

# CopGruen Heide-Verbuschungsdienst – technische Dokumentation

23.10.2024

Marie Lins

Ann-Kathrin Holtgrave

Anke Schroiff

## Inhalt

1. Zweck des Dienstes .....	2
2. Einleitung.....	2
3. Modellberechnung .....	2
3.1 Referenzdaten .....	3
3.2 Verwendete Satellitendaten .....	5
3.2.1 Optische Satellitendaten .....	5
3.2.2 Radardaten .....	5
3.2.3 (Vor-)Prozessierung der Satellitendaten .....	6
3.2.4 Ermittlung typischer Wertebereiche .....	7
3.3 Konzeptionierung eines Klassifikationsmodells .....	8
3.4 Validierung .....	9
4. Umsetzung als Dienst im operationellen Betrieb.....	9
4.1 Eingabegeometrien und -parameter.....	10
4.2 Bearbeitung der Eingangspolygone.....	11
4.3 Prozessierung der Satellitendaten .....	11
4.4 Ableitung der Verbuschungs-Wahrscheinlichkeit .....	11
4.5 Bestimmung der Qualitätsparameter .....	12
4.6 Ergebnisausgabe.....	12
4.6.1 Interpretation der Ergebnisausgabe .....	13
5. Gültigkeitsbereich und Übertragbarkeit .....	14
6. Fehlerquellen und Einschränkungen .....	16
7. Zitierhinweis .....	16
Literatur .....	17

## 1. Zweck des Dienstes

Der Heide-Verbuschungsdienst als eine für die naturräumliche Großlandschaft „Nordostdeutsches Tiefland“ (Naturräumliche Gliederung nach Bundesamt für Naturschutz (BfN), 2022; s. Abbildung 9) verfügbare satellitengestützte Web-Anwendung liefert einen Hinweis auf Verbuschung als Beeinträchtigungsparameter von Flächen trockener Heide-Lebensraumtypen (LRT-Codes 2310, 4030).

## 2. Einleitung

Die Erfüllung der Fauna-Flora-Habitat (FFH)-Berichtspflichten erfordert u.a. die Analyse von Veränderungen bei Flächen eines bestimmten Lebensraumtyps (LRT), wie beispielsweise Trockenen Heiden (LRT-Code 4030) oder Sandheiden mit Besenheide und Ginster auf Binnendünen (LRT-Code 2310). Im Allgemeinen liegt die letzte Kartierung der Flächen mehrere Jahre zurück, wodurch aktuelle Informationen zu deren Zustand fehlen. Für die Priorisierung im Rahmen der Planung einer bevorstehenden Kartiersaison bedarf es aber eines aktuell verfügbaren Veränderungsmaßes. Die Nutzung fernerkundlicher Daten in Kombination mit Cloud-Computing bietet hier eine effiziente Lösung.

Der in diesem Dokument beschriebene und auf der Offenlandmonitor-Website verfügbare Heide-Verbuschungsdienst liefert für betroffene LRT-Flächen Hinweise auf Beeinträchtigungen der Vegetationsstruktur im Hinblick auf Verbuschung durch überschirmende Gehölze. Damit unterstützt er die Nutzenden bei der Priorisierung der FFH-Kartierung. Nutzende laden eigene Polygone von trockenen Heide-LRT (Codes 2310, 4030) hoch und erhalten einen Hinweis, ob diese Flächen von Verbuschung beeinträchtigt sind. Damit ist der Dienst ein niederschwelliges Instrument zur Umsetzung eines landschaftsbezogenen, fernerkundungsbasierten Flächenmonitorings. Er richtet sich insbesondere an die Sachbearbeitenden der Landesumweltbehörden und kann als Planungselement für die terrestrische Geländekartierung, aber auch zur Kontrolle der Bewertung von LRT-Flächen eingesetzt werden.

Das Konzept des Dienstes besteht aus zwei Teilen: der Berechnung eines Verbuschungsmodells als Vorarbeit und der Funktionsweise des Dienstes im operationellen Betrieb.

## 3. Modellberechnung

Hier werden Beschaffung und Aufbereitung aller zur Modellberechnung benötigten Daten, d.h. der Referenzdaten und der satellitenbasierten Modell-Inputdaten, beschrieben. Außerdem wird das verwendete Modell vorgestellt und validiert.

Die Erhebung der Referenzdaten erfolgte in ArcMap 10.6, die restlichen Vorgänge wurden mit der Programmiersprache R (Version 4.3.1) umgesetzt.

Eine grafische Übersicht der Methodik ist Abbildung 1 zu entnehmen. Auf die einzelnen Arbeitsschritte wird in den nächsten Abschnitten genauer eingegangen.

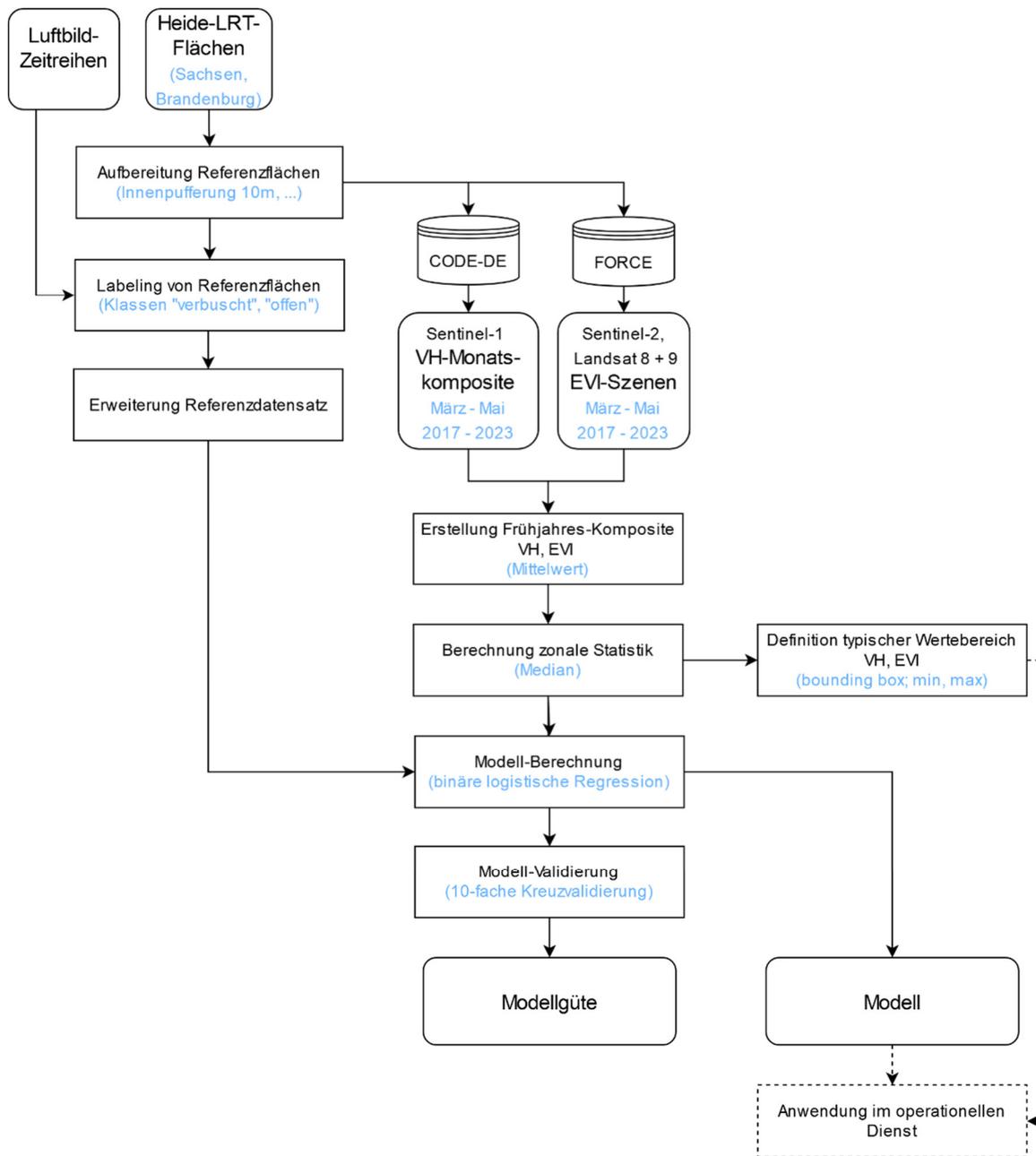


Abbildung 1: Schematische Darstellung der verwendeten Methodik zur Modellberechnung (LRT = Lebensraumtyp, EVI = Enhanced Vegetation Index, VH = Radar-Rückstreuung in vertikal-horizontaler Polarisation).

