

Elementprofil und NIR-Spektroskopie kombiniert mit  
maschinellern Lernen zur Herkunftsbestimmung von Spargel

## **Dissertation**

zur Erlangung des akademischen Grades

*Doctor rerum naturalium*

an der Universität Hamburg

Institut für Lebensmittelchemie, Fachbereich Chemie

Fakultät für Mathematik, Informatik und Naturwissenschaften

vorgelegt von

Bernadette Richter

Hamburg, 2020



Erstgutachter: Prof. Dr. Markus Fischer

Zweitgutachter: Prof. Dr. Bernward Bisping

Disputation & Druckfreigabe: 03.07.2020

Mitglieder der Prüfungskommission:

Prof. Dr. Markus Fischer (Vorsitz)

Prof. Dr. José A. C. Broekaert

Dr. Thomas Hackl



Die vorliegende Arbeit wurde vom 01.06.2015 bis 31.05.2018  
unter der Leitung von Herrn Prof. Dr. Markus Fischer am Fachbereich Chemie,  
Institut für Lebensmittelchemie, der Universität Hamburg angefertigt.



# Publikationen

## Publikationen zur Herkunftsbestimmung von Spargel

Richter, B., Gurk, S., Wagner, D., Bockmayr, M., Fischer, M., **2019**, Food Authentication: Multi-elemental analysis of white asparagus for provenance discrimination. *Food Chemistry*, 286, 475–482. doi: 10.1016/j.foodchem.2019.01.105

Richter, B.\*, Rurik, M.\*, Gurk, S., Kohlbacher, O., Fischer, M., **2019**, Food monitoring: Screening of the geographical origin of white asparagus using FT-NIR and machine learning. *Food Control*, 104, 318–325. doi: 10.1016/j.foodcont.2019.04.032

\* gemeinsame Erstautoren

## weitere Publikationen

Grüner, F., Blumendorf, F., Schmutzler, O., Staufer, T., Bradbury, M., Wiesner, U., Rosentreter, T., Loers, G., Lutz, D., Richter, B., Fischer, M., Schulz, F., Steiner, S., Warmer, M., Burkhardt, A., Meents, A., Kupinski, M., Hoeschen, C., **2018**, Localising functionalised gold-nanoparticles in murine spinal cords by X-ray fluorescence imaging and background-reduction through spatial filtering for human-sized objects. *Scientific Reports*, 8(1), 16561. doi: 10.1038/s41598-018-34925-3

Gu, F., Borewicz, K., Richter, B., van der Zaal, P. H., Smidt, H., Buwalda, P.L., Schols, H. A., **2018**, In Vitro Fermentation Behavior of Isomalto/Malto-Polysaccharides Using Human Fecal Inoculum Indicates Prebiotic Potential. *Molecular Nutrition & Food Research*, 62(12), 1800232. doi: 10.1002/mnfr.201800232

Thielke, M.; Bultema, L.; Brauer, D.; Richter, B.; Fischer, M.; Theato, P., **2016**, Rapid Mercury(II) Removal by Electrospun Sulfur Copolymers. *Polymers*, 8(7), 266. doi: 10.3390/polym8070266



# Inhaltsverzeichnis

Publikationen	I
Abkürzungsverzeichnis	VII
Glossar	XI
Zusammenfassung	1
Abstract	3
1 Einleitung	5
2 Theoretische Grundlagen	7
2.1 Spargel . . . . .	7
2.1.1 Historie und Botanik . . . . .	7
2.1.2 Bedeutung und physiologische Wirkung . . . . .	9
2.1.3 Anbau . . . . .	11
2.1.4 Produktion in Deutschland und weltweit . . . . .	14
2.1.5 Bislang verwendete Strategien zur Herkunftsbestimmung . . . . .	16
2.2 Omics-Techniken . . . . .	19
2.2.1 Isotopolomics . . . . .	19
2.2.2 Metabolomics . . . . .	20
2.3 ICP-MS . . . . .	22
2.3.1 Probeneintrag . . . . .	22
2.3.2 Ionisierung . . . . .	25
2.3.3 Ionentransport . . . . .	27
2.3.4 Kollisions- und Reaktionszelle . . . . .	28

2.3.5	Massenseparation . . . . .	29
2.3.6	Detektion . . . . .	30
2.3.7	Interferenzen . . . . .	32
2.4	FT-NIR . . . . .	34
2.4.1	Grundlagen der NIR-Spektroskopie . . . . .	34
2.4.2	Aufbau eines FT-NIR-Gerätes . . . . .	35
2.4.3	Datenpräprozessierung . . . . .	37
2.5	Multivariate Datenanalyse . . . . .	39
2.5.1	Hierarchische Clusteranalyse und Heatmaps . . . . .	39
2.5.2	Hauptkomponentenanalyse . . . . .	40
2.5.3	Partial Least Squares Discriminant Analysis . . . . .	41
2.5.4	t-SNE . . . . .	41
2.5.5	Feature Selection mithilfe des Boruta-Algorithmus . . . . .	42
2.5.6	Support Vector Machine . . . . .	43
2.5.7	Random Forest . . . . .	44
2.5.8	Kreuzvalidierung . . . . .	45
<b>3</b>	<b>Zielsetzung</b>	<b>49</b>
<b>4</b>	<b>Ergebnisse und Diskussion</b>	<b>51</b>
4.1	Methodenentwicklung für die Multielementanalytik . . . . .	51
4.1.1	Optimierung der Aufschlussmethode . . . . .	51
4.1.2	Entwicklung der Multielementmethode . . . . .	52
4.1.3	Methodenvalidierung . . . . .	58
4.2	Herkunftsbestimmung anhand des Elementprofils . . . . .	63
4.2.1	Datenpräprozessierung . . . . .	63
4.2.2	Probenauswahl . . . . .	64
4.2.3	Datenvisualisierung . . . . .	65
4.2.4	Identifizierung herkunftsspezifischer Elemente . . . . .	72
4.2.5	Klassifikation zur Unterscheidung des Herkunftslandes . . . . .	79
4.2.6	Klassifikation deutscher und importierter Produkte . . . . .	83
4.2.7	Innerdeutsche Probenklassifikation . . . . .	84

4.2.8	Klassifikation zur Standortunterscheidung . . . . .	87
4.2.9	Untersuchung von Einflussfaktoren auf die Herkunfts- bestimmung . . . . .	89
4.3	Vergleich des Elementprofils von Spargel und Boden . . . . .	93
4.3.1	Korrelationen zwischen den Elementen . . . . .	93
4.3.2	Korrelationen der seltenen Erden . . . . .	98
4.3.3	Korrelationen zwischen Spargel und Boden . . . . .	101
4.4	Herkunftsbestimmung mittels FT-NIR . . . . .	105
4.4.1	Datenpräprozessierung . . . . .	105
4.4.2	Probenauswahl . . . . .	107
4.4.3	Datenvisualisierung . . . . .	108
4.4.4	Klassifikation zur Unterscheidung des Herkunftslandes . . . . .	112
4.4.5	Klassifikation deutscher und importierter Produkte . . . . .	116
4.4.6	Eingrenzung des Konfidenzbereiches der Klassifikation . . . . .	118
4.4.7	Untersuchung von Einflussfaktoren auf die Herkunftsbestimmung	124
4.5	Vergleich von ICP-MS- und FT-NIR-basierter Herkunftsbestimmung . .	130
<b>5</b>	<b>Ausblick</b>	<b>133</b>
<b>6</b>	<b>Experimenteller Teil</b>	<b>137</b>
6.1	Probenmaterial . . . . .	137
6.1.1	Akquise der Spargelproben . . . . .	137
6.1.2	Akquise der Bodenproben . . . . .	138
6.1.3	Vorbereitung der Spargelproben . . . . .	138
6.1.4	Vorbereitung der Bodenproben . . . . .	139
6.2	Probenaufschlüsse . . . . .	140
6.2.1	Initialer Aufschluss für pflanzliche Proben . . . . .	140
6.2.2	Aufschluss der Bodenproben . . . . .	140
6.2.3	Reinigungsaufschluss und Gefäßreinigung . . . . .	141
6.3	Messung mittels ICP-MS . . . . .	142
6.3.1	Gerätekalibration . . . . .	142
6.3.2	Messparameter der Multielementmethode . . . . .	143
6.3.3	Externe Kalibrierung . . . . .	144

6.3.4	Qualitätskontrolle . . . . .	145
6.3.5	Blindwert . . . . .	146
6.3.6	Statistische Datenauswertung . . . . .	146
6.4	Messung mittels FT-NIR . . . . .	147
6.4.1	Messparameter . . . . .	147
6.4.2	Statistische Datenauswertung . . . . .	148
6.5	Software . . . . .	148
<b>7</b>	<b>Literatur</b>	<b>149</b>
<b>8</b>	<b>Anhang</b>	<b>175</b>
8.1	Akquirierte Spargelproben . . . . .	175
8.2	Akquirierte Bodenproben . . . . .	186
8.3	Messparameter ICP-MS . . . . .	190
8.4	Kalibrierstandards . . . . .	192
8.5	Geräteoptimierung ICP-MS . . . . .	194
8.6	Methodenvalidierung ICP-MS . . . . .	195
8.7	Konzentrationen der erfassten Elemente . . . . .	202
8.8	Elementkorrelationen . . . . .	206
8.9	FT-NIR-Spektren . . . . .	208
8.10	Klassifikation anhand von FT-NIR-Daten . . . . .	210
8.11	Formelübersicht . . . . .	210
8.12	Chemikalienübersicht . . . . .	211
8.13	Geräteübersicht . . . . .	213
8.14	Verbrauchsmaterialien . . . . .	214
	<b>Abbildungsverzeichnis</b>	<b>215</b>
	<b>Tabellenverzeichnis</b>	<b>219</b>
	<b>Danksagung</b>	<b>223</b>
	<b>Eidesstattliche Versicherung</b>	<b>225</b>

# Abkürzungsverzeichnis

<b>ANOVA</b>	<i>Analysis of Variance</i> ; Varianzanalyse
<b>Av</b>	Avalim; Spargelsorte
<b>Ba</b>	Backlim; Spargelsorte
<b>BB</b>	Brandenburg
<b>BW</b>	Baden-Württemberg
<b>BWN</b>	Baden-Württemberg – Nord
<b>BWS</b>	Baden-Württemberg – Süd
<b>BYAb</b>	Bayern – Abensberg
<b>BYFr</b>	Bayern – Franken
<b>BYSr</b>	Bayern – Schrobenhausen
<b>CHN</b>	China
<b>cps</b>	<i>counts per second</i> ; Counts pro Sekunde
<b>Cu</b>	Cumulus; Spargelsorte
<b>DE</b>	Deutschland
<b>EU</b>	Europäische Union
<b>ES</b>	Spanien
<b>FDA</b>	<i>Food and Drug Administration</i>
<b>FT-NIR</b>	<i>Fourier-Transform Near-Infrared Spectroscopy</i> ; Fourier-Transform-Nahinfrarotspektroskopie
<b>GC</b>	<i>Gas Chromatography</i> ; Gaschromatographie
<b>GC-MS</b>	<i>Gas Chromatography-Mass Spectrometry</i> ; Gaschromatographie mit Massenspektrometrie-Kopplung
<b>Gi</b>	Gijnlim; Spargelsorte

<b>ggA</b>	geschützte geographische Angabe
<b>GHS</b>	<i>Globally Harmonised System</i> ; Global harmonisiertes System zur Einstufung und Kennzeichnung von Chemikalien
<b>Gr</b>	Grolim; Spargelsorte
<b>GR</b>	Griechenland
<b>HCA</b>	<i>Hierarchical Cluster Analysis</i> ; hierarchische Clusteranalyse
<b>He</b>	Herkolim; Spargelsorte
<b>HE</b>	Hessen
<b>HMI</b>	<i>High Matrix Introduction</i>
<b>HPLC</b>	<i>High Performance Liquid Chromatography</i> ; Hochleistungs-Flüssigchromatographie
<b>ICP-MS</b>	<i>Inductively Coupled Plasma Mass Spectrometry</i> ; Massenspektrometrie mit induktiv gekoppeltem Plasma
<b>ICP-OES</b>	<i>Inductively Coupled Plasma Optical Emission Spectrometry</i> ; optische Emissionsspektrometrie mit induktiv gekoppeltem Plasma
<b>IR-MS</b>	<i>Isotope Ratio Mass Spectrometry</i> ; Isotopenverhältnis-Massenspektrometrie/Stabilisotopenanalyse
<b>ISTD</b>	Interner Standard
<b>IUPAC</b>	<i>International Union of Pure and Applied Chemistry</i>
<b>Konz.</b>	Konzentration
<b>LA</b>	<i>Laser Ablation</i> ; Laserablation
<b>LC-MS</b>	<i>Liquid Chromatography-Mass Spectrometry</i> ; Flüssigchromatographie mit Massenspektrometrie-Kopplung
<b>LDA</b>	<i>Linear Discriminant Analysis</i> ; lineare Diskriminanzanalyse
<b>LOD</b>	<i>Limit of Detection</i> ; Nachweisgrenze
<b>LOQ</b>	<i>Limit of Quantification</i> ; Bestimmungsgrenze
<b>MIR</b>	<i>mid-infrared</i> ; mittleres Infrarot
<b>MSC</b>	<i>Multiple Scatter Correction</i> ; Streulichtkorrektur
<b>m/z</b>	Masse-zu-Ladungsverhältnis
<b>N</b>	Norden
<b>NA</b>	<i>not available</i> ; Angabe nicht verfügbar

---

<b>NI</b>	Niedersachsen
<b>NIR</b>	<i>near-infrared</i> ; nahes Infrarot
<b>NL</b>	Niederlande
<b>NRW</b>	Nordrhein-Westfalen
<b>NoGas</b>	<i>NoGas</i> -Modus (ohne Aktivierung der Kollisionszelle)
<b>NW</b>	Nordrhein-Westfalen
<b>NWBo</b>	Nordrhein-Westfalen – Bornheim
<b>NWWa</b>	Nordrhein-Westfalen – Walbeck
<b>O</b>	Osten
<b>p.a.</b>	<i>pro analysi</i> ; analysenrein
<b>PC</b>	<i>Principal Component</i> ; Hauptkomponente
<b>PCA</b>	<i>Principal Component Analysis</i> ; Hauptkomponentenanalyse
<b>PE</b>	Peru
<b>PL</b>	Polen
<b>PLS</b>	<i>Partial Least Squares</i>
<b>PLS-DA</b>	<i>Partial Least Squares Discriminant Analysis</i>
<b>ppb</b>	<i>parts per billion</i> ; ein Milliardstel ( $10^{-9}$ )
<b>QC</b>	<i>Quality Control</i> ; Qualitätskontrolle
<b>R<sup>2</sup></b>	Bestimmtheitsmaß
<b>RBF</b>	<i>Radial Basis Function/Gauß Kernel</i> ; Gauß-Kern
<b>REE</b>	<i>Rare Earth Elements</i> ; seltene Erden
<b>RF</b>	Random Forest
<b>Rfl</b>	Raffaello; Spargelsorte
<b>Rmd</b>	Ramada; Spargelsorte
<b>Rmr</b>	Ramires; Spargelsorte
<b>r</b>	Korrelationskoeffizient
<b>r<sub>P</sub></b>	Korrelationskoeffizient nach Pearson
<b>RP</b>	Rheinland-Pfalz

<b>Rps</b>	Rapsody; Spargelsorte
<b>Rvl</b>	Ravel; Spargelsorte
<b>S</b>	Süden
<b>SEV</b>	Sekundärelektronenvervielfacher
<b>SH</b>	Schleswig-Holstein
<b>SNV</b>	<i>Standard Normal Variate</i> ; Normalisierungsmethode
<b>SO</b>	Südosten
<b>SW</b>	Südwesten
<b>S/C</b>	<i>Spray Chamber</i> ; Sprühkammer
<b>SVM</b>	<i>Support Vector Machine</i>
<b>TDS</b>	<i>Total Dissolved Solids</i> ; gelöste Feststoffpartikel
<b>TOF</b>	<i>Time of Flight</i> ; Flugzeit-Massenseparator
<b>VO</b>	Verordnung
<b>W</b>	Westen

Des Weiteren wurden das internationale Einheitensystem (SI-Einheiten) sowie die Symbole des Periodensystems der Elemente laut Veröffentlichung der IUPAC verwendet [1].

# Glossar

## Beeswarm-Plot

Eindimensionale Darstellung der Verteilung einer Größe, bei der jeder Datenpunkt einzeln gezeigt wird. Die Aneinanderreihung mehrerer Datenpunkte bei einem Messwert verleiht dem Plot ein Bienenschwarm-ähnliches Aussehen und ermöglicht eine Beurteilung der Verteilung aller Datenpunkte [2].

*Abbildungen 28, 59 und 60*

## Bonferroni-Korrektur

Verfahren zur Vermeidung von falsch-positiven Ergebnissen beim multiplen Testen (Testen mehrerer Hypothesen auf demselben Datensatz). Es erfolgt eine Multiplikation der p-Werte mit der Zahl der getesteten Hypothesen [3].

*Kapitel 4.2.4*

## Boruta-Algorithmus

Methode zur Identifizierung der für eine spezielle Fragestellung relevanten Messgrößen innerhalb eines Datensatzes: *Feature Selection*. Algorithmus prüft Gewichtung anhand des Vergleichs eines Messwertes mit einem zufällig kombinierten Messwert.

*Kapitel 2.5.5, 4.4.3 und 4.4.6*

## Complete Linkage

Algorithmus zur Berechnung des Abstands zwischen den Clustern bei hierarchischen Clusterverfahren. Bei der Methode des *Complete Linkage* wird der größtmögliche Abstand zwischen einem Objektpaar aus zwei Clustern berechnet [4].

*Kapitel 2.5.1*

## Heatmap

Matrix zur farblichen Darstellung der Messwerte eines Datensatzes basierend auf einem vorher festgelegten Farbverlauf.

*Kapitel 2.5.1 sowie Abbildungen 24 - 27 sowie 34, 35 und 37*

## Hierarchische Clusteranalyse (HCA)

Distanzbasiertes Verfahren zur Gruppierung von Datenpunkten. Eine Gruppierung erfolgt durch die Bildung von Clustern bestehend aus ähnlichen Objekten.

*Abbildungen 24 - 27, 30, 32 und 33 sowie 49, 52 und 54*

## Hybride

Bezeichnet im biologischen Sinne ein Individuum, welches aus reinerbigen genetisch weit entfernten Elternorganismen durch Kreuzung entstanden ist. Hybridsorten sind deutlich widerstandsfähiger und ertragreicher [5, 6].

*Kapitel 2.1.1*

## Konfusionsmatrix

Matrix zur Gegenüberstellung von tatsächlichen und vorhergesagten Ergebnissen bei der Klassifikation mit Übereinstimmung von wahren und vorhergesagten Zuordnungen auf der Diagonalen.

In der vorliegenden Arbeit wird die Zuordnung zum tatsächlichen, wahren Ergebnis durch eine Beschriftung am linken Rand angegeben, während die mithilfe des Klassifikators vorhergesagten Ergebnisse mit der Beschriftung am oberen Rand gekennzeichnet sind.

*Kapitel 4.2 und 4.4*

## Korrelationskoeffizient nach Pearson ( $r_P$ )

Maß zur Beschreibung des linearen Zusammenhangs zwischen zwei Variablen;

$r_P > 0$ : positive Korrelation;  $r_P \approx 0$ : keine Korrelation;  $r_P < 0$ : negative Korrelation

*Kapitel 4.3, Abbildungen 34, 35 und 37 sowie Tabelle 10*

## Kruskal-Wallis-Test

Statistischer Test für mehrere unabhängige Stichproben zur Überprüfung, ob sich zentrale Tendenzen unterscheiden. Voraussetzung ist das Vorliegen einer mindestens ordinalskalierten abhängigen Variable und einer unabhängigen Variable, die der Gruppenbildung dient [7].

*Kapitel 4.2.4*

## Maschinelles Lernen

Teilbereich der künstlichen Intelligenz; Algorithmen erlernen das Erkennen von Strukturen auf der Grundlage eines Datensatzes, um Vorhersagen auch basierend auf neuen Datensätzen zu ermöglichen.

## Nested cross-validation

Erweiterung der Kreuzvalidierung durch Hinzunahme einer äußeren Testschleife. Die herkömmliche Kreuzvalidierung (innere Schleife) mit einem Großteil der Daten dient der Parameteroptimierung, während ein anderer Teil des Datensatzes zur unabhängigen äußeren Testung herangezogen wird.

*Kapitel 2.5.8, 4.2 und 4.4 sowie Abbildung 17*

## Random Forest (RF)

Überwachtes Verfahren zur Klassifikation bestehend aus einer Sammlung von Entscheidungsbäumen. Das Modell lernt die bekannten Klassenzuordnungen anhand der Entscheidungsbäume, um dies zur Einordnung unbekannter Objekte anzuwenden.

*Kapitel 2.5.7 und 4.4.4*

## RBF-Kern/Gauß-Kern

*Radial Basis Function*; mögliche Kern-Funktion der SVM zur Berechnung der Stützvektoren, welche die Hyperebene zur Klassentrennung bestimmen.

*Kapitel 2.5.6, 4.2 und 4.4*

## Support Vector Machine (SVM)

Überwachtes Klassifikationsverfahren zur nicht-linearen Klassentrennung. Es erfolgt eine Übertragung der Daten in einen höherdimensionalen Raum, in dem eine lineare Trennung mittels Hyperebene möglich ist.

*Kapitel 2.5.6, 4.2 und 4.4*

## t-distributed stochastic neighbor embedding (t-SNE)

Verfahren zur 2D-Visualisierung hochdimensionaler Datensätze durch eine nicht-lineare Dimensionalitätsreduktion.

*Kapitel 2.5.4, 4.2.3 und Abbildung 22*

## Varianzanalyse (ANOVA)

Statistisches Verfahren zum Testen auf Mittelwertunterschiede bei mehreren unabhängigen Stichproben. Voraussetzungen sind neben einer intervallskalierten abhängigen und einer nominal- oder ordinalskalierten unabhängigen Variablen auch Normalverteilung und Varianzhomogenität der Stichproben [8].

*Kapitel 4.2.4*

## Weighted Average Linkage

Algorithmus zur Berechnung des Abstands zwischen den Clustern bei hierarchischen Clusterverfahren. Beim gewichteten Mittelwertverfahren wird der mittlere Abstand aller Objektpaare aus zwei Clustern berechnet [4].

*Kapitel 2.5.1*

## Welch-Test

Signifikanztest zum Prüfen auf signifikante Unterschiede zwischen den Mittelwerten zweier Stichproben; Abwandlung des t-Tests zur Anwendung bei Varianzinhomogenität der Stichproben [9].

*Kapitel 4.2.4*

# Zusammenfassung

Die geographische Herkunft zählt zu den wichtigsten Qualitätsparametern von weißem Spargel und ist maßgeblich für die Kaufentscheidung der Verbraucher. Die Verfälschung von Lebensmitteln, darunter absichtlich fehlerhafte Herkunftsdeklarationen, stellen ein gängiges Problem innerhalb des globalisierten Handels dar. Dies hat eine Täuschung der Verbraucher sowie wirtschaftliche Schäden für Mitbewerber zur Folge, daher werden analytische Methoden zur Überprüfung der deklarierten Herkunft benötigt.

Ziel dieser Arbeit war die Entwicklung einer Methode zur Herkunftsbestimmung von weißem Spargel. Als analytische Ansätze wurden FT-NIR-Messungen sowie eine Untersuchung des Elementprofils mittels ICP-MS gewählt. Für die Methodenentwicklung standen insgesamt 319 Spargelproben aus 7 Ländern (Deutschland, Polen, den Niederlanden, Griechenland, Spanien, China und Peru) und vier Erntejahren zur Verfügung. Um die geographische Herkunft vorherzusagen, wurden *Support Vector Machine*- (SVM) und *Random Forest*-Klassifikatoren als Methoden des maschinellen Lernens eingesetzt.

Ein Klassifikationsmodell zur Unterscheidung des Herkunftslandes basierend auf NIR-Daten liefert eine Genauigkeit von 89.5% bei 5 Klassen. Die Vorhersage des Herkunftslandes anhand des Elementprofils erzielt eine etwas höhere Genauigkeit von 91.2% bei Berücksichtigung zweier zusätzlicher Klassen. Bei beiden analytischen Ansätzen wird durch die Abschätzung des Vertrauensbereiches eine deutliche Verbesserung der Genauigkeit bei der Vorhersage des Herkunftslandes erreicht. Bei diesem Vorgehen werden nur Proben mit hohen SVM-Vorhersagewerten berücksichtigt, was zu Genauigkeiten von bis zu 98% führt. Zudem ermöglichen die Elementdaten eine noch detailliertere Einordnung der geographischen Herkunft: die Unterscheidung von fünf deutschen Regionen mit richtiger Vorhersage für 83.6% der Proben sowie die Zuordnung des Produktionsstandortes mit einer Genauigkeit von 82.6% (13 Klassen), die die hohe Stabilität des Elementprofils über mehrere Erntejahre unterstreicht. Der Vergleich der Elementgehalte von Spargel- und Bodenproben desselben Feldes zeigt für die meisten Elemente starke Korrelationen zwischen beiden Matrices.

Sowohl NIR- als auch ICP-MS-Daten ermöglichen das Erstellen jahresübergreifender Klassifikationsmodelle. Die NIR-basierte Methode bietet schnelle und kostengünstige Messungen, die sich hervorragend für *Screening*-Anwendungen eignen. Darüber hinaus erzielt eine elementbasierte Analyse sehr präzise und robuste Vorhersagen ohne erkennbaren Einfluss des Erntejahres, anhand derer deutliche Verbesserungen der Herkunftsüberwachung in der Routineanalytik erreicht werden können. Die Untersuchung des Isotopoloms eignet sich wegen der exakten Zuordnung zum Produktionsstandort besonders für die Überprüfung von produktbegleitenden Herkunftsangaben. Die Ergebnisse unterstreichen das große Potential von Methoden des maschinellen Lernens für eine leistungsstarke und zuverlässige Untersuchung der geographischen Herkunft von Lebensmitteln.



## Abstract

The geographic origin is an important quality parameter of white asparagus and considered as decisive for the customers' purchase decision. However, adulteration of foodstuff including the intentional falsification of the declared origin is a common problem of globalized supply chains. These practices result in customer deception and economic damage for competitors. Thus, analytical methods are required for reviewing the declared product provenance.

The aim of this work was to develop a method to determine the geographic origin of white asparagus. This issue was addressed by FT-NIR measurements and an investigation of the elemental profile using ICP-MS. In total, 319 asparagus samples from 7 countries (Germany, Poland, the Netherlands, Greece, Spain, China and Peru) were involved in this study, including four years of harvest. To predict the geographic origin, support vector machine (SVM) and random forest classifiers were applied.

A classification model using NIR spectroscopic data enables the distinction of the country of origin, that is correctly predicted for 89.5 % of the considered samples within 5 classes. Based on the elemental profile, the country of origin is classified with a slightly increased accuracy of 91.2 % considering two additional classes. For both analytical techniques, a considerable increase in accuracy is achieved for the prediction of the country of origin when applying a confidence estimation. This approach only considers confidential samples with high SVM prediction scores, resulting in accuracies of up to 98 %. Furthermore, a more detailed prediction of the geographic origin is obtained by means of elemental data, that enable the distinction of five German regions with an accuracy of 83.6 %. A further classification model considering the prediction of the cultivation site achieves an accuracy of 82.6 % including 13 classes and reveals high stability of the elemental profile over several harvest years. Besides, the comparison of element quantities of asparagus and soil samples acquired from the same field displays strong correlations of most elements in both matrices.

FT-NIR and ICP-MS data allow for the construction of classification models that combine samples from several harvest years. The NIR based method provides fast and cost-effective measurements well suited for screening applications. Moreover, the analysis of the elemental profile enables very precise and robust predictions of the geographic origin without a noticeable influence of harvest year and can therefore significantly improve the monitoring of provenance declarations in routine analytics. The study of the isotopologue is particularly well suited for the verification of given geographical indications as it enables the exact characterization of the production site. The results underline the great potential of machine learning techniques for reliable and high-performance investigations of the geographic origin of foodstuffs.



# 1 Einleitung

Die Verfälschung von Lebensmitteln hat eine lange Geschichte und gehört auch heute zum Alltag des globalisierten Handels. Beim Lebensmittelbetrug spricht man von vorsätzlichem Handeln aus finanziellen Beweggründen, meist zur Gewinnmaximierung durch geringere Produktionskosten, was auch unter dem Stichwort *Food Fraud* zusammengefasst wird. Produktfälschungen werden häufig angetrieben durch plötzlich auftretende Preissteigerungen, z. B. aufgrund von Ernteaussfällen und Lieferengpässen. Die Praktiken der Verfälschung sind vielseitig und beinhalten u.a. die Verwendung anderer als der angegebenen Inhaltsstoffe, fehlerhafte Produktdeklarationen, z. B. die Herkunfts- oder Sortenangabe, oder die Nichteinhaltung von Qualitäts- und Reinheitsvorgaben.

Lebensmittelfälschungen täuschen den Verbraucher, was in einigen Fällen auch mit einer Gesundheitsgefahr verbunden sein kann. Zudem entstehen enorme wirtschaftliche Schäden, beispielsweise bei Mitbewerbern, die ihre rechtskonform hergestellten Produkte mit größeren Kosten produzieren und daher zu höheren Preisen anbieten müssen, sodass ein Wettbewerbsnachteil entsteht. Die Europäische Kommission schätzt den wirtschaftlichen Schaden für die Lebensmittelindustrie auf jährlich 8-10 Milliarden Euro [10].

Eine Überprüfung der auf dem Produkt angegebenen Deklaration anhand von Dokumenten innerhalb der Lieferkette hat sich als manipulierbar erwiesen, da auch diese oftmals gefälscht werden. Daher bedarf es belastbarer und zuverlässiger analytischer Methoden, um Betrug effizient aufdecken und zurückverfolgen zu können.

