

UAV-basierte Phänotypisierung von Rapshybriden unter Einsatz von multivariater Statistik und Verfahren des maschinellen Lernens

DENNIS LÖWE¹, SEBASTIAN WARNEMÜNDE², JONATHAN BRASSAC²,
ANDREAS STAHL² & HEIKE GERIGHAUSEN¹

Zusammenfassung: Winterraps ist die bedeutendste Ölfrucht Deutschlands, gerät aufgrund seiner vergleichsweise geringen Stickstoffeffizienz immer wieder in den Fokus von Wissenschaft und Öffentlichkeit. Das Projekt NORA zielt darauf ab mittels digitaler Phänotypisierungstechnologien Modelle zu entwickeln, welche es Züchtern erlauben sollen, schneller und effizienter ressourcenschonende Rapshybridsorten zu selektieren. Unter Zuhilfenahme multispektraler UAV-Aufnahmen und Algorithmen der multivariaten Statistik und des maschinellen Lernens konnten im ersten Versuchsjahr 2023 Modelle entwickelt werden, welche auf ausgewählten Versuchsfeldern einen großen Anteil der Varianz von Frischmasse und Pflanzenstickstoffgehalt erklären können. Aufbauend auf den ersten Zwischenergebnissen sollen im Versuchsjahr 2024 mit einer erweiterten Datenbasis die Ergebnisse des ersten Versuchsjahres validiert und optimiert werden.

1 Einleitung

Angesichts der globalen Veränderungen im Klima, der Wirtschaft und des Bevölkerungswachstums sieht sich die moderne Landwirtschaft mit zunehmenden Herausforderungen konfrontiert. Als Reaktion darauf wurden in der Pflanzenproduktion vermehrt synthetischer Düngemittel eingesetzt und das pflanzenbauliche Management der Ackerkulturen optimiert. Besonders hervorzuheben ist der Einsatz von Stickstoffdüngemitteln, da Stickstoff (N) eine entscheidende Rolle in der Pflanzenentwicklung spielt (TILMAN et al. 2002; SINCLAIR & RUFTY 2012) und zu den am stärksten eingesetzten Düngekomponenten zählt (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO) 2020). Neben der energieintensiven Herstellung und möglichen Nitrateinträgen in das Grundwasser (DI & CAMERON 2002; SEBILO et al. 2013) ist besonders die Entgasung von ungefähr 0,6% (Deutschland, (MATHIVANAN et al. 2021) der N-Düngemenge als Lachgas (N₂O, ~300-faches CO₂-Äquivalent) aus Klimaschutzsicht besorgniserregend. In der Forschung rückt insbesondere der Anbau von Winterraps (*Brassica napus* L.), aufgrund seiner vergleichsweise niedrigen Stickstoffeffizienz (WANG et al. 2020; ZHAN et al. 2023) in den Fokus der Aufmerksamkeit. So nimmt der Winterraps im Herbst besonders viel Bodenstickstoff auf, verliert jedoch nach der Blüte bis zur Ernte einen vergleichsweise hohen Anteil davon wieder. Global gesehen zählt der Raps (Winter-/Sommer-raps) zu den am häufigsten angebauten Ölfrüchten für die Pflanzenölgewinnung und folgt auf Palm- und Sojaöl. Mit einem Anteil von ungefähr 23% des globalen Gesamtvolumens (USDA'S FOREIGN AGRICULTURAL SERVICE 2023) zählt die europäische Union (EU) neben Kanada und China zu den Hauptproduzenten von Raps. Deutschland gehört gemeinsam mit

¹ Julius Kühn-Institut (JKI) – Bundesforschungsinstitut für Kulturpflanzen, Institut für Pflanzenbau und Bodenkunde, Bundesallee 58, D-38116 Braunschweig, E-Mail: [dennis.loewe, heike.gerighausen]@julius-kuehn.de

² Julius Kühn-Institut (JKI) – Bundesforschungsinstitut für Kulturpflanzen, Institut für Resistenzforschung und Stresstoleranz, Erwin-Baur-Str. 27, D-06484 Quedlinburg, E-Mail: [sebastian.warnemuende, jonathan.brassac, andreas.stahl]@julius-kuehn.de

Frankreich und Polen zu den maßgeblichen Rapsproduzenten in der EU. Im deutschen Kontext nimmt der Winterraps eine herausragende Position ein und zählt mit 1.1 Mio. ha (Statistisches Bundesamt) zu den am häufigsten angebaute Fruchtarten (Abb.1). In diesem Kontext ist es hinsichtlich der Einhaltung globaler und nationaler Klimaschutzziele umso wichtiger ressourceneffiziente Sorten zu züchten. Ein Ziel des im Rahmen des Klimaschutzsofortprogramms des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) geförderten Projektes NORA („N optimierter Raps zur Reduktion der Treibhausgasemissionen“) ist es daher, mittels digitaler Phänotypisierungstechnologien die Auswahl und Zucht von ressourcenschonenden Raps-hybridsorten zu erleichtern, um schneller und effektiver N-effiziente Sorten selektieren zu können.