### 3.1 Referenzdaten

Referenzdaten werden benötigt zur Kalibrierung und Validierung des Verbuschungsmodells, das zwischen offenen und verbuschten Heide-LRT-Flächen unterscheidet. Als Flächen-Kulisse dienen daher Polygone der Trockenheiden (Code 4030) und der Sandheiden auf Binnendünen (Code 2310). Sie befinden sich in Sachsen (Trainingsgebiete: Gohrischheide, Königsbrücker Heide und Oberlausitz) und in Brandenburg (Trainingsgebiete: FFH-Gebiete Kleine Schorfheide, Schönower Heide und Forsthaus Prösa). Zur Prüfung der Übertragbarkeit wurden darüber hinaus Testflächen aus Nordrhein-Westfalen ausgewählt. Die Lage der Trainingsgebiete bzw. der Testflächen ist in Abbildung 2 dargestellt. Flächenabgrenzungen basieren auf den aktuellsten verfügbaren Kartierdaten, die größtenteils ab 2015 erhoben wurden (mit Ausnahmen aus dem Jahr 2006).

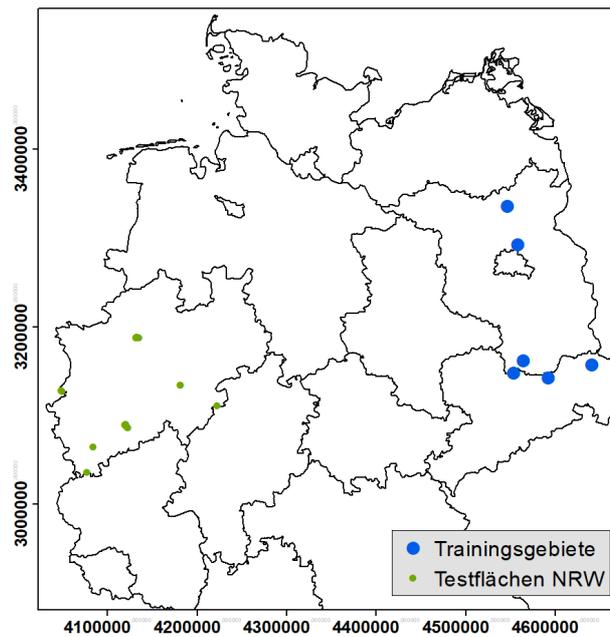


Abbildung 2: Lage der Trainingsgebiete in Brandenburg und Sachsen bzw. der für Übertragbarkeitstests ausgewählten Heide-LRT-Flächen in Nordrhein-Westfalen.

Um Randeffekte zu minimieren, werden die für das Training und die Übertragbarkeitstests ausgewählten Polygone um zehn Meter nach innen gepuffert, d. h. entlang ihrer Innenkanten verkleinert. Die Mindestflächengröße der gepufferten Polygone für die weiteren Analysen beträgt 0,1 ha. Im Anschluss erfolgt eine manuelle Luftbildinterpretation der Polygone anhand der frei verfügbaren aktuellen und historischen Luftbilder der Landesvermessungen (s. Abbildung 3). Es handelt sich um Aufnahmen aus den Referenzjahren 2016-2023, wobei in mehrjährigen Zyklen befliegen wurde. Bei der Klassifizierung werden die Polygone in eine der beiden Struktur-Klassen „verbuscht“ oder „offen“ für das jeweilige Aufnahmejahr eingeordnet. Die Klasse „verbuscht“ enthält durch Verbuschung betroffene Polygone mit lockerer bis dichter Überschirmung. Zu Klasse „offen“ zählen neben offenen Heide-Flächen auch Flächen mit Einzelbäumen, was nach Kartiervorschrift vegetationsstrukturell als lebensraumtypisch gewertet wird. Sind für eine Fläche Luftbildaufnahmen aus mehreren Jahren verfügbar, werden die Klassenzuweisungen für jedes dieser Jahre separat mitgeführt. Dies dient nicht einem Vergleich der Beobachtungen aus verschiedenen Jahren, sondern macht das spätere Modell robuster für eine zeitliche Übertragbarkeit.

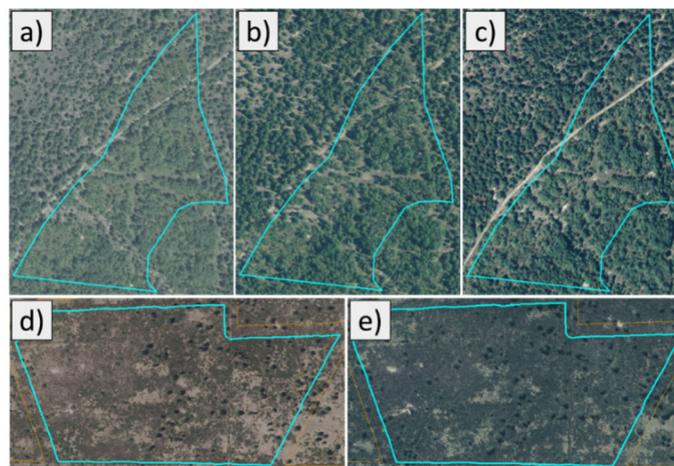


Abbildung 3: Luftbildaufnahmen einer Referenzfläche der Klasse "verbuscht" (oben) aus den Jahren a) 2017, b) 2020, c) 2022 und einer Referenzfläche der Klasse "offen" (unten) (d) 2018 und e) 2022.

Sehr schnelle strukturelle Änderungen innerhalb eines Jahres ergeben sich nur durch Feuer, Maßnahmen wie Entbuschung oder vergleichbaren Ereignissen. Der Aufwuchs von Gehölzen hingegen ist ein langwieriger Prozess. Dies führt zu folgenden fachlich fundierten Annahmen, auf deren Grundlage der Referenzdatensatz auch ohne vorhandenes Luftbild erweitert wird:

- Ist eine Fläche im Luftbild aus Jahr  $x$  offen, so ist sie auch noch im Folgejahr  $x + 1$  offen.
- Ist eine Fläche im Luftbild aus Jahr  $x$  verbuscht, so ist sie auch bereits im Vorjahr  $x - 1$  verbuscht.
- Ist eine Fläche in den Luftbildern aus den Jahren  $x$  und  $x + 3$  offen, so ist sie auch in den Jahren dazwischen offen.
- Ist eine Fläche in den Luftbildern aus den Jahren  $x$  und  $x + 3$  verbuscht, so ist sie auch in den Jahren dazwischen verbuscht.

Insgesamt stehen 1873 Referenzdaten für 361 Flächen aus sieben Referenzjahren zum Training des Modells zur Verfügung. Knapp 40 % der Daten gehören der Klasse „offen“ an.