Der Produktion von hochqualitativen Lebensmitteln wird in der EU ein hoher Stellenwert beigemessen. Zum Schutz dieser Produkte wurde u.a. die geschützte geographische Angabe (ggA) eingeführt. Diese Angabe mit Bezug zu einer speziellen Herkunft legt Anforderungen für die Produktion fest und wird vom Verbraucher als Qualitätsmerkmal betrachtet.

Spargel ist ein wichtiges Anbauprodukt, sowohl in Deutschland als auch weltweit, mit stetig zunehmenden Produktions- und Erntemengen. Unter allen Obst- und Gemüseprodukten ist der Spargel in Deutschland mit sechs geschützten geographischen Angaben zahlenmäßig am stärksten vertreten [11]. Die geographische Herkunft von Spargel stellt ein wichtiges Qualitätskriterium dar und beeinflusst erheblich die

Kaufentscheidung der Verbraucher. Daher werden insbesondere für regionale Produkte oder für Spargel mit einer ggA hohe Preise erzielt, was wiederum Anreize für Fälschungen bei der angegebenen Herkunft schaffen kann. Die vorliegende Arbeit widmet sich daher der Entwicklung analytischer Methoden zur Herkunftsbestimmung von weißem Spargel.

# 2 Theoretische Grundlagen

## 2.1 Spargel

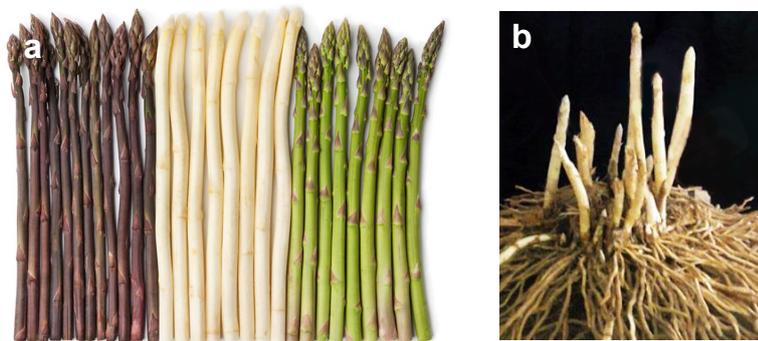
### 2.1.1 Historie und Botanik

Spargel, auch *Asparagus* genannt, bezeichnet eine Pflanzengattung innerhalb der Familie der Spargelgewächse (*Asparagaceae*). Innerhalb dieser Gattung existieren weltweit etwa 300 verschiedene Arten. Neben einigen wild wachsenden oder als Zierpflanzen verwendeten Arten ist die für die Kultivierung wichtigste Art der Gemüsespargel, *Asparagus officinalis* L., der häufig vereinfacht als Spargel bezeichnet wird. Die lateinische Bezeichnung „*Asparagus*“ leitet sich aus dem Griechischen von „*Aspháragos*“ ab, was „junger Trieb“ bedeutet und den essbaren Teil der Pflanze bezeichnet [12, 13]. Der Ursprung des Wortes „*officinalis*“ liegt im Lateinischen und die Bedeutung „arzneilich“ weist auf die heilende Wirkung hin, die dem Spargel zugesprochen wird [14, 15].

Spargel ist ein seit langem bekanntes und beliebtes Gemüse, dessen Ursprung in Vorderasien vermutet wird, von wo aus die Verbreitung nach Europa erfolgte. Es wird angenommen, dass Spargel bereits vor etwa 5000 Jahren in China sowie im antiken Griechenland Anwendung als Heilpflanze fand. Im Römischen Reich erfolgte vermutlich die erste Kultivierung des Spargels [16]. Nach dem Untergang dieser Hochkulturen wurde der Spargel im Mittelalter als geschätztes Gemüse wiederentdeckt, war jedoch zunächst nur wohlhabenden Personen zugänglich, bevor er auch für das normale Bürgertum erschwinglich wurde [17–20].

Man unterscheidet grundsätzlich zwischen weißem, grünem und violetter Spargel, dargestellt in Abbildung 1 a, wovon der grüne Spargel am längsten bekannt ist. Weißer Spargel wurde erst vor etwa 200 Jahren entdeckt, da eher zufällig der besonders feine Geschmack von noch unterirdisch gestochenen Spargel aufgefallen war [21]. Im Gegensatz zum grünen wächst der weiße Spargel ausschließlich unterirdisch, sodass aufgrund der fehlenden Sonneneinstrahlung keine Chlorophyllbildung stattfinden kann. Daher erhält der weiße Spargel seine elfenbeinartige Farbe und den Beinamen „Bleichspargel“. Violetter Spargel wird geerntet, nachdem die Spitze die Erdoberfläche

bereits durchbrochen hat. Durch den Einfluss des Sonnenlichts kommt es zur Bildung von Anthocyanen und somit zur violetten Färbung der Spargelspitze. Mittlerweile existieren jedoch auch Züchtungen, bei denen die gesamte Spargelstange durch intensive Anthocyanbildung tief violett gefärbt ist (vgl. Abb. 1 a). Grüner Spargel wird oberirdisch geerntet, die gesamte Spargelstange ist somit der Sonnenstrahlung ausgesetzt und erhält ihre typische Farbe durch Bildung von Chlorophyll. Aus globaler Sicht kommt dem grünen Spargel mit einem Anteil von 62 % der weltweiten Produktion die größere Bedeutung zu, in Peru beispielsweise liegt der Anteil bei 83 % (2009) [22]. Dagegen spielt in Europa und besonders in Deutschland der Anbau von weißem Spargel eine übergeordnete Rolle mit einem Anteil von etwa 70 % [23, 24].



**Abbildung 1:** (a) Violetter, weißer und grüner Spargel [25];  
(b) Wurzelstock einer Spargelstaude mit Sprossen, adaptiert nach [26].

Die Spargelpflanze ist eine mehrjährige Staude mit einem ausgeprägten Wurzelgeflecht von bis zu 3 Metern Länge, welches sich in horizontaler Richtung im Boden ausbreitet und als Speicherorgan für die Pflanze fungiert. Der Wurzelstock, auch als Rhizom bezeichnet, bildet mit Schuppenblättern besetzte Sprossen aus, die Spargelstangen, welche den essbaren Teil darstellen. Pro Jahr werden etwa sechs bis acht Sprossen pro Wurzelstock ausgebildet, was exemplarisch in Abbildung 1 b dargestellt ist [26, 27].

Wenn die Spargelstangen nicht geerntet werden, können die Sprossen austreiben und Blüten und Früchte bilden, welche in Abbildung 2 gezeigt sind. Spargel gehört zu den zweihäusigen (getrenntgeschlechtigen) Pflanzen, d.h. männliche und weibliche Individuen entwickeln sich auf separaten Spargelpflanzen [20, 28, 29]. Aus den weiblichen Blüten entstehen die tiefrot gefärbten Früchte, aus denen sich im reifen Zustand die Samen entwickeln. Männliche Spargelpflanzen können zwar ebenfalls Früchte ausbilden, aus diesen entstehen jedoch in der Regel keine keimfähigen Samen. Im Spargelanbau werden vorwiegend männliche Hybride eingesetzt, die durch die Kreuzung von genetisch unterschiedlichen Elternorganismen gewonnen werden. Die männlichen Pflanzen müssen keine Energie für die Samenbildung aufwenden, produzieren dickere und längere Stangen und sind somit ertragreicher [30, 31].



**Abbildung 2:** Blüten (a), Früchte (b) und Samen (c) einer Spargelpflanze [32, 33].

### 2.1.2 Bedeutung und physiologische Wirkung

Besonders in Deutschland ist Spargel ein begehrtes und geschätztes Gemüse mit einer durchschnittlichen jährlichen Verzehrsmenge von 1.7 kg pro Kopf [34]. Hierzulande stellt Spargel das Gemüse mit der größten Anbaufläche dar, für dessen Produktion im Jahr 2017 eine Fläche von 23'000 ha bewirtschaftet wurde, was etwa 22 % der deutschen Gemüseanbaufläche entspricht. Der Ernteertrag lag 2017 in Deutschland bei etwa 130'000 t [35, 36]. Die für Deutschland wichtigsten Anbauggebiete liegen in Niedersachsen mit einer Fläche von etwa 5'000 ha, gefolgt von Brandenburg, Nordrhein-Westfalen und Bayern [35]. Des Weiteren liegen mittlerweile für sechs deutsche Gebiete Einträge einer geschützten geographischen Angabe (ggA) im EU-Register der Kommission vor, gezeigt in Abbildung 3: Schrobenshausen, Abensberg und Franken (Bayern), Walbeck und Bornheim (Nordrhein-Westfalen) sowie Beelitz (Brandenburg) [37]. Diese geschützten Herkunftsbezeichnungen werden auf der Grundlage der Verordnung (EU) Nr. 1151/2012 vergeben und sollen dazu beitragen, Qualität und Vielfalt landwirtschaftlicher Produkte als eine der größten Stärken der Europäischen Union zu erhalten. Zudem sollen sie Verbrauchern mit Interesse an regionalen Produkten bzw. Produkten mit einer spezifischen geographischen Herkunft zur Orientierung dienen [38].

Bedingt durch die kurze Saison von etwa April bis zum 24. Juni (Johannistag) ist nur eine eingeschränkte Verfügbarkeit des heimischen Spargels gegeben. Als Qualitätsparameter für den Verbraucher gelten besonders die Frische der Spargelstangen sowie eine regionale Herkunft. Neben deutschem Spargel wird im Zuge von globalisierten Märkten auch importierte Ware in Deutschland angeboten, wodurch über weite Teile des Jahres Spargel im Handel erhältlich ist. Importware stammt zumeist aus Spanien (30 %) und Griechenland (23 %) sowie aus Peru, wo klimabedingt bereits deutlich früher mit der Ernte begonnen werden kann [30].

Spargel ist ein kalorienarmes Lebensmittel mit einem Wasseranteil von 93.5 %. Weitere Bestandteile sind Kohlenhydrate (2 %), Proteine (1.9 %) und Ballaststoffe (1.3 %), darunter Cellulose und Lignin, Fette machen dagegen nur einen Anteil von 0.2 % aus. Zudem enthält Spargel einige wichtige Vitamine und Spurenelemente, darunter Vitamin C,



**Abbildung 3:** Zugelassene geschützte geographische Angaben (ggA) für Spargel in Deutschland [37, 39].

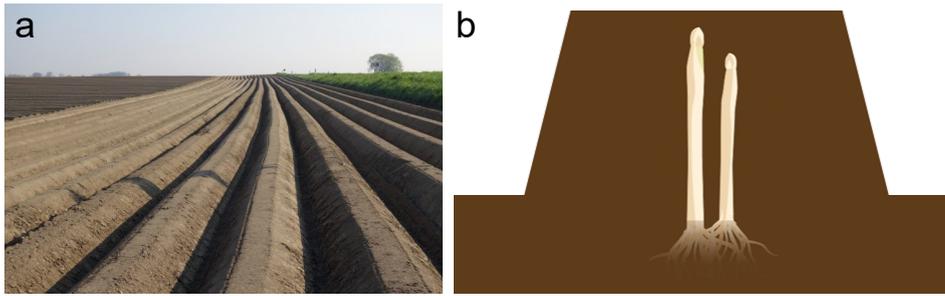
Vitamin E, Vitamin K und Folsäure in signifikanten Mengen [40, 41]. Neben den genannten ernährungsphysiologischen Aspekten sind auch einige gesundheitliche Wirkungen mit dem Spargel assoziiert, die unter anderem mit dem Vorhandensein von sekundären Pflanzeninhaltsstoffen in Verbindung gebracht werden. Der Verzehr von Spargel wirkt nierenanregend, was harntreibende und entwässernde Effekte nach sich zieht [42, 43]. Damit verknüpft ist auch der recht charakteristische Geruch des Urins bei einigen Menschen nach dem Verzehr von Spargel, der auf den Abbau von Asparaginsäure zurückzuführen ist. Diese schwefelhaltige Substanz dient vermutlich dem Schutz der Pflanze vor Pilz- und Bakterienbefall und wird im menschlichen Körper zu verschiedenen leicht flüchtigen Schwefelverbindungen wie etwa Methanthiol verstoffwechselt [44–46]. Die Bildung dieser geruchsaktiven Substanzen ist vermutlich ursächlich auf den Einsatz der Schwefeldüngung im Spargelanbau zurückzuführen [47]. Darüber hinaus sind eine Reihe von positiven medizinisch assoziierten Effekten postuliert worden, die insbesondere auf die im Spargel enthaltenen Saponine oder auf extrahierte Kohlenhydrate zurückgeführt werden, darunter antimykotische Wirkungen, Antitumor-Aktivität, Antimutagenese-Aktivität und eine Wirkung gegen Diabetes [42, 48–52].

Unerwünschte Wirkungen werden hingegen durch die in Abbildung 2 b gezeigten roten Früchte der Spargelpflanze verursacht. Diese weisen eine geringe Giftigkeit auf, sodass nach dem Verzehr von etwa fünf der reifen Früchte Übelkeit und Erbrechen auftreten können [33].

### 2.1.3 Anbau

Für den Spargelanbau sind leichte und sandige Böden besonders geeignet, da diese eine gute Durchwurzelbarkeit aufweisen, was für die Ausbildung des Rhizoms von Bedeutung ist. Zudem neigen Sandböden nicht zu Staunässe, sind leicht zu bearbeiten und erleichtern das Stechen der Spargelstangen. Darüber hinaus erwärmen sich sandige Böden im Frühjahr schneller als festere Böden, was zu einem frühzeitigen Austrieb der Spargelstangen führt. Vor der Neuanlage eines Spargelfeldes ist eine Vorbereitung des Bodens erforderlich, welche sich je nach Zustand des Feldes über einen Zeitraum von ein bis drei Jahren erstreckt. Während dieser Zeit wird die Nährstoffsituation durch Gründüngung oder das Ausbringen von Gülle und Stallmist verbessert. Zur Gründüngung eignen sich beispielsweise Lupine, Senf und Ölrettich, die den Boden mit Nährstoffen versorgen und ihn locker halten. Des Weiteren sollte der Boden frei von Steinen und Wurzelunkräutern sein und einen pH-Wert zwischen 5.5 und 6.0 aufweisen. Vor Beginn der Spargelpflanzung ist eine Bodenanalyse ratsam, um den Nährstoffstatus zu überprüfen. Dabei werden insbesondere der pH-Wert, der Humusgehalt sowie die Konzentration an Kalium, Magnesium, Calcium, Bor und Phosphor untersucht, welche für das Wachstum der Spargelpflanzen eine besondere Bedeutung haben. Basierend auf den Ergebnissen der Bodenuntersuchung kann anschließend eine zielgerichtete Düngung des Bodens mit den erforderlichen Nährstoffen erfolgen [53–56].

Bei der Anlage eines Neufeldes werden meist einjährige Jungpflanzen verwendet, deren Pflanzung sehr früh im Jahr (Februar bis März) erfolgt, bevor Knospenaustrieb und Wurzelwachstum einsetzen, was das Anwachsen im Boden verbessert. Die Spargelpflanzen werden in Reihen in etwa 20 cm Tiefe gesetzt, mit einem Pflanzenabstand von etwa 40 cm und einem Reihenabstand von etwa 2 m. Um zu verhindern, dass der austreibende Spargel die Erdoberfläche durchbricht und somit seine weiße Farbe verliert, ist das Anhäufeln der Erde über den Spargelreihen zu einem Damm von etwa 40 cm Höhe notwendig, was in Abbildung 4 gezeigt ist. Die Ernte der Spargelstangen ist ab dem 2. oder 3. Jahr möglich, jedoch in geringerem Umfang. Ab dem 4. Jahr steht eine Spargelanlage in vollem Ertrag und kann für etwa 10 Jahre beerntet werden. Eine längere Erntezeit ist nicht sinnvoll, da mit zunehmendem Alter der Anlage der Ertrag sinkt und die Bewirtschaftung somit nicht mehr rentabel ist. Anschließend ist eine Wartezeit von mindestens 10-15 Jahren erforderlich, bis auf demselben Feld erneut eine Spargelkultur angelegt werden kann. Ohne diese Wartezeit sind erhebliche Ernteeinbußen zu erwarten, was auf Pilzsporen im Boden zurückzuführen ist, die sich über die Jahre hinweg im Boden anreichern, die Wurzeln der Spargelpflanzen angreifen und bislang nicht wirksam bekämpft werden können [53, 57–59].



**Abbildung 4:** (a) Spargelfeld mit aufgehäufelten Dämmen [60],  
(b) schematische Darstellung des Querschnitts durch einen Spargeldamm [61].

Im Spargelanbau werden häufig Kombinationsdünger zur Versorgung des Bodens mit Stickstoff, Phosphor und Kalium eingesetzt, auch NPK-Dünger genannt, die mengenmäßig die wichtigsten Nährstoffe im Spargelanbau darstellen. Des Weiteren sind auch Mischdünger zur Versorgung mit Spurennährstoffen üblich, die insbesondere Kupfer, Eisen, Mangan, Bor, Magnesium und Schwefel enthalten [53, 62]. Vereinzelt wird auch Mist oder Kompost zur Düngung eingesetzt [55]. Darüber hinaus ist mitunter eine umfangreiche Anwendung von Pflanzenschutzmitteln notwendig, um die Spargelpflanzen vor verschiedenen Schädlingen zu schützen [63]. Dazu gehören Schadinsekten wie die Spargelfliege sowie verschiedene Spargelkäfer [64–67]. Eine bedeutende Schädigung entsteht zudem durch Spargelrost, eine Pilzkrankheit, die durch überwinterrfähige Sporen langfristig wirkt und Ernteaufälle von bis zu einem Drittel hervorrufen kann. Hinzu kommen die sogenannte Grauschimmelfäule, die durch parasitäre Pilze der Gattung *Botrytis* verursacht wird, sowie die Wurzelfäule, ausgelöst durch Pilze der Gattung *Fusarium* [68–71]. Eine Reihe weiterer Schädigungen sind beim Spargel bekannt, darunter auch Viruserkrankungen [72–75]. Die beschriebenen Schädlinge wirken auf unterschiedliche Weise, greifen meist die Wurzel des Spargels oder den jungen Trieb an und vermindern so das Pflanzenwachstum entweder direkt an der Wurzel oder indirekt, indem durch erschwerte Nährstoffbildung nach der Ernte die Pflanze langfristig und nachhaltig geschädigt wird.

Der Erntezeitraum erstreckt sich von ca. April bis Ende Juni, wobei der Beginn der Ernte von Jahr zu Jahr je nach Wetterlage variiert. Erntebeginn und Ertrag werden besonders von der Bodentemperatur bestimmt: ab 12 °C fangen die Spargelstangen an auszutreiben, die Ernte ist ab einer Bodentemperatur von 16 °C möglich [76]. Die Ausrichtung des Spargelfeldes mit den nebeneinander angeordneten Reihen von Spargeldämmen erfolgt idealerweise in Nord-Süd-Richtung, um eine möglichst gleichmäßige Sonneneinstrahlung zu erzielen [77]. Zur Steuerung der Bodentemperatur werden im Spargelanbau mittlerweile standardmäßig verschiedene Foliensysteme eingesetzt. Am weitesten verbreitet ist eine schwarz-weiße Taschenfolie, mit der der Spargeldamm der Länge nach bedeckt wird. Je nachdem, welche Seite der Folie obenauf liegt, kann durch verstärkte Absorption

(schwarz) oder Reflexion (weiß) der Sonnenstrahlung die Temperatur im darunter liegenden Damm beeinflusst und die Erntezeit somit gezielt gesteuert werden. Natürlicherweise ist der Ertrag zu Beginn der Saison aufgrund niedriger Temperaturen sehr gering, steigt stark an, wenn die optimale Bodentemperatur erreicht ist, und sinkt zum Ende der Saison wieder deutlich ab. Der Einsatz von Foliensystemen trägt maßgeblich dazu bei, Erntespitzen zu vermeiden, sodass die Nachfrage über die gesamte Saison bedient werden kann [77, 78]. Darüber hinaus verhindert die Folienabdeckung durch Abschirmung der Sonneneinstrahlung eine Grünfärbung der Spargelspitzen, wenn diese die Erdoberfläche durchbrechen, sowie eine Überhitzung des Dammes, was zur Bildung von Hohlräumen innerhalb der Spargelstangen und damit zu einem Qualitätsverlust führt [79]. Neben der erwähnten Taschenfolie finden verschiedene Tunnelsysteme wie Einreihen- oder Mehrreihen-Minitunnel, Thermo- und Antitaufolien, befüllbare Folien und Abdeckflies Anwendung [80–82].

Bei der Beeinflussung des Erntezeitpunktes spielt ebenfalls die Wahl der angebauten Spargelsorte eine wichtige Rolle. Die einzelnen Sorten unterscheiden sich unter anderem hinsichtlich ihres Erntezeitpunktes, des Ertrags, der Anfälligkeit gegenüber Schädlingen, der bevorzugten Bodenart und dem Aussehen der Spargelstangen. Die Sorten Gijnlim, Ramires, Avalim und Ramada zählen zu den sehr frühen Sorten, gefolgt von Grolim und Cumulus, die ebenfalls früh geerntet werden können [83–88]. Rapsody und Herkolim stellen mittelfrühe bis späte Sorten dar, wohingegen Backlim eher zum Ende der Saison hin geerntet wird [89–91]. Für den großtechnischen Anbau von Spargel ist eine Kombination von frühen und späten Sorten angebracht, um ausreichende Erntemengen über die gesamte Saison hinweg zu erzielen. Bei der Wahl der Sorte ist zudem die vorliegende Bodenart zu beachten, um ein optimales Wachstum zu erreichen.

Das traditionelle Ende der Erntesaison ist am 24. Juni, dem Johannistag. Danach sollte keine Ernte weiterer Spargelstangen erfolgen, damit die Pflanze austreiben und dabei Nährstoffreserven für die nächste Saison bilden kann. Andernfalls ist mit deutlich verminderten Erträgen im Folgejahr zu rechnen [53, 55, 92].

Die Ernte des Spargels erfolgt bislang fast ausschließlich von Hand, indem der Damm geöffnet wird, sobald eine Spargelstange die Erdoberfläche durchbricht, der Spargel bei einer Länge von etwa 30 cm gestochen und die Erde anschließend wieder zum Spargeldamm angehäufelt wird. Nach der Ernte der Spargelstangen werden diese einer Schockkühlung mit Eiswasser von 1–2 °C unterzogen, um Enzyme zu inaktivieren und somit nachteilige Veränderungen zu vermeiden. Durch metabolische Aktivitäten kommt es andernfalls schnell zu strukturellen Veränderungen wie einer Verholzung der Spargelstangen durch Lignifizierung und dem Verlust löslicher Kohlenhydrate. Aufgrund einer hohen Atmungsaktivität werden Kohlenhydrate abgebaut, es folgt der Abbau von weiteren Inhaltsstoffen wie Proteinen. All diese Prozesse haben negative

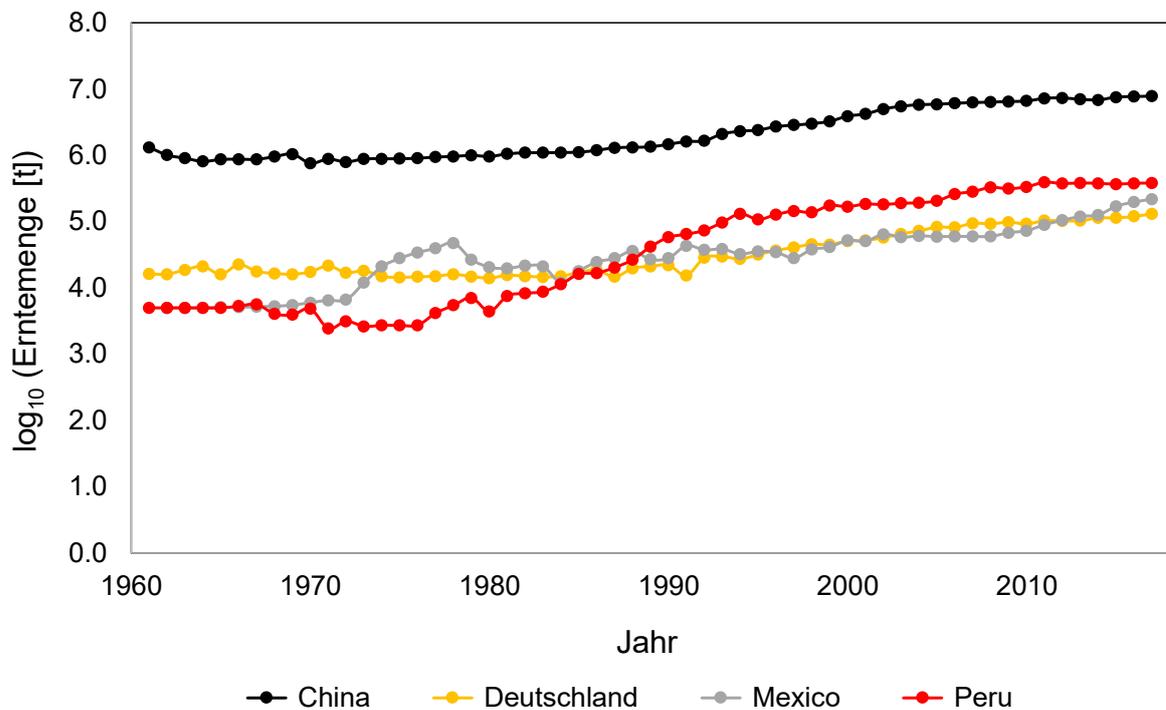
Auswirkungen auf Textur und Geschmack und somit auf die Genusstauglichkeit des Spargels [93]. Die Lagerung der Spargelstangen erfolgt ebenfalls bei 1-2 °C und einer hohen Luftfeuchtigkeit, meist in Kühllhäusern mit automatischer Beregnung, was auch mikrobiellem Verderb, Wasserverlust und einer Violettverfärbung vorbeugt [93, 94].

#### 2.1.4 Produktion in Deutschland und weltweit

Im Vergleich der wichtigsten Länder für die Spargelproduktion belegt Deutschland den vierten Platz hinter China, Peru und Mexiko. China verfügt dabei über die mit Abstand größten Produktionsmengen. Eine Übersicht über Erntemengen, Anbauflächen und Ertrag der zehn wichtigsten Anbauländer im Erntejahr 2017 ist in Tabelle 1 dargestellt. Abbildung 5 zeigt zudem die Entwicklung der jährlich geernteten Mengen an Spargel von 1961 bis 2017. Anhand der Abbildung ist eine kontinuierliche Zunahme der Erntemengen der vier größten Anbauländer in den letzten Jahrzehnten zu verzeichnen. Dies trifft auch für die weltweite Erntemenge zu, die sich von 1.7 Mio.t im Jahr 1961 auf 8.9 Mio.t im Jahr 2017 erhöht hat [36]. Eine sehr ähnliche Entwicklung wurde auch für die bewirtschafteten Anbauflächen erfasst [36]. Hauptunterschiede zwischen den einzelnen Ländern liegen, neben dem schwerpunktmäßigen Anbau von weißem oder grünem Spargel, besonders im Erntezeitraum, bedingt durch die unterschiedlichen klimatischen Bedingungen. Während die Erntesaison in Deutschland von etwa April bis Ende Juni andauert, kann Spargel in den Ländern Südeuropas, wie Griechenland und Spanien, aufgrund des milderen Klimas bereits

**Tabelle 1:** Erntemenge, Anbaufläche und Ertrag der zehn wichtigsten Anbauländer für Spargel (2017) [36].

Land	Erntemenge [t]	Anbaufläche [ha]	Ertrag [t/ha]
China	7'845'162	1'411'978	5.6
Peru	383'098	32'365	11.8
Mexico	245'681	26'139	9.4
Deutschland	130'881	23'190	5.6
Spanien	63'433	13'755	4.6
Italien	46'419	6'687	6.9
USA	34'100	9'430	3.6
Japan	29'981	5'378	5.6
Thailand	25'000	1'800	13.9
Niederlande	20'800	3'322	6.3



**Abbildung 5:** Entwicklung der Erntemengen der vier wichtigsten Anbauländer für Spargel von 1961 bis 2017 [36].

im Februar geerntet werden [30, 95, 96]. In Peru ist darüber hinaus auch eine zweite Erntesaison möglich, die sich in etwa von September bis Dezember erstreckt, da die Temperaturen dort über das Jahr hinweg nur sehr geringen Schwankungen unterworfen sind [97–99]. Dies spiegelt sich auch in einem deutlich höheren Ertrag wider (vgl. Tabelle 1). Im Handel verfügbarer frischer Spargel stammt in den Wintermonaten meist aus Peru, während ab Februar größere Mengen an griechischem und spanischem Spargel importiert werden, bis ab März auch deutscher Spargel verfügbar ist. Der importierte Spargel ist im Vergleich zum deutschen in der Regel deutlich preisgünstiger, der heimische Spargel ist dagegen beim Verbraucher sehr viel beliebter. Gerade zu Beginn der Saison in Deutschland ist der Konkurrenzdruck zu importierten Waren besonders hoch, da hierzulande nur kleine Mengen verfügbar sind. Gleichzeitig ist die Nachfrage nach deutschem Spargel am Anfang der Saison hoch, was dazu führt, dass mit der frühen Ware sehr hohe Preise von über 20 EUR/kg erzielt werden können. Der griechische Spargel erreicht dagegen in Spitzenzeiten Preise von etwa 8 EUR/kg, die mit zunehmendem Angebot des deutschen Spargels teilweise bis unter 2 EUR/kg absinken [100]. Die große Preisdifferenz zwischen deutschem und importiertem Spargel schafft zudem einen Anreiz für vorsätzliche Fälschungen der Herkunftsangabe beim Verkauf von Spargel, indem importierte Produkte als deutsche ausgegeben und somit deutlich teurer verkauft werden können. Auch Spargel aus den östlichen Nachbarländern, besonders aus Polen, stellt ein denkbares und vermutetes Ausgangsprodukt solcher Umetikettierungen dar, da Spargel dort zu deutlich geringeren Kosten produziert,

gleichzeitig aber auch nur zu niedrigeren Preisen verkauft werden kann. Diese liegen im Schnitt bei etwa 2.50 EUR/kg, die Preisspanne reicht dabei von 1.20 EUR bis 5 EUR/kg [100–104].

