\text{Green} - \text{SWIR2})/(\text{Green} + \text{SWIR2}) * 10000$
- MTCI (MERIS Terrestrial Chlorophyll Index)
 - $(\text{Red Edge 2} - \text{Red Edge 1})/(\text{Red Edge 1} - \text{Red}) * 1000$
- NDREI1 (Normalized Difference Red Edge Index 1)
 - $(\text{Red Edge 2} - \text{Red Edge 1})/(\text{Red Edge 2} + \text{Red Edge 1}) * 10000$
- NDREI2 (Normalized Difference Red Edge Index 2)
 - $(\text{Red Edge 3} - \text{Red Edge 1})/(\text{Red Edge 3} + \text{Red Edge 1}) * 10000$
- SLAVI (Specific Leaf Area Vegetation Index)
 - $\text{NIR}/(\text{Red} + \text{SWIR 2}) * 1000$
- NDWI1 (Normalized Difference Water Index 1)
 - $(\text{NIR} - \text{SWIR 1})/(\text{NIR} + \text{SWIR 1}) * 10000$
- NDWI2 (Normalized Difference Water Index 2)
 - $(\text{NIR} - \text{SWIR 2})/(\text{NIR} + \text{SWIR 2}) * 10000$
- IRECI (Inverted Red Edge Chlorophyll Index)
 - $(\text{NIR} - \text{Red})/(\text{Red Edge 1}/\text{Red Edge 2}) * 1000$

C Ableitung von Baumarteninformationen aus Sentinel-2-Satellitendaten

Zur Erstellung einer Karte, die das Vorkommen von Baumarten bzw. Baumartengruppen zeigt, existieren verschiedene geeignete Methoden. Im Folgenden beschreiben wir ein Verfahren, das im Rahmen des Projektes F³ erprobt wurde und sich für die Ziele des Projektes F³ bewährt hat. Datengrundlage sind (a) Sentinel-2-Satellitendaten, die wie in Abschnitt A beschrieben, aufbereitet wurden und (b) Referenzdaten, die das Vorkommen von Baumarten bzw. Baumartengruppen räumlich verorten. Das hier beschriebene Verfahren ist im R-Skript *„F3-Baumarten_Sentinel2.R“*¹¹ implementiert, welches folgende Schritte durchführt:

- Einlesen der Sentinel-2-Satellitendaten
- Einlesen der Referenzdaten
- Verschneiden der Sentinel-2-Satellitendaten und der Referenzdaten zur Extraktion der Sentinel-2-Pixelwerte
- Trainieren eines Random Forest-Klassifikators
- Erstellen einer Baumarten- bzw. Baumartengruppenkarte
- Erstellen einer Genauigkeitsmatrix und Berechnung der Kartengenauigkeiten

Aufgrund der Verwendung eines Random Forest-Modells, das problemlos weitere Datensätze bei der Modellierung berücksichtigen kann, ist das Verfahren flexibel erweiterbar. Sollten den Nutzenden weitere Datensätze zur Verfügung stehen, die das Potential haben die Baumartendifferenzierung zu verbessern (z. B. Vegetationshöhenmodelle als Indikator für das Baumalter), können auch diese Datensätze in das Verfahren integriert und für die Modellierung verwendet werden.

C.1 Referenzdaten

Als Quelle für Referenzdaten kommen Datensätze in Frage, die das räumliche Vorkommen von Baumarten abbilden. Bei der Auswahl der Referenzdaten sollte darauf geachtet werden, dass die Daten (a) alle zu erfassenden Baumarten abdecken und (b) für jede Baumart eine ausreichend große Menge an Datenpunkten enthalten. Des Weiteren müssen (c) alle Datenpunkte reine Klassen darstellen, also keine Mischklassen enthalten, und (d) sollten Aufnahmezeitpunkt der Referenzdaten und Aufnahmezeitpunkt der Satellitendaten nicht zu weit auseinanderliegen. Im Idealfall handelt es sich bei den Referenzdaten um Daten von terrestrischen Erhebungen, welche auch aus Stichprobeninventuren abgeleitet werden können. Zur Nutzung des hier beschriebenen Verfahrens und insbesondere zur Durchführung der Quantifizierung der Kartengenauigkeit ist es wichtig, dass die Referenzdatenpunkte unabhängig voneinander ausgewählt wurden und jeder Datenpunkt einem einzelnen Pixel der Sentinel-2-Satellitendaten zugeordnet werden kann. Werden pro Datenpunkt mehrere Pixel ausgewählt, resultieren die hier implementierten Berechnungen der Kartengenauigkeiten in falschen Ergebnissen!

Im Rahmen des Projekts F³ wurden Forsteinrichtungsdaten (Bestandes- und Betriebsinventuren) der Landesforstbetriebe als Referenzdaten verwendet.

Exemplarisch soll kurz erläutert werden, wie die Daten der Betriebsinventur der Niedersächsischen Landesforsten für die Klassifikation aufbereitet wurden. Für die niedersächsischen Projektgebiete in

¹¹ R-Skript *„F3-Baumarten_Sentinel2.R“*: https://www.waldwissen.net/assets/FVA/Technik_und_Planung/F3-Projekt/F3_Dokumentation_und_Skripte_01.zip [letzter Zugriff 11.01.2021].