## 3.2 Verwendete Satellitendaten

### 3.2.1 Optische Satellitendaten

Für die Berechnung des Verbuschungsmodells werden die optischen Daten der Sentinel-2-Satelliten des Copernicus-Programms der ESA sowie der Satelliten Landsat 8 und 9 aus dem Landsat-Programm der NASA genutzt (Siehe auch <https://sentiwiki.copernicus.eu/web/sentiwiki>, European Space Agency, 2024 bzw. <https://landsat.gsfc.nasa.gov/satellites/>, National Aeronautics and Space Administration, 2024). Nach Soubry & Guo (2022) stellen diese Daten geeignete Datenquellen dar, um Indikatoren für Verbuschung abzuleiten. Die genannten Daten verfügen über vier Kanäle, welche die Wellenlängenbereiche Rot, Grün, Blau und Nahes Infrarot (NIR) abdecken. Die räumliche Auflösung dieser Kanäle beträgt bei Sentinel-2 10 m und bei Landsat 8 und 9 30 m. Bei der Auswertung von Satellitendaten sind Indizes, also die Kombination verschiedener Kanäle zu einer neuen, synthetischen Aufnahme, ein gängiges Analysemittel. Sie profitieren von den charakteristischen spektralen Unterschieden zwischen verschiedenen Oberflächen. Photosynthetisch aktive Vegetation weist eine charakteristische, stark zunehmende Reflexion vom roten Wellenlängenbereich bis hin zum NIR auf. Dies wird beispielsweise durch den Enhanced Vegetation Index (EVI) erfasst, der folgendermaßen durch die Verrechnung der Kanäle Rot, Blau und NIR definiert ist (Heute et al., 2002):

$$EVI = \frac{2,5 * NIR - Rot}{NIR + 6 * Rot - 7,5 * Blau + 1}$$

Mit zunehmender Biomasse vitaler Vegetation steigt in der Regel auch der gemessene Indexwert. Auf offenen Heideflächen mit Zwergsträuchern werden vergleichsweise niedrige EVI-Werte beobachtet, während überschirmte Heideflächen höhere EVI-Werte aufweisen. Daher wird der EVI zur Berechnung des Verbuschungsmodells eingesetzt. Der Aufnahmezeitraum der verwendeten EVI-Aufnahmen wird auf den Beginn der Vegetationsperiode, d.h. auf das Frühjahr eines Jahres beschränkt. In diesem Zeitraum ist die Unterscheidung zwischen überschirmten Bereichen und offenen Flächen mit Gräsern, die tendenziell später aufwachsen, besonders gut möglich. Eine Einschränkung bei der Verwendung optischer Daten ist der Einfluss von Wolken und Wolkenschatten auf die Aufnahmen. Bei Wolken können von optischen Sensoren keine spektralen Reflexionen der Erdoberfläche detektiert werden.

### 3.2.2 Radardaten

Zur Berechnung des Verbuschungsmodells werden außerdem Radardaten der Sentinel-1-Satelliten des Copernicus-Programms der ESA verwendet (Siehe auch <https://sentiwiki.copernicus.eu/web/sentiwiki>, European Space Agency, 2024). Sie liefern in der Regel wetterunabhängig durchgehend nutzbares Bildmaterial. Sie erfassen die Intensität der Rückstreuung („Backscatter“, in Dezibel), mit der ein vom Sensor aktiv ausgesendetes Signal von einer bestimmten Oberfläche zum Sensor reflektiert wird. Solche Daten ermöglichen Aussagen über die Textur einer betrachteten Fläche. So weisen raue Oberflächen, die mit Vegetation (insb. mit Gehölzen) bedeckt und

strukturell komplex sind, tendenziell höhere Rückstreuungswerte auf als ebene Oberflächen. Die Aufnahmen von Sentinel-1 stellen damit ebenfalls eine Datenquelle dar, aus der Indikatoren für Verbuschung abgeleitet werden können (Soubry & Guo, 2022). Zur Berechnung des Verbuschungsmodells wird als Polarisation „VH“ (vertikal-horizontal) verwendet, d.h. es wird vertikal polarisierte Strahlung ausgesendet und horizontal polarisierte Strahlung empfangen. Es werden ausschließlich Frühjahres-Aufnahmen genutzt.

### 3.2.3 (Vor-)Prozessierung der Satellitendaten

Analysefähige optische Satellitendaten, werden durch einen FORCE-Datenwürfel bereitgestellt (Framework for Operational Radiometric Correction for Environmental Monitoring) (Frantz, 2019). FORCE ist ein umfassendes Open-Source-Toolset zur Verarbeitung von Satellitenbildern (siehe auch: <https://force-eo.readthedocs.io/en/latest/about.html>). Die Sentinel-2- und Landsat-Szenen werden mit Hilfe des FORCE-Verarbeitungssystems zunächst radiometrisch und geometrisch korrigiert, neu projiziert und wolkenmaskiert. Während dieses Prozesses wird die Landsat-Auflösung von 30m auf 10m mittels Nearest-Neighbor-Resampling angepasst. Für die so vorprozessierten Szenen wird über FORCE jeweils der Index EVI berechnet. Als Inputdaten zur Kalibrierung des Verbuschungsmodells werden hier alle EVI-Szenen der drei Frühjahresmonate März, April und Mai aus den Jahren 2017 bis 2022 bezogen. Diese Rasterdaten werden anschließend in zwei Schritten pixelweise zeitlich aggregiert: Zuerst wird pro Jahr und Monat der durchschnittliche EVI berechnet, auch EVI-Monatskomposit genannt. Im zweiten Schritt wird wieder pro Jahr nun wiederum aus diesen drei Monatskompositen der durchschnittliche EVI berechnet. Damit liegt pro Jahr ein Frühjahreskomposit als Raster vor. Zuletzt erfolgt die räumliche Aggregation dieser Rasterdaten durch Berechnung des Flächenmedians pro Polygon der Referenzdaten pro Referenzjahr.

Die Vorprozessierung von Radardaten erfordert einen hohen Rechenaufwand. Daher werden die über die Plattform CODE-DE (Copernicus Data and Exploitation Platform – Deutschland) als Monatskomposite verfügbaren Sentinel-1-Backscatter-Daten in VH-Polarisation verwendet. CODE-DE ist eine cloudbasierte Plattform, die für die Verarbeitung und Nutzung von Copernicus-Satellitendaten an deutschen Behörden entwickelt wurde (siehe auch: <https://code-de.org/de/>). Die dort vorgehaltenen Monatskomposite basieren auf den geländekorrigierten Backscatter-Daten, die sowohl in Ascending- als auch in Descending-Flugrichtung während eines Monats aufgenommen wurden. Die Monatskomposite der Jahre 2017 bis 2023 für die Frühjahresmonate März, April und Mai werden wie folgt weiterverarbeitet: Sie werden pixelweise gemittelt und zu einem jährlichen Frühjahres-Komposit zusammengefasst. Dann wird für diese jährlich berechneten Raster, wie bereits beim EVI, eine räumliche Aggregation per Median pro Polygon der Referenzdaten pro Referenzjahr durchgeführt.

Das Ergebnis der Satellitendatenprozessierung sind das polygonweise berechnete EVI-Frühjahreskomposit und das polygonweise berechnete VH-Backscatter-Frühjahreskomposit je Referenzjahr. Beide dienen als Inputdaten zum Kalibrieren des Verbuschungsmodells. Die Werteverteilung dieser Fernerkundungs-Inputdaten im zweidimensionalen Merkmalsraum ist in Abbildung 4 dargestellt.