### 2.1.5 Bisher verwendete Strategien zur Herkunftsbestimmung

Die Untersuchung der geographischen Herkunft von weißem Spargel ist in den letzten Jahren immer weiter in den Fokus gerückt, sodass bereits erste Arbeiten dazu veröffentlicht wurden.

Swoboda *et al.* untersuchten die geographische Herkunft von Spargel aus den Ländern Österreich, Ungarn, der Slowakei, Deutschland, den Niederlanden und Peru anhand des Strontium-Isotopenverhältnisses mittels Multikollektor-ICP-MS [105]. Der Vergleich von Strontium-Konzentrationen und Isotopenverhältnissen der einzelnen Länder zeigt eine gute Abgrenzung der Proben aus Österreich von denen aus der Slowakei und Ungarn, wobei eine große Ähnlichkeit der beiden letztgenannten Länder zu beobachten war. Die Proben aus Deutschland, den Niederlanden und Peru wiesen eine große Übereinstimmung in der Strontiumverteilung zu den drei zuvor genannten Ländern auf und waren daher nicht von diesen unterscheidbar. Insgesamt wurden 137 Proben aus 2 Erntejahren untersucht, wobei nur eine kleine Probenanzahl für die Länder Deutschland und Peru mit 6 bzw. 3 Proben zur Verfügung stand. Bei der Auswertung der Daten wurde ein Vergleich der Strontium-Isotopenverhältnisse und die Gegenüberstellung mit den Strontiumkonzentrationen vorgenommen.

In der Arbeit von Menzel *et al.* wurde untersucht, inwieweit die Herkunft von weißem Spargel aus zwei verschiedenen Anbaugebieten in Deutschland mithilfe einer NMR-basierten Methode unterschieden werden kann [106]. Die Korrelation der Relaxationszeiten der durchgeführten NMR-Experimente lieferte erste Anhaltspunkte für eine Unterscheidung der Herkunftsgebiete, besonders bei der Analyse des Saftes, der aus den Spargelstangen gewonnen wurde. Für die Untersuchung wurden mehrere Spargelstangen von fünf verschiedenen Standorten aus demselben Erntejahr verwendet, die Probenzahl war somit stark limitiert, weshalb die Arbeit von den Autoren als Pilotprojekt eingestuft wird.

In einer weiteren Untersuchung von Zannella *et al.* wurden Spargel- und Bodenproben von fünf Bauern aus einer italienischen Region mit geschützter geographischer Angabe, *Bassano del Grappa*, analysiert [107]. Die Proben wurden hinsichtlich ihrer Multielementzusammensetzung sowie der Strontium-Isotopenverhältnisse untersucht. Basierend auf

dem Variationskoeffizienten wurden fünf Elemente identifiziert, die als potentiell charakteristisch für die genannte Region angesehen werden. Das Strontium-Isotopenverhältnis in den Spargelproben wies eine hohe Ähnlichkeit zu dem im bioverfügbaren Anteil der zugehörigen Bodenproben auf, war jedoch stark abweichend zum Isotopenverhältnis im gesamten Boden. Ein Vergleich mit den zuvor genannten Untersuchungen von Swoboda *et al.* [105] zeigte eine teilweise Abgrenzbarkeit von den dort untersuchten Spargelproben.

Insgesamt war bei den bisherigen Arbeiten zur geographischen Herkunftsbestimmung von Spargel die Zahl der zur Verfügung stehenden Proben recht limitiert. Zudem stammten die untersuchten Proben zumeist aus einem Erntejahr, mit Ausnahme der Arbeit von Swoboda *et al.*, bei der zwei Erntejahre berücksichtigt, jedoch kein Vergleich zwischen den einzelnen Jahren vorgenommen wurde. Des Weiteren war auch das Gebiet, aus dem Proben zur Analyse bezogen wurden, mitunter sehr klein. So wurden von Zannella *et al.* nur Spargelproben aus einer eng umgrenzten Anbauregion im Radius von 8 km untersucht [107]. Im Falle der Arbeit von Menzel *et al.* wurden Proben aus zwei deutschen Regionen (Rheinland und Franken) herangezogen [106], sodass kein direkter Vergleich mit anderen Anbauregionen bzw. Herkunftsländern möglich war. Die Datenauswertung basierte dabei auf dem Vergleich der Konzentrationen, Stabilisotopenverhältnisse bzw. Relaxationszeiten sowie der Gegenüberstellung von Variationskoeffizienten der untersuchten Elementkonzentrationen. In keiner der gezeigten Arbeiten wurden Verfahren zur multivariaten Datenanalyse bzw. Methoden des maschinellen Lernens herangezogen.

In einer neueren Arbeit von Creydt *et al.* wird die Herkunftsbestimmung von weißem Spargel auf der Grundlage von LC-ESI-qTOF-MS-basierten Untersuchungen thematisiert [108]. Dabei stand mit etwa 400 Spargelproben bereits eine deutlich größere Probenzahl zur Verfügung, wobei vier Erntejahre und sechs Herkunftsländer, darunter auch Proben aus allen relevanten deutschen Herkunftsgebieten, einbezogen wurden. Eine Klassifikation des Herkunftslandes mittels *Random Forest* basierend auf den Daten des *non-targeted Metabolomics*-Ansatzes lieferte eine Genauigkeit von 82% bei einem Vorhersagemodell für sechs Herkunftsländer und zwei Erntejahre bei der Klassifikation von 115 Proben. Ein Teil der untersuchten Metabolite zeigten Abweichungen zwischen den einzelnen Erntejahren, sodass eine gewisse Jahresabhängigkeit beobachtet wurde. Die Gegenüberstellung der Ergebnisse mit denen einer Untersuchung der Stabilisotopen (IR-MS) lieferte vergleichbare Aussagen wie der LC-MS-Ansatz, wobei auch hier eine jahreszeitliche Abhängigkeit auftritt und somit eine jährliche Neuberechnung erforderlich ist.

In der amtlichen Lebensmittelüberwachung wird zur Überprüfung der deklarierten Herkunft aktuell in mehreren Bundesländern in Deutschland standardmäßig die Analyse der Stabilisotopenverhältnisse (IR-MS) im Spargel oder daraus isolierten Fraktionen

herangezogen [109–112]. Dafür werden die Elemente Wasserstoff, Sauerstoff, Kohlenstoff, Stickstoff und Schwefel analysiert, deren Isotopenzusammensetzung sich je nach Standort unterscheidet. Beispielsweise weist die Masse verschiedener Wassermoleküle kleine Unterschiede auf, die auf die natürliche Verteilung der stabilen Isotope von Wasserstoff und Sauerstoff zurückzuführen sind. Wasser, welches aus dem Meer verdunstet, wird über die Regenwolken ins Landesinnere transportiert. Dabei reichern sich leichte Wassermoleküle eher in der Gasphase an, während sich die schwereren Moleküle in der flüssigen Phase sammeln und schneller abregnen. Daraus folgt ein Konzentrationsprofil von Nord-Nord-West nach Süd-Süd-Ost mit zunehmend leichteren Wassermolekülen je weiter das Meer entfernt ist. Gebiete auf der Ost-West-Achse mit etwa gleicher Entfernung zum Meer sind dagegen schwer zu unterscheiden [112–114]. Die Anwendung dieser Methode von Rossier *et al.* auf weißen Spargel aus verschiedenen Herkunftsländern lieferte kein spezifisches Muster für jede Region, sodass eine eindeutige Zuordnung zum Herkunftsland nicht möglich war. In der Arbeit wurde zudem eine jährliche Schwankung des Signals identifiziert [115].

---

## 2.2 Omics-Techniken

Unter dem Begriff der Omics-Techniken werden verschiedene Verfahren zur Charakterisierung eines biologischen Systems zusammengefasst. Diese beleuchten die Bestandteile eines Organismus, deren Interaktionen sowie den funktionellen Status als übergeordnetes Erscheinungsbild einer Zelle oder eines zellulären Netzwerkes [116–119]. Die Omics-Kaskade wird häufig zur Beschreibung der Komponenten eines Systems herangezogen. Die klassische Omics-Kaskade fußt auf der genetischen Information, dem Genom, dessen Untersuchung als *Genomics* bezeichnet wird [118, 120, 121]. Während dem Vorgang der Transkription innerhalb einer Zelle wird die DNA in RNA umgeschrieben, die Untersuchung der RNA wird daher auch als *Transcriptomics* bezeichnet [121–123]. Die gebildete RNA wird wiederum zur Proteinsynthese benötigt. Die Gesamtheit aller Proteine wird im Begriff des Proteoms zusammengefasst, dessen Untersuchung entsprechend als *Proteomics* bezeichnet wird [119, 124]. Als letztes Element der Omics-Kaskade wird häufig die Disziplin der *Metabolomics* angesehen, die sich mit der Erfassung des Metaboloms beschäftigt, der Gesamtheit aller Stoffwechselprodukte [116, 125, 126]. Neben den vier genannten Teilbereichen sind jedoch mittlerweile viele weitere Disziplinen etabliert, die ebenfalls der Beschreibung der Bestandteile eines zellulären Systems dienen, darunter die Gebiete *Epigenomics*, *Lipidomics* und *Glycomics* [116, 127]. In den letzten Jahren ist mit der Disziplin der *Isotopolomics* ein weiterer Begriff entstanden, der die Untersuchung der Elementzusammensetzung und der Isotopenverteilung eines Organismus (Isotopolom) beschreibt [128, 129]. Da die in dieser Arbeit angewandten analytischen Ansätze, ICP-MS und NIR, in die Bereiche *Isotopolomics* und *Metabolomics* einzuordnen sind, sollen diese im Folgenden etwas genauer beschrieben werden.

### 2.2.1 Isotopolomics

Unter dem Begriff *Isotopolomics* wird die Untersuchung der Isotopenzusammensetzung eines biologischen Systems zusammengefasst, die unterschiedlich motiviert sein kann. So können verschiedene Isotope eines Elementes quantifiziert und ins Verhältnis zueinander gesetzt werden, etwa um unterschiedliche Standorte beschreiben zu können, wie dies bei der erwähnten Herkunftsunterscheidung mittels Stabilisotopenanalyse angewendet wird. Zudem wird die Isotopenuntersuchung zur Quantifizierung möglichst vieler Elemente genutzt, um letztlich das Elementprofil eines Organismus so präzise wie möglich charakterisieren zu können [128, 129].

Die in einem pflanzlichen Organismus enthaltenen Elemente sind in großem Maße abhängig von der Elementzusammensetzung des Bodens, auf bzw. in dem sie wachsen.

Die Frage, ob ein im Boden enthaltenes Element von der Pflanze aufgenommen werden kann, wird von Mobilität und Verfügbarkeit der Elemente bestimmt, die wiederum von verschiedenen Faktoren beeinflusst werden. Dazu zählen die Löslichkeit der Elemente und ihrer Verbindungen, der pH-Wert im Boden, Bodenart und -beschaffenheit, die Konzentration an gelösten organischen Substanzen, aber auch klimatische Einflüsse wie Niederschläge, Temperatur und Sonneneinstrahlung [130–136]. Einige Elemente liegen zudem intern oder extern gebunden an bestimmte Bodenbestandteile wie Silikatstrukturen, organische Verbindungen oder Eisen-Manganoxide vor, sodass sie in dieser Form nicht für die Aufnahme durch Pflanzen verfügbar sind [137–139]. Entscheidend ist außerdem, ob das entsprechende Element eine Funktion im pflanzlichen Organismus ausübt und die Pflanze daher über Aufnahmewege für dieses Element verfügt. Solche Transportsysteme können mitunter auch die Aufnahme von für die Pflanze nicht notwendigen Elementen ermöglichen, wenn diese eine chemische Ähnlichkeit zu essentiellen Elementen aufweisen [140]. Auch andere Organismen, wie etwa symbiotisch lebende Pilze, können die Aufnahme von Elementen aus dem Boden beeinflussen [141]. Die Elementaufnahme steht zudem im Zusammenhang mit der Genetik, da verschiedene genetische Varianten mit unterschiedlichen Transportproteinen einhergehen können [142]. Des Weiteren sind in Bezug auf die Elementzusammensetzung des Bodens auch anthropogene Einflüsse zu beachten, wie etwa das Ausbringen von Dünge- und Pflanzenschutzmitteln, aber auch industriebedingte Kontaminationen.

Im Gegensatz zum Isotopolom werden bei der Analyse des Metalloms, auch *Metallomics* genannt, das Zusammenspiel von Metallen bzw. Halbmetallen mit verschiedenen Biomolekülen im Organismus und seinen Kompartimenten, darunter DNA, RNA und Metabolite, sowie die physiologischen und toxikologischen Effekte von Metallen untersucht [143, 144].

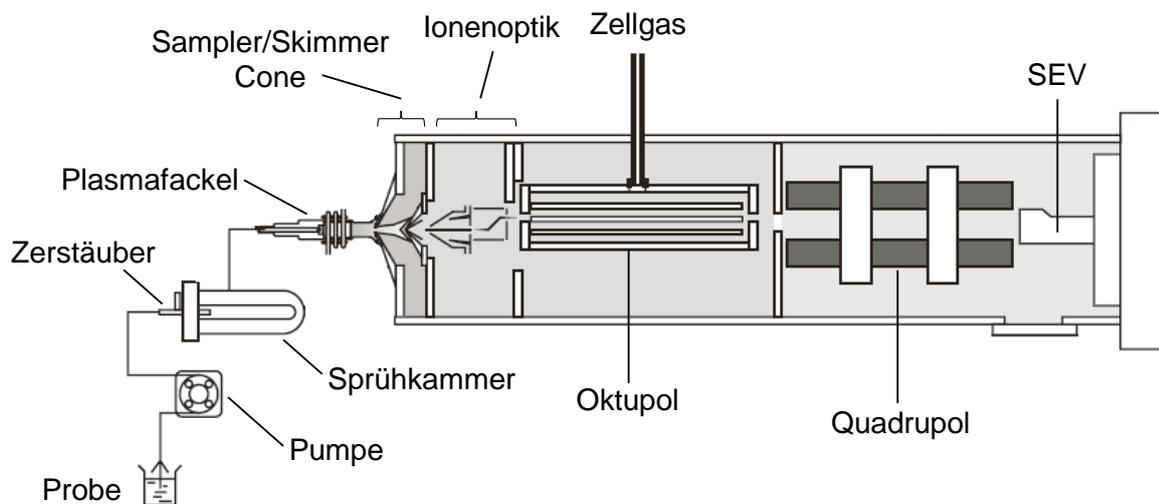
### 2.2.2 Metabolomics

Unter dem Begriff *Metabolomics* wird die umfassende Untersuchung der Metaboliten eines biologischen Systems zusammengefasst [145–148]. Man unterscheidet dabei häufig zwischen den beiden Ansätzen der *targeted*- und der *non-targeted*-Analyse. Die *non-targeted*-Strategie zielt auf die möglichst vollständige Erfassung der Stoffwechselprodukte eines Organismus ab, was für diesen eine Art Fingerabdruck darstellt, weshalb man auch von *Fingerprinting*-Techniken spricht. Identifizierung und Quantifizierung der Metaboliten stehen dabei zunächst im Hintergrund [146, 149]. Der *non-targeted*-Ansatz zeigt zugleich einen wesentlichen Vorteil der *Metabolomics*, da so eine Analyse auch ohne Kenntnis der Metaboliten möglich ist [150]. Um der Erfassung aller Metaboliten nahe

zu kommen, bedarf es der Anwendung von möglichst hochauflösenden Analysetechniken, mit denen auch kleine Unterschiede wahrgenommen werden können. Dabei finden besonders LC-MS, GC-MS und NMR-basierte Methoden Anwendung. Des Weiteren kommen auch infrarotspektroskopische Methoden wie FT-IR und FT-NIR zum Einsatz, die zwar eine deutlich niedrigere Auflösung liefern, jedoch sehr schnelle Messungen erlauben und weitestgehend ohne Probenaufarbeitung auskommen, wodurch ein hoher Probendurchsatz ermöglicht wird. Zudem handelt es sich um zerstörungsfreie Methoden, sodass das Probenmaterial auch zur späteren Analyse mit einer hochauflösenden Methode zur Verfügung steht [149, 151, 152]. *Targeted*-Methoden zielen dagegen auf die Identifizierung und Quantifizierung von ausgewählten Metaboliten ab, die als Markersubstanzen für die zu unterscheidenden Merkmale innerhalb der gewählten Fragestellung dienen [152–155]. Zu den Einflussfaktoren für die von einem Organismus produzierten Metaboliten zählen einerseits äußere Faktoren wie Klima- und Wetterbedingungen, besonders Niederschlag, Temperatur und Sonneneinstrahlung, aber auch die Bodenbeschaffenheit, der Einsatz von Dünge- und Pflanzenschutzmitteln sowie die umgebenden Organismen, wie etwa Phytopathogene [156–158]. Daneben spielen auch Art und Sorte einer Pflanze als innere Faktoren eine wichtige Rolle, da sich die zugrundeliegende Genetik über die zuvor beschriebenen Abläufe innerhalb der Omics-Kaskade unmittelbar auch auf die Bildung von Stoffwechselprodukten auswirkt [159–161].

## 2.3 ICP-MS

Bei der *Inductively Coupled Plasma Mass Spectrometry*, kurz ICP-MS, handelt es sich um eine massenspektrometrische Methode zur Elementanalytik. Die Proben werden meist in flüssiger Form aufgenommen und in ein Aerosol überführt, aus dem im Plasma positiv geladene Analytteilchen gebildet werden, die im Massenanalysator aufgetrennt und anschließend detektiert werden [162, 163]. Der Geräteaufbau ist exemplarisch in Abbildung 6 gezeigt. Die wesentlichen Schritte der Messung werden im Folgenden näher erläutert.

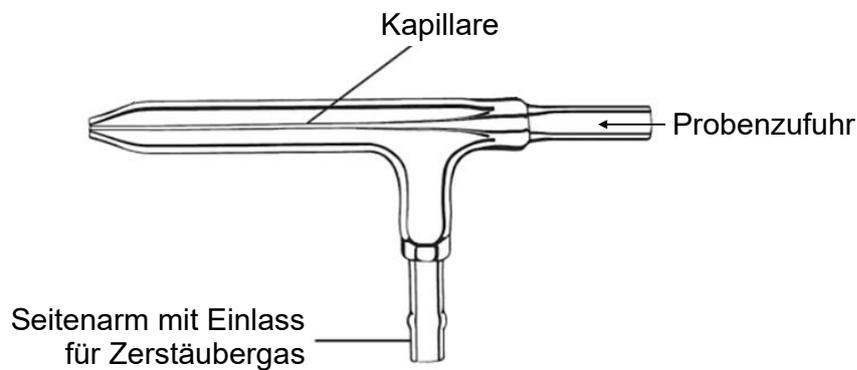


**Abbildung 6:** Schematischer Aufbau eines Quadrupol-basierten ICP-MS-Gerätes, modifiziert nach [164].

### 2.3.1 Probeneintrag

Der Probeneintrag stellt einen entscheidenden Schritt im Hinblick auf die folgende Ionisierung der Analyten dar. Die zentralen Bauteile innerhalb des Probeneintragungssystems sind Zerstäuber und Sprühkammer, die für die Überführung der Probe in ein Aerosol und die anschließende Selektion der gebildeten Tröpfchen sorgen [162, 163]. Die flüssigen Proben werden meist mithilfe einer peristaltischen Pumpe angesaugt und mit einer Geschwindigkeit von etwa 1 mL/min bis zum Zerstäuber transportiert. Die peristaltische Pumpe besteht aus mehreren Rollen, die sich mit gleicher Geschwindigkeit bewegen und über die die Probenschläuche gespannt werden. Durch die Bewegung und den Druck der Rollen wird die Flüssigkeit im Schlauch transportiert. Häufig wird zwischen Pumpe und Zerstäuber zusätzlich ein Dreiecksventil geschaltet, über welches der interne Standard gleichmäßig zugegeben wird. Es existieren diverse Formen von Zerstäubern mit jeweils etwas unterschiedlicher Funktionsweise, von denen der konzentrische Zerstäuber

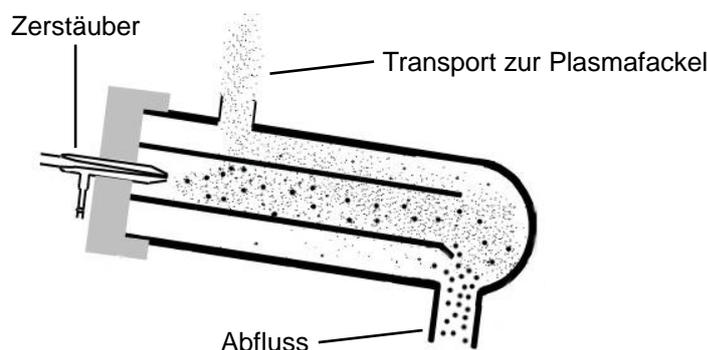
aus Borosilikatglas oder Quarz am häufigsten Anwendung findet [162, 165]. Dieser Zerstäubertyp besteht aus einer sehr feinen Kapillare, die von einem Hohlraum umgeben ist, was schematisch in Abbildung 7 dargestellt ist. Über einen Seitenarm wird das Zerstäubergas, meist Argon, zugeführt, welches den Hohlraum mit einer Flussrate von etwa 1 - 3 mL/min durchströmt und dabei einen Druck von 2 - 3.5 bar erreicht. Die Probe durchläuft die enge Kapillare und nimmt aufgrund der Verengung an Geschwindigkeit zu, was auch als „Venturi-Effekt“ bezeichnet wird [165, 166]. An der Spitze des Zerstäubers, die schmal zuläuft, treffen Probe und Zerstäubergas aufeinander, was eine sofortige Aufspaltung der flüssigen Probe in ein feines Aerosol zur Folge hat. Aufgrund der sehr schmalen Kapillare des Zerstäubers ist die Verwendung von feststofffreien Proben erforderlich, da andernfalls die Gefahr einer Verstopfung besteht.



**Abbildung 7:** Schematischer Aufbau eines konzentrischen Zerstäubers, modifiziert nach [162].

Da die Plasmaentladung bei dissoziierenden großen Tröpfchen sehr ineffizient verläuft, erfolgt anschließend eine Selektion des Aerosols in der Sprühkammer, damit nur kleine Tröpfchen ins Plasma weitergeleitet werden. Ziel ist die Bildung eines Aerosols mit homogener verteilter Tröpfchengröße und möglichst kleinem Durchmesser der Tröpfchen, was zu einer gleichmäßigen Ionisierung der Analyten im Plasma führt und die Robustheit des Plasmas verbessert. Zusätzlich sollen pulsierende Bewegungen des Aerosols, bedingt durch die Bewegung der Rollen der peristaltischen Pumpe, abgeschwächt werden. In der Praxis werden zwei Typen von Sprühkammern verwendet, die *double pass* und die zyklonische Sprühkammer, wobei die erstgenannte die größte Relevanz hat. Am häufigsten wird die sogenannte *Scott-like double pass*-Sprühkammer eingesetzt, welche in Abbildung 8 gezeigt ist. Diese Sprühkammer verfügt über ein zentrales Rohr, in welches die Spitze des Zerstäubers ragt, sodass das gebildete Aerosol direkt in die Sprühkammer abgegeben wird. Das Aerosol durchläuft das zentrale Rohr, wobei größere Tröpfchen mit einem Durchmesser von mehr als 10  $\mu\text{m}$  aufgrund der Schwerkraft nach unten fallen, sich an der Wandung des Rohres absetzen und in Richtung Abfluss fließen,

wo sie abgesaugt werden. Durch die Sammlung der größeren Tröpfchen im Abflussrohr entsteht ein Druck in der Sprühkammer, der dazu führt, dass die kleinen Tröpfchen in Richtung der Plasmafackel transportiert werden können. Insgesamt gelangen somit nur etwa 2 % des Aerosols ins Plasma. Die Sprühkammer wird häufig von außen gekühlt um die thermische Stabilität der Probe zu gewährleisten, was eine verbesserte Signalstabilität und geringere Oxidraten zur Folge hat. Bei der Verwendung von organischen Lösungsmitteln trägt eine Kühlung zudem zu deren Abtrennung und damit verbunden zur Plasmastabilität bei. Die *double pass*-Sprühkammer besteht oft aus Quarz, aber auch verschiedene Kunststoffmodelle etwa aus Polysulfid sind bekannt [162, 163, 165, 167–170].



**Abbildung 8:** Schematischer Aufbau einer *Scott-like double pass*-Sprühkammer, modifiziert nach [171].

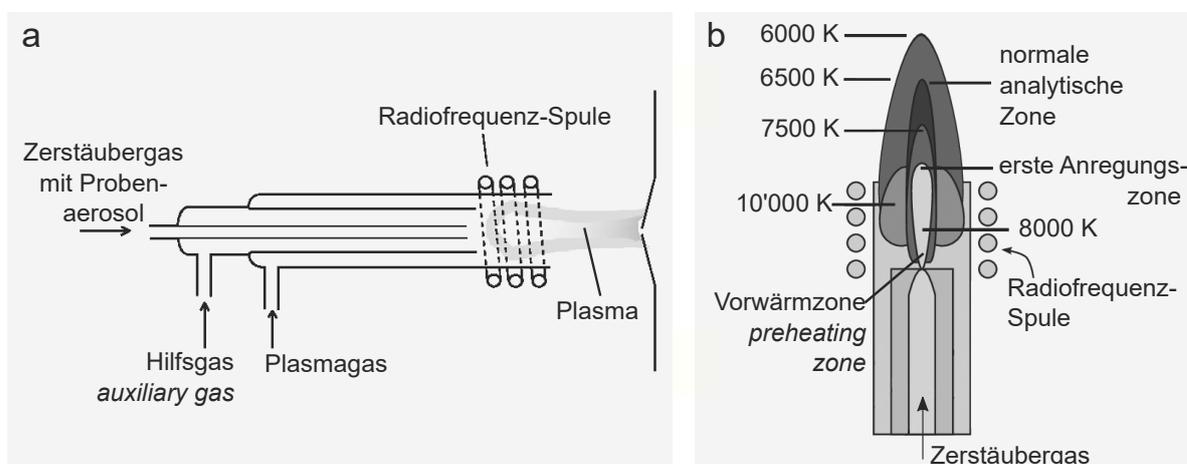
Neben Feststoffpartikeln, die zur Verstopfung des Zerstäubers führen können, sind auch hohe Konzentrationen an gelösten Feststoffen (*Total Dissolved Solids*, TDS) problematisch, da sich diese nach Verdampfung des Lösungsmittels im Plasma auf den *cones* ablagern und auch hier zum Verstopfen führen können. Um eine Verdünnung der Proben vor der Messung zu vermeiden, was zu einer erhöhten Variabilität der Proben durch Kontamination und Verdünnungsfehler führt, verfügen viele Geräte über einen Mechanismus der Aerosolverdünnung. Mithilfe eines Verdünnungsgases (*dilution gas*), ebenfalls Argon, welches an der Verbindungsstelle zwischen Sprühkammer und Plasmafackel eingeleitet wird, kann das gebildete Aerosol gleichmäßig und reproduzierbar verdünnt werden. Auf diese Weise wird die Toleranz gegenüber gelösten Feststoffen deutlich erhöht, sodass auch Proben mit hohem Matrixanteil ohne Verdünnung injiziert werden können. Daher spricht man auch von *High Matrix Introduction* (HMI) [172–174].

Neben dem Eintrag flüssiger Proben über einen Autosampler sind auch verschiedene Kopplungen der ICP-MS möglich. Verbreitet sind mittlerweile Kopplungen mit einer HPLC oder GC, aber auch Kopplungen von Ionenchromatographie, Kapillar- oder Gelelektrophorese, einer *Scanning Flow Cell* oder der *Flow Injection Analysis* mit der ICP-MS wurden bereits angewendet [175–180].

Zusätzlich zur Analyse von Proben in flüssiger Form, der ein Aufschluss vorangestellt ist, stehen mittlerweile eine Reihe von Lasergeräten zur Verfügung, mit denen Partikel mittels Laserablation (*Laser Ablation*, LA) aus einem Feststoff herausgelöst werden. Je nach Wellenlänge des verwendeten Lasers können Proben unterschiedlichster Beschaffenheit und Härte ablatiert werden. Die Oberfläche der Probe wird dabei mit dem Laserstrahl abgetastet und die herausgelösten Partikel werden anschließend mithilfe eines Trägergases, Argon oder Helium, direkt zur Plasmafackel des ICP-MS geleitet. Durch die Verweildauer des Laserstrahls auf einer Position der Probe kann die Eindringtiefe des Lasers variiert werden, sodass auch unterschiedliche Schichten einer Probe untersucht werden können. Die Laserablation ermöglicht zudem eine nahezu zerstörungsfreie Analyse, für die nur sehr geringe Materialmengen benötigt werden [162, 181–183].

### 2.3.2 Ionisierung

Die Ionisierung der Analyten erfolgt mithilfe eines Argonplasmas, welches an der Öffnung der Plasmafackel gebildet wird. Die Plasmafackel, schematisch in Abbildung 9 a dargestellt, besteht aus drei konzentrischen Röhren meist aus Quarz, seltener aus Keramik [162, 184]. In die mittlere Röhre wird das Zerstäubergas mit dem Proben-aerosol mit einer Flussrate von etwa 1 L/min eingeleitet. In die äußere Röhre wird das Plasmagas (Argon) mit einer Flussrate von 12-17 L/min eingeführt, sodass ein tangentialer Gasstrom um die innen liegenden Röhren entsteht. Dazwischen befindet sich ein weiterer Argongasstrom, das Hilfsgas (*auxiliary gas*), mit einem Fluss von etwa 1 L/min. Das Ende der Fackel umgibt eine Radiofrequenz-Spule, welche über 3-5 Windungen verfügt und mit dem Radiofrequenzgenerator als Spannungsquelle verbunden ist [162].