Heide, Solling und Harz wurden Daten der Betriebsinventur aus den Jahren 2015 (Heide), 2012 (Harz) und 2011 (Solling) genutzt. An den Stichprobenpunkten der Betriebsinventur erfolgt die Aufnahme konzentrischer Probekreise mit einem Durchmesser von (a) 26 m für Bäumen mit einem Brusthöhendurchmesser (BHD) > 30 cm und einem Durchmesser von (b) 12 m für die Aufnahme von Bäumen mit einem BHD \geq 7 cm und einem BHD \leq 30 cm. Um möglichst baumartenreine Probekreise zu identifizieren, wurde für jeden Probepunkt ermittelt, welche Baumart den nach Grundfläche gewichteten größten Anteil stellt. Probeflächen auf denen eine Baumart mehr als 75% der Grundfläche repräsentiert, wurden als Referenzdaten für die Baumartenklassifikation ausgewählt.

Neben den Daten der Betriebsinventur können auch Bestandesdaten der Forsteinrichtung als Referenzdaten genutzt werden. Dies wurde für die Klassifikation der Baumarten in Baden-Württemberg zusätzlich getan. Hier wurden alle Reinbestände der Zielbaumarten ausgewählt und mit einem Abstand von 20 m zu Bestandesgrenzen und Infrastruktur abgeschnitten, um die Überschneidung mit anderen Beständen, Straßen und Waldwegen auszuschließen. Für jeden Mittelpunkt eines Sentinel-2-Pixels, der in eines der verbleibenden Polygone fällt, wurde eine Punktgeometrie erstellt. Die Polygone (und die dazugehörigen Punkte) wurden in Trainings- und Testpolygone unterteilt. Hierdurch konnte eine geographische Abhängigkeit zwischen Trainings- und Testpixeln vermieden werden.

C.2 Baumartenerfassung

Benötigte Daten

Das R-Skript `„Baumarten_Sentinel2_F3.R“`¹¹ erstellt eine Baumartenkarte und eine Genauigkeitsmatrix, auf dessen Grundlage die Genauigkeiten der Karte berechnet werden. Um das Skript zu nutzen, werden folgende Eingangsdaten benötigt:

- (a) Sentinel-2-Satellitendaten (aufbereitet nach dem Verfahren zur Vorprozessierung von Sentinel-2-Satellitendaten, siehe Abschnitt A).
- (b) Referenzdaten: Punkt-Shapefile der Referenzdatenpunkte, z. B. Mittelpunktkoordinaten von Inventurstichprobenpunkten. Die Attributtabelle der Referenzdaten darf nur die Spalten (1) ID und (2) Baumart enthalten.
- (c) Waldmaske im Rasterformat, um Nicht-Waldflächen zu maskieren. Die Waldmaske muss die Werte 1 = Wald und NA = Nicht-Wald enthalten. Falls Auflösung und/oder Lage der Pixel der Waldmaske nicht mit der Sentinel-2-Kachel übereinstimmen, wird die Waldmaske durch ein Nearest Neighbor Resampling an die Sentinel-2-Daten angepasst.

Vorbereitung des R-Skripts

Bevor das Skript `„Baumarten_Sentinel2_F3.R“` gestartet werden kann, müssen die folgenden Angaben gemacht werden:

- (a) **Zeile 42:** Pfad zum Arbeitsverzeichnis (Working Directory). In diesem Ordner werden die Ausgabedateien (Baumartenkarte und Genauigkeitsmatrix) gespeichert.
- (b) **Zeile 45:** Anzahl der Prozessoren, die für die Berechnung genutzt werden soll.



- (c) **Zeile 48:** Ausdehnung des Gebiets, für das eine Baumartenkarte erstellt werden und auf das sich die Kartenvalidierung beziehen soll. Die Gebietsausdehnung ist über die Angabe der Koordinaten Xmin, Xmax, Ymin und Ymax zu spezifizieren.
- (d) **Zeile 51:** Pfad zur Waldmaske (Raster).
- (e) **Zeile 54:** Pfad zu den Referenzdaten (Shape-Datei).
- (f) **Zeile 57:** Name der Attributtabellenspalte mit der Baumarteninformation
- (g) **Zeile 60:** Pfad zu den Sentinel-2-Satellitendaten (nach Vorprozessierung).
- (h) **Zeile 63:** Name der Ausgabedatei (Baumartenkarte). Neben der Baumartenkarte mit 10 m-Auflösung wird unter dem gleichen Namen die auf 20 m generalisierte Variante der Baumartenkarte gespeichert (Namenszusatz „_20m.tif“). Die Ergebnisse der Kartenvalidierung werden als CSV-Datei ebenfalls mit dem hier spezifizierten Namen exportiert (Namenszusatz „_accuracy.csv“).

Ausgabedateien

Die Ausgabe des Skripts besteht aus drei Dateien, die im in Zeile 42 angegebenen Verzeichnis abgelegt werden. Dies sind (a) die Baumartenkarte mit 10 m-Auflösung, (b) die Baumartenkarte mit 20 m-Auflösung und (c) die Tabelle mit den Ergebnissen der Validierung.