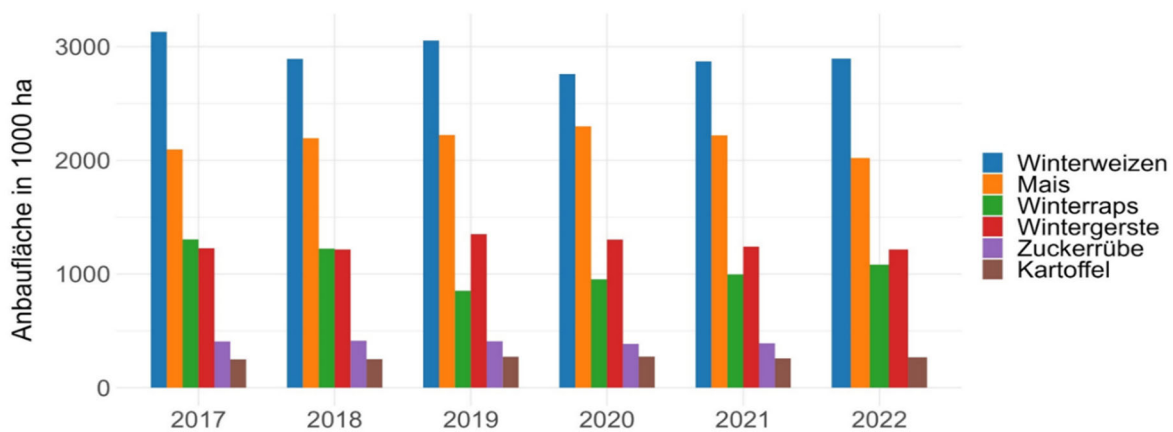


Abb. 1: Anbauflächen der verbreitetsten Fruchtarten Deutschlands im Vergleich zwischen 2017 und 2022 (BUNDESMINISTERIUM FÜR ERNÄHRUNG UND LANDWIRTSCHAFT (BMEL) 2022).

2 Methodik

2.1 Feldversuche

Im Rahmen des Versuchsjahres 2022/2023 wurden Stickstoffsteigerungsversuche auf dem Gelände des Julius-Kühn-Instituts (JKI) in Quedlinburg (QLB) (5 Genotypen, 5 Düngestufen, 150 Parzellen) und am Standort des JKI in Braunschweig (BS) (4 Genotypen, 5 Düngestufen, 80 Parzellen) angelegt. Am Standort QLB wurde Winterraps ausgesät, in BS wurden Sommer-raps-genotypen untersucht. Eine Hälfte der jeweiligen Versuche blieb für die Ertragsermittlung unangetastet um Auskunft über den Effekt der jeweils zusätzlichen N-Düngung auf Ertrag, N-Entzug, N-Verluste (durch Bilanzierung erfasst) und N-Effizienz geben zu können. Die zweite Hälfte der Versuche wurde während der Vegetationsperiode wiederholt manuell bonitiert und zeitgleich mittels UAV-gestützter Multispektralkamerasysteme überflogen. Zu den erhobenen Parametern zählen: BBCH, Biomasse (frisch/trocken), SPAD-Messungen, Leaf area index (LAI), Pflanzen- und Blattstickstoffgehalt sowie die Pflanzenhöhe. Die Beprobungsfläche pro jeweiligen Beprobungszeitpunkt bezifferte 0,64 m² (QLB) und 1 m² (BS). Zusätzlich wurde eine Kollektion von 200 Testhybriden, die auf vier Tester (MSL-Mütter) und 50 unterschiedliche Väter zurückgeht (4x50), an mehreren Versuchsstandorten (Quedlinburg, Weddegast, Hohenlieth, Salzkotten, Asendorf, Rauschholzhausen; Abb. 2) in verschiedenen typischen

Rapsanbauregionen Deutschlands unter zwei kontrastierenden N-Stufen (100 kg N/ha-Differenz) in einem p-Rep-Design angebaut. Die Testhybridversuche wurden nur unregelmäßig und in geringen Umfang beprobt und überflogen.

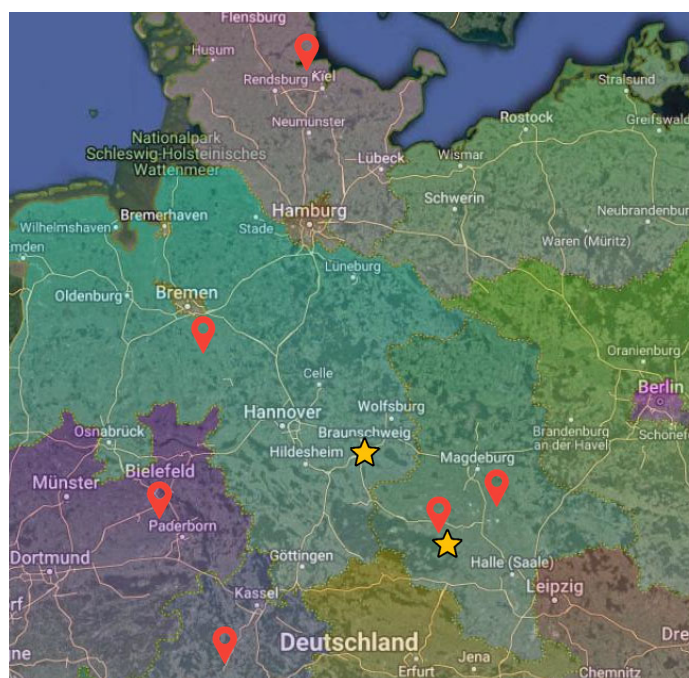


Abb. 2: Raps-Anbaustandorte im Projekt NORA in der Saison 2022/2023. Die Feldversuche der Projektpartner sind in Rot, die N-Steigerungsversuche (QLB und BS) in Gelb dargestellt. Bundesländer sind farbig hinterlegt