**Abbildung 9:** (a) Schematischer Aufbau der Plasmafackel, modifiziert nach [185]; (b) Temperaturzonen des Plasmas, modifiziert nach [186].

Wird Radiofrequenzenergie, meist mit einer Leistung von 750-1500 W, an die Spule angelegt, oszilliert ein Wechselstrom innerhalb der Spule mit der Frequenz des Generators. Es werden Radiofrequenzgeneratoren mit einer Frequenz von 27 oder 40 MHz verwendet, da diese speziell für die genannte Anwendung reserviert sind. Der oszillierende Strom erzeugt ein starkes elektromagnetisches Feld an der Spitze der Fackel. In diesem Bereich wird zum Entzünden des Plasmas ein Hochspannungsfunke erzeugt, der das Herausschlagen von Elektronen aus einigen Argonatomen des Plasmagases bewirkt. Die Elektronen werden im Magnetfeld beschleunigt und kollidieren mit weiteren Argonatomen, aus denen wiederum Elektronen herausgeschlagen werden, sodass eine Kettenreaktion ausgelöst wird. Diese kollisionsinduzierte Ionisierung führt zur sogenannten ICP-Entladung: die Aufspaltung des Gases in Argonatome, Argonkationen und Elektronen, die das Plasma bilden. Die Plasmaentladung wird so lang aufrechterhalten, wie Radiofrequenzenergie an der Spule anliegt. Zur Ionisierung von Argon wird eine Energie in Höhe der ersten Ionisierungsenergie benötigt, welche 15.8 eV beträgt. Dies ist ausreichend zur Ionisierung fast aller Elemente des Periodensystems, von denen 84 % eine erste Ionisierungsenergie von unter 10 eV aufweisen. Lediglich die ersten Ionisierungsenergien der Elemente Helium, Neon und Fluor liegen oberhalb der von Argon, sodass diese Elemente der Analytik mittels ICP-MS unter Verwendung eines Argonplasmas nicht zugänglich sind. Einige Geräte verfügen zusätzlich über eine *shield torch*, eine dünne zylinderförmige Platinschicht, die das in der Spule liegende Ende der Fackel umgibt. Diese sorgt für eine enge Energieverteilung der Ionen, was im Hinblick auf die Interferenzabtrennung in der Kollisionszelle von besonderer Bedeutung ist [162, 187–189].

Das Plasma besteht aus verschiedenen Temperaturzonen, die in Abbildung 9 b gezeigt sind. Das Probenaerosol verlässt die mittlere Röhre der Fackel, welche auch als Probeninjektor bezeichnet wird, mit einer so hohen Geschwindigkeit, dass es eine Art Tunnel in der Mitte des Plasmas entstehen lässt. Die Analyten liegen zunächst in Tröpfchenform, beispielsweise als gelöstes Salz an das Lösungsmittel gebunden, vor. Gelangen sie ins Plasma, verdampft das Lösungsmittel, sodass kleine Feststoffpartikel entstehen. Diese dringen weiter ins Plasma ein und gehen dabei in den gasförmigen Zustand über, bevor sie atomisiert werden. In der analytischen Zone des Plasmas findet schließlich die Ionisierung der Analyten statt, indem nach Kollision mit den im Magnetfeld beschleunigten Elektronen das Herausschlagen von Elektronen aus den Analytatomen folgt, welche sodann als Kation vorliegen. Die Effizienz der Ionisierung reicht je nach Höhe der Ionisierungsenergie des jeweiligen Elementes von 40 % bis über 95 % und liegt im Mittel bei etwa 80 %. Zu einem geringen Anteil findet auch eine Ionisierung durch Kollision mit Argonionen statt [162, 169, 187, 190].

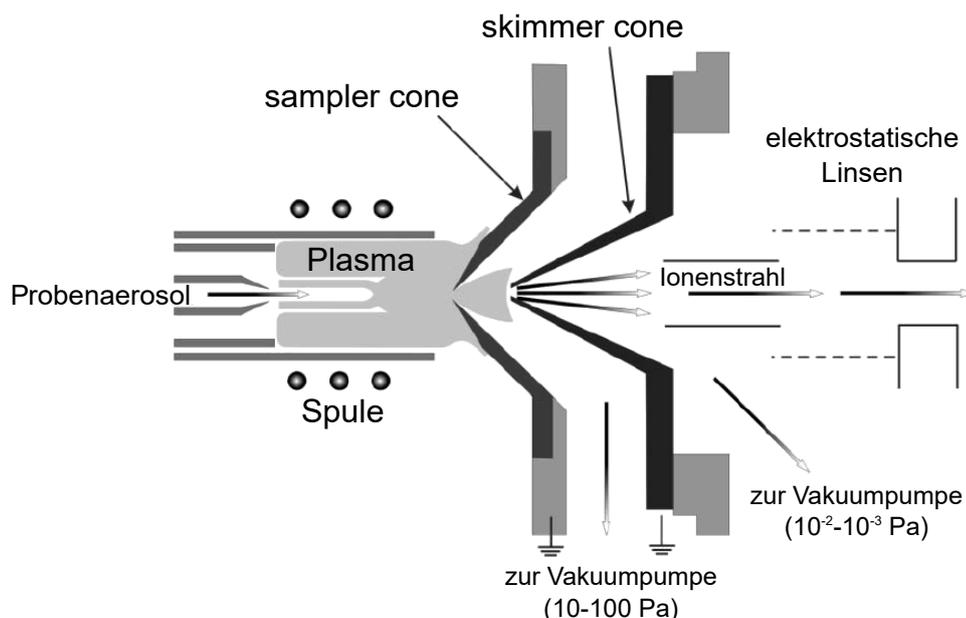
Die Ionisierung der Analyten im Plasma unterscheidet die ICP-MS zentral von der ICP-OES, bei der im Plasma lediglich die Anregung der Analytatome und die

anschließend wieder abgegebene Energie in Form von Photonen (Emission) untersucht wird [191, 192].

### 2.3.3 Ionentransport

Die im Plasma gebildeten Ionen werden anschließend zum Massenanalysator transportiert, wobei sie das *Interface* sowie die Ionenoptik (Linsensystem) passieren. Herausforderung dabei ist der Übergang von Normaldruck (Plasma) zum Hochvakuum (Massenanalysator) [162, 163].

Das *Interface* besteht aus zwei oder drei Konen (*cones*), der *sampler* und *skimmer cone* und ggf. einer *hyper skimmer cone*, die an der Spitze eine sehr kleine Öffnung aufweisen (Abbildung 10). Die *cones* bilden eine Art Zwischenstufe für die zu überwindende Druckdifferenz von etwa einem bar ( $10^5$  Pa) im Bereich des Plasmas über etwa 400 Pa im Bereich zwischen *sampler* und *skimmer cone* und dem Hochvakuumbereich mit einem Druck von bis zu  $10^{-6}$  Pa. Ziel ist es, einen Teil der Ionen als Ionenstrahl effizient und ohne Veränderung der elektrischen Eigenschaften durch das *Interface* hin zum Linsensystem zu transportieren. Die *cones* bestehen meist aus Nickel, für besonders korrosive Proben wie flusssäure- oder königswasserhaltige Lösungen werden dagegen *cones* aus Platin eingesetzt. Zudem werden die *cones* wassergekühlt, um eine Schädigung zu verhindern, da sie aufgrund des direkt anliegenden Plasmas den hohen Temperaturen unmittelbar ausgesetzt sind [162, 163, 193, 194].



**Abbildung 10:** Schematische Darstellung des Übergangs der Analyten vom Plasma durch die *cones* hin zum Linsensystem, modifiziert nach [195].

Die Ionenoptik, die sich direkt hinter der *skimmer cone* anschließt, dient der Fokussierung des Ionenstrahls, wobei Ionen über den gesamten Massenbereich hinweg in einer stabilen Flugbahn gehalten werden müssen. Sie besteht aus einer Kombination verschiedener elektrostatischer Ionenlinsen, die die Ionen zunächst aus dem *Interface* „extrahieren“ und sie beschleunigen. Eine Fokussierung des Ionenstrahls wird durch das Anlegen eines positiven Potentials erreicht. Da die Analytione positiv geladen sind, werden sie abgestoßen und im Zentrum gebündelt. Eine weitere Aufgabe der Ionenoptik ist die Abtrennung von Partikeln, Neutralteilchen und Photonen, welche als Nebenprodukte im Plasma gebildet werden, da diese zur Verstärkung des Hintergrundrauschens führen. In der Praxis wird dies häufig durch den Einsatz einer weiteren Linse (*off-axis lens*) umgesetzt, mithilfe derer eine Ablenkung des Ionenstrahls erreicht wird, sodass dieser direkt in die Kollisions-/Reaktionszelle geführt wird. Die Zelle ist daher etwas außerhalb zur Achse des Ionenstrahls angeordnet (*off-axis*). Die nicht geladenen Störteilchen werden dagegen nicht abgelenkt und können somit effizient abgetrennt werden [162, 196–198].

#### 2.3.4 Kollisions- und Reaktionszelle

An die Ionenoptik schließt sich häufig ein Oktupol an, welcher als Kollisions- oder Reaktionszelle eingesetzt wird und der Abtrennung von polyatomischen Interferenzen dient. Dabei handelt es sich um Störionen, die aus mehr als einem Atom hervorgehen und annähernd dieselbe Masse wie ein einatomiges Analytium aufweisen [162].

Der Oktupol besteht aus acht metallischen Stäben, die sich in einer gasgefüllten Kammer befinden, welche je nach verwendetem Gas als Kollisions- oder Reaktionszelle betrieben werden kann. Zur Verwendung als Kollisionszelle wird diese mit einem inerten Gas, häufig Helium, gefüllt. Durch das Einleiten von Helium in die Zelle kommt es zur Kollision der Ionen mit den Heliumatomen, was zur Verringerung der kinetischen Energie der Teilchen führt. Die Häufigkeit des Zusammenstoßes mit einem Heliumatom ist bei polyatomischen Teilchen höher, da diese eine größere Oberfläche aufweisen. Im Vergleich zu den Analytione verlieren sie daher in größerem Maße an kinetischer Energie, sodass eine Unterscheidung beider Teilchen aufgrund ihrer kinetischen Energie ermöglicht wird. Nur die Analytione mit einer höheren kinetischen Energie werden anschließend weiter in Richtung des Massenanalysators transportiert. Dies wird auch als *Kinetic Energy Discrimination* bezeichnet. Die Häufigkeit der Kollisionen kann über die Flussrate des Zellgases gesteuert werden, wobei eine zu häufige Kollision auch zum Verlust von Analytione und somit zu sinkender Sensitivität führt. Durch die Abtrennung der Störteilchen aufgrund ihrer kinetischen Energie ist eine dichte Energieverteilung der Analytione von enormer Bedeutung, um eine effiziente Unterscheidung zu den

Interferenzteilchen zu gewährleisten. Der Einsatz der Kollisionszelle unter Verwendung von Helium als Zellgas wird auch als Helium-Modus bezeichnet. Im Gegensatz dazu wird im Reaktionsmodus ein reaktives Gas ( $\text{H}_2$ ,  $\text{NH}_3$ ,  $\text{O}_2$ ,  $\text{CH}_4$ ) eingesetzt, welches spezifische Reaktionen entweder mit dem Analyten oder einer Interferenz eingeht. Dies setzt die genaue Kenntnis der vorliegenden Interferenzen voraus. Durch die Reaktion mit dem eingeleiteten Gas entsteht ein Teilchen veränderter Masse, sodass Analyt und Interferenz anschließend über das Masse-zu-Ladungsverhältnis aufgetrennt werden können. Im Vergleich zum Quadrupol ist der Oktupol deutlich kleiner, die Zelle kann somit unter höherem Druck betrieben werden, was eine größere Zahl an Teilchenkollisionen ermöglicht. Die beschriebene Kollisions- und Reaktionszelle wird auch als *Octopole Reaction System* bezeichnet [162, 197, 199–202].

### 2.3.5 Massenseparation

Als Massenanalysator für die gebildeten Analytionen wird in der Regel entweder ein Quadrupol oder ein doppelt fokussierendes Sektorfeld verwendet.

Der Quadrupol findet in etwa 90 % aller ICP-MS-Geräte Anwendung und ist damit der wichtigste Massenanalysator [162]. Er besteht aus vier identischen metallischen Stäben meist zylindrischer Form, die parallel zueinander im Quadrat angeordnet sind. An die beiden gegenüberliegenden Stäbe wird jeweils eine positive bzw. negative Gleichspannung angelegt, die von einer Wechselspannung mit Radiofrequenz überlagert ist und jeweils zeitversetzt vorliegt (Phasenverschiebung um  $180^\circ$ ). Ein Analytion kann den Quadrupol nur dann durchlaufen, wenn es eine stabile Flugbahn aufweist. Bei zu großer Amplitude der Schwingungsbewegung eines Ions kollidiert dieses mit den Stäben des Quadrupols, wird dabei entladen und durch die Vakuumpumpe abgesaugt. Daher wird das Verhältnis aus Gleich- und Wechselstrom so eingestellt, dass eine stabile Flugbahn durch das Zentrum der Stäbe ermöglicht wird. Da sich die Ionen je nach Masse sehr unterschiedlich verhalten, muss dieses Verhältnis für jede zu untersuchende Masse separat eingestellt werden. Bei der Analyse einer Probe werden daher die spezifischen Einstellungen für alle vorgegebenen Massen nacheinander durchlaufen, sodass alle Analyten den Quadrupol passieren und anschließend detektiert werden können. Die meisten Quadrupol-basierten Geräte weisen eine Auflösung von 300-400 auf, sodass eine Massendifferenz von 0.7 bis 1.0 amu unterschieden werden kann [162, 197]. Eine Weiterentwicklung der beschriebenen Technik stellen die Triple-Quadrupol-Geräte dar, die seit 2012 eingesetzt werden [203]. Sie verfügen über einen zweiten Quadrupol, welcher sich vor der Kollisions-/Reaktionszelle befindet und als vorgeschalteter Massenfilter fungiert. Dies bietet die Möglichkeit, den Ionenstrahl bereits vor dem Passieren der

Zelle auf Objekte mit der Zielmasse zu begrenzen, sodass Interferenzen mit der gleichen Masse im Oktupol gezielt ausgeschaltet werden können. Aufgrund des vorangegangenen Massenfilters kann der Hintergrund dabei enorm reduziert werden, was zu einer deutlich verbesserten Sensitivität sowie Selektivität der Messungen führt.

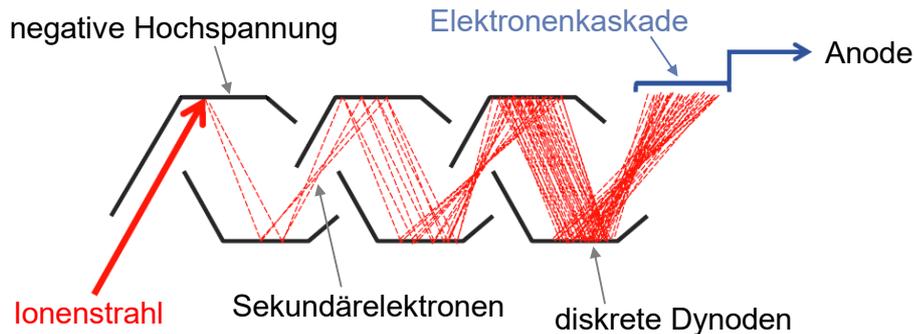
Darüber hinaus findet ein doppelt fokussierendes Sektorfeld Anwendung zur Massenseparation bei hochauflösenden ICP-MS-Geräten. Bei diesem Massenanalysator wird ein Elektromagnet mit einem elektrostatischen Analysator kombiniert. Die Ionen, die zuvor in der Ionenoptik auf einige Kilovolt beschleunigt wurden, passieren den Analysator durch den Eintrittsspalt und gelangen in das Magnetfeld, wo ihre Flugbahn abhängig vom Masse-zu-Ladungsverhältnis unterschiedlich abgelenkt wird. Es findet somit eine Auftrennung der Analyten nach dem  $m/z$ -Verhältnis statt. Die Ionen durchlaufen anschließend den elektrostatischen Analysator, wo durch das Anlegen eines Potentials eine Angleichung der kinetischen Energie der Ionen eines  $m/z$ -Verhältnisses stattfindet. Dies führt zu sehr schmalen Peaks und somit zu einer deutlich verbesserten Auflösung. Die Ionen verlassen den Massenanalysator schließlich über den Austrittsspalt und gelangen zum Detektor. Die Auflösung kann über die Breite von Eintritts- und Austrittsspalt variiert werden. Da eine höhere Auflösung (schmalere Spalt) auch eine geringere Sensitivität mit sich bringt, muss die verwendete Auflösung auf die jeweilige Nachweisgrenze abgestimmt sein. Geräte dieser Bauart ermöglichen die Messung mit einer Auflösung von bis zu 10'000 [162, 197, 204, 205].

Neben Quadrupol und Sektorfeld existieren vereinzelt auch Geräte, bei denen eine TOF-Röhre als Massenanalysator eingesetzt wird [206].

#### 2.3.6 Detektion

Die Detektion der im Massenanalysator aufgetrennten Ionen erfolgt mithilfe eines Sekundärelektronenvervielfachers (SEV). Dieser besteht aus 12-24 einzelnen Dynoden mit einer Oberflächenbeschichtung aus Metall oder einem Halbleiter, die in zwei gegenüberliegenden Reihen oder in einer Röhre angeordnet sind, was in Abbildung 11 dargestellt ist. Da der Detektor auch sensitiv gegenüber Photonen ist, wird er nicht in der gleichen Achse wie der Massenanalysator, sondern etwas versetzt zu diesem, angeordnet, um das Hintergrundrauschen möglichst gering zu halten. An die erste Dynode liegt ein hohes negatives Potential an, sodass die positiv geladenen Analytionen von dieser angezogen werden. Beim Auftreffen der Ionen auf der Dynode werden Elektronen aus den Atomen der Dynodenoberfläche herausgeschlagen, die sogenannten Sekundärelektronen. An die letzte Dynode liegt kein Potential an, sodass sich ein elektrisches Feld aufgrund der Spannungsdifferenz zwischen erster und letzter Dynode ausbildet. Die Elektronen

werden daher zur zweiten Dynode hin beschleunigt und beim Auftreffen werden weitere Sekundärelektronen emittiert. Dieser Prozess wiederholt sich bis zur letzten Dynode, sodass eine Elektronenkaskade entsteht und das relativ schwache Signal des Ionenstrahls in ein elektrisches Signal umgewandelt und effektiv verstärkt wird. Der elektrische Impuls wird schließlich von einer Anode aufgenommen. Die Messung dieses Impuls-Signals wird auch als Puls-Modus oder *pulse-counting mode* bezeichnet, wovon sich die Maßeinheit des detektierten Signals als *counts per second* (cps) ableitet [162, 207–210].



**Abbildung 11:** Schematische Darstellung der Funktionsweise eines Sekundärelektronenvervielfachers (SEV), modifiziert nach [211].

Der Detektor im *pulse-counting mode* ist für Messungen im Bereich von 0 bis  $10^6$  cps geeignet. Um den Detektionsbereich hin zu höheren Konzentrationen zu erweitern, wird der beschriebene Aufbau um die Messung im sogenannten Analog-Modus ergänzt. Dabei wird das Ionensignal zusätzlich an der mittleren Dynode gemessen. Überschreitet das gemessene Signal einen bestimmten Schwellenwert, wird es bereits an dieser Stelle verarbeitet. Liegt es unter dem Schwellenwert, so verläuft die Signalkaskade weiter wie oben beschrieben. Dieser analoge Modus ist für Signale im Bereich von  $10^4$  bis  $10^9$  cps geeignet, sodass insgesamt eine Detektion über 9 Größenordnungen möglich ist. Bei Messungen im Analog-Modus wird die jeweilige Masse im Puls-Modus übersprungen, um den Detektor zu schonen. Beide Messmodi verfügen über unterschiedliche Steigungen bei der Kalibrierung, sodass eine Kreuz-Kalibrierung vorgenommen werden muss, um beide Bereiche aufeinander abzustimmen und Linearität über den gesamten Messbereich zu gewährleisten [162, 212, 213].

Bei neueren Sektorfeld-Geräten wird darüber hinaus ein zusätzlicher Faraday-Detektor eingesetzt, mit dem der Messbereich nochmals um den Bereich von  $10^9$  bis  $10^{12}$  cps auf 12 Größenordnungen erweitert wird. Dabei wird der Ionenstrahl in einen Faraday-Käfig (Faraday-Becher) geleitet, wo er auf eine Elektrode trifft. Durch den Elektronenfluss an der Elektrode werden die Ionen entladen und es kommt zum Spannungsabfall, der proportional zur Zahl der auftreffenden Ionen ist [205, 209, 211].

### 2.3.7 Interferenzen

Bei der Analytik mittels ICP-MS sind eine Reihe von Interferenzen bekannt, die die Quantifizierung des jeweiligen Zielelementes stören können [214]. Man unterscheidet dabei häufig zwischen spektralen und nicht-spektralen Interferenzen [215].

Zu den spektralen Interferenzen zählen alle Störteilchen mit etwa derselben Masse wie der Zielanalyt. Dies ist der Fall bei isobaren Interferenzen, d.h. wenn Isotope unterschiedlicher Elemente eine fast übereinstimmende Masse aufweisen. Die meisten Elemente kommen in Form verschiedener Isotope vor, sodass das analysierte Isotop vor dem Hintergrund solcher Überlagerungen und im Hinblick auf die natürliche Häufigkeit ausgewählt werden muss. Eine Ausnahme bilden die Reinelemente, die nur in Form eines Isotops vorliegen und daher besonders anfällig gegenüber Massenüberlagerungen sind, wie etwa durch polyatomische bzw. Molekül-Interferenzen. Dabei handelt es sich um Ionen, die aus mehreren Atomen bestehen. Häufige Beispiele sind die Bildung von Argiden, Oxiden, Hydriden, Hydroxiden und Chloriden, da diese Elemente bedingt durch das Argonplasma, die umgebende Luft bzw. durch die Zusammensetzung biologischer Proben in hohen Konzentrationen vorkommen. Des Weiteren werden kohlenstoffhaltige Dimere beobachtet, die aufgrund des in der Aufschlusslösung enthaltenen Restkohlenstoffs entstehen. Diese Verbindungen können mit der Masse eines Isotops überlagern, was als positive Interferenz bezeichnet wird. Andererseits können auch negative Interferenzen auftreten, wenn das Analytisorotop an der Bildung einer Molekülinterferenz beteiligt ist und somit ein Teil des Analyten nicht mehr bei der eigentlichen Masse detektiert werden kann. Interferenzen dieser Art können mithilfe der im Abschnitt 2.3.4 beschriebenen Kollisions-/Reaktionszelle effektiv ausgeschaltet werden. Zudem ist das Verhältnis von Analyt und Interferenz ausschlaggebend, welches anhand der Oxidrate (Verhältnis Ce/CeO) beispielhaft im Zuge des *Tunings* vor jeder Messung überprüft wird. In der Regel werden Oxidraten von  $< 1.5\%$  beobachtet, sodass eine Interferenz das Analytsignal nur geringfügig beeinflusst und mittels Kollisionszelle beseitigt werden kann. Problematisch ist zudem die Bildung von doppelt geladenen Ionen im Plasma, da diese durch das detektierte  $m/z$ -Verhältnis ein Signal bei ihrer halben Masse verursachen. Dies tritt besonders bei Elementen auf, deren zweite Ionisierungsenergie sehr niedrig und unterhalb von 15.8 eV liegt. Die verschiedenen spektralen Interferenzen können auch additiv auftreten. Die hier beschriebenen Massenüberlagerungen sind in erster Linie auf die verhältnismäßig geringe Auflösung von Quadrupol-Geräten zurückzuführen, wohingegen bei hochauflösenden Geräten häufig eine Unterscheidung möglich ist [210, 215, 216].

Unter dem Begriff der nicht-spektralen Interferenzen werden Störungen aufgrund von Matrixeffekten zusammengefasst, insbesondere durch Ablagerungen, Verschleppung vorheriger Proben (*memory effect*) oder bedingt durch die physikalischen Eigenschaften

der Matrixkomponenten. Hohe Matrixkonzentrationen können die Analyse in vielfältiger Weise beeinflussen, insbesondere die Aerosolbildung, das Ionisierungsgleichgewicht im Plasma sowie der Probenransport durch *Interface* und Ionenoptik sind betroffen. Je nach vorhandenen Matrixbestandteilen kann es zur Abschwächung oder Verstärkung des Probensignals kommen. Ein hoher Salzgehalt in der Probe kann beispielsweise zu Ablagerungen an den *cones* führen, sodass sich deren Öffnung verkleinert, was eine Abschwächung des Analytsignals zur Folge hat, bis hin zur vollständigen Verstopfung der *cones*. Eine Signalverstärkung wird beispielsweise in Anwesenheit hoher Konzentrationen an Alkali- und Erdalkalimetallen beobachtet. Solche Effekte beeinflussen nicht einzelne Isotope, sondern meist einen größeren Massenbereich. Zur Kompensation von Matrixeffekten werden daher interne Standardelemente bei jeder Messung einbezogen, die eine dem Analyten ähnliche Masse sowie vergleichbare Ionisierungseigenschaften aufweisen sollten [215–219].

---

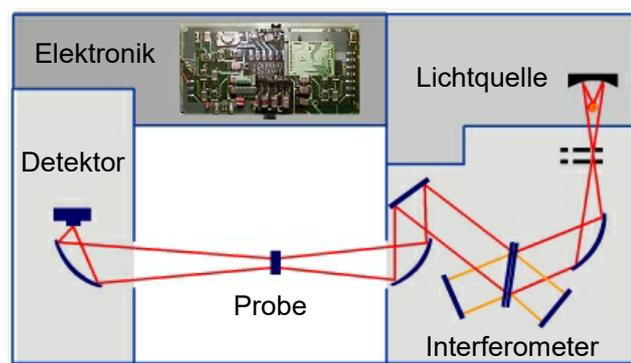
## 2.4 FT-NIR

### 2.4.1 Grundlagen der NIR-Spektroskopie

Bei der NIR-Spektroskopie werden Molekülschwingungen untersucht, die durch Strahlung aus dem nahen Infrarotbereich verursacht werden. Die nahe Infrarotstrahlung (NIR) erstreckt sich über den Wellenlängenbereich von 800-2'500 nm bzw. den Bereich der Wellenzahlen (reziproke Wellenlänge) von 4'000-12'500  $\text{cm}^{-1}$  und grenzt somit an den Bereich der mittleren Infrarotstrahlung (MIR, > 2'500 nm) sowie der sichtbaren Strahlung (< 800 nm). Trifft Infrarotstrahlung auf das zu untersuchende Material, wird ein Teil der Strahlung von der Probe absorbiert, was zur Anregung von Molekülschwingungen führt. Grund dafür ist eine periodische Änderung des Dipolmomentes, sodass nur Moleküle mit einem permanenten elektrischen Dipolmoment für die Infrarotspektroskopie zugänglich sind. Gemäß den Prinzipien der Quantenmechanik können im Molekül nur diskrete Energiebeträge aufgenommen werden. Durch Absorption eines Lichtquants ändert sich der Schwingungszustand eines Moleküls und geht vom Grundzustand (Quantenzahl  $n=0$ ) in einen höheren Zustand über, was eine Schwingung im Molekül auslöst. Beim Übergang zum nächst höheren Zustand mit  $n=1$  spricht man von einer Grundschiwingung. Der Übergang nach  $n=2$  wird als erste Oberschiwingung bezeichnet, der Übergang nach  $n=3$  entsprechend als zweite Oberschiwingung. Der erreichte Zustand wird durch die Energie des Lichtquants beeinflusst, daher werden im Nahinfrarotbereich Oberton- und Kombinationsschwingungen beobachtet, deren Grundschiwingungen im MIR-Bereich liegen. Man unterscheidet zwischen Valenzschwingungen in Richtung der Bindung mit Änderung der Bindungslänge, welche symmetrisch oder asymmetrisch verlaufen können, und Deformationsschwingungen in anderer als der Bindungsrichtung, sowohl innerhalb als auch außerhalb einer Ebene, die mit einer Änderung des Bindungswinkels verbunden sind. Nahinfrarotbanden sind häufig auf Streckschwingungen von C-H-, O-H- und N-H-Bindungen zurückzuführen (Valenzschwingungen). Der nicht von der Probe absorbierte Teil der Nahinfrarotstrahlung wird reflektiert, am Detektor erfasst und anschließend in ein Absorptionssignal umgerechnet [220–226]. Die mittels NIR-Strahlung hervorgerufenen Absorptionsbanden weisen im Vergleich zum MIR-Bereich eine breitere Form auf und überlappen großflächig, sodass eine Zuordnung der Banden zu Molekülen bzw. funktionellen Gruppen nur begrenzt möglich ist. Die Auswertung von NIR-Daten erfordert daher den Einsatz geeigneter statistischer Verfahren.