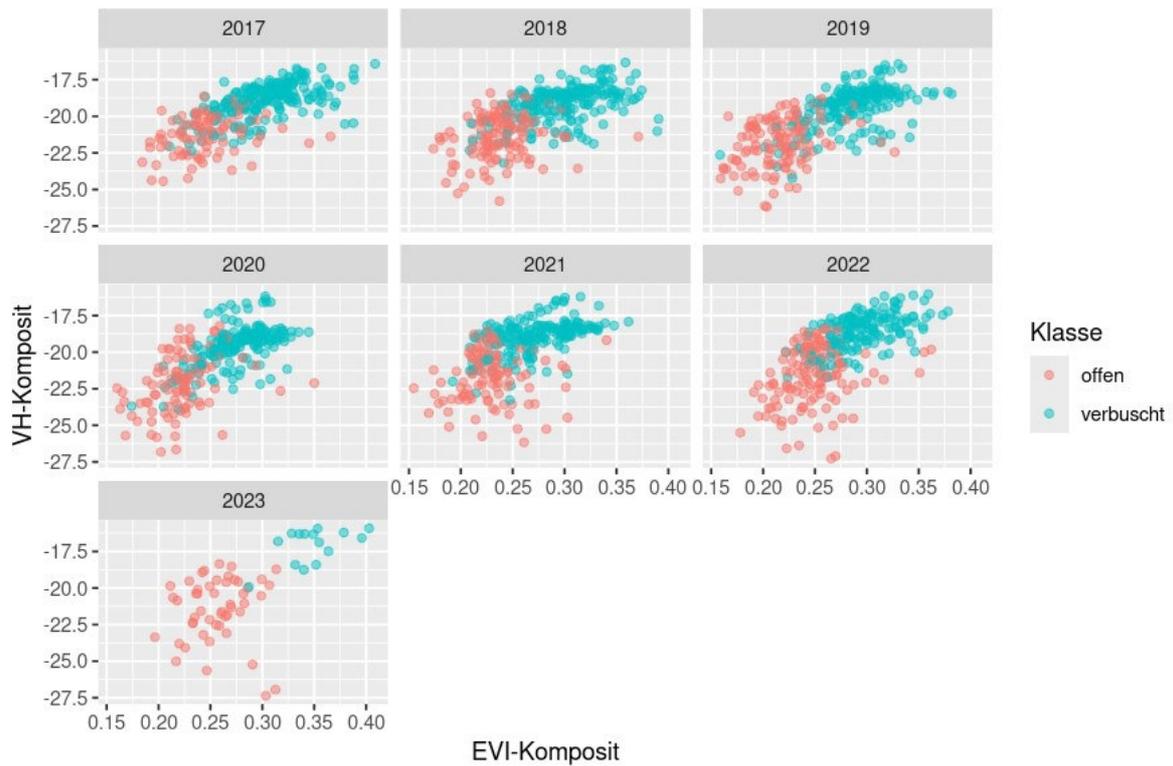


Abbildung 4: Werteverteilung von EVI-Komposit und VH-Backscatter-Komposit der Referenzdaten pro Referenzjahr. Es sind deutliche Cluster der beiden Klassen zu erkennen.

### 3.2.4 Ermittlung typischer Wertebereiche

Auf Basis der Referenzdaten wird der typische Wertebereich im zweidimensionalen Merkmalsraum für die beiden satellitenbasierten Inputdaten – das Frühjahres-Komposit von EVI und VH-Backscatter – ermittelt. Dieser Bereich wird als das minimale, nicht achsenparallele Rechteck definiert, das alle Datenpunkte enthält (engl.: minimal bounding box) und ist in Abbildung 5 zu sehen.

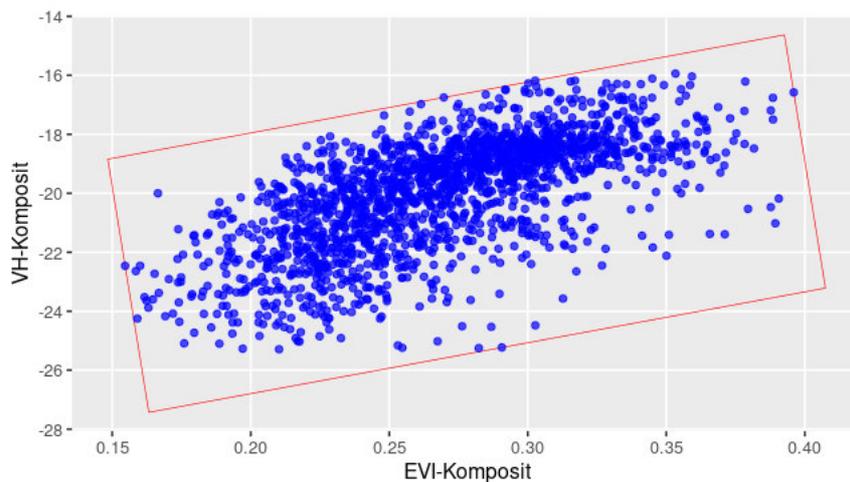


Abbildung 5: Minimale Bounding Box (rot), die alle EVI- und VH-Backscatter-Komposit-Wertekombinationen der Referenzdaten (blau) enthält.

Darüber hinaus werden die minimalen und maximalen Werte der EVI- und VH-Backscatter-Komposite, die von den Referenzdaten angenommen werden, separat erfasst (s. Tabelle 1: Minimale und maximale Werte der EVI- und VH-Backscatter-Komposite der Referenzdaten.).

Tabelle 1: Minimale und maximale Werte der EVI- und VH-Backscatter-Komposite der Referenzdaten.

	Minimum	Maximum
EVI-Komposit	0,15	0,4
VH-Komposit	-25,29	-15,94

Diese Informationen werden im operationellen Dienst genutzt, um einen Qualitätsparameter zur Beurteilung der Plausibilität der Modellergebnisse zu liefern.

### 3.3 Konzeptionierung eines Klassifikationsmodells

Für die Konzeptionierung des Dienstes wird ein Verbuschungsmodell als überwachte Klassifikation trainiert. Als Trainingsdaten (in Kapitel „Referenzdaten“ beschrieben) dienen Polygone, bei denen die Strukturklassen „verbuscht“ und „offen“ bekannt sind. Im Klassifikationsprozess wird auf Basis dieser Trainingsdaten eine Funktion abgeleitet, die für neue Daten modellierte Klassenzuordnungen liefern kann. Datengrundlage für das Klassifikationsmodell sind dabei die räumlich aggregierten Frühjahrskomposite der EVI Indizes und der VH-Backscatter Daten. Da für die Modellierung nur zwei Klassen unterschieden werden müssen, wurde ein logistisches Regressionsmodell (Cheng et al. 2006) gewählt, das für eine zu analysierende Fläche Wahrscheinlichkeiten einer Klassenzugehörigkeit erhebt. Ab einer Wahrscheinlichkeit von 50% wird die Klasse „verbuscht“ vorhergesagt. Abbildung 6 zeigt, für welche Werte des EVI-Komposits und des VH-Backscatter-Komposits eine Verbuschungswahrscheinlichkeit von über 50 % vorhergesagt werden kann.

Zur Validierung des Modells wird der Datensatz in Kalibrierungs- und Validierungsdaten aufgeteilt. Für den operationellen Betrieb des Dienstes wird das Verbuschungsmodell auf Basis des vollständigen Referenzdatensatzes berechnet.

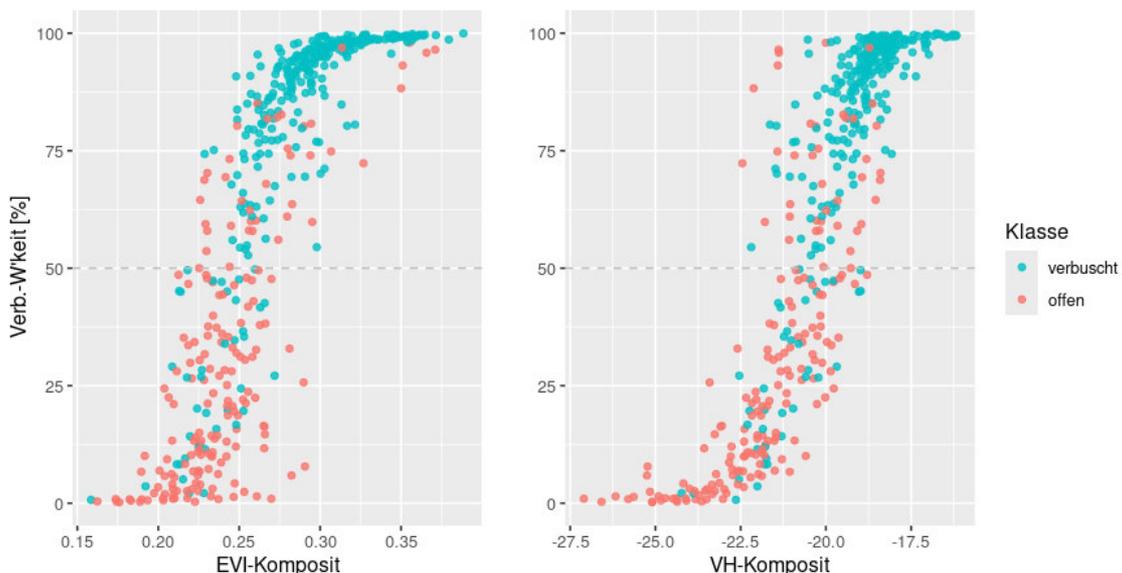


Abbildung 6: Vorhergesagte Verbuschungswahrscheinlichkeit für den Validierungsdatensatz nach EVI- und VH-Backscatter-Komposit. Die Trennlinie bei 50 % markiert die Grenze zwischen den Klassen „verbuscht“ und „offen“.