- (a) Die Baumartenkarte mit 10 m-Auflösung enthält die mit Random Forest modellierten Baumarten.
- (b) Die Baumartenkarte mit 20 m-Auflösung wurde durch ein Resampling der Baumartenkarte mit 10 m-Auflösung erstellt. D. h. die Auflösung wurde um den Faktor 2 vergrößert, wobei die jeweils häufigste Baumart pro 20 m-Pixel übernommen wurde. Sollten mehrere Baumarten gleich häufig vorkommen, wird per Zufallszahl entschieden, welche Baumart ausgewählt wird.
- (c) Die Ergebnisse der Validierung umfassen zwei Konfusionsmatrizen sowie die daraus abgeleiteten Kennwerte der Kartengenauigkeit (vgl. Tabelle 2). Die erste Matrix enthält die Anzahl der korrekt bzw. falsch klassifizierten Pixel (siehe Tabellen 1 & 3 in Olofsson et al. 2013¹²). Die Matrix wird durch eine im Zuge des Modelltrainings durchgeführte Kreuzvalidierung des Random Forest-Modells erstellt. Zur Erstellung der zweiten Matrix (Tabellen 2 & 4 in Olofsson et al. 2013¹²) wurden die Einträge der ersten Matrix mit den Flächenanteilen der Baumarten in der Karte gewichtet. Durch die Gewichtung erfolgt die Berechnung der Genauigkeiten entsprechend der kartierten Flächenanteile und nicht gewichtet nach der Verteilung der Referenzdatenpunkte auf die verschiedenen Baumarten. Die Tabelle enthält die im Folgenden aufgeführten Spalten (für eine detaillierte Beschreibung der Berechnung siehe Olofsson et al. 2014¹³):
 - a. **maparea:** [ha] Kartierte Fläche je Baumart in Hektar.

¹² Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment, Vol. 129: 122 – 131.

¹³ Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment, Vol. 148: 42 – 57.

- b. **prop_maparea:** [%] Prozentualer Anteil je Baumart an der kartierten Waldfläche.
- c. **adj_proparea:** [%] Mittels der Genauigkeitsmatrix korrigierter Anteil je Baumart an der kartierten Waldfläche.
- d. **CI_adj_proparea:** [%] Konfidenzintervall des mittels der Genauigkeitsmatrix korrigierten Anteils je Baumart an der kartierten Waldfläche (adj_proparea).
- e. **adj_area:** [ha] Mittels der Genauigkeitsmatrix korrigierte Fläche je Baumart.
- f. **CI_adj_area:** [ha] Konfidenzintervall der mittels der Genauigkeitsmatrix korrigierten Fläche je Baumart an der kartierten Waldfläche (adj_area).
- g. **UA:** [%] Nutzergenauigkeit.
- h. **CI_UA:** [%] Konfidenzintervall der Nutzergenauigkeit.
- i. **PA:** [%] Produzentengenauigkeit.
- j. **CI_PA:** [%] Konfidenzintervall der Produzentengenauigkeit.
- k. **OA:** [%] Gesamtgenauigkeit.
- l. **CI_OA:** [%] Konfidenzintervall der Gesamtgenauigkeit.

C.3 Validierungsergebnisse aus dem F³-Projekt

Das Ziel der Baumartenklassifizierung im F³-Projekt ist die Herleitung zusätzlicher Eingangsdaten für die Verbesserung der Modellierung von Holzvorrat und oberirdischer Biomasse (siehe auch Dokumentation zu Meilenstein M8 „Verfahren zur Modellierung von Holzvorrat und oberirdischer Biomasse“¹). Die Genauigkeit der Baumartenklassifizierung ist zum einen von der Qualität der Satellitendaten abhängig, zum anderen aber auch stark von Qualität und Quantität der verfügbaren Referenzdaten beeinflusst. Die im Folgenden dargestellten Validierungsergebnisse der Baumartenklassifikation der Untersuchungsgebiete des F³-Projektes können daher nur als illustrierende Beispiele dienen. Mit anderen Satellitendaten und anderen Referenzdaten durchgeführte Klassifikationen können höhere, aber auch niedrigere Genauigkeiten erreichen. Dies wird auch daran deutlich, dass die Baumartenkarten der F³-Untersuchungsgebiete keine einheitliche Genauigkeit aufweisen, sondern die Gesamtgenauigkeiten zwischen 57,1% (Baden-Württemberg) und 86% (Harz) schwanken. Auch Produzenten und Nutzergenauigkeiten der einzelnen Baumarten variieren mitunter stark von Gebiet zu Gebiet.

Tabelle 2 bis Tabelle 3 zeigen die Validierungsergebnisse der Baumartenklassifikation der niedersächsischen Projektgebiete in Heide, Solling und Harz (die zugehörigen Karten finden Sie im Anhang). Hier wurden jeweils die Baumarten Eiche (EI), Buche (BU), Fichte (FI), Douglasie (DGL), Kiefer (KI) und Lärche (LAE) klassifiziert. Die Validierung ergab Gesamtgenauigkeiten von 74,1% (Heide, Tabelle 2), 76,7% (Solling, Tabelle 3) und 85,8% (Harz, Tabelle 4). In Solling und Harz erreichen nur Buche und Fichte Produzentengenauigkeiten über 80%, die Produzentengenauigkeiten der anderen Baumarten liegen unter 50%. In der Heide werden nur Eiche (Produzentengenauigkeit 69,8%) und Kiefer (Produzentengenauigkeit 91,4%) gut erfasst. Die Produzentengenauigkeiten der anderen Baumarten liegen bei knapp über 50% (Douglasie) oder darunter.