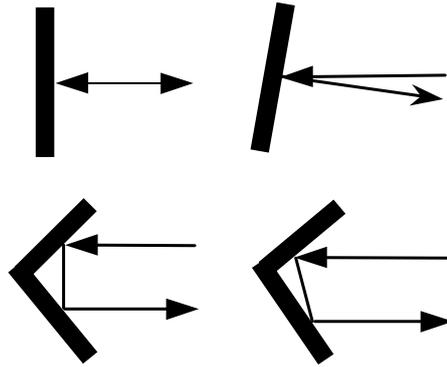
## 2.4.2 Aufbau eines FT-NIR-Gerätes

Für die NIR-Messungen im Rahmen der vorliegenden Arbeit wurde das FT-NIR-Gerät TANGO der Firma Bruker zur Untersuchung des Probenmaterials in fester Form im Reflexionsmodus verwendet. Der Aufbau eines FT-IR-Gerätes ist schematisch in Abbildung 12 dargestellt. Herzstück eines jeden IR-Gerätes ist das Interferometer, welches aus einem System verschiedener Spiegel besteht. Zunächst wird die von der Lichtquelle eintreffende Strahlung am Strahlenteiler (halbdurchlässiger Spiegel) in zwei Einzelstrahlen aufgeteilt, indem ein Teil der Strahlung reflektiert und ein anderer durchgelassen wird. Beide Teilstrahlen treffen nun auf weitere Spiegel und anschließend wiederum auf den Strahlenteiler, an dem sie rekombinieren.



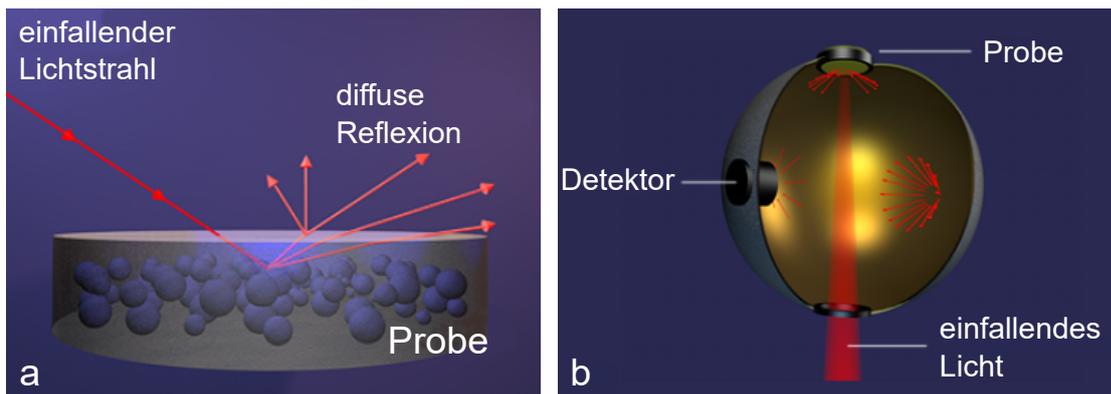
**Abbildung 12:** Prinzipieller Aufbau eines FT-IR-Spektrometers, modifiziert nach [227].

Anders als bei älteren Geräten, bei denen ein klassisches Michelson-Interferometer verwendet wird, verfügen neuere Geräte über ein ROCKSOLID<sup>TM</sup>-Interferometer [228]. Bei diesem werden keine flachen Spiegel sondern sogenannte *CubeCorner*-Spiegel verwendet, die würfelförmig über Eck angeordnet sind und den Vorteil bieten, dass die Richtung der reflektierten Strahlen auch bei Neigung des Spiegels konstant bleibt und dieser daher im Gegensatz zu den flachen Spiegeln nicht exakt ausgerichtet werden muss (vgl. Abbildung 13). Dies führt zu einer deutlich höheren Genauigkeit der Messungen [228, 229]. Während einer dieser Spiegel fest angeordnet ist, bleibt der zweite Spiegel beweglich, sodass die zurückgelegte Weglänge des hier auftreffenden Teilstrahls variiert werden kann. Bei der Rekombination am Strahlenteiler kommt es zur Überlagerung beider Teilstrahlen mit unterschiedlicher Intensität in Abhängigkeit von der Weglänge und der vorhandenen Frequenz, was als Interferenz bezeichnet wird. Der rekombinierte Strahl wird schließlich zur Probe geleitet, wo ein Teil der Strahlung vom Probenmaterial absorbiert, ein anderer reflektiert wird. Die reflektierte Strahlung wird vom Detektor als Interferogramm erfasst und mithilfe der verwendeten Software in ein Absorptionssignal umgewandelt [225, 226].



**Abbildung 13:** Vergleich des Strahlengangs an flachen Spiegeln (oben) und *CubeCorner*-Spiegeln (unten) bei Neigung des Spiegels [228].

Bei der Messung von festen Partikeln wird diffuse Reflexion beim Auftreffen der Strahlung auf die Probe beobachtet. Dabei wird der einfallende Strahl aufgrund der unterschiedlichen Anordnung der einzelnen Feststoffpartikel und des Fehlens einer glatten Oberfläche in verschiedenen Winkeln reflektiert (Abbildung 14 a). Die diffus reflektierte Strahlung wird, wie in Abbildung 14 b dargestellt, mithilfe einer goldbeschichteten Integrationskugel (*integrating sphere*) gesammelt und zum Detektor, einer InGaAs-Dynode, geführt [230–232]. Das gemessene Interferogramm wird anschließend mittels Fourier-Transformation (FT) von einem zeitabhängigen Verlauf in ein frequenzabhängiges Spektrum umgewandelt [226].



**Abbildung 14:** (a) Diffuse Reflexion an Partikeloberflächen, modifiziert nach [233]; (b) Schematischer Schnitt durch eine Integrationskugel, modifiziert nach [234].

### 2.4.3 Datenpräprozessierung

Spektroskopische Messungen basieren auf dem Lambert-Beer'schen Gesetz, wonach die Intensität eines Lichtstrahls beim Probendurchgang durch Absorption des Probenmaterials abgeschwächt wird und Absorption und Konzentration in einem linearen Zusammenhang stehen [235]. Da in der NIR-Spektroskopie verschiedene Phänomene beobachtet werden, die Abweichungen dieses linearen Zusammenhangs verursachen, ist eine geeignete Datenprozessierung fundamental wichtig, um vergleichbare und aussagekräftige Ergebnisse zu erhalten. Den Abweichungen können verschiedenste physikalische und chemische Phänomene zugrunde liegen, wie beispielsweise die Streuung an Probenpartikeln, inhomogene Partikelgrößen, die Packdichte der Partikel, Temperaturschwankungen oder molekulare Interaktionen [222, 236]. Im Spektrum äußert sich dies etwa durch eine Verschiebung einzelner Spektren in Richtung der y-Achse (additive Effekte) oder eine kurvig ansteigende Grundlinie (multiplikative Effekte). Die genannten Phänomene sorgen für eine erhebliche Zunahme der Varianz in den Spektren durch systematisches Rauschen, welches bis zu 99 % der Varianz ausmachen kann, sodass die eigentlichen Probenunterschiede enorm überlagert werden [237]. Daher ist eine Kompensation dieser Effekte im Zuge der Datenvorbehandlung erforderlich. Häufig verwendete Methoden zur Korrektur der Spektren sind Streulichtkorrektur (*Multiplicative Scatter Correction*; MSC), *Standard Normal Variate* (SNV) sowie Ableitungen und Verfahren zur Glättung von Spektren [224, 235, 238].

Die Streulichtkorrektur (MSC) zählt zu den wichtigsten Methoden der Präprozessierung von NIR-Spektren, da die Lichtstreuung an festen Partikeln großen Einfluss auf den Verlauf der Spektren hat [237]. Sie basiert auf der unterschiedlichen Wellenlängenabhängigkeit von Streulicht und absorbiertem Licht. Bei der MSC wird eine Korrektur jedes einzelnen Spektrums unter Zuhilfenahme eines Referenzspektrums, welches die „ideale“ Probe beschreibt, vorgenommen. Als Referenzspektrum dient häufig das Mittelwertspektrum über alle verwendeten Spektren. Das Probenspektrum wird dabei als Linearkombination aus dem „idealen“ Spektrum und einem Korrekturkoeffizienten für additive und multiplikative Effekte ausgedrückt, anhand dessen das korrigierte Spektrum berechnet wird [235, 237, 239]. Die SNV dagegen beruht auf der Standardnormalverteilung und stellt ebenfalls eine Normalisierungsmethode zum Entfernen von Streulichteffekten dar, die ähnliche Ergebnisse wie die MSC erzielt. Die Korrektur jedes einzelnen Spektrums erfolgt, indem von jedem Datenpunkt des Spektrums der jeweilige Mittelwert über alle Spektren subtrahiert und anschließend durch die Standardabweichung aller Spektren geteilt wird [235, 240]. Zusätzlich werden bei der Präprozessierung standardmäßig Ableitungen angewendet, da auf diese Weise sowohl additive als auch multiplikative Effekte erfolgreich korrigiert werden können. Der häufig beobachtete beinahe lineare Basislinienanstieg kann mit der 1. Ableitung entfernt werden, da die

1. Ableitung einer linearen Funktion zu einer Funktion mit einer Steigung von 0 führt. Mithilfe der 2. Ableitung kann die Verschiebung der Spektren gegeneinander in Richtung der y-Achse aufgehoben werden [235]. Die Anwendung von Ableitungen kann zudem dazu beitragen, dass überlappende Peaks besser getrennt werden können. Nachteil der Ableitungen ist jedoch, dass auch spektrales Rauschen in den Spektren verstärkt wird. Aus diesem Grund wird die Ableitung häufig mit einer Methode zur Glättung der Spektren kombiniert, wobei oft der Algorithmus nach Savitzky-Golay angewandt wird. Dabei muss beachtet werden, dass die Glättung (*Smoothing*) das Rauschen effektiv entfernt, eine zu starke Glättung jedoch zum Informationsverlust führt [224].

---

## 2.5 Multivariate Datenanalyse

Sowohl bei den ICP-MS-Messungen als auch bei der Analytik mittels FT-NIR wurde bei jeder Probe eine Vielzahl an Messgrößen erfasst. Das Resultat sind große Datensätze mit zahlreichen untersuchten Variablen, den Isotopen bzw. Wellenzahlen. Bei der simultanen Untersuchung vieler Variablen spricht man von multivariaten Analysemethoden. Für deren Auswertung werden Methoden der multivariaten Datenanalyse benötigt, um Besonderheiten, Zusammenhänge und Muster innerhalb der oft unübersichtlichen Datensätze identifizieren zu können [241]. Die Grundlagen der in dieser Arbeit verwendeten statistischen Methoden seien daher nachfolgend kurz dargestellt.

### 2.5.1 Hierarchische Clusteranalyse und Heatmaps

Unter dem Begriff der Clusteranalyse werden verschiedene Verfahren zusammengefasst, mit denen eine Gruppierung von Datenpunkten erzielt werden soll, wobei ähnliche Objekte in einem gemeinsamen Cluster angeordnet werden. Die Clusteranalyse dient somit der Erkennung von Strukturen innerhalb eines Datensatzes.

Neben anderen Clusterverfahren stellt die hierarchische Clusteranalyse eine nicht überwachte (*unsupervised*) Methode dar, d.h. es werden keine Informationen bezüglich der Anzahl der zu bildenden Cluster vorgegeben. Man unterscheidet dabei zwischen divisiven Algorithmen, bei denen sich anfangs alle Datenpunkte in einem Großcluster befinden, welches anschließend in immer kleinere Cluster aufgeteilt wird, und agglomerativen Verfahren. Letztere gehen von einem einzelnen Datenpunkt aus und clustern diesen mit einem weiteren Datenpunkt mit der geringsten Distanz zum ersten. Dem zugrunde liegt eine Distanzmatrix aller Datenpunkte, zu deren Berechnung verschiedenste Distanzmaße herangezogen werden können, darunter der euklidische Abstand und der Korrelationskoeffizient nach Pearson. Die zwei geclusterten Datenpunkte werden nun als ein neues Objekt betrachtet, die Distanzmatrix darauf basierend neu berechnet und anschließend wird wiederum der Datenpunkt mit der geringsten Distanz zum Cluster kombiniert. Diese Vorgehensweise wird so lang fortgesetzt, bis alle Datenpunkte in ein großes Cluster einbezogen wurden, das Ergebnis wird als Dendrogramm bezeichnet [4, 242, 243].

Neben der Wahl des Distanzmaßes existieren verschiedene Herangehensweisen zur Durchführung des Clusterings und der Anzahl der resultierenden Cluster. Man unterscheidet dabei zwischen *Linkage*- und Varianz-Methoden. Eine häufig verwendete *Linkage*-Methode ist die des *Weighted Average Linkage*, bei der der Mittelwert zwischen den Distanzen aller Paare von Datenpunkten der verschiedenen Cluster betrachtet wird. Eine weit verbreitete Methode stellt zudem das *Complete Linkage* dar, bei dem der

maximale Abstand aller Datenpunkte der verschiedenen Cluster herangezogen wird [4, 244–246].

Die Clusteranalyse wird häufig mit der Darstellung einer *Heatmap* kombiniert, was auch als *Cluster Heatmap* bezeichnet wird. Dabei werden die Objekte eines Datensatzes als unterschiedlich gefärbte Kacheln in einer Matrix dargestellt und an zwei rechtwinkligen Rändern wird ein Clustering der Zeilen bzw. der Spalten des Datensatzes hinzugefügt. Die Färbung der einzelnen Objekte anhand eines Farbverlaufs erfolgt je nach Art der Daten beispielsweise auf der Grundlage eines Konzentrationsprofils. Die *Cluster Heatmap* stellt ein weit verbreitetes Verfahren zur Darstellung großer Datensätze in übersichtlicher und leicht zu interpretierender Form dar [247, 248].

### 2.5.2 Hauptkomponentenanalyse

Die Hauptkomponentenanalyse (*Principal Component Analysis*; PCA) stellt eine weitere Möglichkeit der Gruppierung von Daten sowie der bildlichen Darstellung hochdimensionaler Datensätze dar, indem diese in einen Raum mit geringerer Dimensionalität übertragen werden. Dieser Raum wird durch die Hauptkomponenten (*Principal Components*; PC) beschrieben. Dabei handelt es sich um unkorrelierte lineare Variablen, die Linearkombinationen der ursprünglichen Variablen darstellen und auf der Varianz innerhalb des Datensatzes basieren. Mathematisch betrachtet wird dafür die ursprüngliche Datenmatrix in ein Produkt aus *Scores*- und *Loadings*-Matrix zerlegt. Die Hauptkomponenten werden jeweils so gebildet, dass sie die maximale Varianz beschreiben, die nicht bereits durch eine vorangehende Hauptkomponente beschrieben wurde. Somit ergibt sich ein absteigender Anteil der beschriebenen Varianz von PC1 hin zu den höheren PC. Es existieren so viele Hauptkomponenten wie die Zahl der Variablen im ursprünglichen Datensatz, die Zahl der benötigten PC ist jedoch abhängig vom Anteil der Varianz, der durch sie beschrieben werden soll. Da ein Großteil der Varianz und auch der relevanten Informationen bereits durch die ersten drei Hauptkomponenten beschrieben wird, ermöglicht deren Anwendung eine zwei- oder dreidimensionale Visualisierung und somit eine deutliche Vereinfachung des Datensatzes [4, 241, 249–252].

Bei der PCA handelt es sich um eine Methode des unüberwachten Lernens (*unsupervised learning*), für deren Berechnung keine Klasseninformationen benötigt werden, da die Gruppierung der Daten ausschließlich aufgrund der Varianz erfolgt. Somit ist eine Clusterbildung im *Scores*-Plot, in dem die untersuchten Objekte dargestellt sind, auf Unterschiede zwischen diesen zurückzuführen. Hinweise darauf, welche Variablen den größten Einfluss auf eine beobachtete Gruppierung haben, liefert der *Loadings*-Plot, der die untersuchten Merkmale abbildet. Die Position

der einzelnen *Loadings* geht dabei mit der Bedeutung für die Objekte an gleicher Position im *Scores*-Plot einher. Neben *Scores*- und *Loadings*-Plot ist auch eine Darstellung als Biplot möglich, in dem beide Einzelplots kombiniert werden [4, 250, 253].

### 2.5.3 Partial Least Squares Discriminant Analysis

Bei der *Partial Least Squares Discriminant Analysis* (PLS-DA) handelt es sich um ein überwachtes Lernverfahren (*supervised learning*), welches eine Erweiterung der Hauptkomponentenanalyse darstellt. Sie kann sowohl zur bildlichen Darstellung eines Datensatzes als auch zur Klassifikation herangezogen werden. Im Gegensatz zur PCA sind die Gruppen der Objekte eines Datensatzes bereits bekannt und die Klasseninformationen werden bei der Berechnung miteinbezogen. Aus diesem Grund wird nicht nur die Datenmatrix berücksichtigt, sondern auch eine zweite Matrix, die die abhängigen Variablen, die Informationen über die Gruppierung der Objekte, beinhaltet. Analog zur PCA werden PLS-Komponenten zur Transformation der Daten gewählt, die die maximale Varianz bzw. Restvarianz der Datenmatrix beschreiben, zusätzlich müssen sie jedoch eine Korrelation zur Matrix der abhängigen Variablen aufweisen. Diese abhängigen Variablen sind anders als die gemessenen Datenpunkte Größen, die nicht direkt gemessen werden können und daher mithilfe eines Modells vorhergesagt werden sollen. Die Unterscheidung der Gruppen mittels PLS-DA erfolgt mithilfe einer linearen Trennung des Raumes entsprechend der Position der jeweiligen Objekte [4, 254–259].

### 2.5.4 t-SNE

Das *t-distributed Stochastic Neighbor Embedding* (t-SNE) bezeichnet eine Methode zur nicht-linearen Dimensionalitätsreduktion. Sie eignet sich daher besonders für die Visualisierung hochdimensionaler Datensätze, da eine Darstellung des gesamten Datensatzes in zwei Dimensionen ermöglicht wird. In diesem Punkt unterscheiden sich t-SNE und die zuvor beschriebenen Methoden PCA und PLS, da diese in zwei Dimensionen nur einen Teil des Datensatzes widerspiegeln [260].

Zur Dimensionalitätsreduktion wird der Datensatz in eine Matrix mit paarweisen Ähnlichkeiten bzw. Distanzen (z. B. dem euklidischen Abstand) überführt, die anschließend visualisiert wird. Dazu wird die Ähnlichkeitsmatrix in eine Matrix bedingter Wahrscheinlichkeiten übersetzt, der, im Unterschied zum *Stochastic Neighbor Embedding* (SNE), keine Gauß'sche Wahrscheinlichkeitsverteilung, sondern eine Student-t-Verteilung zugrunde liegt. Ähnliche Objekte weisen dabei hohe Wahrscheinlichkeitswerte auf. Durch die Anwendung der Student-t-Verteilung wird das sogenannte „*Crowding problem*“ ver-

ringert, wonach die korrekte Darstellung von im hochdimensionalen Raum äquidistant verteilten Objekten im Zweidimensionalen in einigen Fällen nicht möglich ist. Die Wahrscheinlichkeitsmatrix wird für die Visualisierung herangezogen und zur bildlichen Darstellung in einen zweidimensionalen Plot übertragen. Dabei werden sowohl lokale als auch globale Strukturen des Datensatzes, wie etwa das Vorhandensein von Clustern, miteinbezogen. Allerdings muss bei der Interpretation der Plots beachtet werden, dass durch den zugrunde liegenden Algorithmus die Größe der einzelnen Cluster und der Abstand zwischen den Clustern verändert werden, sodass diese, im Gegensatz zur PCA, nicht direkt aus dem Plot abzulesen sind [260, 261].

Bei der Berechnung eines t-SNE-Plots muss zum einen die Perplexität (*perplexity*) optimiert werden. Dieser Parameter dient der Abwägung zwischen lokalen und globalen Aspekten der Daten und spiegelt eine Schätzung der Anzahl an unmittelbaren Nachbarn für jeden Datenpunkt wider. Ein weiterer zu berücksichtigender Parameter ist die Anzahl an Durchläufen (*steps*), in denen die Darstellung der Daten optimiert wird [260, 262, 263].

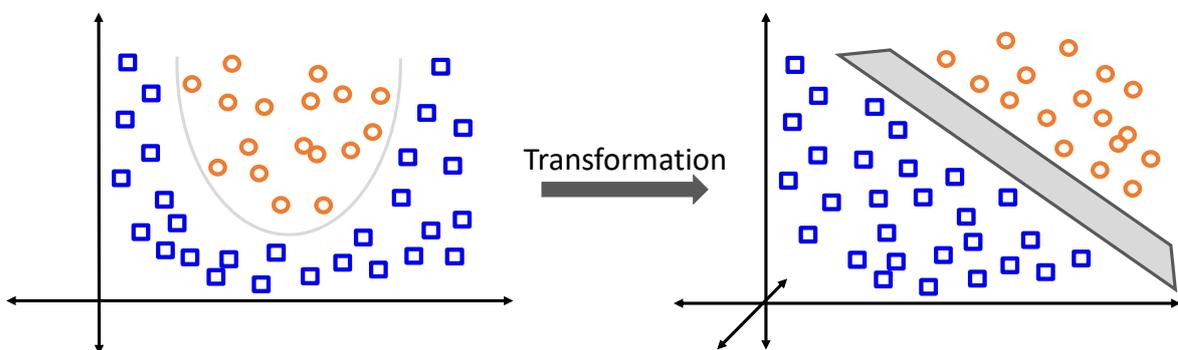
### 2.5.5 Feature Selection mithilfe des Boruta-Algorithmus

Eine große Anzahl untersuchter Messgrößen (*Features*) in einem Datensatz und damit verbunden eine hohe Dimensionalität kann eine Reihe von Herausforderungen mit sich bringen. Die Datenauswertung und Berechnung von Klassifikationsmodellen erfordert mitunter hohe Rechenleistungen, insbesondere wenn die Zahl der *Features* die der untersuchten Proben weit überschreitet. Häufig sind jedoch bei weitem nicht alle *Features* für das jeweilige Trennproblem von Bedeutung, sodass eine Reduzierung des Datensatzes auf die relevanten *Features* eine deutlich kürzere Berechnungszeit mit sich bringen kann. Weitere Vorteile liegen in der Möglichkeit zur Visualisierung sowie im verbesserten Verständnis der Kausalität von Klassifikationsentscheidungen. Durch die *Feature Selection* kann die Genauigkeit eines Klassifikationsmodells verbessert werden, zudem wird das Risiko für ein *Overfitting* des Modells reduziert [264, 265].

Eine weit verbreitete Methode der *Feature Selection* ist der Boruta-Algorithmus. Die Identifizierung der relevanten *Features* gelingt in diesem Fall mithilfe sogenannter „Schattenattribute“. Dem Messwert für jedes *Feature* wird ein solches Schattenattribut hinzugefügt, welches zufällig aus den gemessenen Werten aller Proben für das jeweilige *Feature* entnommen wird. Der Algorithmus prüft nun die Gewichtung des *Features* und vergleicht diese mit dem jeweiligen Schattenattribut. Alle *Features*, die, verglichen mit ihrem Schattenattribut, nicht signifikant wichtiger für die Klassifikationsentscheidung sind, werden entfernt. Die verbleibenden *Features* stellen dementsprechend die für die Klassifikation relevanten Merkmale dar [266].

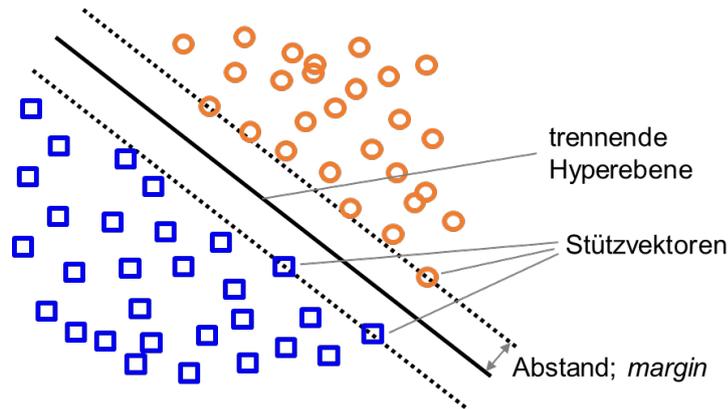
## 2.5.6 Support Vector Machine

Viele Trennprobleme lassen sich nicht mit linearen Methoden lösen, da die fraglichen Klassen eine Überlagerung zeigen und somit nicht durch einfaches Hinzufügen einer Hyperebene vollständig trennbar sind. Eine Möglichkeit der Trennung solcher Klassen bietet die *Support Vector Machine* (SVM). Dieses Klassifikationsverfahren aus dem Bereich des maschinellen Lernens ermöglicht eine Trennung von nicht-linearen Grenzflächen verschiedener Klassen durch die Anwendung unterschiedlicher Kerne. Es handelt sich zudem um eine überwachte Lernmethode (*supervised learning*), da die Klasseninformationen der zur Verfügung stehenden Daten einbezogen werden. Die Grundidee der SVM ist die Übertragung der zu trennenden Daten in einen Raum höherer Dimension, in dem eine lineare Trennung mittels Hyperebene möglich ist, was vereinfacht in Abbildung 15 gezeigt ist. Werden die Daten und die Hyperebene anschließend in den ursprünglichen Raum zurückgeführt, so wird die Hyperebene zu einer nicht-linearen Grenzfläche, in der gezeigten Abbildung nimmt sie beispielsweise die Form einer Hyperbel an. Der ursprüngliche SVM-Algorithmus ist für die Trennung von zwei Klassen vorgesehen, es existieren jedoch verschiedene Ansätze der Erweiterung für Multiklassen-Klassifikationen [267–270].



**Abbildung 15:** Nicht-lineare Klassentrennung mittels *Support Vector Machine*.

Ist eine Trennung von zwei Klassen durch eine lineare Hyperebene möglich, ggf. nach Transformation in einen höherdimensionalen Raum, so wird diese so gewählt, dass sie zu beiden Klassen den größtmöglichen Abstand (*margin*) aufweist, wie in Abbildung 16 dargestellt. Man spricht daher auch von einer *large margin classification*. Auf diese Weise wird eine optimale Einordnung neuer Datenpunkte in eine der zu untersuchenden Klassen gewährleistet. Der Abstand von der Hyperebene zu dem am nächsten gelegenen Datenpunkt bezeichnet man auch als Rand, Punkte, die auf dem Rand liegen, werden als Stützvektoren (*Support Vectors*) bezeichnet. Diese Stützvektoren bestimmen die Funktion der Hyperebene, die der Entscheidung bei der Klassifikation neuer Datenpunkte zugrunde liegt. Zur Berechnung der Stützvektoren werden die Datenpunkte in Form



**Abbildung 16:** Beispielhafte Darstellung von Hyperebene und Stützvektoren bei der *Support Vector Machine*.

von Skalarprodukten ausgedrückt, was in Räumen höherer Dimension jedoch enorme Rechenleistungen erfordert. Um dies zu umgehen, bedient man sich des sogenannten *Kernel-Tricks*, d.h. statt der Berechnung der Skalarprodukte greift man auf eine Kern-Funktion zurück, die sich wie ein Skalarprodukt verhält. Typischerweise verwendete *Kernels* sind der lineare, sigmoide, polynomiale oder der Gauß-Kern (*RBF-Kernel*; *Radial Basis Function*). Neben den Parametern der jeweiligen Funktion wurden weitere Variablen eingeführt, die bewirken, dass der Einfluss einzelner Ausreißer innerhalb des Datensatzes auf die Wahl der Hyperebene abgeschwächt wird [4, 267, 268, 271].

Ein großer Vorteil der SVM liegt in der hohen Generalisierbarkeit der erstellten Modelle, die somit gut geeignet sind für die Anwendung auf neue Daten. Hinzu kommt, dass durch den Algorithmus *Overfitting* vermieden werden kann, was zu robusten Modellen führt. Zudem ist die SVM auch für Datensätze geeignet, bei denen nur kleine Stichproben vorliegen. Als Schwächen der Methode ist zum einen die Wahl des Kerns zu nennen, welche empirisch erfolgt, zum anderen erfordert die Erweiterung des Datensatzes immer auch eine Neuberechnung des Modells [271–273].

### 2.5.7 Random Forest

Der *Random Forest* (RF) ist ein überwachtes (*supervised*) Verfahren des maschinellen Lernens, welches zur Klassifikation genutzt wird. Es beschreibt eine Sammlung von Klassifikatoren mit einer baumartigen Struktur, was den Namen der Methode begründet. Grundlage für die Erstellung eines Klassifikationsmodells ist ein Datensatz mit bekannten Zuordnungen für die Klassen, die mithilfe des Modells vorhergesagt werden sollen. Auf diesen Daten aufbauend lernt das Modell, die Zusammenhänge innerhalb des Datensatzes zu erkennen und die jeweils zugehörigen bekannten Klassen richtig

zuzuordnen, was allgemein allen Methoden des maschinellen Lernens zugrunde liegt. Beim *Random Forest* lernt das Modell anhand von Entscheidungsbäumen, die so erstellt werden, dass sie die bekannte Klassenzuordnung bestmöglich wiedergeben. Jeder einzelne Entscheidungsbaum wird mithilfe von Zufallsvektoren konstruiert und erweitert. Dieser bestimmt an jedem Verzweigungspunkt eine zufällig gewählte Teilmenge der Variablen des Datensatzes, die für die Unterscheidung der Klassen herangezogen wird. Die jeweils für die Klassenunterscheidung am besten geeignete Variable aus dieser Teilmenge wird als neue Verzweigungsstelle in den Entscheidungsbaum eingefügt. Die Größe der Teilmenge kann im Zuge der Modelloptimierung variiert werden, wobei häufig mit der Wurzel aus der Anzahl an Variablen begonnen wird. Auf diese Weise wird eine Vielzahl an Entscheidungsbäumen erstellt, die aufgrund der jeweils ausgewählten Variablen untereinander eine große Varianz aufweisen können. Die Kombination der vielen auf der Grundlage zufällig ausgewählter Variablen erstellten Entscheidungsbäume wird als *Random Forest* bezeichnet [274–277].