2.2 Multispektrale Datenerfassung und photogrammetrische Datenverarbeitung

Als Trägerplattform für die multispektrale Kameratechnik wurden die DJI Matrice 600 (*DJI Innovations Science and Technology Co., Ltd, Shenzhen, Guangdong, China*) (BS) und die DJI Matrice 300 (QLB, BS) genutzt. Zur Bild-Datenerhebung wurde das RedEdge-MX Dual Camera System (*AgEagle Aerial Systems Inc., Wichita, Kansas, USA*) mit 10 spektralen Kanälen (444 – 842 nm, Tab. 1) gewählt.

Tab. 1: Zentrumswellenlängen und Full-Width-Half-Mean (FWHM) der Spektralkanäle des RedEdge-MX Dual Camera Systems

Kanalbezeichnung	Zentrumswellenlänge (nm)	FWHM (nm)
coastal blue	444	28
blue	475	32
green 1	531	14
green 2	560	27
red 1	650	16
red 2	668	14
red edge 1	705	10
red edge 2	717	12
red edge 3	740	18
NIR	842	57

Die Befliegungen an weiteren Teststandorten fanden durch UAV-Systeme der jeweiligen Projektmitglieder statt (Tab. 3). Die Datenerfassung erfolgte, wenn möglich unmittelbar vor der manuellen Bonitur, frühestens jedoch einen Tag zuvor. Die Befliegung fanden immer im Zeitraum von zweieinhalb Stunden vor bzw. zweieinhalb Stunden nach dem lokalen Sonnenhöchststand statt, um die Variation des Sonnenwinkels zu minimieren und Abschattungseffekte so gering wie möglich zu halten. Gleichzeitig wurde versucht die Flüge nicht zum Sonnenhöchststand durchzuführen, um sog. Sunspots in den Bildaufnahmen zu vermeiden. Bei der Flugdurchführung wurde auf eine durchgehend gleichbleibende Beleuchtung geachtet. Vor einem Flug wurde stets ein kalibriertes Reflexionsziel mit der Kamera aufgenommen. Zusätzlich besitzt das genutzte RedEdge-MX Dual Kamerasystem einen Sonneneinstrahlungssensor (Lichtsensor). Für die Durchführung der radiometrischen Kalibrierung wurde, nach der Empfehlung des Herstellers, im Falle eines Fluges bei sonnigen und klaren Bedingungen nur das Reflexionsziel verwendet. Bei bedecktem Himmel oder vollständiger Bewölkung wurden sowohl Lichtsensor- als auch Reflexionszieldaten für die Kalibrierung genutzt. Da keine onlinefähige RTK-Korrektur mit dem benutzten Kamerasystem möglich war, wurden sog. Ground-Control-Points (GCPs) über die beflogenen Flächen verteilt, mittels RTK-Messtechnik eingemessen und für die präzise Georeferenzierung (Positionierungsgenauigkeit vertikal und horizontal < 2cm) genutzt. Weiterhin wurde eine frontale Überlappung von $\geq 80\%$, eine seitliche Überlappung von $\geq 75\%$ und eine Boden-Pixel-Auflösung von ≤ 2 cm für alle Flüge festgelegt. Weitere Flugparameter sind Tab. 2 zu entnehmen.

Tab. 2: Aufnahmeparameter und Ausstattung der Kalibrierungsversuche an den Standorten Quedlinburg und Braunschweig für das Versuchsjahr 2022/2023

Flugparameter	Quedlinburg (JKI-RS)	Braunschweig (JKI-PB)
Frontale Überlappung	80 %	80 %
Seitliche Überlappung	75 %	75 %
Boden-Pixel-Auflösung	≤ 2 cm/px	1.88 cm/px
Flughöhe	30 m	27 m
Anzahl Flüge (vor Blüte)	5	4
Anzahl Flüge (nach Blüte)	6	4
Anzahl GCP	9	17
Trägerplattform	DJI M300	DJI M600*, DJI M300
Sensor	RedEdge-MX Dual Camera System	RedEdge-MX Dual Camera System

*UAV wurde nur für einen Flug verwendet.

Für die Prozessierung der UAV-Daten wurde die Software Pix4DMapper (Version 4.8.0, Pix4D S.A., Route de Renens, 241008 Prilly, Schweiz) verwendet. Ergebnis der Bildzusammensetzung durch Pix4D sind (in der Namensgebung vom Hersteller) „Orthomosaik“ sowie „Reflektanzkarten“ der einzelnen spektralen Bänder. Die Orthomosaik sind nicht radiometrisch korrigiert, sondern erfahren eine Farbbalancierung, welche das Ziel hat, optisch „gutausschende“ Bilder zu generieren. Die Reflektanzkarten hingegen versuchen den Reflektionsgrad der erfassten Szene möglichst real abzubilden. Dabei werden die Nadir-nahen Pixel der Originalbilder im gewichteten Durchschnitt zu einer Gesamtkarte verrechnet. Genauere Informatio-

nen zur Berechnungen werden von Pix4D nicht offengelegt. In Anbetracht der geringen Flughöhe der UAVs und des daher angenommenen begrenzten atmosphärischen Einflusses wurde keine zusätzliche atmosphärische Korrektur durchgeführt. Die gewählte vertikale und horizontale Bildüberlappung (>75%) sicherte, ein stabiles Bildmatching, sowie große Bildbereiche mit einer Nadir-nahen Blickrichtung. Unter Nutzung aller beschriebenen Maßnahmen konnten insgesamt zuverlässige Rekonstruktionen und hochwertige Reflektanzkarten generiert werden.