Flächendeckende Fernerkundungsbasierte Forstliche Strukturdaten

Tabelle 2: Validierungsergebnis zur Baumartenkartierung im Untersuchungsgebiet Heide. In den vorangestellten Konfusionsmatrizen entsprechen die Spalten den Referenzdaten und die Zeilen der Klassifikation.

	EI	BU	FI	DGL	KI	LAE		EI	BU	FI	DGL	KI	LAE	maparea	prop_maparea	adj_proparea	Cl_adj_proparea	adj_area	Cl_adj_area	UA	Cl_UA	PA	Cl_PA	OA	Cl_OA
EI	23	4	3	0	3	1	EI	3,6	0,6	0,5	0	0,5	0,2	758,65	5,3	5,1	1,4	73644,357	208,071	67,6	16	69,8	16,8	74,1	4,4
BU	2	10	0	1	4	0	BU	0,3	1,5	0	0,1	0,6	0	356,1	2,5	3,5	1,5	50540,245	216,793	58,8	24,1	41,4	19,1		
FI	2	0	27	3	18	1	FI	0,5	0	6,4	0,7	4,3	0,2	1753,69	12,1	15,9	3,4	229597,113	498,107	52,9	13,8	40,6	10		
DGL	1	0	4	20	1	1	DGL	0,2	0	0,7	3,6	0,2	0,2	700,9	4,9	6,8	2	98192,476	288,567	74,1	16,8	52,7	15,2		
KI	2	4	27	8	198	12	KI	0,6	1,2	8	2,4	58,6	3,6	10727,31	74,3	64,1	4,2	925608,487	601,633	78,9	5,1	91,4	2,6		
LAE	0	1	1	0	0	2	LAE	0	0,2	0,2	0	0	0,5	143,42	1	4,6	2,1	66424,322	310,043	50	56,6	10,7	11,9		

Tabelle 3: Validierungsergebnis zur Baumartenkartierung im Untersuchungsgebiet Solling. In den vorangestellten Konfusionsmatrizen entsprechen die Spalten den Referenzdaten und die Zeilen der Klassifikation.

	EI	BU	FI	DGL	KI	LAE		EI	BU	FI	DGL	KI	LAE	maparea	prop_maparea	adj_proparea	Cl_adj_proparea	adj_area	Cl_adj_area	UA	Cl_UA	PA	Cl_PA	OA	Cl_OA
EI	133	44	18	4	2	14	EI	3,2	1,1	0,4	0,1	0	0,3	2237,31	5,2	7,9	0,7	337345,879	301,934	61,9	6,5	41,2	4,1	76,9	1,2
BU	162	1543	250	10	4	137	BU	3,9	37,5	6,1	0,2	0,1	3,3	21868,88	51,2	41,3	1,1	1763593,013	463,084	73,3	1,9	90,8	1,1		
FI	37	162	2299	62	33	129	FI	0,6	2,5	35,5	1	0,5	2	17957,92	42,1	42,3	0,9	1806295,023	402,497	84,5	1,4	83,9	1,5		
DGL	0	0	3	7	0	0	DGL	0	0	0	0,1	0	0	67,06	0,2	1,4	0,3	59782,814	129,269	70	29,9	7,7	3,4		
KI	0	0	0	0	0	1	KI	0	0	0	0	0	0	14,41	0	0,8	0,2	34161,608	100,757	0	NA	0	0		
LAE	5	10	11	1	6	21	LAE	0,1	0,2	0,3	0	0,1	0,5	556,43	1,3	6,2	0,7	264752,462	290,589	38,9	13,1	8,2	2,7		

Tabelle 4: Validierungsergebnis zur Baumartenkartierung im Untersuchungsgebiet Harz. In den vorangestellten Konfusionsmatrizen entsprechen die Spalten den Referenzdaten und die Zeilen der Klassifikation.

	EI	BU	FI	DGL	KI	LAE		EI	BU	FI	DGL	KI	LAE	maparea	prop_maparea	adj_proparea	Cl_adj_proparea	adj_area	Cl_adj_area	UA	Cl_UA	PA	Cl_PA	OA	Cl_OA
EI	41	9	10	0	1	2	EI	1,2	0,3	0,3	0	0	0,1	1274,32	1,9	3,1	0,5	208980,3	307,555	65,1	11,9	39,2	6,6	85,8	1
BU	56	869	176	12	3	52	BU	1,3	20,9	4,2	0,3	0,1	1,3	18934,17	28,1	24,5	0,9	1651618,5	582,941	74,4	2,5	85,3	1,8		
FI	30	181	3434	57	14	59	FI	0,6	3,3	63,4	1,1	0,3	1,1	47019,81	69,7	68	0,9	4584084	593,211	91	0,9	93,2	0,8		
DGL	0	0	2	7	0	0	DGL	0	0	0,1	0,2	0	0	169,86	0,3	1,5	0,3	101119,5	218,774	77,8	28,8	12,7	4,9		
KI	0	0	0	0	0	0	KI	0	0	0	0	0	0	4,05	0	0,4	0,2	26965,2	113,624	NA	NA	0	0		
LAE	1	0	1	0	0	0	LAE	0	0	0	0	0	0	10,79	0	2,4	0,4	161791,2	296,509	0	0	0	0		