Zur Klassifikation einer unbekannt Probe mit den gemessenen Variablen durchläuft diese nacheinander alle Entscheidungsbäume des Modells und wird in jedem Baum einer Klasse zugeordnet. Da die einzelnen Entscheidungsbäume sehr verschieden sein können, sind auch unterschiedliche Klassenzuordnungen wahrscheinlich. Als Klassifikationsergebnis wird daher die Klasse ausgegeben, die in der Summe aller Entscheidungsbäume am häufigsten genannt wurde [274, 277].

Vorteil dieser Methode ist, dass die Entscheidungsbäume leicht zu generieren und zu interpretieren sind, allerdings weist ein einzelner Entscheidungsbaum nur eine geringe Vorhersagegenauigkeit für unabhängige Proben auf. Das Modell profitiert daher von der Vielzahl an Entscheidungsbäumen: je größer die Zahl der Entscheidungsbäume desto robuster ist das Modell und die resultierende Vorhersagegenauigkeit. Bei einer ausreichend großen Anzahl an Entscheidungsbäumen konvergiert zudem der Fehler des Modells, weshalb *Random Forests* recht robust gegenüber *Overfitting* sind [274, 277, 278].

### 2.5.8 Kreuzvalidierung

Nach jeder Erstellung eines Klassifikationsmodells ist eine Validierung vorzunehmen, welche der Überprüfung der Vorhersageleistung und der Abschätzung des Fehlers des Modells dient. Man unterscheidet grundsätzlich zwischen interner und externer Validierung, je nachdem ob die zur Validierung herangezogenen Testdaten aus dem Datensatz zur Modellerstellung oder aus einem gänzlich separaten Datensatz stammen. Zu beachten ist jedoch, dass dieselben Daten nie gleichzeitig zur Modellerstellung (Lernstichprobe; *training fold*) und zur Überprüfung eines Modells (Kontrollstichprobe; *test fold*) heran-

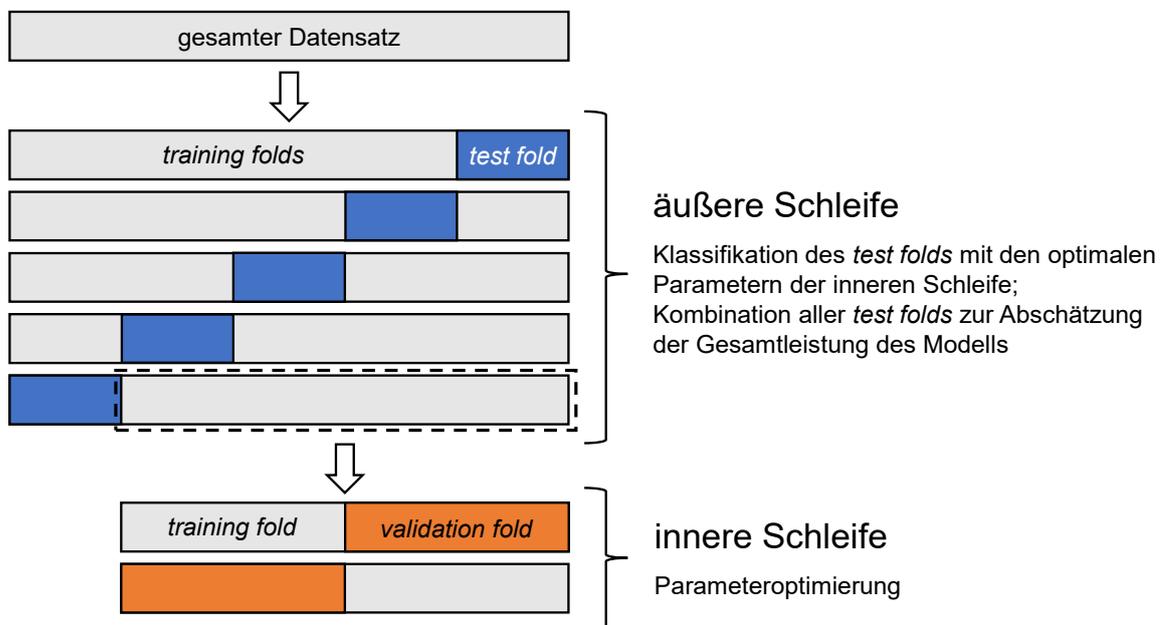
gezogen werden dürfen, um eine echte Validierung des Modells zu gewährleisten [252, 279].

Die  $k$ -fache Kreuzvalidierung ist eine der am häufigsten verwendeten Herangehensweisen zur Validierung von Vorhersagemodellen. Sie ist ein Verfahren der internen Validierung, d.h. die zur Verfügung stehenden Daten werden effizient genutzt, da alle Daten zur Modellerstellung herangezogen werden. Dafür wird der Datensatz zufällig in  $k$  Teile gleicher Größe unterteilt.  $k-1$  Teile werden zur Modellerstellung genutzt, welches anschließend mit dem übrigen Teil getestet wird, indem diese Proben mithilfe des Modells klassifiziert werden. Dieser Prozess wird  $k$ -fach wiederholt, sodass abwechselnd jeder der Teile einmal zur Validierung verwendet wird. Der Anteil der richtig vorhergesagten Proben an der Gesamtprobenzahl wird als Vorhersagegenauigkeit (*prediction accuracy*) bezeichnet [280, 281]. Je nach Unterteilung des Datensatzes können zufällige Effekte auftreten, die das Ergebnis der Validierung beeinflussen. Daher kann die Kreuzvalidierung mehrmals wiederholt werden, jeweils mit wechselnder Unterteilung des Datensatzes, um mit dem Mittelwert der Wiederholungen ein robusteres Ergebnis zu erhalten [282, 283]. Bei der Unterteilung des Datensatzes ist zudem die Anwendung einer Stratifizierung ratsam, um in jedem der  $k$  Teile eine Verteilung der zu untersuchenden Klassen analog zum Gesamtdatensatz abbilden zu können. Nur wenn bei der Modellbildung genügend Informationen über eine Klasse zur Verfügung stehen, kann sichergestellt werden, dass, insbesondere bei einer kleineren Klassengröße, hohe Vorhersagegenauigkeiten für alle Klassen erreicht werden. Die Datensatzteilung sollte daher randomisiert, jedoch mit der Randbedingung einer gleichmäßigen Verteilung der Klassen auf die einzelnen Teile, durchgeführt werden (*stratified random sampling*) [280, 284, 285].

Neben der Überprüfung der Leistungsfähigkeit eines Klassifikationsmodells wird die Kreuzvalidierung auch zur Optimierung von Modellen herangezogen. Dabei werden verschiedene Zahlenwerte für die benötigten Parameter eingesetzt und das resultierende Modell jeweils validiert und die Ergebnisse miteinander verglichen. So können die Parameter identifiziert werden, die die besten Vorhersagegenauigkeiten liefern. Problematisch ist jedoch, dass dabei sowohl die Parameteroptimierung als auch die Validierung mit denselben Daten durchgeführt wird. Dies führt häufig zum *Overfitting* und zur Verzerrung des Modells durch die sehr präzise Ausrichtung an die zugrundeliegenden Daten, was auch als *Bias* bezeichnet wird. Zieht man dagegen unabhängige Daten für die Vorhersage mithilfe des erstellten Klassifikationsmodells heran, werden in der Regel deutlich niedrigere Genauigkeiten erzielt. Die  $k$ -fache Kreuzvalidierung mit integrierter Parameteroptimierung ist daher nur bedingt geeignet, um den Fehler eines Modells korrekt darzustellen [286, 287].

Um dieses Problem zu umgehen und das Modell an unabhängigen Daten zu validieren, gleichzeitig jedoch alle Daten effektiv nutzen zu können, wurde das Konzept der *nested*

*cross-validation* entwickelt. Diese stellt eine Art verschachtelte Kreuzvalidierung dar, die mitunter auch als doppelte Kreuzvalidierung bezeichnet wird. Die Validierung wird dazu in zwei Teilprozesse unterteilt, eine innere und eine äußere Kreuzvalidierung, gezeigt in Abbildung 17. Der Datensatz wird zunächst in  $m$  gleichgroße Teile zerlegt, ein Teil bildet die äußere Schleife (*outer loop*), die übrigen Teile werden in der inneren Schleife (*inner loop*) verwendet. In der inneren Schleife wird zunächst eine Optimierung der Modellparameter mittels  $k$ -facher Kreuzvalidierung, wie zuvor beschrieben, durchgeführt. Das so erstellte Modell wird anschließend mit den Daten der äußeren Schleife validiert, d.h. die Proben der äußeren Schleife werden anhand des in der inneren Schleife optimierten Modells klassifiziert. Da die Daten der äußeren Schleife zu keinem Zeitpunkt zur Modelloptimierung herangezogen wurden, stellen sie unabhängige Daten dar, mit denen die Genauigkeit und der Fehler des Modells wahrheitsgemäß dargestellt werden können. Diese Vorgehensweise wird  $m$ -fach wiederholt, sodass jeder Teil einmal in der äußeren Schleife verwendet und somit jede Probe einmal klassifiziert wurde. Nimmt man alle erhaltenen Klassifikationsergebnisse zusammen, erhält man für jede Probe eine Vorhersage, anhand derer die Genauigkeit des Modells berechnet werden kann. Auch bei diesem Verfahren kann eine mehrmalige Wiederholung des gesamten Prozesses durchgeführt werden, um zufällig auftretende Effekte auszugleichen. Mithilfe der *nested cross-validation* werden sehr robuste Modelle erhalten, die eine hohe Generalisierbarkeit, d.h. Anwendbarkeit auf neue Daten, aufweisen [286–288].



**Abbildung 17:** Ablauf einer *nested cross-validation*; modifiziert nach [289].



## 3 Zielsetzung

Ziel dieser Arbeit war die Methodenentwicklung zur Untersuchung der geographischen Herkunft von weißem Spargel. Als Grundlage der Methode sollten zwei analytische Ansätze dienen, zum einen die Untersuchung des Elementprofils, zum anderen die Messung mittels Nahinfrarotspektroskopie. Im Vordergrund der Untersuchungen standen deutsche Anbauprodukte zum Zweck des Schutzes der Verbraucher sowie der einheimischen Produzenten. Daher wurden Spargelproben aus allen Ländern mit wirtschaftlicher Relevanz für den deutschen Markt einbezogen.

Zur Unterscheidung der geographischen Herkunft von Spargel sollten verschiedene Fragestellungen untersucht werden, darunter die Vorhersage des Herkunftslandes sowie eines deutschen oder nicht deutschen Ursprungs, aber auch eine Zuordnung zu kleinflächigen Herkunftsgebieten. Im Fokus sollte dabei auch die möglichst präzise Herkunftsvorhersage von Spargel aus den Regionen mit einer geschützten geographischen Angabe stehen.

An die zu entwickelnden Verfahren wurden verschiedene Anforderungen gestellt. Im Mittelpunkt standen einerseits präzise Vorhersageergebnisse, zudem waren robuste Methoden gefragt, die eine breite Anwendbarkeit ermöglichen, wozu insbesondere die jahresübergreifende Nutzung zählt. Da bislang wenig Literatur zur Untersuchung der Herkunft von Nutzpflanzen unter Berücksichtigung verschiedener Erntejahre existiert, sollte diese Fragestellung in der vorliegenden Arbeit miteinbezogen werden. Des Weiteren war die Wirtschaftlichkeit der Verfahren, insbesondere bezogen auf Analysenzeiten und -kosten, zu beachten, um eine Übertragung in die routinemäßige Kontrolle der deklarierten Herkunft zu ermöglichen.

Eine besondere Bedeutung sollte der statistischen Auswertung der generierten Daten beigemessen werden. Bislang etablierte Methoden zur Herkunftsbestimmung von Spargel lieferten unzureichende Ergebnisse, wobei das Potential von Methoden des maschinellen Lernens kaum genutzt wurde. Daher galt es zu untersuchen, ob diese Herangehensweise zu einer Verbesserung der Vorhersage der geographischen Herkunft führt.

Im Anschluss an die parallele Entwicklung von Methoden basierend auf ICP-MS- bzw. FT-NIR-Daten sollte zudem ein Vergleich und eine Bewertung beider Herangehensweisen vorgenommen werden.



## 4 Ergebnisse und Diskussion

Teile der hier dargestellten Ergebnisse wurden bereits im Rahmen von wissenschaftlichen Publikationen veröffentlicht [290, 291]:

Richter, B., Gurk, S., Wagner, D., Bockmayr, M., Fischer, M., **2019**, Food Authentication: Multi-elemental analysis of white asparagus for provenance discrimination. *Food Chemistry*, 286, 475–482.

Richter, B., Rurik, M., Gurk, S., Kohlbacher, O., Fischer, M., **2019**, Food monitoring: Screening of the geographical origin of white asparagus using FT-NIR and machine learning. *Food Control*, 104, 318–325.

### 4.1 Methodenentwicklung für die Multielementanalytik

#### 4.1.1 Optimierung der Aufschlussmethode

Zur Vorbereitung der Proben für die Untersuchung des Elementprofils war ein Aufschluss notwendig, da die Proben in flüssiger Form analysiert werden sollten. Die im Absatz 6.2.1 dargestellte initiale Aufschlussmethode für pflanzliche Proben wurde dazu an die genannte Fragestellung angepasst. Ziel war es, einen Totalaufschluss des Probenmaterials zu erreichen und gleichzeitig das Endvolumen und damit verbunden die Verdünnung der aufgeschlossenen Probe zu minimieren, da einige der Zielelemente in sehr kleinen Konzentrationsbereichen zu erwarten waren.

Als Ergebnis der Aufschlussoptimierung wurde das Volumen der zugegebenen Reagenzien von 21 mL, wie in der initialen Methode vorgesehen, auf 10 mL reduziert, wobei 8 mL konzentrierte Salpetersäure (69 %) und 2 mL Wasserstoffperoxid (30 %) eingesetzt wurden. Dies ging einher mit einer geringen Reduktion der Einwaagemenge auf 500 mg, um eine zu heftige Reaktion zu vermeiden. Die in der initialen Methode enthaltenen Zeiten und Temperaturen wurden beibehalten, da diese für einen Aufschluss der gesamten Probe erforderlich waren, sodass die Methode mit einer Dauer von 30 min

unter den genannten Bedingungen einen vollständigen und effizienten Probenaufschluss ermöglichte. Daher konnte auf den Einsatz anderer Säuren, wie Salzsäure oder Schwefelsäure, die mit Schwierigkeiten bei Aufschluss und Messung durch hohe Siedepunkte bzw. die Chlorid-Matrix verbunden ist, verzichtet werden. Nach dem Spülen des Aufschlussgefäßes wurde die Probenlösung auf ein Volumen von 20 mL aufgefüllt und stand direkt für die Messung am ICP-MS zur Verfügung. Durch den erzielten Totalaufschluss und den geringeren Verdünnungsfaktor im Vergleich zur initialen Methode wurde die Voraussetzung für eine sensitive Messung geschaffen, insbesondere für die Elemente im sub-ppb-Bereich wie etwa den seltenen Erden. Nach Abschluss der Methodenoptimierung wurden ausschließlich hochreine Reagenzien verwendet, um das Hintergrundrauschen möglichst gering zu halten.

#### 4.1.2 Entwicklung der Multielementmethode

Zu Beginn der Methodenentwicklung stand die Wahl der zu untersuchenden Analyten. Dafür wurde ein breites Spektrum an Elementen in Betracht gezogen, darunter verschiedene Metalle, Alkali- und Erdalkalimetalle, Halbmetalle und die Metalle der seltenen Erden. Diese Vielzahl von Elementen wurde in eine Multielementmethode aufgenommen, die anschließend optimiert wurde, um möglichst jedes dieser Elemente mit ausreichender Selektivität und Sensitivität messen zu können. Eine Zusammenfassung aller initial untersuchten Elemente ist Tabelle 24 im Anhang im Abschnitt 8.3 zu entnehmen. Die Mengenelemente Natrium, Kalium, Calcium und Magnesium sowie das Element Yttrium wurden nicht einbezogen, da diese eine weit höhere Konzentration als alle übrigen Elemente aufwiesen. Dies hätte eine separate Messung dieser Elemente nach vorangegangener Verdünnung erfordert und zur Verdopplung der Messzeit geführt, was im Hinblick auf die Wirtschaftlichkeit der Methode nicht angemessen ist. Durch Gespräche mit den beteiligten Spargelerzeugern zeigte sich zudem, dass diese die Elemente Kupfer, Mangan, Bor und Eisen sowie die zuvor genannten Mengenelemente in Form von Dünge- und Pflanzenschutzmitteln auf den Spargelfeldern ausbrachten [292–294]. Daraus ergibt sich ein bedeutender Einflussfaktor, der die regional abhängige Elementverteilung verfälschen kann. Ziel dieser Arbeit war es jedoch, ein Modell für die Herkunftsbestimmung von Spargel zu entwickeln, welches möglichst robust und allgemeingültig sein soll. Die Veränderung des Elementprofils durch zusätzlichen Elementeintrag im Anbau stellt dagegen einen anthropogenen Einfluss dar, der in diesem Zusammenhang schwer zu handhaben ist und einem robusten Modell im Wege steht. Daher wurden diese Elemente aus der Methodik ausgeschlossen.

Nach der Wahl der Elemente erfolgte die Auswahl eines oder mehrerer Isotope, die für die Messung herangezogen werden sollten. Entscheidungskriterien waren dabei die natürliche Häufigkeit der einzelnen Isotope, die im Hinblick auf eine hohe Sensitivität ausreichend groß sein sollte, sowie mögliche isobare Interferenzen durch Überlappung der Massenzahlen von Isotopen unterschiedlicher Elemente. Mit dem für die Messungen verwendeten Quadrupol-basierten ICP-MS-Gerät können ganzzahlige Massenzahlen unterschieden werden. Isotope, deren Massenzahlen nur in den Nachkommastellen abweichen, können dagegen nicht getrennt werden.

Neben isobaren Interferenzen können auch Molekülinterferenzen die Quantifizierung der verschiedenen Elemente stören. Solche polyatomischen Interferenzen werden durch das Plasmagas Argon, die in der Umgebungsluft enthaltenen Gase Stickstoff und Sauerstoff, die verwendeten Reagenzien während des Aufschlusses, den Restkohlenstoffgehalt und die spezifischen Matrixbestandteile einer Probe bedingt. Wird eine Interferenz beobachtet, ist das Messen im *NoGas*-Modus nicht mehr ausreichend und das jeweilige Isotop muss im Helium-Modus gemessen werden, d.h. mit aktivierter Kollisionszelle. Während der Methodenentwicklung war daher auch die Optimierung des Heliumflusses notwendig, welches als Kollisionsgas verwendet wurde. Dabei ist es von Bedeutung, ein Gleichgewicht zwischen effizienter Interferenzauslöschung und der Erhaltung ausreichend hoher Countzahlen für alle Analyten zu finden. Während eine erhöhte Heliumflussrate dazu führt, dass die Interferenzteilchen durch häufigere Kollision mit dem Helium stark an kinetischer Energie verlieren und somit gut abgetrennt werden können, wird als Folge der Kollision auch ein Teil der Analyten so stark verlangsamt, dass auch diese Teilchen keine stabile Flugbahn durch den Oktupol aufweisen, mit den Stäben kollidieren und abgesaugt werden. Folglich wird bei steigender Heliumflussrate auch das Analytsignal immer schwächer und unterschreitet ab einem gewissen Punkt die Countzahl, die für eine präzise und sensitive Messung erforderlich ist. Der optimale Gasstrom variiert für jedes Element, sodass in der Multielementmethode eine Näherung für alle Analyten gefunden werden muss. Auf dieser Grundlage wurde eine Heliumflussrate für den Kollisionsmodus von 5.5 mL/min ausgewählt, da bei diesem Fluss sowohl eine gute Interferenzauslöschung als auch eine sensitive Messung möglich war.

Zur Quantifizierung der Analyten wurde eine externe Kalibrierung angewandt. Da sich die Konzentrationen der gewählten Elemente über einen Bereich von sub-ppb bis etwa 1 ppm erstreckten, musste der Kalibrierbereich entsprechend groß gewählt werden. Angepasst an die sehr niedrigen Konzentrationen der seltenen Erden wurde der kleinste Konzentrationspunkt mit 0.001 µg/L gewählt und die höchste Konzentration bei 1000 µg/L festgesetzt, sodass mit 20 Kalibrierpunkten insgesamt ein Konzentrationsbereich von sechs Größenordnungen abgedeckt werden konnte. Zur Erstellung der Kalibrierstandards wurden zwei Multielementstandard-Stammlösungen

käuflich erworben. Im Vergleich zu den übrigen Elementen waren die Analyten Arsen, Beryllium, Bor, Eisen, Selen und Zink in der erworbenen Stammlösung bereits um den Faktor 10 höher konzentriert, was die gleichzeitige Messung unterschiedlicher Konzentrationsbereiche ohne zusätzliche Verdünnungsschritte ermöglicht. Der Kalibrierbereich wurde anschließend für jedes Element separat angepasst, in Abhängigkeit von Nachweisgrenze und linearem Bereich sowie den Elementkonzentrationen in den Proben. Die Quantifizierung der Proben erfolgte anhand der linearen Funktion der Kalibriergerade, welche aus der Auftragung des Verhältnisses aus der Countzahl der Probe und der des internen Standards gegen die Konzentration des Standards resultierte. Einen Überblick über die Zusammensetzung der verwendeten Multielementstandards geben die Tabellen 25 und 26 im Anhang im Abschnitt 8.4.

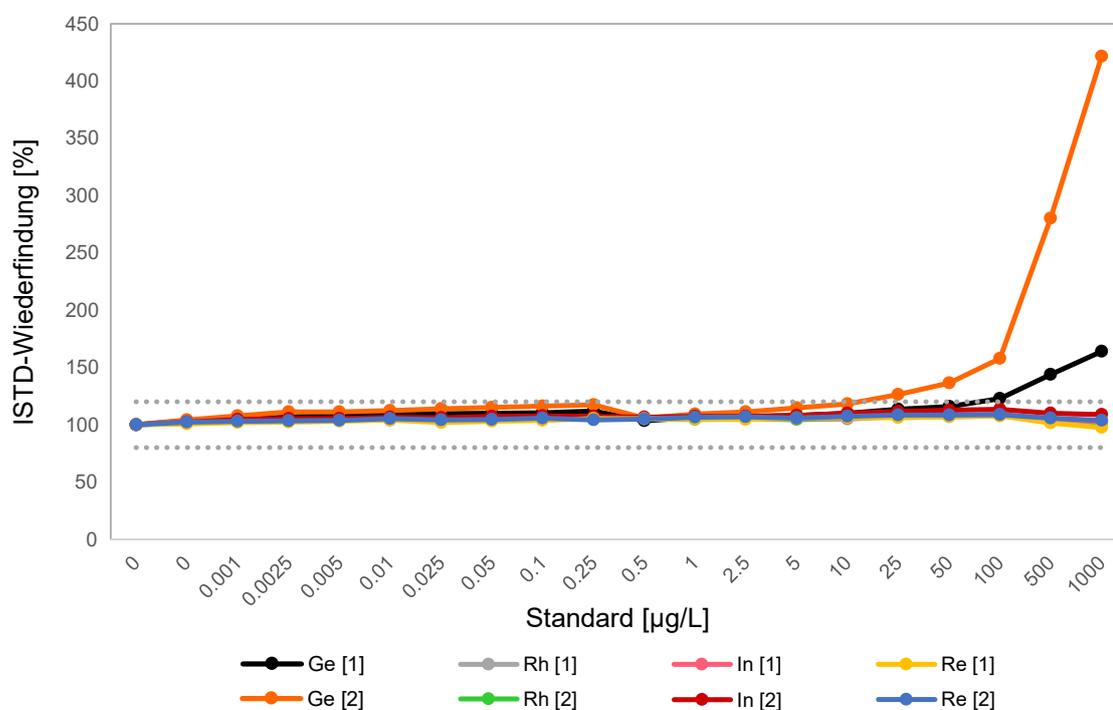
Zum Ausgleich einer möglichen Gerätedrift während der Messung und von Matrixeffekten bedingt durch die Probenmatrix wurden interne Standards (ISTD) verwendet. Dafür wurden die Elemente Germanium, Rhodium, Indium und Rhenium mit den Isotopen  $^{72}\text{Ge}$ ,  $^{103}\text{Rh}$ ,  $^{115}\text{In}$  und  $^{185}\text{Re}$  ausgewählt, da diese natürlicherweise nur in sehr geringen Mengen vorkommen. Jedem Isotop der Multielementmethode wurde nun ein Element als interner Standard zugeordnet. Analyt und interner Standard sollten sich während der Messung und insbesondere während der Ionisierung im Plasma sehr ähnlich verhalten, daher war darauf zu achten, dass beide etwa im gleichen Massenbereich liegen und ähnliche Ionisierungsenergien aufweisen. Die in der Methode enthaltenen Elemente erstreckten sich über einen Massenbereich von Lithium mit einer Masse von 7 als leichtestes bishin zu Uran der Masse 238 als schwerstes Isotop. Die Ionisierungsenergien lagen dabei in einem Bereich von 4.2 - 9.8 eV mit einer mittleren Energie von 6.8 eV. Wie Tabelle 2 zu entnehmen ist, ermöglichen es die gewählten Isotope, den Massenbereich der Analyten und die unterschiedlich hohen Ionisierungsenergien abzubilden. Lediglich für die sehr leichten Elemente wäre ein interner Standard mit niedrigerer Masse wünschenswert gewesen. Der dafür häufig angewandte interne Standard Scandium konnte hier jedoch nicht verwendet werden, da dieses Element bereits als Analyt in der Methode und ebenso in der Multielementstandard-Stammlösung enthalten war. Die internen Standards

**Tabelle 2:** Massenzahl und erste Ionisierungsenergien der als interne Standards verwendeten Elemente.

Element	Massenzahl	1. Ionisierungsenergie
Germanium	72	7.9 eV
Rhodium	103	7.5 eV
Indium	115	5.8 eV
Rhenium	185	7.9 eV

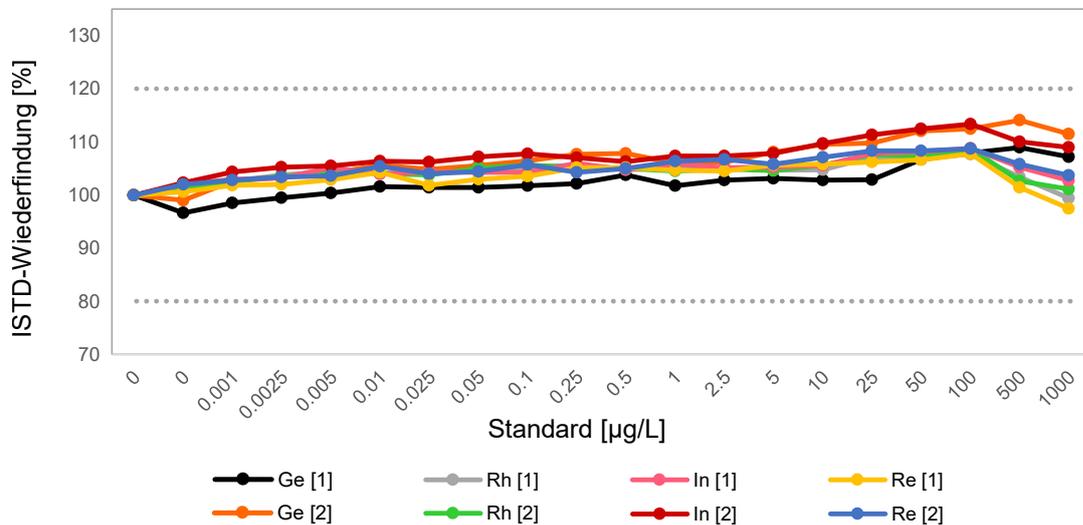
wurden mit einer Konzentration von 1 mg/L über ein Dreiecksventil mit der Analytlösung vereint. Die Konzentration wurde dabei recht hoch gewählt, um ein stabiles Signal des internen Standards zu gewährleisten. Nach den beschriebenen Kriterien wurde für jeden Analyten einer der internen Standardelemente ausgewählt, wobei zumeist ein Abwägen zwischen den Anforderungen erforderlich war, da selten sowohl Massenzahl als auch Ionisierungsenergie exakt zueinander passen. Die Zuordnung für alle Analyten ist im Detail in Tabelle 24 im Anhang dargestellt.

Bei der Messung der Kalibrierreihe war etwa ab dem Kalibrierpunkt der Konzentration 5 µg/L eine Zunahme des Signals für den internen Standard Germanium mit der Masse 72 zu beobachten, was in Abbildung 18 dargestellt ist. Dies war jedoch nur dann zu beobachten, wenn die Metalle der seltenen Erden anwesend waren; in einem weiteren Testlauf ohne diese Elemente war kein derartiger Anstieg zu verzeichnen. Somit war es naheliegend, dass dieser Effekt im Zusammenhang mit den seltenen Erden stehen muss. Da der Standard der seltenen Erden kein Element mit einem Isotop der Masse 72 enthält, kommt als mögliche Erklärung die Bildung von doppelt geladenen Ionen eines Isotops der Masse 144 in Frage, welche dasselbe  $m/z$ -Verhältnis wie das Germaniumisotop der Masse 72 aufweisen. Für die Masse 144 existieren Isotope der Elemente Neodym und Samarium, die beide in der Standardlösung enthalten waren. Die natürliche Häufigkeit dieser Isotope liegt bei 23.8 % (Nd) und 3.1 % (Sm), was eine in diesem Zusammenhang relevante Menge darstellt, um einen solchen Effekt auszulösen.



**Abbildung 18:** Verlauf der internen Standards (ISTD) bei der Messung der Kalibrierreihe in *NoGas*-Modus [1] und Helium-Modus [2] in Anwesenheit aller Standardelemente.