Tab. 3: Übersicht über Trägerplattformen und Kamerasysteme je Standort im Projekt NORA.

Standort	Trägerplattform	Kamerasystem
Hohenlieth	DJI Phantom 4 RTK	RGB 20 MP
	DJI Matrice 300 RTK	P1 Zenmuse (RGB) 45 MP
Weggegast (Biendorf)	DJI M300*	RedEdge-MX Dual Camera System
Rauschholzhausen	DJI Phantom P4 RTK	5-Kanal Kamerasystem + RGB
	DJI M300*	RedEdge-MX Dual Camera System
Asendorf	DJI Matrice 600	Micasense Altum
Thüle/Salzkotten	DJI Matrice 600	Micasense Altum
Braunschweig	DJI Matrice 600	RedEdge-MX Dual Camera System
	DJI M300	RedEdge-MX Dual Camera System
Quedlinburg	DJI Matrice 300	RedEdge-MX Dual Camera System

2.3 Datentransformation & Modellbildung

Vor der Modellierung durchlaufen die spektralen Reflektanzkarten je Aufnahmezeitpunkt mehrere Vorprozessierungsschritte. Dazu zählen die Generierung eines virtuellen Rasterstapels aus den Reflektanzkarten der einzelnen Wellenlängenkanäle, die einheitliche Benennung der Einzelbanden, die Projektion in ein einheitliches Koordinatensystem (EPSG:32632), das räumliche „Resampling“ auf 2 cm Auflösung, sowie die Berechnung spektraler Indizes (Tab. 4) und die Datenextraktion für die Erstellung einer finalen Datenmatrix. Für die Einzelschritte wurde ein Skript in der Programmiersprache R (Version 4.3.2) erstellt, welches hauptsächlich auf dem Paket „terra“ (Version 1.7-65 (HIJMANS et al. 2023)) basiert.

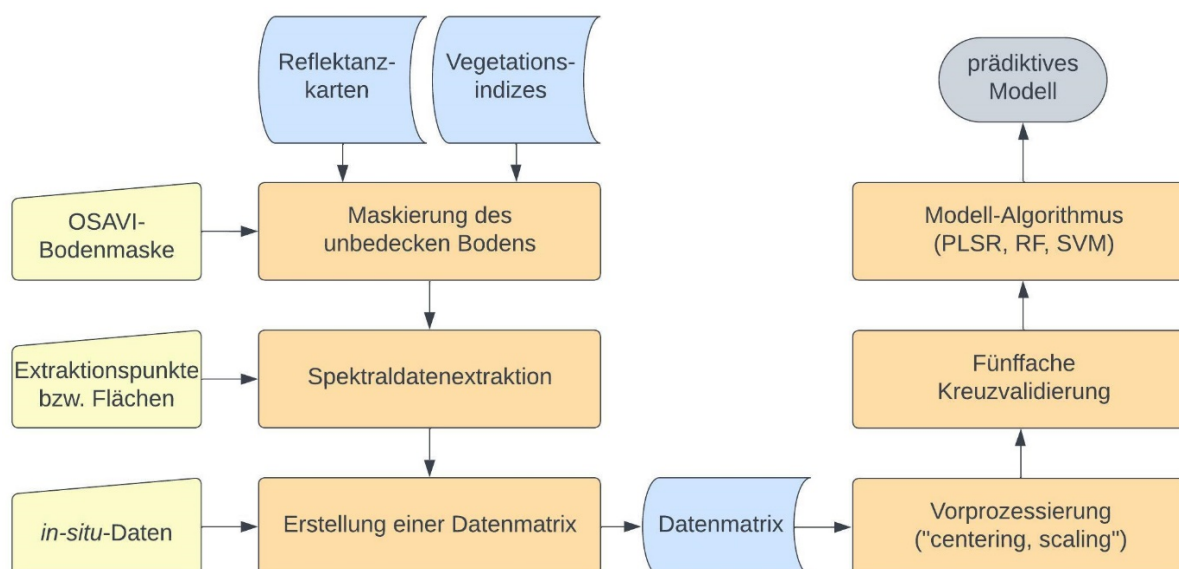


Abb. 3: Schematische Darstellung der Arbeitsschritte zur Generierung eines prädiktiven Modells. Blaue Kästen stellen Basisdatensätze dar, gelbe Kästen externe Daten und orange Kästen Prozesse

Um Referenzdaten zu den *in-situ*-Beprobungen zu extrahieren wurden zunächst zwei verschiedene Sampling-Strategien verfolgt. Ein Ansatz basiert auf einem Punktsampling im Zentrum der tatsächlichen Beprobungseinheit, welcher um einen Puffer (0,5 m Radius) erweitert wurde, ein Weiterer extrahiert den Mittelwert der Spektraldaten der gesamten Parzelle. Als zusätzlicher Zwischenschritt wurde eine auf dem Optimized Soil-Adjusted Vegetation Index (OSAVI, HABOUDANE et al. 2002) basierende Maskierung unbedeckten Bodens (Schwellwert $\leq 0,2$) durchgeführt. Die extrahierten Daten (spektrale Einzelbänder, Indizes) werden in Folgeschritten mit den *in-situ*-Beprobungen verschnitten, um eine Datenmatrix für die Modellierung zu generieren (Abb. 3).