In Baden-Württemberg wurden landesweite Baumartenreferenzdaten verwendet, welche, wie oben beschrieben, auf den Stichprobenplots der Betriebsinventuren und Forsteinrichtungspolygonen basieren. Die Stichprobenpunkte wurden zusätzlich anhand der Vegetationshöhenmodelle automatisiert gefiltert und dabei Punkte mit Vegetation niedriger als 7 m aus dem Datensatz entfernt. Im Rahmen einer Masterarbeit (Költzow 2020¹⁴) wurde nach den oben beschriebenen Methoden eine landesweite Baumartenkarte erzeugt. Die Validierungsergebnisse sind in Tabelle 5 dargestellt. Zusätzlich zu den in den niedersächsischen Projektgebieten klassifizierten Baumarten, wurden in Baden-Württemberg auch die Baumarten Tanne (TA) und Roteiche (REI) sowie die Klasse sonstige Laubbäume (SLb) klassifiziert. Dies war durch die regional größere Bedeutung der Baumart Tanne und der spektralen Besonderheit der Roteiche bedingt. Im Rahmen der Masterarbeit wurde der Einfluss unterschiedlicher Sampleproportionen untersucht. Die in Tabelle 5 dargestellten Ergebnisse basieren auf einem balancierten Referenzdatensatz mit je 100 Datenpunkten pro Baumart für die Validierung und 1000 Punkten pro Baumart für das Trainieren des Modells.

Mit dem landesweiten Ansatz wurde eine Gesamtgenauigkeit von 57,1% erzielt (Konfusionsmatrix in Tabelle 5). Die höchste Nutzergenauigkeit erreicht die Baumart Kiefer (83,3%), gefolgt von der Roteiche (79,5%). Die Baumartengruppe der sonstigen Laubbäume weist die insgesamt geringste Genauigkeit auf. Dies ist durch die hohe Varianz innerhalb dieser Klasse zu erklären, die mehrere verschiedene Baumarten verschiedener Spektralsignaturen umfasst. Die beiden Nadelbaumarten Douglasie und Kiefer werden oft fälschlicherweise als Fichten klassifiziert.

Diese Klassifizierung wurde für die gesamte Waldfläche Baden-Württembergs entwickelt und bildet somit eine sehr große Variabilität von Bestandessituationen ab. Hier zeigt sich besonders, dass für eine gute Qualität der Baumartenkarte sehr gute Trainingsdaten notwendig sind, welche die gesamte Variabilität der Baumartenvorkommen abbilden. Wird eine große Fläche klassifiziert, so beeinflusst neben der Anzahl und Güte der Trainingsdaten auch die Qualität der gewählten Satellitenszenen das Ergebnis der Klassifikation.

Die Validierung fand nur an Reinbeständen statt. Deshalb ist, über die hier angegebene Gesamtgenauigkeit der berechneten räumlich expliziten Baumartenkarte hinaus, zu beachten, dass die Genauigkeit im Bereich von Mischbeständen reduziert sein kann. Mischbestände führen auf der Ebene der Sentinel-2-Satellitendaten zu Mischpixeln, welche verschiedene Baumarten umfassen und deren Klassifikation eine große Herausforderung darstellt.

¹⁴ Jonathan Költzow: Tree Species Classification using Random Forest and Sentinel-2 Multispectral Satellite Imagery. MSc-Arbeit, Hochschule für nachhaltige Entwicklung Eberswalde 2020.



Flächendeckende Fernerkundungsbasierte Forstliche Strukturdaten

Tabelle 5: Validierungsergebnis zur Baumartenkartierung in Baden-Württemberg. In den vorangestellten Konfusionsmatrizen entsprechen die Spalten den Referenzdaten und die Zeilen der Klassifikation.

	TA	BU	LAE	SLB	FI	KI	DGL	EI	REI		TA	BU	LAE	SLB	FI	KI	DGL	EI	REI
TA	91	3	18	9	3	1	8	3	1	TA	8,8	0,3	1,7	0,9	0,3	0,1	0,8	0,3	0,1
BU	0	86	13	28	3	1	8	22	2	BU	0	17,3	2,6	5,6	0,6	0,2	1,6	4,4	0,4
LAE	0	1	50	12	0	5	2	2	0	LAE	0	0,1	5	1,2	0	0,5	0,2	0,2	0
SLB	0	1	3	10	0	1	2	0	4	SLB	0	0,2	0,6	1,9	0	0,2	0,4	0	0,7
FI	0	1	11	8	92	45	33	0	0	FI	0	0,2	1,8	1,3	14,7	7,2	5,3	0	0
KI	0	3	0	2	0	40	1	2	0	KI	0	0,2	0	0,1	0	2,5	0,1	0,1	0
DGL	4	0	0	1	2	5	39	0	0	DGL	0,2	0	0	0	0,1	0,2	1,9	0	0
EI	4	3	5	12	0	2	5	69	0	EI	0,2	0,1	0,2	0,6	0	0,1	0,2	3,4	0
REI	0	2	0	18	0	0	2	2	93	REI	0	0	0	0,3	0	0	0	0	1,6