### 3.4 Validierung

Zur Bewertung des erstellten Verbuschungsmodells wurde eine 10-fache Kreuzvalidierung durchgeführt. Der Referenzdatensatz wurde in zehn Teile, sogenannte Resamples, aufgeteilt, die jeweils 187 oder 188 Testdaten enthalten. Dieses Resampling erfolgt stratifiziert, d.h. das Verhältnis der Anzahl an Beobachtungen der Klassen „offen“ und „verbuscht“ bleibt in jedem Resample konstant. In jedem der zehn Durchläufe („Folds“) wird ein Resample als Validierungsdatensatz verwendet, während die restlichen Daten zur Kalibrierung des Modells dienen. Während der Validierung jedes Folds werden die vom Modell vorhergesagten Klassen mit den tatsächlichen Klassen verglichen. Die Klasse „verbuscht“ wird grundsätzlich ab einer Wahrscheinlichkeit von 50% vorhergesagt. Am Ende wird die durchschnittliche Modellgüte über alle Folds hinweg berechnet. Die Ergebnisse der Kreuzvalidierung sind wie folgt:

- Gesamtgenauigkeit: 84% (mit einer Standardabweichung von 2% zwischen den Folds)
- Kappa-Wert: 0,66 (mit einer Standardabweichung von 0,05 zwischen den Folds)

Für detailliertere Informationen zur Güte des Modells wird auf Basis einer zufälligen Datenpartition von Kalibrierungs- und Validierungsdaten im Verhältnis 3:1 eine Konfusionsmatrix berechnet. Die Genauigkeiten sind Tabelle 2 zu entnehmen.

Tabelle 2: Konfusionsmatrix für eine zufällige Datenpartition der Referenzdaten.

		Referenz		User's Accuracy	Commission Error
		offen	verbuscht		
Vorhersage	offen	138	38	78%	22%
	verbuscht	40	252	86%	14%
Producer's Accuracy		78%	87%	<b>Overall Acc.: 83%</b>	
Omission Error		22%	13%		

## 4. Umsetzung als Dienst im operationellen Betrieb

In diesem Teil wird beschrieben, nach welchem Konzept der Heide-Verbuschungsdienst für einen operationellen Betrieb umgesetzt ist. Eine grafische Übersicht des Konzepts ist Abbildung 7 zu entnehmen.

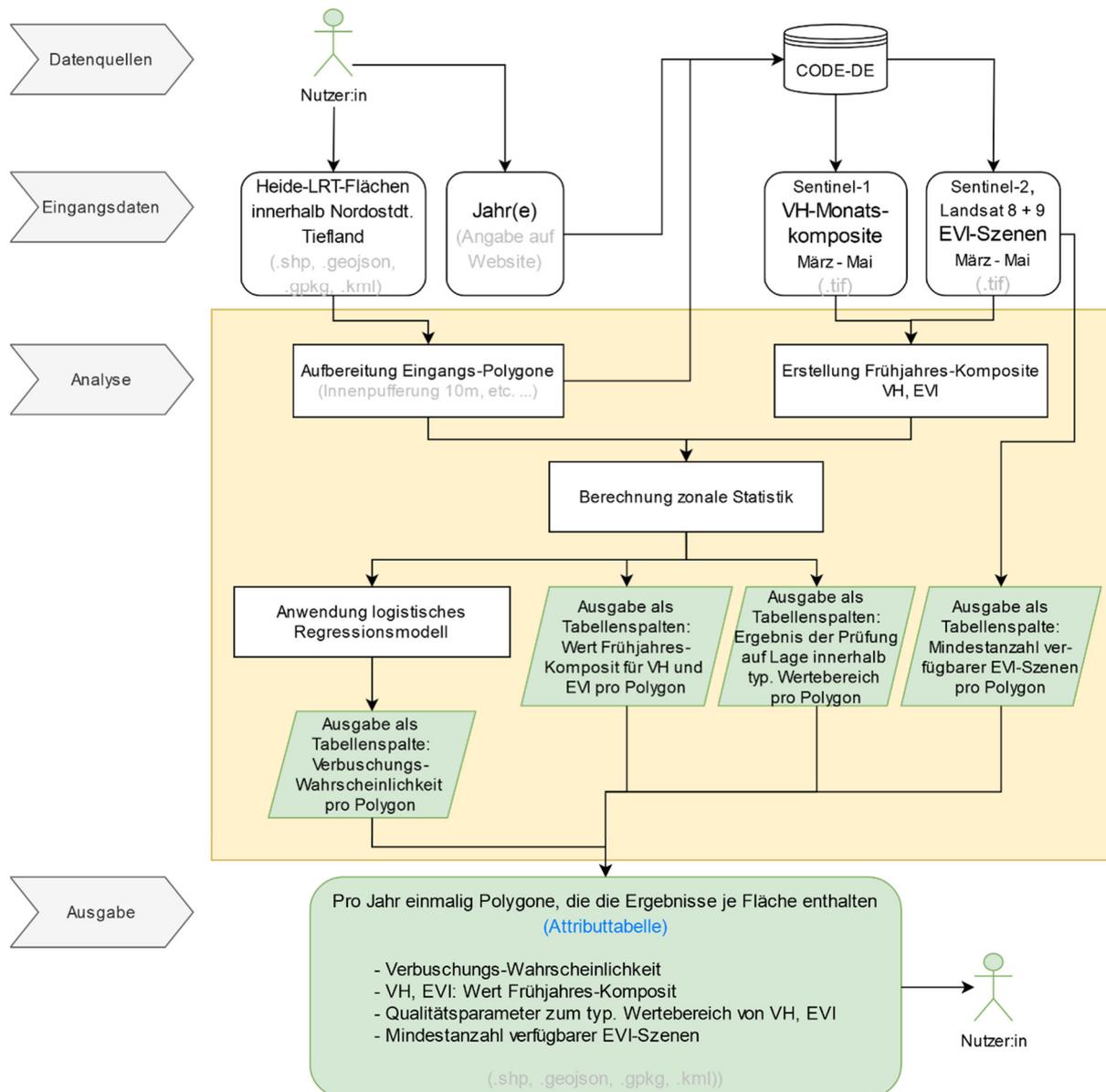


Abbildung 7: Schematische Darstellung des Konzepts des Heide-Verbuschungsdienstes im operationellen Betrieb.

#### 4.1 Eingabegeometrien und -parameter

Der Dienst setzt voraus, dass bei den Nutzenden raumbezogene Flächen-Geometrien über die zu prüfenden Flächen vorliegen. Diese werden über eine Eingabemaske hochgeladen.

Polygone:

- **Formate:** Die Geometrien bzw. Vektordaten oder Polygone können in den gängigen Geodatenformaten ESRI Shapefile, GeoJSON oder KML in das System eingespeist werden.
- **Eindeutige Zuweisung der Polygone:** Eingangspolygone müssen mit einer eindeutigen ID versehen sein. Der zugehörige Attributspaltenname wird nach dem Upload auf der Nutzeroberfläche abgefragt
- **Projektion:** Die Polygone sollten in einem gültigen Koordinationssystem vorliegen.
- **Hinweise:**
  - o Nur Polygone innerhalb der naturräumlichen Großlandschaft „Nordostdeutsches Tiefland“ (Naturräumliche Gliederung nach BfN 2022) werden analysiert

- Nur für Polygone mit FFH-LRT-Codes 2310 und 4030 sind sinnvolle Ergebnisse zu erwarten.
- Es findet keine Plausibilitätsprüfung der verwendeten Polygone statt. Für die Korrektheit der Polygone und die Einhaltung des Datenschutzes sind die Nutzenden selbst verantwortlich.