Bei den Standards niedrigerer Konzentration kam dies nicht zum Tragen, was vermutlich darauf zurückzuführen ist, dass der interne Standard mit 1 mg/L eine um Größenordnungen höhere Konzentration aufwies, sodass er dort zu vernachlässigen ist. Im weiteren Verlauf wurde der Standard der seltenen Erden daher nur bis zu einer Konzentration von  $\leq 5 \mu\text{g/L}$  zugegeben, was zu einer sehr guten Signalstabilität für den internen Standard führte (Abb. 19). Bei der Quantifizierung der Analyten wurde kein Isotop der Masse 144 verwendet. Für die Messung des Elementes Neodym wurde dagegen das Isotop  $^{146}\text{Nd}$ , für Samarium  $^{147}\text{Sm}$  herangezogen, da diese Isotope keine isobare Interferenz aufweisen und zudem im Fall des Samariums eine größere natürliche Häufigkeit gegeben ist (14.99%). Da der beschriebene Effekt nur in Anwesenheit großer Konzentrationen an seltenen Erden zu erkennen war, die in den untersuchten Proben jedoch nicht auftraten, und zudem andere Isotope zur Messung herangezogen wurden, gibt es keine Anhaltspunkte für eine fehlerhafte Quantifizierung der Analyten aufgrund des verwendeten Germanium-Standards. Alle übrigen Elemente des internen Standards, Rhodium, Indium und Rhenium, zeigten keine vergleichbare Signalerhöhung, wie den Abbildungen 18 und 19 zu entnehmen ist. Zudem lag die Rate an doppelt geladenen Teilchen, welche vor jeder Messung im Rahmen des *Tunings* ermittelt wurde, im Mittel bei 1.44% im *NoGas*-Modus und bei 1.57% im Helium-Modus, was die Anforderungen an eine selektive Messung erfüllt (vgl. Abschnitt 8.5).



**Abbildung 19:** Verlauf der internen Standards (ISTD) bei der Messung der Kalibrierreihe im *NoGas*-Modus [1] und Helium-Modus [2] in Abwesenheit der Lanthanoide in den Standards  $> 5 \mu\text{g/L}$ .

Ein weiterer bei der Methodenentwicklung zu berücksichtigender Parameter ist die Temperatur der Sprühkammer, welche die feinen Tröpfchen des vom Zerstäuber erzeugten Aerosols weiter in Richtung Plasma transportiert. Wird die Sprühkammer bei

Raumtemperatur betrieben, sind Temperaturschwankungen durch Veränderungen der Raumtemperatur nicht auszuschließen. Dies führt ebenfalls zu Schwankungen beim Probentransport vom Zerstäuber bis zum Plasma und damit verbunden zu Einbußen bei der Signalstabilität und zu einer höheren Oxidrate. Da bei der Aufarbeitung und Analyse der Proben keine organischen Lösungsmittel oder sonstige leicht flüchtige Reagenzien verwendet wurden, war die Kühlung auf Temperaturen unter 0 °C nicht erforderlich, die Sprühkammer wurde daher auf eine Temperatur von 2 °C eingestellt.

Bei der Optimierung der Methode ist zudem die Analysenzeit pro Probe zu beachten, da diese bei einer möglichen späteren Übertragung in den Routinebetrieb einen bedeutsamen wirtschaftlichen Faktor darstellt. Ziel war es daher, die Analysenzeit so kurz wie möglich zu halten. Auf der anderen Seite muss eine ausreichende Spülzeit gegeben sein, um Kontaminationen durch zuvor gemessene Proben zu vermeiden. Zusätzlich sind Stabilisierungszeiten zu beachten, etwa beim Wechsel von *NoGas*- zu Helium-Modus und umgekehrt. Die Dauer der Analyse einer Probe ist zudem abhängig von der Zahl der gewählten Analyten und der Messmodi, da alle Massen in jedem Messmodus nacheinander am Quadrupol aufgetrennt und anschließend detektiert werden. Dazu muss auch die *Integration time*, die Zeit in der die Daten des jeweiligen Isotops akquiriert werden, für jeden Analyten einzeln so gewählt werden, dass eine ausreichende Erfassung des Signals ermöglicht wird (vgl. Tabelle 24 im Anhang). Zur Verkürzung der Analysendauer wurde der Spülvorgang bereits begonnen, während am Detektor die letzten Massen abgetastet wurden. Dies ist möglich, da die Analyten von der Autosamplernadel zum Zerstäuber bis hin zum Plasma einen recht weiten Weg zurücklegen müssen, bevor sie nach dem Passieren des Quadrupols am Detektor erfasst werden können. Dies erstreckt sich über den Zeitraum von etwa einer Minute, sodass eine anfängliche Überlappung von Detektion und Spülvorgang nicht zu einer Verfälschung der Ergebnisse führt. Nach der Optimierung der einzelnen Schritte von Probentransport und Parametereinstellung für den jeweiligen Messmodus konnte eine Analysendauer von 7 min pro Probe erreicht werden, wobei 45 Isotope und 38 Elemente erfasst wurden.

Der HMI-Modus (*High Matrix Introduction*), der eine zusätzliche Verdünnung der Probe durch Zufuhr eines Argon-Gasstroms zwischen Sprühkammer und Plasma ermöglicht, brachte während der Methodenoptimierung keine Verbesserung der Analytik. Dieser ist notwendig, wenn es durch eine hohe Konzentration an gelösten Feststoffen bedingt durch die Probenmatrix zu Ablagerungen bis hin zu Verstopfungen der Konen kommt. Dies wurde bei den Spargelproben nicht beobachtet. Zudem war eine Verdünnung der Probe wegen der niedrigen Gehalte einiger Elemente nicht anzustreben, weshalb der HMI-Modus in der finalen Methode nicht verwendet wurde.

### 4.1.3 Methodenvalidierung

Die zuvor entwickelte Methode wurde einer umfangreichen Validierung unterzogen, um Leistungsfähigkeit und Zuverlässigkeit beurteilen zu können. Die Methodenvalidierung erfolgte in Anlehnung an die DIN 32645, die Richtlinien zur Bioanalytischen Methodenvalidierung der *Food and Drug Administration* (FDA) sowie die Leitlinie zur Methodenvalidierung des Umweltbundesamtes [295–298].

Um die Richtigkeit der Methode bewerten zu können, wurden zunächst zertifizierte Referenzmaterialien mit einer dem Probenmaterial ähnlichen Matrix untersucht. Im Hinblick auf die Spargelproben wurden zwei Materialien pflanzlicher Matrix analysiert, ein Wasserliliengewächs (BCR 670) und ein Gerstengraspulver (DLA 41/2015) [299, 300]. Diese lagen ebenso wie das Probenmaterial in lyophilisierter Form vor, sodass eine Vergleichbarkeit gegeben ist. Die Referenzmaterialien wurden analog zu den Spargelproben aufgeschlossen (s. 4.1.1) und mit den Parametern der optimierten Methode analysiert. Eine detaillierte Gegenüberstellung der zertifizierten mit den analysierten Elementgehalten ist im Anhang in den Tabellen 27 und 28 im Abschnitt 8.6 dargestellt. Die untersuchten Gehalte lagen bis auf wenige Ausnahmen innerhalb der angegebenen Unsicherheit der zertifizierten Werte und zeigten Wiederfindungen zwischen 0.83 und 1.11. Abweichend davon wies Nickel eine Wiederfindung von 0.63 auf, was jedoch im Bereich der Unsicherheit lag. Im Mittel lag die Wiederfindung etwas unterhalb von 1, was mit dem kleinen verbliebenen Rückstand nach dem Aufschluss des Wasserliliengewächses in Zusammenhang gebracht wird. Auch im Falle des Gerstengrases verblieb ein geringer unlöslicher Rückstand, da dieses Silikatstrukturen enthält, die mit der hier gewählten Aufschlussmethodik nicht aufgebrochen werden können [301, 302]. Dies trifft auf den Spargel jedoch nicht zu, da dieser stets vollständig aufgeschlossen wurde. Abgesehen von einer geringfügigen Unterschreitung bei vier der seltenen Erden konnten die zertifizierten Werte trotz eines kleinen Aufschlussrückstands bei allen Elementen erreicht und somit die Anforderungen an die Richtigkeit der Methode erfüllt werden. Ausnahmen stellen dagegen die Elemente Scandium und Selen dar, die stärkere Abweichungen vom zertifizierten Wert und eine Wiederfindung von 1.2 bzw. 2.1 zeigten. Mögliche Gründe dafür können matrixspezifische Interferenzen sein, die die analysierten Massen überlagern und mit den verwendeten Methodenparametern nicht eliminiert werden konnten. Dies wird besonders deutlich bei dem ermittelten Gehalt an Selen, der mit einer Wiederfindungsrate von über 200 % deutlich zu hoch lag. Als mögliche Interferenzen kommen  $^{40}\text{Ar}^{37}\text{Cl}^+$ ,  $^{36}\text{Ar}^{40}\text{Ar}^1\text{H}^+$  oder  $^{38}\text{Ar}_2^1\text{H}^+$  für das Isotop  $^{77}\text{Se}$  bzw.  $^{40}\text{Ar}_2^1\text{H}_2^+$  oder  $^{12}\text{C}^{35}\text{Cl}_2^+$  für das Isotop  $^{82}\text{Se}$  in Frage, da die Elemente Wasserstoff, Kohlenstoff, Argon und Chlor unvermeidlich anwesend sind durch Plasmagas, Umgebungsluft und Analysenlösung einschließlich der pflanzlichen Matrixelemente. Auch für

das Isotop  $^{45}\text{Sc}$  sind die Interferenzen  $^{12}\text{C}^{16}\text{O}_2^1\text{H}^+$ ,  $^{14}\text{N}_2^{16}\text{O}^1\text{H}^+$  und  $^{13}\text{C}^{16}\text{O}_2^+$  durch die bei der Messung anwesenden Elemente denkbar [214]. Da die Gehalte bereits unter Einsatz des Helium-Modus ermittelt wurden und dennoch zu hoch sind, war die verwendete Heliumflussrate von 5.5 mL/min zur vollständigen Interferenzabtrennung nicht ausreichend. Ein möglicher Lösungsansatz wäre daher die Erhöhung der Flussrate. Im Rahmen der beschriebenen Multielementmethode war dies jedoch nicht möglich, da mit dem genannten Heliumfluss für andere Analyten bereits Countzahlen an der Grenze der für eine sensitive Analyse notwendigen Größenordnung erreicht wurden und eine weitere Verringerung der Countzahlen somit nicht vertretbar war. Aus diesem Grund wurden die Elemente Scandium und Selen aus der Multielementmethode ausgeschlossen und blieben bei der Datenauswertung unberücksichtigt. Im Gegensatz dazu konnten für das Element Arsen mit einer Wiederfindung von 100 % sehr genaue Werte ermittelt werden. Bei Arsen handelt es sich um ein sehr interferenzanfälliges Element, da zur Messung dieses Reinelementes ausschließlich das Isotop der Masse 75 zur Verfügung steht, welches stark durch die  $^{40}\text{Ar}^{35}\text{Cl}^+$ -Interferenz gestört wird. Die sehr gute Wiederfindung zeigt hier, dass mit den gewählten Methodenparametern und insbesondere der genannten Heliumflussrate eine sehr gute Interferenzauslöschung möglich ist.

Zur Bewertung der Analyse der Bodenproben wurde außerdem ein Bodenreferenzmaterial (BCR 667, *estuarine sediment*) untersucht [303]. Eine Vergleichbarkeit von zertifizierten und analysierten Werten ist hier jedoch nur bedingt gegeben, da beim Probenaufschluss im Gegensatz zum Referenzmaterial auf einen Totalaufschluss verzichtet wurde. Stattdessen erfolgte eine Extraktion des Bodens zum Aufschluss des bioverfügbaren Anteils, wobei ein silikathaltiger Rückstand verblieb. Im Bodenmaterial liegt ein Teil der Elemente an Silikate gebunden oder in Silikate eingeschlossen vor, die bei einem unvollständigen Aufschluss nicht komplett erfasst werden. Der Vergleich der analysierten mit den zertifizierten Werten zeigte daher wie erwartet Wiederfindungen unterhalb von 1, die bei den Elementen mit zertifizierten Gehalten zwischen 0.49 und 0.90 und im Mittel bei 0.73 lagen (Tabelle 29). Die Wiederfindungen der Elemente mit Richt- und Hinweiswert lagen etwas höher zwischen 0.86 und 1.06. Die Richtigkeit der Methode in Bezug auf die Matrix Boden kann anhand des gezeigten Referenzmaterials aufgrund der abweichenden Probenaufarbeitung nicht abschließend bewertet werden. Die gewählte Methodik eignet sich somit nicht zur Quantifizierung des Elementgehaltes im gesamten Boden. Im Kontext dieser Arbeit war jedoch nicht der absolute Gehalt, sondern vielmehr der bioverfügbare Elementanteil in Relation zu den Gehalten im Spargel von Interesse, weshalb dieser analytische Ansatz dennoch angewandt werden konnte.

Zur weiteren Methodenevaluation wurden mithilfe der Leerwertmethode Nachweis- und Bestimmungsgrenze ermittelt, wobei die dreifache bzw. neunfache Standardabweichung

einer zehnfachen Blindwertmessung als Berechnungsgrundlage diente [295]. Nachweis- und Bestimmungsgrenzen (*limit of detection*; LOD und *limit of quantification*; LOQ) sind im Anhang in Tabelle 30 im Abschnitt 8.6 zusammengefasst. Die Mehrzahl der Elemente zeigte eine Nachweisgrenze im niedrigen bis mittleren ng/L-Bereich. Die höchste Nachweisgrenze lag bei 322 ng/L für das Element Aluminium, der Median über alle Elemente lag bei 2.2 ng/L. Die Elemente der seltenen Erden wiesen sehr geringe Nachweisgrenzen im mittleren pg/L-Bereich auf, die zwischen 78 und 440 pg/L lagen. Insgesamt wurden somit sehr niedrige Nachweisgrenzen erreicht, die für alle Elemente im Hinblick auf den zu erwartenden Konzentrationsbereich ausreichend waren, so dass die gezeigte Methode eine sensitive Bestimmung der gewählten Analyten ermöglicht.

Weiterhin wurde das Bestimmtheitsmaß ( $R^2$ ) zur Beurteilung der Güte der Regression bei der Kalibrierung herangezogen. Zur Erstellung der Regressionsgeraden wurde das Verhältnis aus den Countzahlen von Kalibrierstandard und internem Standard gegen die jeweilige Konzentration aufgetragen. Die Kalibriergeraden aller Elemente wiesen Bestimmtheitsmaße zwischen 0.9993 und 1.0000 auf. Diese hohen Werte zeigen, dass ein sehr guter linearer Zusammenhang zwischen dem Countzahlverhältnis und der Konzentration über einen weiten Konzentrationsbereich besteht und das genannte Verhältnis daher zur Bestimmung der Konzentration in den Proben sehr gut geeignet ist. Eine Auflistung der Bestimmtheitsmaße für alle Elemente findet sich im Anhang in Tabelle 30.

Zur genaueren Untersuchung des linearen Zusammenhangs sowie des Arbeitsbereiches wurde die Linearität der Kalibriergeraden aller Elemente berechnet. Dazu wurde in Anlehnung an die DIN 38402 ein Linearitätstest über die Punkt-zu-Punkt-Steigung durchgeführt [304, 305]. Bei diesem Verfahren wird abschnittsweise aus den Datenpaaren von zwei benachbarten Messpunkten der Kalibriergerade jeweils die Steigung berechnet. Bei der Auftragung der Differenz von Steigung und Median für jedes benachbarte Datenpaar sollte sich ein Verlauf mit Streuung um den Median ergeben, wobei eine Schwankungsbreite von 10 % um den Median akzeptiert wird. Nichtlineare Bereiche an den Enden der Kalibriergerade können dabei anhand einer systematisch zunehmenden Abweichung zum Median identifiziert werden. Die so ermittelten linearen Bereiche, dargestellt in Tabelle 30, erstrecken sich bei allen Elementen über weite Teile des Messbereiches von 0.001 - 1000 µg/L, die die gemessenen Konzentrationen aller Elemente mit einschließen. Abweichend davon endete der Messbereich der seltenen Erden wie zuvor beschrieben bereits bei 5 µg/L (vgl. Abschnitt 4.1.2).

Des Weiteren wurden verschiedene Stabilitätsparameter der Messmethodik beleuchtet. Dies umfasste zum einen die Kurzzeitstabilität, welche auch als Wiederholpräzision oder *repeatability* bezeichnet wird. Diese wurde bestimmt als Standardabweichung bei der wiederholten Messung einer Kalibrierlösung der Konzentration 5 µg/L nach jeder

10. Probe über den Verlauf eines Messtages (ca. 8 Stunden). Die Wiederholpräzision betrug für alle Analyten im Mittel 3 % und reichte von 1.9 % bis 5.9 % (Tabelle 31), sodass man für alle Elemente von stabilen Wiederholmessungen innerhalb eines kurzen Zeitraums sprechen kann.

Zum anderen wurde die Langzeitstabilität für beide untersuchten Matrices bestimmt, die auch als interne Reproduzierbarkeit (*internal reproducibility*) bezeichnet wird. Dafür wurden eine zufällig ausgewählte Spargelprobe und das Boden-Referenzmaterial verwendet (BCR 667), die bei jedem Probenaufschluss erneut mit aufgeschlossen und anschließend vermessen wurden. Über den Verlauf mehrerer Monate wurde daraus die Wiederholbarkeit über einen längeren Zeitraum berechnet. Diese lag für die Matrix Spargel im Durchschnitt über alle Elemente bei 10.2 % und schwankte zwischen 3.7 % und 14.5 % bzw. lag bei 15.9 % für das Element Thorium. Dabei nicht berücksichtigt sind die Elemente Beryllium, Arsen, Silber, Tellur und Bismut, deren Wiederholbarkeit zwischen 28 % und 98 % lag, was durch die sehr niedrigen Gehalte dieser Elemente in der ausgewählten Spargelprobe, insbesondere für Beryllium, zu begründen ist, die unterhalb oder im Bereich der Nachweisgrenze lagen. Das verwendete Material war daher für die Beurteilung der Langzeitstabilität dieser Elemente nicht geeignet. Alle übrigen Elemente lagen im Bereich von  $\pm 15$  %, der nach der Richtlinie der FDA als Akzeptanzkriterium empfohlen wird [296], lediglich die Abweichung des Elementes Thorium lag knapp außerhalb dieses Bereiches. Dies ist ebenfalls auf den niedrigen Thorium-Gehalt zurückzuführen, da die Variabilität im Bereich kleiner Konzentrationen zunimmt. Da diese jedoch nur knapp oberhalb des empfohlenen 15 %-Intervalls lag, wurde die vorliegende Langzeitvariabilität akzeptiert.

Die Wiederholbarkeit für das untersuchte Bodenmaterial lag im Mittel bei 7.6 % und schwankte im Bereich von 5.0 % bis 13.6 %. Einzige Ausnahme bildet das Element Tellur, welches mit 17.8 % eine höhere Abweichung aufweist. Auch hier lag eine geringe Konzentration an Tellur in der Messlösung vor, welche sich im Konzentrationsbereich zwischen Nachweis- und Bestimmungsgrenze befand. Eine Verdünnung der aufgeschlossenen Bodenproben war aufgrund der höheren Konzentrationen der übrigen Elemente notwendig, sodass die Abweichung etwas oberhalb der empfohlenen 15 % an dieser Stelle zu Gunsten der anderen Elemente akzeptiert wurde. Bis auf Tellur erfüllen alle Elemente die genannte Empfehlung, 33 der insgesamt 36 Elemente zeigten sogar eine Abweichung unterhalb von 10 %. Mit Ausnahme von Tellur weisen auch die in der zuvor erwähnten Spargelprobe sehr gering konzentrierten Elemente Beryllium, Arsen, Silber und Bismut Abweichungen unter 15 % auf (vgl. Tabelle 31).

Insgesamt zeigte die Untersuchung der Stabilitätsparameter zufriedenstellende Ergebnisse, sowohl bei der Wiederholung in kurzen Zeitabständen als auch bei Messungen über einen längeren Zeitraum, sodass die Methode nach den zugrunde gelegten Richt- und Leitlinien als geeignet zur Quantifizierung der untersuchten Elemente angesehen wird.

Darüber hinaus konnte die Leistungsfähigkeit der entwickelten Methode im Rahmen von zwei Laborvergleichsuntersuchungen überprüft werden. Als teilnehmendes Labor für die Laborvergleichsuntersuchungen DLA 32/2017 und DLA 46/2017 wurden zwei pflanzliche Materialien, Kartoffelpulver und Wurzelgemüse, nach der gezeigten Methode analysiert und die Ergebnisse mit denen der weiteren Teilnehmer verglichen. Eine Übersicht der Ergebnisse ist den Tabellen 32 und 33 im Anhang zu entnehmen. Für die jeweils untersuchten Elemente wurden aus den Daten aller Teilnehmer der robuste Mittelwert sowie die robuste Standardabweichung ermittelt. Zusätzlich wurde für den analysierten Wert jedes Teilnehmers der entsprechende z-Score berechnet. Dieser ist ein Maß für die Abweichung eines Wertes zum Mittelwert in einer Gruppe aus mehreren Messwerten. Im besten Fall liegt der z-Score bei 0 und ist mit dem Mittelwert identisch. Das Vorzeichen des z-Scores gibt an, in welche Richtung der Wert vom Mittelwert abweicht. Der z-Score wird herangezogen, um das Abschneiden der einzelnen Teilnehmer bewerten zu können, wobei ein Betrag des z-Scores  $\leq 2$  im Allgemeinen als zufriedenstellend erachtet wird [306]. Für die mit der gezeigten Methode analysierten Elementgehalte beider Laborvergleichsuntersuchungen ergaben sich z-Scores von  $0.1 \leq |z| \leq 0.6$ . Dies zeigt eine sehr gute Vergleichbarkeit der vorliegenden Methode mit den Ergebnissen anderer Labore, da die niedrigen z-Scores Ergebnisse mit niedrigen Abweichungen zum Mittelwert aller Teilnehmer anzeigen. Die vorgestellte Methode konnte somit auch in der praktischen Anwendung und im Vergleich zu anderen Laboren eine hohe Leistungsfähigkeit unter Beweis stellen.

Auf der Grundlage der Methodvalidierung erfolgte die Auswahl der Elemente für die finale Multielementmethode. Es wurden dabei nur die Elemente einbezogen, die die Anforderungen an die Validierungsparameter ausreichend erfüllten. Aus diesem Grund wurden die Elemente Selen und Scandium hier nicht weiter berücksichtigt. Mit der endgültigen Methode konnten 36 Elemente mit einer Analysenzeit von 6.5 Minuten erfasst werden, wobei Aufnahme- und Spülzeiten inbegriffen sind. Ein Überblick über die final gewählten Methodenparameter einschließlich der vor jeder Messung im *Tuning* optimierten Parameter ist in Tabelle 17 im Abschnitt 6.3.2 zusammengefasst.

---

## 4.2 Herkunftsbestimmung anhand des Elementprofils

### 4.2.1 Datenpräprozessierung

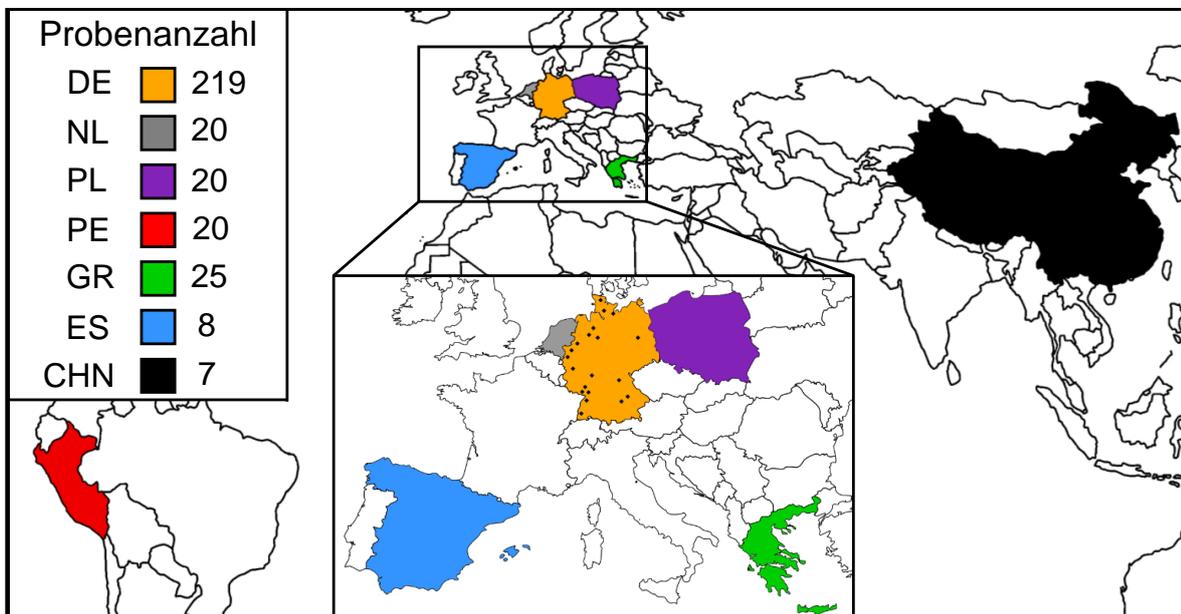
Für die Auswertung der Elementkonzentrationen im Hinblick auf die Herkunftsunterscheidung des Spargels war zunächst eine Präprozessierung der Daten notwendig. Durch die Analyse sehr unterschiedlicher Elemente wurden auch große Unterschiede in den jeweiligen Konzentrationen beobachtet, die von etwa 0.1 µg bis 100 mg je Kilogramm Probenmaterial reichten. Um die Gehalte aller Elemente in verschiedenen Größenordnungen miteinander vergleichen zu können, wurde eine Logarithmierung ( $\log_{10}$ ) vorgenommen.

Für die spätere Erstellung einer *Heatmap* wurde zusätzlich eine z-Standardisierung der Daten durchgeführt. Dazu wurden Mittelwert und Standardabweichung der logarithmierten Daten eines jeden Elementes berechnet. Für jede Probe wurde vom Gehalt eines Elementes zunächst der Mittelwert über alle Proben abgezogen und anschließend durch die Standardabweichung geteilt, was zum z-Score führte. Dieses Vorgehen wurde analog für die Konzentrationen aller Elemente durchgeführt, das Resultat waren z-Scores mit dem Mittelwert 0 und einer Standardabweichung von 1 für jedes Element [252].

Bei der überwiegenden Mehrheit der Datenpunkte lagen die gemessenen Konzentrationen deutlich oberhalb der ermittelten Nachweis- und Bestimmungsgrenzen der Multielementmethode, sodass eine sensitive Bestimmung und Quantifizierung möglich war. Bei wenigen Elementen (As, Be, Bi, Ag, Te, Tl) wurden jedoch Gehalte unterhalb der Nachweisgrenze für einige Proben beobachtet, was in Abbildung 58 im Anhang, Abschnitt 8.6, dargestellt ist. Da ein anderer Teil der Proben jedoch deutlich oberhalb der Nachweisgrenze lag und das jeweilige Element somit durchaus relevant für die genannte Fragestellung sein kann, wurden diese Analyten nicht eliminiert. Die niedrigen Werte unterhalb der Nachweisgrenze können analytisch als eine Konzentration von 0 betrachtet werden. Da Nullwerte jedoch zu Schwierigkeiten bei der Logarithmierung der Daten geführt hätten, wurden diese Werte stattdessen mit der Nachweisgrenze gleichgesetzt.

### 4.2.2 Probenauswahl

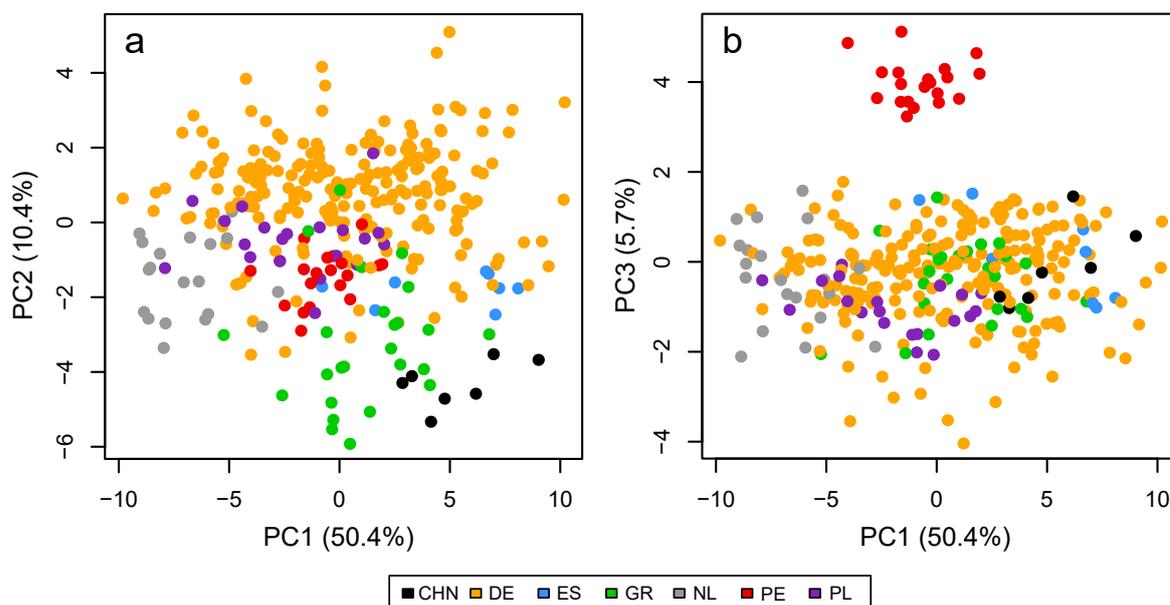
Um Probenunterschiede auf anderer Ebene als dem geographischen Herkunftsgebiet zu vermeiden, wurden alle Proben möglichst gleich behandelt. Trotz dieser Bemühungen konnten Unterschiede bei der Probenkühlung nicht vermieden werden. So erfolgte der Transport der deutschen Proben während der Probenakquise im Jahr 2014 bei  $-20\text{ }^{\circ}\text{C}$ , in den Jahren 2015 und 2016 dagegen bei  $+3\text{ }^{\circ}\text{C}$ . In Bezug auf die Elementzusammensetzung sind jedoch keine Unterschiede zwischen den Proben aufgrund der abweichenden Lagerungstemperatur zu erwarten. In beiden Fällen wurden die Proben auf allen Stufen der Lagerung und Verarbeitung gekühlt, sodass kein Austritt von Flüssigkeit mit darin gelösten Elementen stattfand, wie dies etwa beim Auftauen zu beobachten wäre. Ein möglicher Wasserverlust des bei  $+3\text{ }^{\circ}\text{C}$  gelagerten Spargels aufgrund von Verdunstung ist ebenfalls nicht von Bedeutung, da alle Proben später lyophilisiert und die ermittelten Konzentrationen auf die Trockenmasse bezogen wurden. Daher wurden die Spargelproben aller Erntejahre in die folgende Auswertung miteinbezogen. Eine Übersicht der Proben aus den verschiedenen Herkunftsländern ist in Abbildung 20 gegeben. Insgesamt standen 319 Proben zur Untersuchung der geographischen Herkunft basierend auf dem Elementprofil zur Verfügung, davon 219 aus Deutschland, jeweils 20 aus den Niederlanden, Polen und Peru, 25 aus Griechenland sowie 8 aus Spanien und 7 aus China.