	maparea	prop_maparea	adj_proparea	CI_adj_proparea	adj_area	CI_adj_area	UA	CI_UA	PA	CI_PA	OA	CI_OA
TA	198584,8	13,3	9,2	1,1	13725968	16256,2	66,4	7,9	95,7	2,8	57,1	3,7
BU	489691,2	32,8	18,4	2,6	27451937	38955,09	52,8	7,7	94	3,5		
LAE	107835,8	7,2	11,9	2,1	17754242	31532,67	69,4	10,7	42	7,9		
SLB	58184,04	3,9	11,9	2,4	17754242	36462,45	47,6	21,9	15,6	6,7		
FI	451969,9	30,3	15,7	2,3	23423663	34169,02	48,4	7,1	93,7	4,7		
KI	45177,31	3	11	2	16411484	30000,18	83,3	10,7	22,9	4,7		
DGL	36879,75	2,5	10,5	2,1	15665507	32069,87	76,5	11,8	18,1	4,3		
EI	74496,89	5	8,5	1,8	12681601	27531,02	69	9,1	40,4	9,1		
REI	29133,48	2	2,8	0,9	4177469	13484,76	79,5	7,3	55,5	17,9		

ANHANG

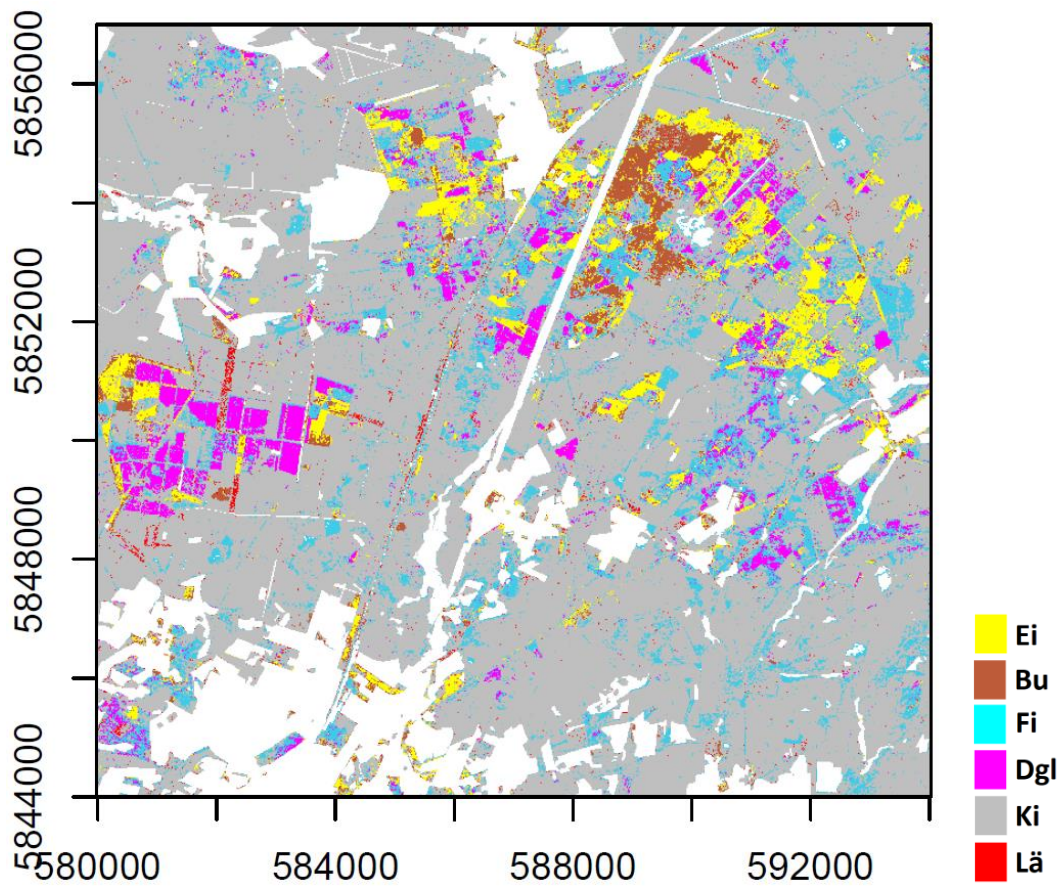


Abbildung 9: Klassifikation von Baumarten im Untersuchungsgebiet Heide. Die zugehörigen Validierungsergebnisse befinden sich in Tabelle 2.



Abbildung 10: Ausschnittsvergrößerung aus Abbildung 9. Schwarze Linien kennzeichnen die Bestandesabgrenzungen der Forsteinrichtung.