Andere Eingabeparameter:

- Analysejahr(e): Die Konfiguration des Dienstes erfordert die Angabe eines Zeitraums, für den die Analyse durchgeführt werden soll (Analysezeitraum ab 2017, Aktualisierung fortlaufend und abhängig von der Datenbereitstellung). Die Angabe einzelner oder mehrerer Jahre ist möglich. Bei Angabe mehrerer Jahre werden die Ergebnisse für jedes Jahr gesondert ausgegeben. Analysen beziehen sich auf den Zeitraum März bis Mai der angegebenen Jahre.

## 4.2 Bearbeitung der Eingangspolygone

Bevor die Satellitendaten auf die zu analysierenden Inputpolygone angewendet werden, erfolgt deren Aufbereitung in mehreren Einzelschritten. Zunächst wird für jedes Polygon eine interne eindeutige ID („cg\_id“) vergeben. Multipart-Polygone aus den Originaldaten erhalten für jeden Teil eine eigene interne ID, um später separate Analysen zu ermöglichen. Anschließend werden die Polygone auf ihre Lage innerhalb des Naturraums „Nordostdeutsches Tiefland“ geprüft und auf ein internes einheitliches Koordinatensystem umprojiziert (EPSG: 3035). Durch eine Pufferung von zehn Metern nach innen werden mögliche Mischeffekte an den Polygongrenzen verhindert. Sehr kleine oder schmale Flächen werden dadurch aus der Analyse ausgeschlossen. Schließlich erfolgt eine Filterung nach Mindestgröße, sodass nur Polygone mit einer Fläche von mindestens 0,1 ha nach der Innenpufferung in die weiteren Analysen einbezogen werden. Dies ist eine geeignete Grenze, wenn Sentinel-1- und Sentinel-2-Daten mit einer räumlichen Auflösung von 10 Metern verwendet werden.

## 4.3 Prozessierung der Satellitendaten

Die Aufbereitung der Satellitendaten im operationellen Betrieb des Dienstes erfolgt analog zu der in Kapitel „Modellberechnung – Verwendete Satellitendaten“ detailliert beschriebenen Prozessierung der Satellitendaten für die Referenzdaten.

D.h. für die gewählten Jahre werden alle EVI-Szenen der Sensoren Sentinel-2 und Landsat 8 und 9 aus dem Zeitraum März bis Mai über FORCE (Frantz, 2019) prozessiert. Anschließend werden sie per Durchschnitt pixelweise zunächst zu Monatskompositen und anschließend zu jährlichen Frühjahres-Kompositen gemittelt. Der Median als zonale Statistik des EVI-Frühjahres-Komposits wird jedem Polygon als Attribut hinzugefügt („evi\_komp“).

Mit den VH-Backscatter-Daten wird ähnlich verfahren, wobei hier nicht die Einzelaufnahmen, sondern die bereits zeitlich aggregierten Monatskomposite aus dem CODE-DE-Portfolio zur Berechnung der Frühjahres-Komposite herangezogen werden. Jedes Polygon erhält im Ergebnis dieses Schritts den Flächenmedian des VH-Backscatter-Frühjahres-Komposits als ein weiteres Attribut („vh\_komp“).

## 4.4 Ableitung der Verbuschungs-Wahrscheinlichkeit

Die zuvor erläuterten, pro Polygon aufbereiteten Satellitendaten („evi\_komp“ und „vh\_komp“) dienen als Eingabe für ein bereits trainiertes binäres logistisches Regressionsmodell, beschrieben in Kapitel „Modellberechnung – Verwendetes Modell“. Dieses Modell unterscheidet auf Grundlage des EVI-Frühjahres-Komposits und des VH-Backscatter-Frühjahres-Komposits zwischen offenen und verbuschten Flächen. Es wird verwendet, um die Verbuschungs-Wahrscheinlichkeit für jedes Polygon vorherzusagen („verb\_wkeit“). Das Modell ist so konzipiert, dass eine Wahrscheinlichkeit von 50 % als geeigneter Schwellenwert zur Unterscheidung zwischen den beiden Strukturklassen dient.

## 4.5 Bestimmung der Qualitätsparameter

Im Rahmen einer Dienstanfrage wird geprüft, ob die Satellitendaten, die für den angefragten Zeitraum prozessiert und als aktueller Modell-Input verwendet werden, innerhalb des typischen Wertebereichs liegen. Dieser Wertebereich wird in Abschnitt „Modellberechnung – Ermittlung typischer Wertebereich“ durch Minimum- und Maximum-Statistiken sowie die minimale Bounding Box der Referenzdaten definiert. Es erfolgen drei Prüfungen: Zum einen wird überprüft, ob die aktuelle Wertekombination der beiden Input-Daten, EVI- und VH-Backscatter-Frühjahres-Komposit, innerhalb der vorgegebenen minimalen Bounding Box im zweidimensionalen Merkmalsraum liegt („wert\_pruef“). Zum anderen wird für beide Input-Daten getrennt geprüft, ob die aktuellen Werte innerhalb des jeweils durch Minimum und Maximum definierten Wertebereichs liegen („evi\_pruef“, „vh\_pruef“). Ein aktueller Wert gilt damit als „typisch“, wenn er sich im von den Referenzdaten aufgespannten Wertebereich bewegt.

Zusätzlich wird über FORCE (Frantz, 2019) die Anzahl der wolkenfreien EVI-Szenen pro Pixel im Frühjahres-Zeitraum der angegebenen Jahre ermittelt. Diese Information wird polygonweise auf das Minimum aggregiert („min\_cso“). An Sentinel-1-Daten, die unabhängig von Bewölkung regelmäßig verfügbar sind, liegen deutschlandweit durchschnittlich mindestens neun Aufnahmen pro Monat vor (nach Missionsende von Sentinel-1B im Dezember 2021 sind es vier). Diese Angabe ist nicht variabel und wird daher nicht als zusätzlicher Parameter ausgegeben.

## 4.6 Ergebnisausgabe

Die Ausgabe des Dienstes erfolgt im selben Format und Koordinatensystem, in dem die Polygone von den Nutzenden für den Dienst bereitgestellt wurden.

Pro (ggf. aufgetrennter) Eingangs-Geometrie (und Jahr) werden zwei Ausgaben generiert:

A: Name\_results\_<Jahr>

- Polygone mit Angabe der Verbuschungs-Wahrscheinlichkeit pro Polygon („verb\_wkeit“), den Modell-Inputdaten („evi\_komp“, „vh\_komp“), sowie Qualitätsparametern („groesse“, „wert\_pruef“, „evi\_pruef“, „vh\_pruef“, „min\_cso“); Ausgabe der nach innen gepufferten Polygone nach Ausschließen von Polygonen, die nicht prozessiert werden

B: Name\_orig\_features

- Originalgeometrien zzgl. der Angabe, welche Polygone nicht weiter analysiert wurden („cg\_proz\_JJ“)

Tabelle 3 gibt eine Übersicht der Felder, die vom Heide-Verbuschungsdienst generiert werden. Werden mehrere Jahre angefordert, so werden diese Dateien für jedes Jahr gesondert ausgegeben.

Tabelle 3: Übersicht über alle Attribute, die für der Ergebnisausgabe des Heide-Verbuschungsdienstes generiert werden.