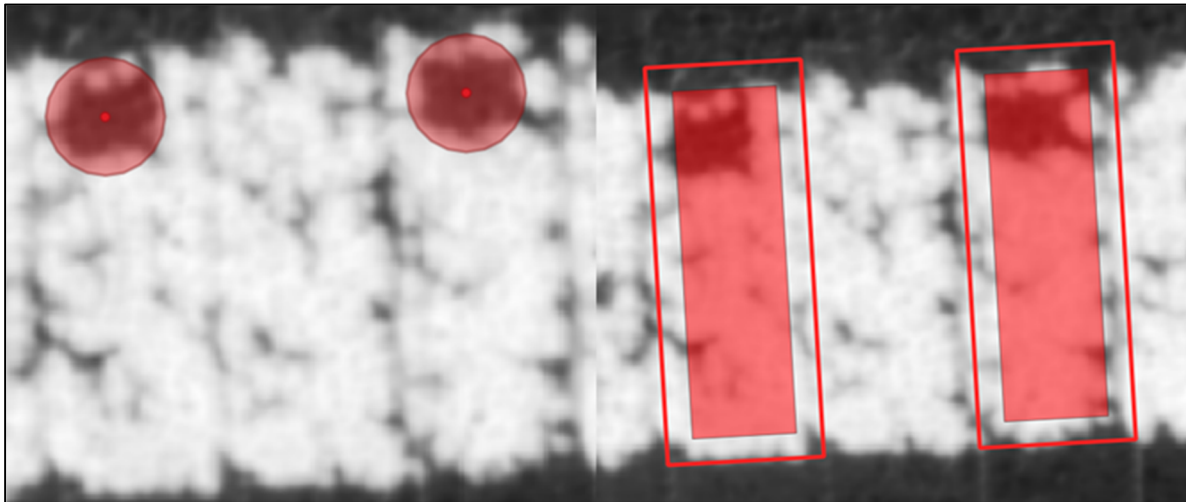


Abb. 4: Gegenüberstellung der Datenextraktionsstrategien. Links: Punktbasierter Ansatz mit 0,5 m Puffer nach außen. Rechts: Parzellenbasierter Ansatz mit 0,25 m Pufferzone nach innen

Ziel der Modellierung ist es mittels multispektralen Eingangsdaten (Einzelbänder, Indizes) und Algorithmen der multivariaten Statistik (Partial Least Squares Regression (PLSR), WOLD et al. 2001) und des maschinellen Lernens (Random Forest (RF), BREIMAN 2001), Support Vector Machine (SVM), EVGENIOU & PONTIL 2001), Modelle zu generieren, mit denen es möglich ist, die Zielvariablen Frischmasse, Trockenmasse und Stickstoffgehalt in Rapspflanzen zu erfassen. Für die Versuche in QLB und BS stehen im Versuchsjahr 2022/2023 *in-situ*-Daten von Frischmasse [t/ha] und Pflanzenstickstoffgehalt [N, in %] zur Verfügung. Daher liegt der Fokus auf diesen Parametern. Aufgrund von noch nicht vorliegenden Ergebnissen der Stickstoffanalytik für den Standort BS liegt der Fokus der hier vorgestellten Modellierung auf dem Standort QLB. Da sich außerdem der Wachstumszeitraum vor der Rapsblüte bezüglich der Stickstoffdynamik als bedeutend darstellt (BOUCHET et al. 2016), wurde dieser für die erste Modellbildung ausgewählt. Insgesamt lagen 149 Datenpunkte im Betrachtungszeitraum vor. Im Zuge der Modellbildung wurde sich aufgrund der relativ geringen Größe des Datensatzes bei jedem der genutzten Verfahren für eine fünffache Kreuzvalidierung entschieden. Des Weiteren wurden Datentransformationen („centering“, „scaling“) und eine Hyperparameteroptimierung mittels Rastersuche („grid search“) durchgeführt. Die Centering- und Scaling-Methoden dienen dazu, die Variablen in Trainingsdaten zu zentrieren und zu skalieren, um eine bessere Vergleichbarkeit und Modellleistung zu gewährleisten, indem sie sicherstellen, dass Variablen mit unterschiedlichen Maßeinheiten oder Skalen gleichmäßig zum Modell beitragen. Die Grid-Search-Methode durchsucht systematisch eine vordefinierte Kombination von Hyperparameter-Werten. Diese einfache und transparente Methode gewährleistet eine umfassende Exploration des

Hyperparameter-Raums, ist jedoch aufgrund ihrer Rechenintensivität nicht immer die effizienteste Wahl. Alle Einzelbänder und Indizes wurden genutzt. Als Fehlerkennwerte dienen der „Root-Mean-Square-Error“ (RMSE) und das Bestimmtheitsmaß (R^2). Die gesamte Modellbildung erfolgte mit dem Paket „caret“ (Version 6.0-94, KUHN 2008) in RStudio.