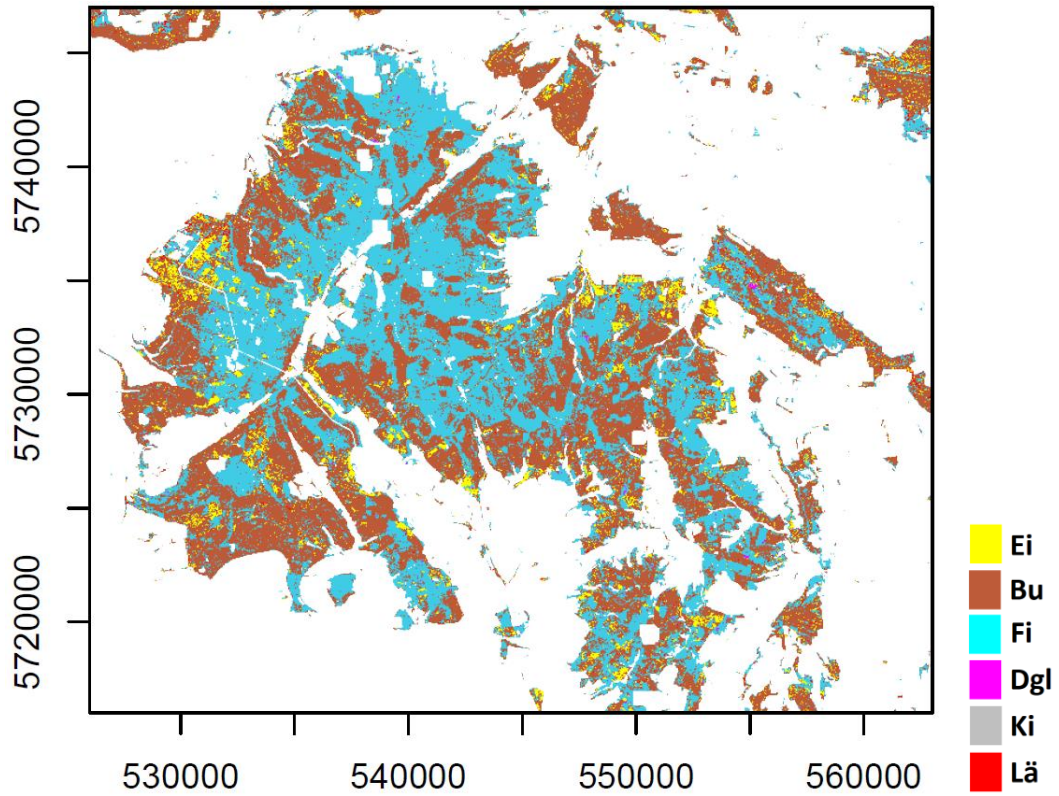


Abbildung 11: Klassifikation von Baumarten im Untersuchungsgebiet Solling. Die zugehörigen Validierungsergebnisse befinden sich in Tabelle 3.



Abbildung 12: Ausschnittsvergrößerung aus Abbildung 11. Schwarze Linien kennzeichnen die Bestandesabgrenzungen der Forsteinrichtung.

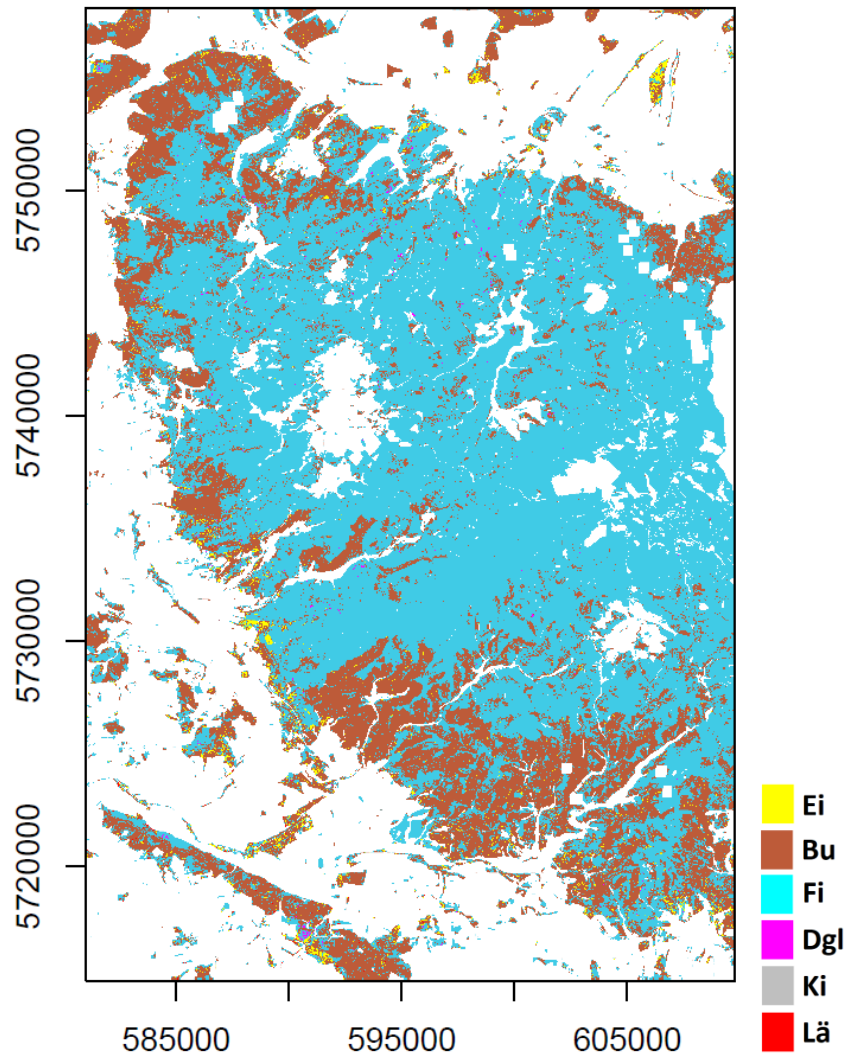


Abbildung 13: Klassifikation von Baumarten im Untersuchungsgebiet Heide. Die zugehörigen Validierungsergebnisse befinden sich in Tabelle 4.