Attributname	Attributname im WebGIS	Beschreibung	Werte(-bereich)	Ausgabe
Originales ID-Attribut	-	Flächenkennung entsprechend der Eingangs-Datei		A, B
cg_id*	-	Eindeutige ID Zuweisung pro Polygon		A, B
verb_wkeit	Verbuschungs-Wahrscheinlichkeit in %	Wahrscheinlichkeit, dass das Polygon verbuscht ist	0 bis 100 [%]	A

		Hinweis: Nicht zu verwechseln mit prozentualem Verbuschungsgrad!		
jahr	Analysejahr	Analysejahr	ab 2017	A
evi_komp	EVI-Komposit	Wert EVI-Frühjahres-Komposit im Analysejahr für das Polygon	ca. 0,1 bis 0,4 *	A
vh_komp	VH-Komposit	Wert Frühjahres-Komposit der Rückstreuung (VH) im Analysejahr für das Polygon	ca. -27 bis -16 [Dezibel] *	A
groesse	Bewertung Flächengröße	Einteilung der Fläche nach ihrer Größe	- 0,1 bis 0,25 ha: „klein“ - 0,25 bis 0,5 ha: „ok“ - über 0,5 ha: „gut“	A
wert_pruef	Prüfung Wertekombination EVI, VH	Kategorischer Qualitätsparameter zur Überprüfung, ob die Frühjahrs-Komposite von EVI und VH-Backscatter typisch im Merkmalsraum liegen	- „Wertkombination typisch“ - „Wertkombination untypisch“	A
evi_pruef	Prüfung EVI-Wert	Kategorischer Qualitätsparameter zur Überprüfung, ob das EVI-Frühjahres-Komposit innerhalb des typischen Wertebereichs liegt	- „im typischen EVI Bereich“ - „untypisch niedriger EVI“ - „untypisch hoher EVI“	A
vh_pruef	Prüfung VH-Wert	Kategorischer Qualitätsparameter zur Überprüfung, ob das VH-Backscatter-Frühjahres-Komposit innerhalb des typischen Wertebereichs liegt	- „im typischen VH Bereich“ - „untypisch niedriger VH“ - „untypisch hoher VH“	A
min_cso	Min. Anzahl EVI-Beobachtungen	Minimale Anzahl an Satellitenbildern, die in die Berechnung der EVI-Komposite mit eingegangen sind („Clear Sky Observation“)	0 bis 26 *	A
cg_proz_JJ		Information ob dieses Polygon prozessiert wird. JJ steht für das jeweilige Analysejahr.	- „Polygon wird prozessiert“ - „Polygon durch Puffer verloren“ - „Polygon nicht im Zielgebiet“	B

\* Wertebereich basierend auf Referenzdatensatz

#### 4.6.1 Interpretation der Ergebnisausgabe

Zur Interpretation der ausgegebenen Attribute gibt es folgende Hinweise:

- „verb\_wkeit“: Grundsätzlich ist das Modell zur Berechnung der Verbuschungswahrscheinlichkeit so konzipiert, dass die Wahrscheinlichkeit von 50 % als geeigneter Schwellenwert zur Unterscheidung von verbuschten und offenen Heide-Flächen gilt. Dies kann

z.B. je nach angefragter Region und Einfluss der vorherrschenden meteorologischen Bedingungen auf die Modell-Inputdaten variieren. Achtung: Wert ist nicht zu verwechseln mit prozentualem Verbuschungsgrad!

- „evi\_komp“, „vh\_komp“: Die Werte stammen aus den aufbereiteten Satellitendaten (EVI- und VH-Backscatter-Frühjahres-Komposite), auf deren Grundlage die Verbuschungswahrscheinlichkeit berechnet wird. Die Ausgabe dieser Attribute erfolgt lediglich zur Kenntnis. Die Daten können bei Bedarf auch von den Nutzenden selbst weiterverarbeitet werden, beispielsweise zur Erstellung von Zeitreihen.
- „groesse“: Für kleine Flächen ist das Ergebnis der Verbuschungswahrscheinlichkeit unter Umständen nicht verlässlich.
- „wert\_pruef“, „evi\_pruef“, „vh\_pruef“: Die Qualität der Ergebnisse hängt davon ab, wie ähnlich die aktuellen Modell-Input-Daten der Dienst-Anfrage den ursprünglichen Daten zur Kalibrierung des Modells sind. Abweichungen der Frühjahreskomposite von EVI und VH-Backscatter können insbesondere dann auftreten, wenn der angefragte Zeitraum meteorologisch oder klimatisch ungewöhnlich ist (z.B. ein Jahr mit einem ungewöhnlich frühen Beginn der Vegetationsperiode) oder wenn ein Gebiet angefragt wird, das sich stark von den Gebieten unterscheidet, die in den Referenzdaten vertreten sind. Daher wird im Rahmen einer Dienstanfrage geprüft, ob die prozessierten Satellitendaten innerhalb des typischen Wertebereichs liegen, der in Abschnitt „Modellberechnung – Ermittlung typischer Wertebereich“ durch Minimum- und Maximum-Statistiken sowie die minimale Bounding Box der Referenzdaten definiert wird. Diese Prüfungen geben Aufschluss darüber, ob das Modell auf den aktuellen Werten trainiert wurde oder ob eine Extrapolation stattfindet. Abweichungen („untypische“ Werte/Wertkombinationen) können darauf hindeuten, dass das Modell ggf. nicht ohne Weiteres auf den gewünschten Zeitraum oder das gewünschte Gebiet übertragbar ist. Mögliche Erklärungen für Abweichungen sind:
  - „untypisch hoher EVI“: besonders viel photosynthetisch aktive Biomasse, z.B. durch überschirmende Gehölze
  - „untypisch niedriger EVI“: besonders wenig photosynthetisch aktive Biomasse, z.B. wegen weitläufiger Sandbereiche
  - „untypisch hoher VH“: besonders raue Oberfläche, z.B. durch überschirmende Gehölze
  - „untypisch niedriger VH“: besonders ebene Oberfläche, z.B. wegen weitläufiger Sandbereiche
- „min\_cso“: Bei der Verwendung von Satellitendaten spielt deren Verfügbarkeit eine wichtige Rolle für die Aussagekraft der zu produzierenden Ergebnisse. In Zeitreihen optischer Daten können aufgrund von Wolken größere Datenlücken entstehen. „min\_cso“ gibt an, wie viele wolkenfreie Satellitenbilder im Zeitraum März bis einschließlich Mai im Analysejahr pro Polygon mindestens zur Verfügung standen. Ist diese Anzahl sehr niedrig (nahe Null), ist das Ergebnis der Verbuschungswahrscheinlichkeit unter Umständen nicht verlässlich.

## 5. Gültigkeitsbereich und Übertragbarkeit

Die Übertragbarkeit des Verbuschungsmodells lässt sich grundsätzlich nur für jene Gebiete abschätzen, für die Referenzdaten vorhanden sind. Für die Entwicklung des Heide-Verbuschungsdienstes wurden Daten aus den Bundesländern Sachsen (SN), Brandenburg (BB) und Nordrhein-Westfalen (NRW) zusammengetragen. Die Testgebiete in Sachsen und Brandenburg sind aus fernerkundlicher und phänologischer Sicht relativ ähnlich (s. Abbildung 8).

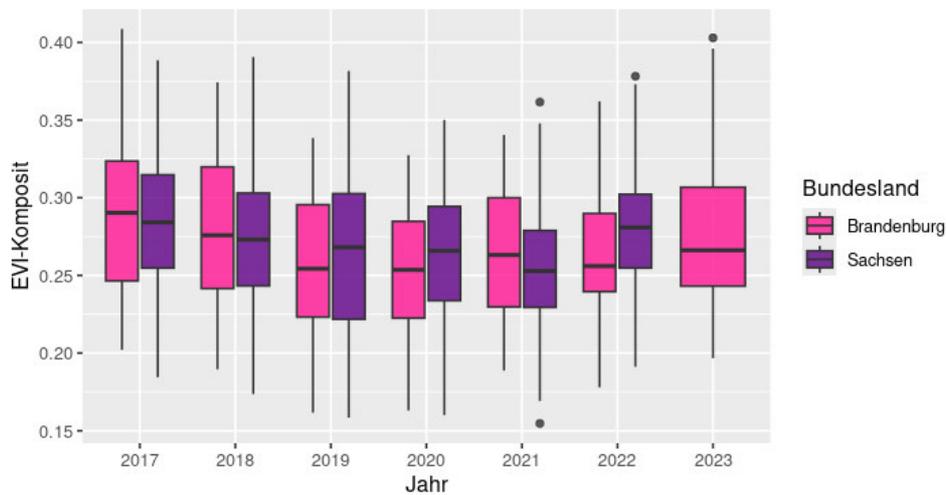


Abbildung 8: Werteverteilung der EVI-Komposite der Referenzdaten für Sachsen und Brandenburg nach Referenzjahr.