Tab. 4: Auswahl an spektralen Indizes zur Abschätzung von Biomasse und Chlorophyllgehalt.

Bezeichnung	Abkürzung	Formel	Referenz
Normalized Difference Vegetation Index	NDVI	$\frac{NIR - Red}{NIR + Red}$	ROUSE JR et al. 1973
Normalized Difference Red Edge Index	NDRE	$\frac{NIR - Red\ Edge}{NIR + Red\ Edge}$	BARNES et al. 2000
Improved Modified Chlorophyll Absorption Ratio Index	MCARI	$((NIR - RED) - 0.2 * (NIR - Green)) * \left(\frac{NIR}{Red}\right)$	DAUGHTRY 2000
Ratio Vegetation Index	RVI	$\frac{NIR}{Red}$	XUE et al. 2004
Optimized Soil-Adjusted Vegetation Index	OSAVI	$\frac{NIR - Red}{NIR + Red + 0.16}$	HABOUDANE et al. 2002
Visible Atmospherically Resistant Index	VARI	$\frac{Green - Red}{Green + Red - Blue}$	GITELSON et al. 2002
Red-Edge Chlorophyll Index	CL _{re}	$\left(\frac{NIR}{Red\ Edge}\right) - 1$	GITELSON et al. 2003

3 Ergebnisse

Die Modellanalysen für den Zeitraum vor der Rapsblüte zeigen, dass in Abhängigkeit der Extraktionsmethode teilweise deutliche Unterschiede in den Fehlerkennwerten (Tab. 5) erkennbar sind. Generell ist festzustellen, dass der punktbasierte Datenextraktionsansatz genauere Ergebnisse produziert als die parzellenbasierte Strategie. Mit einem RMSE von 2,01 t/ha weist das punktbasierte RF-Modell zur Schätzung der Frischmasse die geringste Abweichung zu den *in-situ*-Daten (Abb. 5) auf. Das Bestimmtheitsmaß mit 0,97 deutet darauf hin, dass das genutzte Modell einen Großteil der vorhandenen Variabilität erklären kann. Im Vergleich dazu zeichnen sich die SVM- und PLSR-Modelle durch ein geringeres Bestimmtheitsmaß und einen höheren RMSE aus. Die Unterschiede zwischen PLSR und SVM sind gering.

Bezogen auf den Parameter des Pflanzenstickstoffgehaltes lassen sich zahlreiche Parallelen zur Modellierung der Frischmasse erkennen (Tab. 5). So zeigt der punktbasierte RF-Ansatz erneut mit einem RMSE von 0,18 % N-Gehalt und einem Bestimmtheitsmaß von 0,92 die genauesten Ergebnisse. Abb. 6 zeigt, dass das genutzte RF-Modell die höchste Genauigkeit im mittleren Wertebereich aufweist. Bei hohen bzw. niedrigen Stickstoffgehalten ist ein Trend zur Unter- bzw. Überschätzung zu beobachten. Im Gegensatz zur Modellierung der Frischmasse zeigen die weiteren verwendeten Modelle hier jedoch deutlich geringere Genauigkeiten auf. So können die PLSR- und SVM-Modelle ungefähr 40% weniger Varianz des Datensatzes erklären und weisen ungefähr doppelt so hohe Abweichungen zu den Referenzwerten vor. Die präziseren Ergebnisse des Random Forest (RF) können auf eine Überanpassung (overfitting) zurückzuführen sein. Darauf deuten erhebliche Unterschiede in den Fehlerkennwerten im Trainingsdatensatz und im Validierungsdatensatz in der Kreuzvalidierung hin.

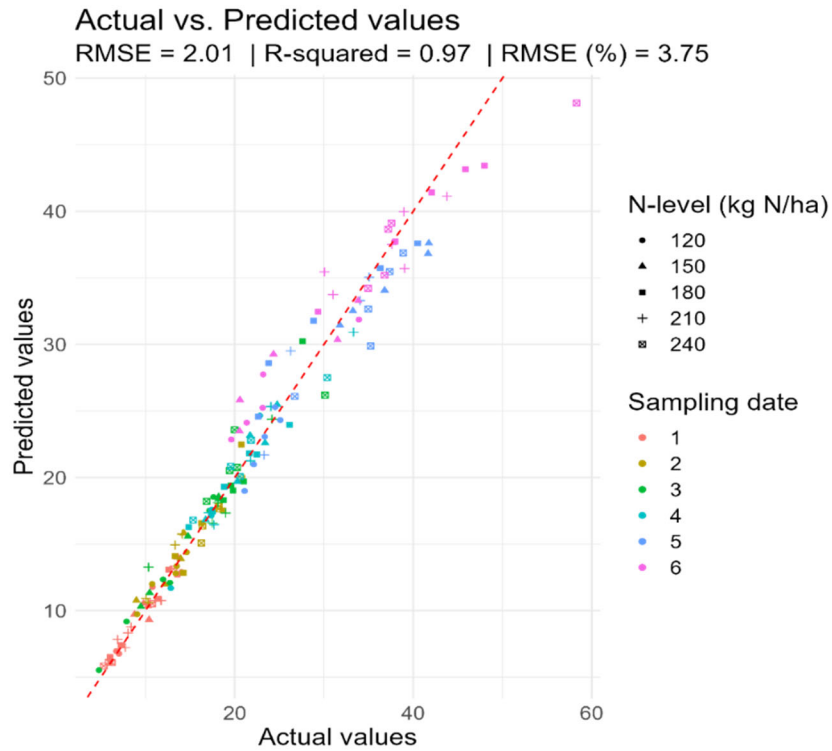


Abb. 5: Vergleich der vorhergesagten und tatsächlichen Frischmasse [t/ha] des RF-Algorithmus basierend auf der punktbasierten Datenextraktion (n = 149). Die Beprobungszeitpunkte wurden durch unterschiedliche Farben kenntlich gemacht. Die N-Stufen über Symbole