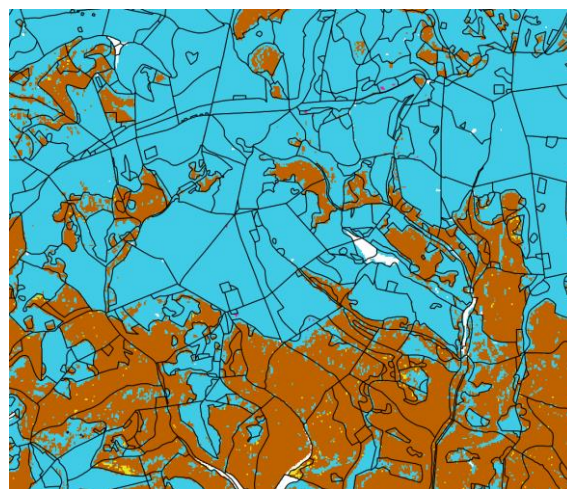


Abbildung 14: Ausschnittsvergrößerung aus Abbildung 13. Schwarze Linien kennzeichnen die Bestandesabgrenzungen der Forsteinrichtung.

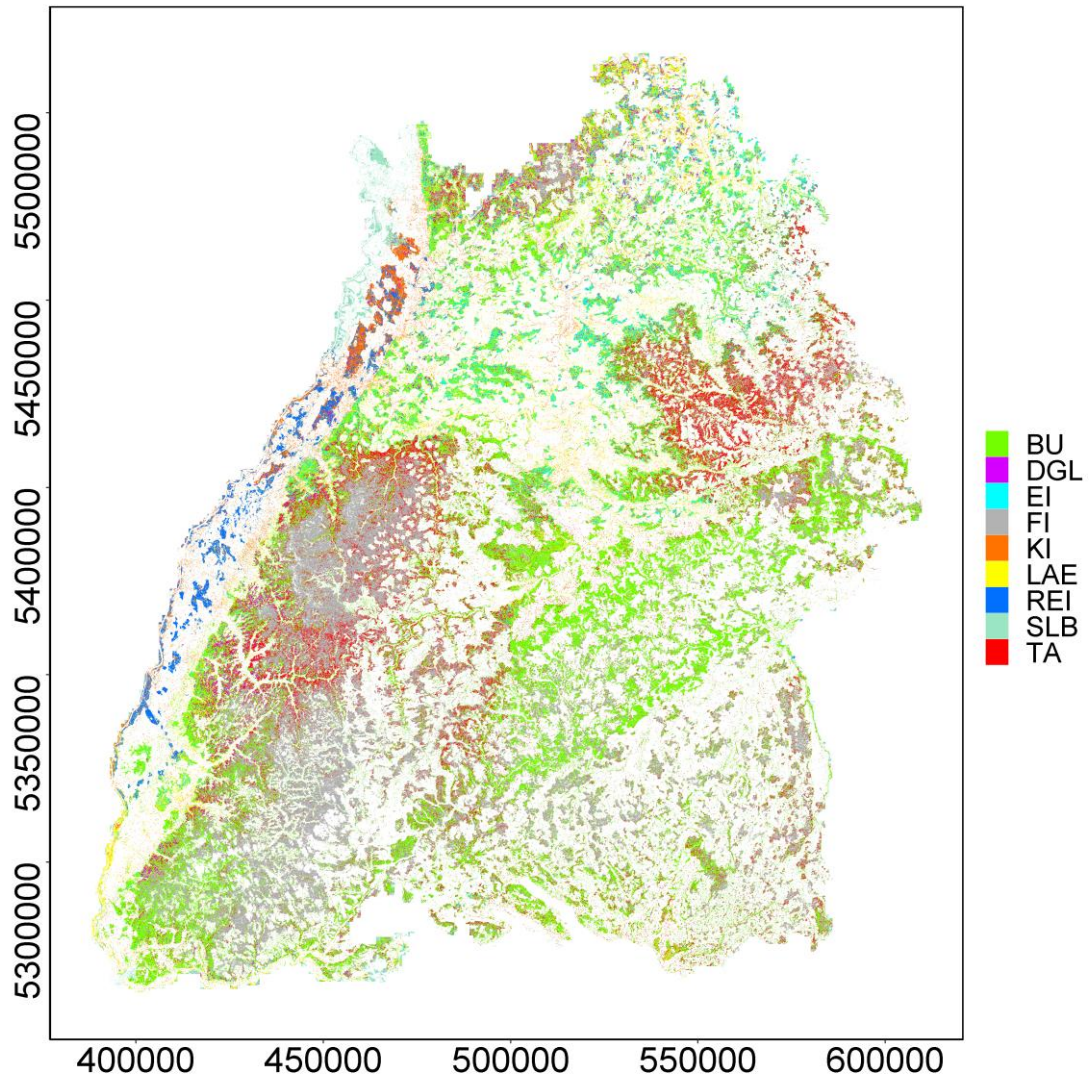


Abbildung 15: Klassifikation von Baumarten in Baden-Württemberg. Die zugehörigen Validierungsergebnisse befinden sich in Tabelle 5.