Daher eignen sich diese Daten als Trainingsdaten für ein gemeinsames Modell. Fachlich begründet ist die Gültigkeit der Modellergebnisse an die naturräumlichen Gegebenheiten und nicht an administrative Landesgrenzen gebunden. Daher ist die naturräumliche Großlandschaft „Nordostdeutsches Tiefland“ (s. Abbildung 9) als Gültigkeitsbereich für das Verbuschungsmodell festgelegt (Naturräumliche Gliederung nach BfN, 2022).

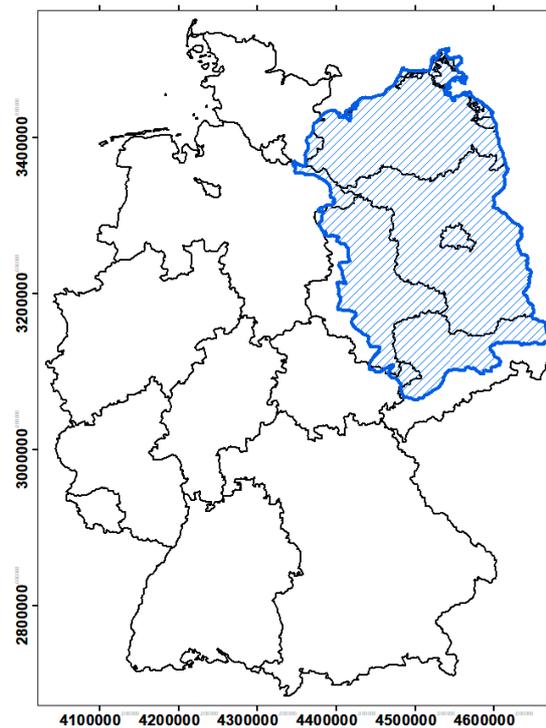


Abbildung 9: Naturräumliche Großlandschaft „Nordostdeutsches Tiefland“ (blau) als Gültigkeitsbereich des Heide-Verbuschungsdienstes (Naturräumliche Gliederung nach BfN, 2022).

Referenzdaten aus Nordrhein-Westfalen zeigen, dass dieselben Parameter, die im nordostdeutschen Tiefland zur Unterscheidung von offener und verbuschter Heide verwendet werden, hier nicht anwendbar sind (s. Abbildung 10). Dies bedeutet, dass das Modell für Nordrhein-Westfalen nicht übertragbar ist. Für andere Regionen standen keine entsprechenden Referenzdaten zur Verfügung, weshalb eine Aussage zur Übertragbarkeit des Modells auf diese Gebiete nicht getroffen werden kann.

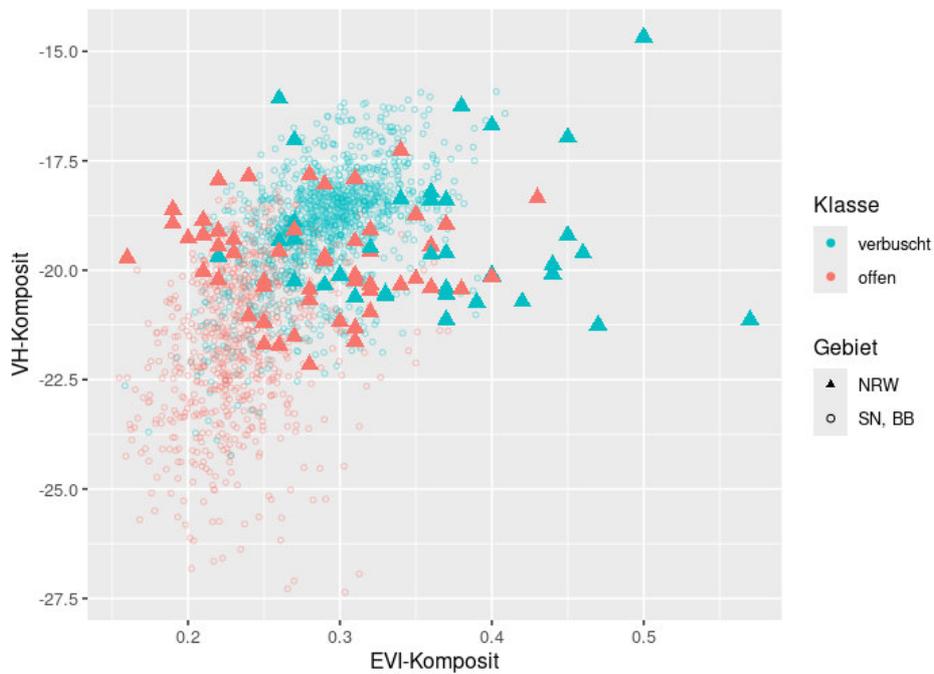


Abbildung 10: Werteverteilung von EVI-Komposit und VH-Backscatter-Komposit der Referenzdaten aus NRW in Bezug zu Werten für Sachsen und Brandenburg. Insb. anhand des VH-Backscatter-Komposits können die Klassen „verbuscht“ und „offen“ nicht unterschieden werden.

## 6. Fehlerquellen und Einschränkungen

Die Ergebnisse sind nur so genau wie die von den Nutzenden bereitgestellten Eingabepolygone. Der Dienst ist nur für Flächen der LRT-Codes 2310 und 4030 konzipiert. Für andere LRT können keine sinnvollen Ergebnisse erwartet werden. Flächen außerhalb der naturräumlichen Großlandschaft „Nordostdeutsches Tiefland“ werden nicht analysiert.

Die ausgewerteten Satellitendaten weisen systembedingt gewisse Grenzen hinsichtlich der Erfassungsgenauigkeit von Verbuschung auf. Die räumliche Auflösung, d.h. die Pixelkantenlänge beschränkt sich auf 10 m (Sentinel-1 und -2) bzw. 30 m (Landsat 8 und 9). Die Erkennungsgenauigkeit hängt auch ab von der Verfügbarkeit und Qualität der Satellitenbilder, z.B. können größere Lücken in den Satellitendaten oder in der Wolkenmaskierung die Ergebnisse beeinflussen.

Auf Jahre, in denen meteorologische Ausnahmebedingungen herrschen (z.B. ein besonders früher oder besonders später Beginn der Vegetationsperiode), ist der Dienst ggf. nicht ohne Weiteres übertragbar. Ob dies bei einer konkreten Dienstanfrage der Fall ist, kann den kategorischen Qualitätsparametern entnommen werden.

## 7. Zitierhinweis

Lins, M., Holtgrave, A., Schroiff, A. (2024): CopGruen – Heide-Verbuschungsdienst. Ein Verbundprojekt von BfUL, LANUV, LfU Schleswig-Holstein, RLP AgroScience, Thünen Institut, TU Berlin.

Bei Verwendung des Dienstoutputs muss folgende Quelle zitiert werden:

Frantz, D. (2019). FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. *Remote Sensing*, 11, 1124

## Literatur

Bundesamt für Naturschutz (BfN) (2022). Naturräumliche Gliederung (WMS) [Data set]. <https://nokis.mdi-de-dienste.org/trefferanzeige?docuuid=685e36cd-87a7-4e7c-a780-ab7b19d54115>

Cheng, Q., Varshney, P. K., & Arora, M. K. (2006). Logistic regression for feature selection and soft classification of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 3(4), 491-494.

Frantz, D. (2019). FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. *Remote Sensing*, 11, 1124

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, 83(1-2), 195-213.

Soubry, I., & Guo, X. (2022). Quantifying woody plant encroachment in grasslands: A review on remote sensing approaches. *Canadian Journal of Remote Sensing*, 48(3), 337-378.