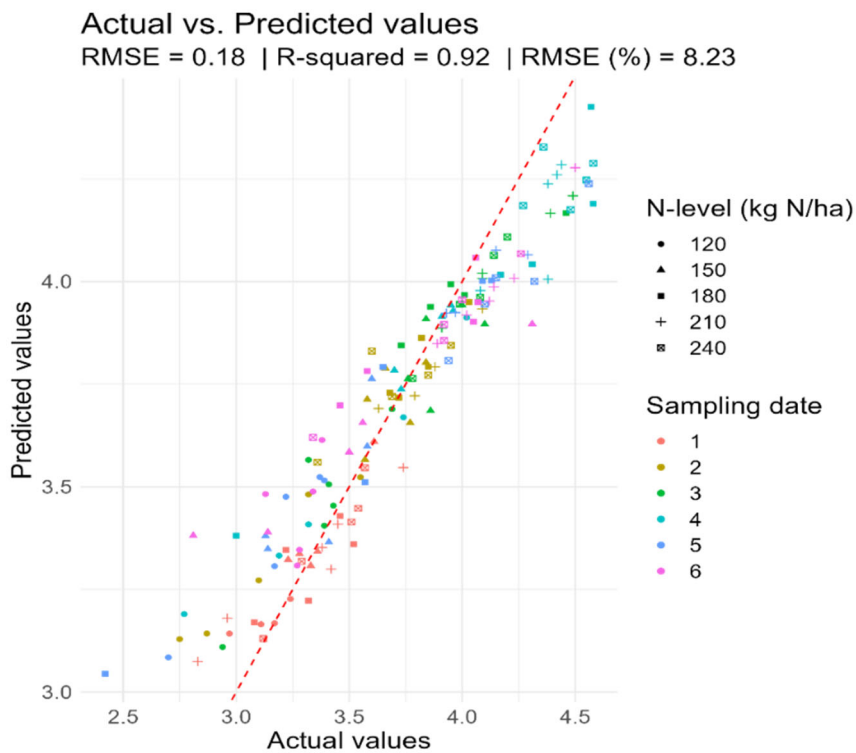


Abb. 6: Vergleich des vorhergesagten und tatsächlichen Pflanzenstickstoffgehalts [N in %] des RF-Algorithmus basierend auf der punktbasierten Datenextraktion (n = 149). Die Beprobungszeitpunkte wurden durch unterschiedliche Farben kenntlich gemacht. Die N-Stufen über Symbole

Tab. 5: Vergleich verschiedener Methoden zur Abschätzung von Frischmasse [t/ha] und Pflanzenstickstoffgehalt [N, in %] im Winterraps am Standort QLB vor der Rapsblüte.

Modell	Methode	Zielvariable	RMSE	RMSE (%)	R ²
PLSR	Punkt ¹	Frischmasse [t/ha]	3,72	6,94	0,88
	Parzelle ²		4,63	8,65	0,81
RF	Punkt		2,01	3,75	0,97
	Parzelle		2,28	4,27	0,96
SVM	Punkt		3,49	6,52	0,9
	Parzelle		3,89	7,26	0,87
PLSR	Punkt	Pflanzenstickstoffgehalt [N, in %]	0,32	14,62	0,54
	Parzelle		0,32	15,92	0,46
RF	Punkt		0,18	8,32	0,92
	Parzelle		0,18	8,52	0,91
SVM	Punkt		0,36	16,44	0,45
	Parzelle		0,39	18,09	0,32

¹Punktbasierte Datenextraktion, ²Parzellenbasierte Datenextraktion

Die im ersten Versuchsjahr erzielten Ergebnisse sind vielversprechend. Dennoch müssen die vorhandenen Modelle im Rahmen einer unabhängigen Validierung auf Stabilität und Übertragbarkeit geprüft werden.

4 Fazit & Ausblick

Die Ergebnisse des ersten Versuchsjahres zeigen, dass UAV-basierte, digitale Technologien ein großes Potenzial für eine Phänotypisierung in der Züchtungsforschung aufweisen. Weitere Analysen sind erforderlich um die Robustheit der Modelle zu sicherzustellen. Hierfür werden im Versuchsjahr 2023/24 erneut UAV-Befliegungen an den Versuchsstandorten im Projekt NORA einschließlich in-situ-Datenerhebungen durchgeführt, um so die vorhandene Datengrundlage zu erweitern, die bestehenden Modelle zu vervollkommen und deren Übertragbarkeit auf verschiedene Umwelten zu prüfen. Ein Schwerpunkt wird zudem auf der Implementierung einer "Feature Selection" (z.B. per „Recursive Feature Elimination“) liegen, wobei die Auswahl der relevanten Merkmale auf ihre Wirksamkeit für den spezifischen Anwendungsfall hin evaluiert wird. Darüber hinaus werden Methoden der Hyperparameteroptimierung (z.B. Bayessche Optimierung, Ensemble Methoden) integriert, um die Effizienz der Algorithmen zu maximieren und eine Überanpassung (overfitting) zu minimieren. Weitere Verfahren aus dem Bereich des Machine Learning (z.B. Gaussian Process) werden den bereits etablierten Modellen vergleichend gegenübergestellt.

5 Literaturverzeichnis

BARNES, E., CLARKE, T., RICHARDS, S., COLAIZZI, P., HABERLAND, J. & KOSTRZEWSKI, M., 2000: Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. <https://api.semanticscholar.org/CorpusID:128773162>, letzter Zugriff 26.01.24.

- BREIMAN, L., 2001: Random Forests. *Machine Learning* **45** (1), 5-32.
- BOUCHET, A. S., LAPERCHE, A., BISSUEL-BELAYGUE, C., SNOWDON, R., NESI, N. & STAHL, A., 2016: Nitrogen use efficiency in rapeseed. A review. *Agron. Sustain. Dev.*, **36**, 38.
- BUNDESMINISTERIUM FÜR ERNÄHRUNG UND LANDWIRTSCHAFT (BMEL), 2022: Landwirtschaft: Tabellen Kapitel C, H.II und H.III des Statistischen Jahrbuchs. <https://www.bmel-statistik.de/landwirtschaft/tabellen-kapitel-c-hii-und-hiii-des-statistischen-jahrbuchs>, letzter Zugriff 26.01.24.
- DAUGHTRY, C., 2000: Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment*, **74**(2), 229-239.
- DI, J. & CAMERON, K., 2002: Nitrate leaching in temperate agroecosystems: sources, factors and mitigating strategies. *Nutrient Cycling in Agroecosystems*, **64**(3), 237-256.
- EVGENIOU, T. & PONTIL, M., 2001: Support Vector Machines: Theory and Applications. *Machine Learning and Its Applications*, **2049**, 249-257.
- GITELSON, A., GRITZ, Y. & MERZLYAK, M., 2003: Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, **160**(3), 271-282.
- GITELSON, A., ZUR, Y., CHIVKUNOVA, O. & MERZLYAK, M., 2002: Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. *Photochemistry and Photobiology*, **75**(3), 272-281.
- HABOUDANE, D., MILLER, J., TREMBLAY, N., ZARCO-TEJADA, P. & DEXTRAZE, L., 2002: Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, **81**(2-3), 416-426.
- HIJMANS, R., BIVAND, R., PEBESMA, E. & SUMMER, M., 2023: terra: Spatial Data Analysis. <https://cran.r-project.org/web/packages/terra/index.html>, letzter Zugriff 26.01.24.
- KUHN, M., 2008: Building Predictive Models in R Using the caret Package. *J. Stat. Soft.*, **28**(5).
- MATHIVANAN, G., EYSHOLDT, M., ZINNBAUER, M., RÖSEMANN, C. & FUB, R., 2021: New N₂O emission factors for crop residues and fertiliser inputs to agricultural soils in Germany. *Agriculture, Ecosystems & Environment*, **322**, 107640.
- ROUSE JR, J. W., HAAS, R. H., SCHELL, J., & DEERING, D., 1973: Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. <https://ntrs.nasa.gov/api/citations/19730017588/downloads/19730017588.pdf>, letzter Zugriff 26.01.24.
- SEBILO, M., MAYER, B., NICOLARDOT, B., PINAY, G., & MARIOTTI, A., 2013: Long-term fate of nitrate fertilizer in agricultural soils. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(45), 18185-18189.
- SINCLAIR, T. & RUFTY, T., 2012: Nitrogen and water resources commonly limit crop yield increases, not necessarily plant genetics. *Global Food Security*, **1**(2), 94-98.
- STATISTISCHES BUNDESAMT (DESTATIS), 2024: Anbaufläche, Erntemenge, Ertrag je Hektar (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten. <https://www-genesis.destatis.de/>, letzter Zugriff 26.01.24.
- TILMAN, D., CASSMAN, K., MATSON, P., NAYLOR, R. & POLASKY, S., 2002: Agricultural sustainability and intensive production practices. *Nature*, **418**(6898), 671-677.
- USDA'S FOREIGN AGRICULTURAL SERVICE, 2023: Rapeseed 2023: World Production. USDA's Foreign Agricultural Service. <https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=2226000>, letzter Zugriff 26.01.2024.

- WANG, L., LU, P., REN, T., LIU, T., GEILFUS, C. & LU, J., 2020: Improved nitrogen efficiency in winter oilseed rape hybrid compared with the parental lines under contrasting nitrogen supply. *Industrial Crops and Products*, **155**, 112777.
- WOLD, S., SJÖSTRÖM, M. & ERIKSSON, L., 2001: PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**(2), 109-130.
- XUE, L., CAO, W., LUO, W., DAI, T. & ZHU, Y., 2004: Monitoring Leaf Nitrogen Status in Rice with Canopy Spectral Reflectance. *Agron.j.*, **96**(1), 135-142.
- ZHAN, N., XU, K., JI, G., YAN, G., CHEN, B., WU, X. & CAI, G., 2023: Research Progress in High-Efficiency Utilization of Nitrogen in Rapeseed. *International Journal of Molecular Sciences*, **24**(9).