**Abbildung 20:** Anzahl und Herkunft der Proben, die für die Herkunftsbestimmung anhand des Elementprofils herangezogen wurden; schwarze Punkte: Markierung der Ernteorte der deutschen Proben [290].

### 4.2.3 Datenvisualisierung

Zu Beginn der Datenauswertung wurde eine Visualisierung vorgenommen, um einen Überblick über den Datensatz und erste Anhaltspunkte bezüglich der Herkunftsunterscheidung zu gewinnen. Eine anschauliche Untersuchung aller Proben im Hinblick auf deren geographische Herkunft stellt die in Abbildung 21 gezeigte Hauptkomponentenanalyse (PCA) dar. Bei Betrachtung des *Scores Plots* der ersten beiden Hauptkomponenten (Abb. 21 a) ist eine gute Abgrenzung der chinesischen Proben zu erkennen. Zudem sind auch die griechischen und niederländischen Proben jeweils weitestgehend abgetrennt. Die Proben aus Polen, Spanien und Peru überlagern dagegen vollständig mit einem Teil der deutschen Proben. Nimmt man die dritte Hauptkomponente hinzu (Abb. 21 b), so ist eine sehr deutliche Abtrennung der peruanischen Proben sichtbar, während die Proben aus allen übrigen Ländern stark überlappen. Daraus geht hervor, dass die Informationen, die durch die ersten beiden Hauptkomponenten abgedeckt werden, charakteristisch insbesondere für die Länder China und Griechenland sind. Zur Beschreibung der peruanischen Proben ist dagegen die dritte Hauptkomponente von besonderer Bedeutung.

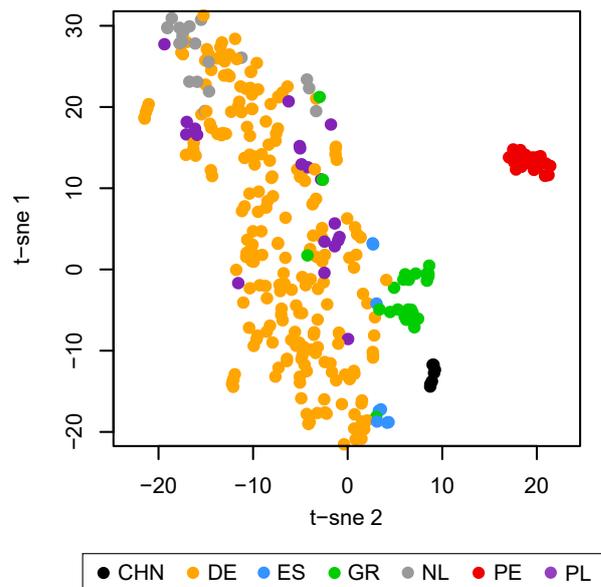


**Abbildung 21:** Hauptkomponentenanalyse mit (a) der ersten und zweiten PC und (b) der ersten und dritten PC zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten.

Des Weiteren ist die Varianz hervorzuheben, die von den drei gezeigten Hauptkomponenten beschrieben wird. Diese beträgt 50.4% (PC1), 10.4% (PC2) und 5.7% (PC3), sodass mit den ersten drei Hauptkomponenten insgesamt 66.5% der Varianz berücksichtigt werden. Dies stellt einen vergleichsweise niedrigen Wert dar, da ein Drittel der Varianz

auf diese Weise nicht erfasst wird. Um 95% der Varianz abbilden zu können, werden dagegen die ersten 16 Hauptkomponenten benötigt, was jedoch schwer darzustellen ist, da die Dimensionalität deutlich zu hoch ist. Um die Varianz innerhalb des beschriebenen Datensatzes möglichst vollständig sichtbar zu machen, erscheint die PCA daher nicht als die geeignetste Methode.

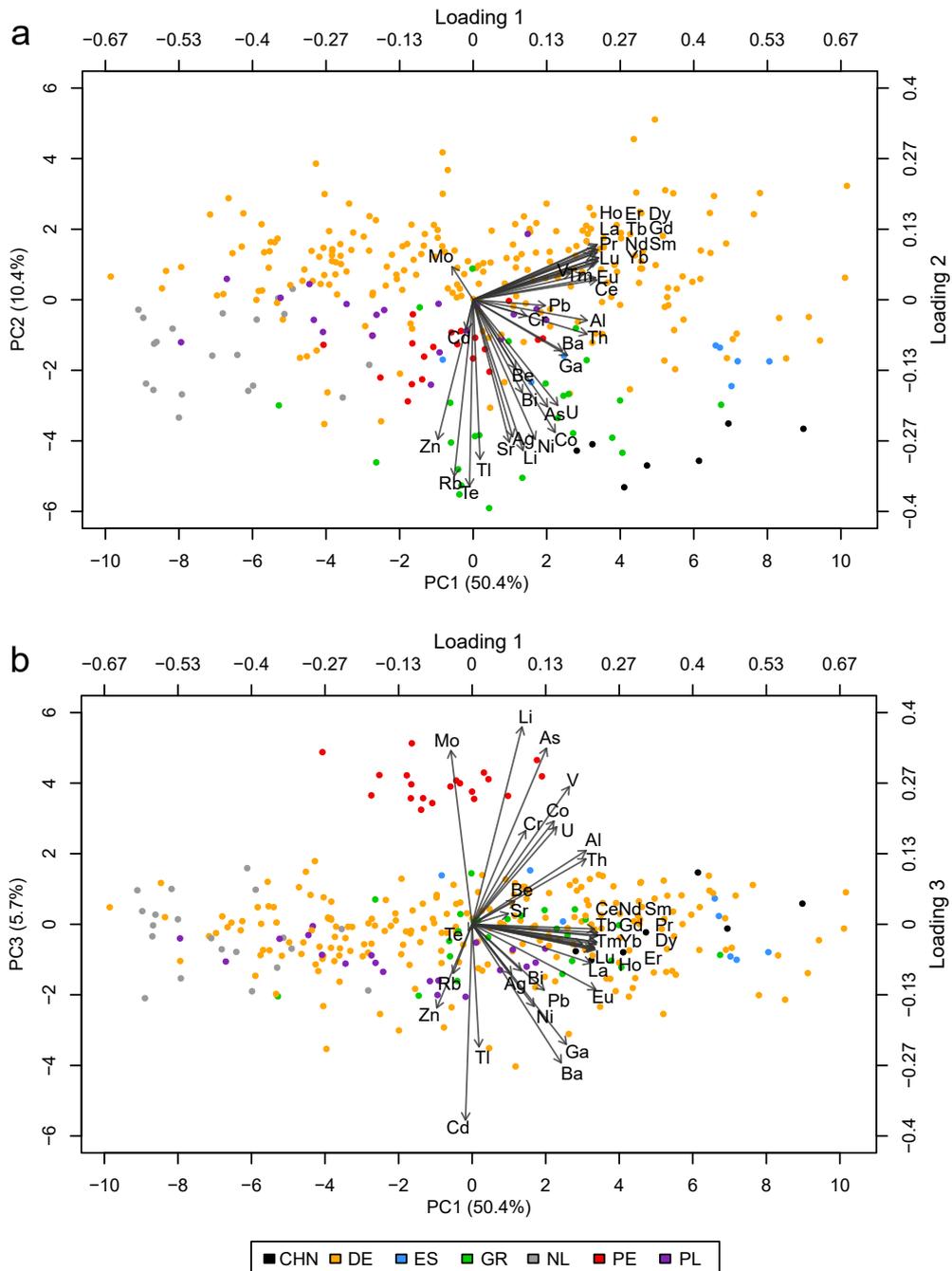
Aus diesem Grund wurde das t-SNE als weiteres Verfahren der Datenvisualisierung eingesetzt. Beim *t-distributed Stochastic Neighbor Embedding* (t-SNE) handelt es sich um eine Methode für die zweidimensionale Darstellung hochdimensionaler Daten, bei der alle Dimensionen des ursprünglichen Datensatzes berücksichtigt werden. Bei der Auswertung ist zu beachten, dass bedingt durch die Datentransformation dichte Cluster ausgedehnt und dünner besetzte eingeeengt werden, sodass im Gegensatz zur PCA die optisch wahrgenommene Größe und der Abstand der Cluster nicht direkt auf den Datensatz übertragen werden können [260, 262]. Abbildung 22 zeigt den t-SNE-Plot für die Unterscheidung der Spargelproben anhand des Herkunftslandes. Noch deutlicher als bei der zuvor gezeigten PCA sind hier isolierte Cluster für die Länder Peru und China sowie Spanien und Griechenland zu erkennen, wobei sich bei den beiden letztgenannten Ländern wenige Proben außerhalb des Clusters befinden. Zudem sind die niederländischen Proben am oberen linken Rand konzentriert, zeigen jedoch eine Überlappung mit einigen deutschen Proben. Einzig die Proben polnischer Herkunft sind etwas weiter verteilt und nur schwer von den deutschen Proben abzugrenzen, was in ähnlicher Weise auch der PCA zu entnehmen ist. Der t-SNE-Plot zeigt somit bereits im zweidimensionalen Raum Möglichkeiten zur Unterscheidung fast aller Herkunftsländer auf, was in der PCA auch mit drei Dimensionen



**Abbildung 22:** Unterscheidung des Herkunftslandes mittels t-SNE basierend auf Multielementdaten.

weniger gut möglich war. Mit dem Wissen der bedingt durch die Transformation verzerrten Abstandsmaße bietet das t-SNE somit eine aussagekräftige und leicht zu interpretierende Verbildlichung des Datensatzes.

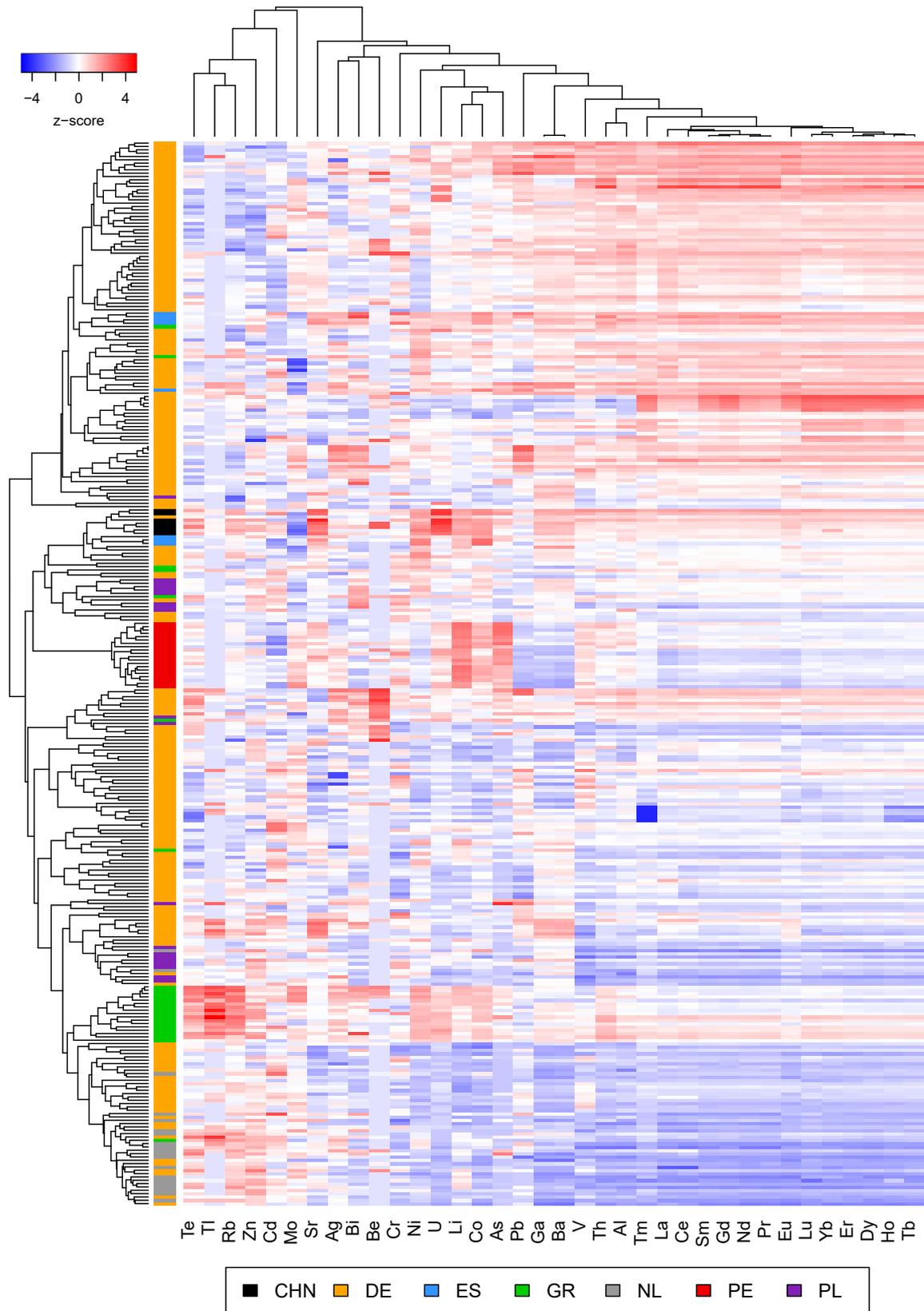
Um erste Rückschlüsse auf die zugrunde liegenden Elemente für die in der PCA und dem t-SNE-Plot beobachtete Auftrennung ziehen zu können, wurde der *Scores Plot* der PCA um den jeweils zugehörigen *Loadings Plot* ergänzt, was in Abbildung 23 in Form von Biplots dargestellt ist. Dabei weist die Pfeilrichtung der *Loadings* (Elemente)



**Abbildung 23:** Biplots der PCA zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten für (a) die erste und zweite sowie (b) die erste und dritte Hauptkomponente.

auf hohe Konzentrationen in den Proben im jeweiligen Areal des Plots hin. Daher können auf der Grundlage der ersten beiden Hauptkomponenten (Abb. 23 a) hohe Konzentrationen an Zink, Rubidium, Tellur, Thallium, Strontium und Lithium in den Proben mit griechischer Herkunft angenommen werden. Die Positionen der Elemente Gallium, Uran, Arsen und Cobalt verweisen dagegen auf charakteristisch hohe Konzentrationen in den chinesischen Proben. Bei Betrachtung der ersten und dritten Hauptkomponente (Abb. 23 b) können zudem Molybdän, Lithium und Arsen als charakteristisch für die Herkunft Peru vermutet werden. Insgesamt sind jedoch nur schwer eindeutige Zusammenhänge zwischen den Herkunftsländern und den jeweils charakteristischen Elementen erkennbar, was sich beispielsweise anhand der unterschiedlichen Positionen des Elementes Lithium in beiden Plots zeigt. Ursache dafür ist zum einen die fehlende scharfe Abgrenzung der einzelnen Probencluster. Hinzu kommt, dass die gezeigten Plots der Hauptkomponentenanalyse nur einen Teil der vielschichtigen Dimensionalität widerspiegeln, was bei der Darstellung als zweidimensionale Abbildung zu Schwierigkeiten bei der Interpretation führt. Zur Identifizierung charakteristischer Elemente ist der *Loadings Plot* der PCA daher nicht ausreichend.

Im nächsten Schritt wurde eine *Heatmap* basierend auf den normalisierten Daten angefertigt. Diese setzt sich zusammen aus einer hierarchischen Clusteranalyse (HCA) der Parameter, hier der Proben (vertikal) und der Elemente (horizontal), und dem zentralen Konzentrationsprofil für alle Elemente und Proben, was in Abbildung 24 dargestellt ist. Das Konzentrationsprofil zeigt einen Verlauf von blau nach rot, d.h. von niedriger zu hoher Konzentration. Die HCA wurde mit dem Korrelationskoeffizienten nach Pearson als Distanzmaß erstellt. Bei Betrachtung der Clusteranalyse der Elemente am oberen Rand fällt eine große Ähnlichkeit zwischen den Metallen der seltenen Erden auf, da diese Elemente ein gemeinsames Cluster bilden, welches sich von allen anderen Elementen stärker unterscheidet. Dies wird ebenfalls im Konzentrationsprofil deutlich, da die seltenen Erden in einer Probe fast immer im gleichen Verhältnis konzentriert sind, was sich durch eine vergleichbare Färbung erkennen lässt. Eine große Ähnlichkeit ist zudem zwischen den Elementen Barium und Gallium erkennbar, die sich ebenfalls durch einen fast identischen Farbverlauf im Konzentrationsprofil widerspiegelt. Daher wird eine starke Korrelation dieser Elemente untereinander vermutet. Anhand der Korrelationskoeffizienten der Elemente, welche in Abbildung 25 als *Heatmap* gezeigt sind, konnte dies bestätigt werden, da die erwähnten Elemente die stärksten Korrelationen untereinander aufweisen (rote Färbung). Ein möglicher Grund für die starke Korrelation von Gallium und Barium könnte eine Interferenz aufgrund von doppelt geladenen Bariumionen sein. Die Bildung von zweifach positiven Ionen des Hauptisotops  $^{138}\text{Ba}$  führt zu einem Signal bei halber Masse, welches mit dem des Isotops  $^{69}\text{Ga}$  überlagert. Da Barium mit 10.0 eV eine sehr geringe zweite Ionisierungsenergie aufweist, ist prinzipiell eine Neigung zur doppelten Ionisierung zu erwarten, was den beobachteten Zusammenhang zwischen Gallium und



**Abbildung 24:** *Heatmap* kombiniert mit hierarchischem Clustering zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten.

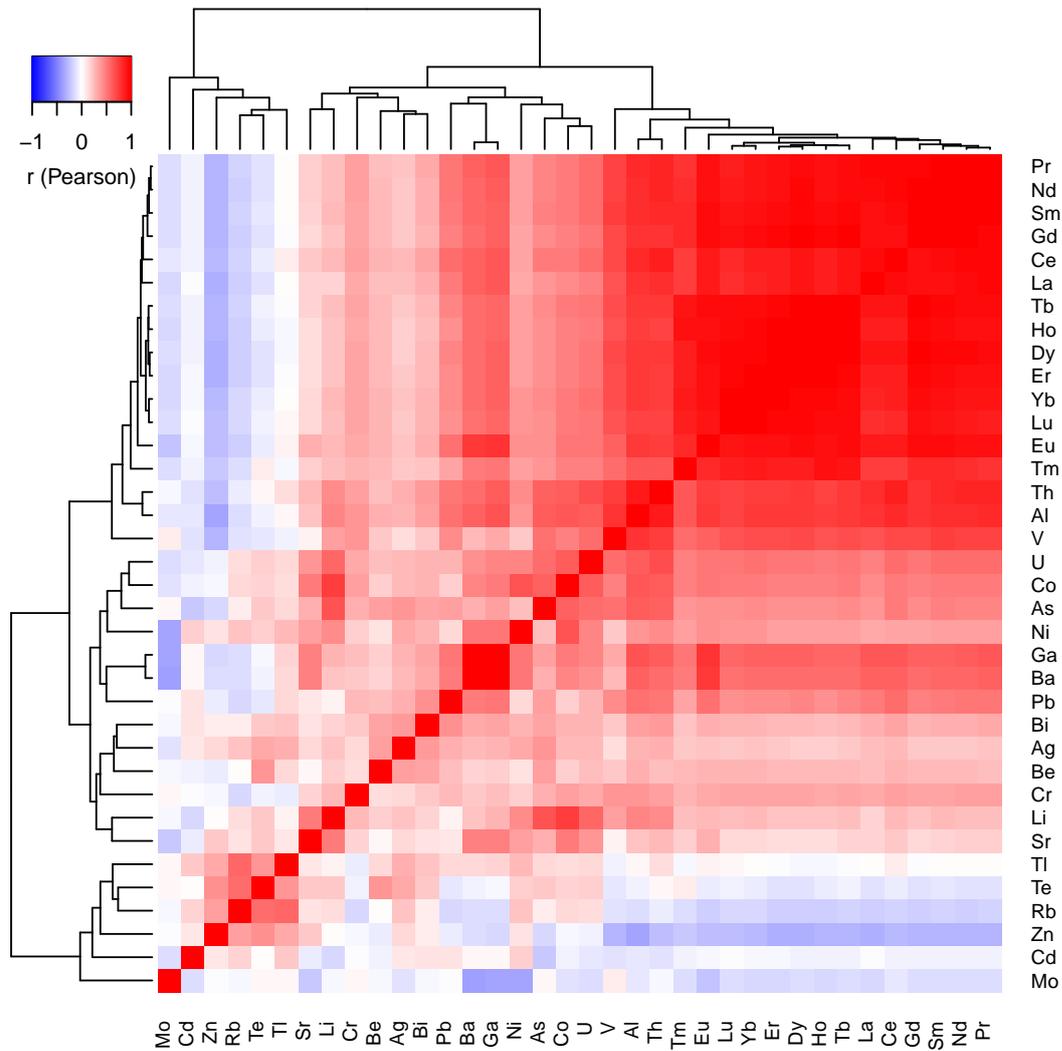
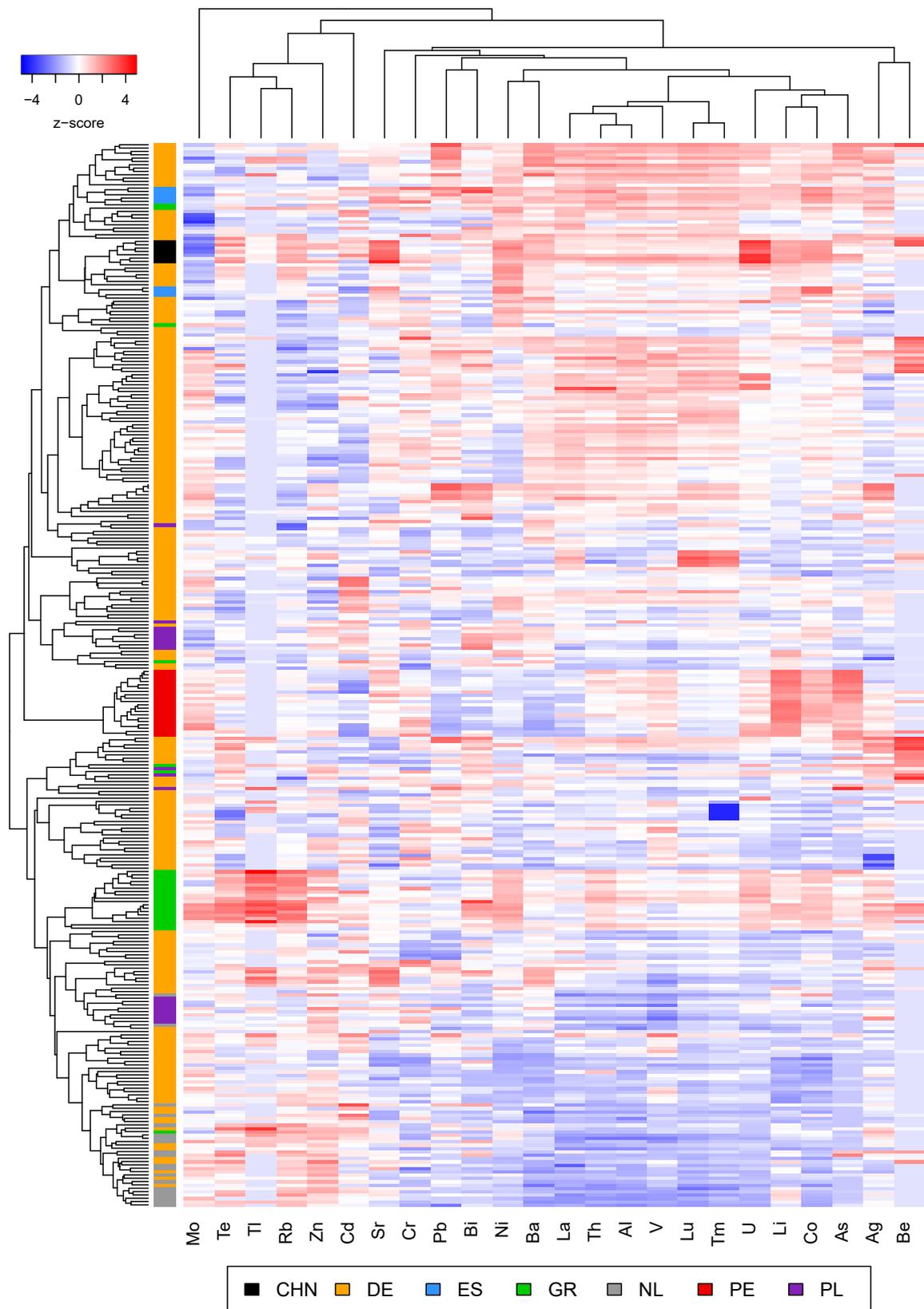


Abbildung 25: Korrelationen nach Pearson aller untersuchten Elemente.

Barium erklären könnte.

Um redundante Daten zu vermeiden, wurde die *Heatmap* neu berechnet, wobei von den beiden erwähnten Gruppen stark korrelierender Elemente (seltene Erden und Gallium/Barium) jeweils nur ein Vertreter berücksichtigt wurde (Abbildung 26). In der HCA der Proben (linker Rand) wird bereits eine gute Auftrennung der Herkunftsländer deutlich. So bilden die in rot dargestellten peruanischen Proben und die schwarz markierten chinesischen Proben jeweils ein separates Cluster, welches alle Proben aus diesem Herkunftsland beinhaltet. Betrachtet man die peruanischen Proben im Elementprofil, so fallen dort recht niedrige Gehalte an Barium, Gallium und Blei auf, während die Elemente Arsen, Cobalt und Lithium hoch konzentriert sind. Im Vergleich mit den übrigen Proben ist der hohe Gehalt an Lithium als besonders charakteristisch hervorzuheben. Auffallend bei den chinesischen Proben sind die sehr niedrigen Molybdän-Gehalte, die jedoch auch bei Proben anderer Herkunft beobachtet werden. Im Gegensatz dazu sind die Strontium- und Uran-Gehalte stark erhöht, was bei den übrigen Proben



**Abbildung 26:** *Heatmap* kombiniert mit hierarchischem Clustering ohne korrelierende Elemente zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten.

nicht der Fall ist. Die griechischen Proben (grün) sind in der HCA nicht komplett isoliert, die Mehrzahl der Proben befindet sich jedoch in einem eigenen Cluster. Dem Elementprofil der griechischen Proben ist eine teils sehr hohe Thalliumkonzentration zu entnehmen, während dieses Element in fast allen anderen Proben nur sehr gering konzentriert ist. Die spanischen Proben (blau) befinden sich im oberen Bereich des Dendrogramms, aufgetrennt in zwei Gruppen, wobei hier zunächst kein Rückschluss auf charakteristische Konzentrationsverteilungen gezogen werden kann. Die in grau gezeigten niederländischen Proben sind überwiegend in einem Cluster am unteren Rand angeordnet, gemischt mit einigen deutschen Proben. Bei den Proben niederländischer Herkunft ist im Konzentrationsprofil ein sehr niedriger Gehalt an seltenen Erden auffällig. Die Proben polnischer Herkunft (lila) lassen sich am wenigsten gut abgrenzen, ein Teil befindet sich in der unteren Hälfte, ein anderer ist in der Mitte des Dendrogramms zu finden. Auch im Konzentrationsprofil sind keine charakteristischen Elementverteilungen zu erkennen. Anhand der gezeigten *Heatmaps* können somit bereits erste eindeutige Zusammenhänge zwischen Elementkonzentration und der geographischen Herkunft abgeleitet werden.

### 4.2.4 Identifizierung herkunftsspezifischer Elemente

Auf der Grundlage erster Hinweise, die aus dem *Loadings Plot* der PCA sowie der *Heatmap* gewonnen wurden, erfolgte nun die Identifizierung von Elementen, die für die unterschiedlichen Herkunftsgebiete charakteristisch sind und die möglicherweise später für eine Herkunftsbestimmung herangezogen werden können.

Die zuvor erwähnten Indizien für charakteristische Elemente wurden mit statistischen Verfahren überprüft und es wurde getestet, ob weitere Elemente für die Herkunftsunterscheidung von Bedeutung sind. Bei der Wahl eines dafür geeigneten statistischen Testverfahrens musste darauf geachtet werden, dass die jeweiligen Voraussetzungen erfüllt sind. Durch die große Zahl an deutschen Proben war für diese Gruppe im Vergleich zu den übrigen Ländern meist auch eine deutlich höhere Varianz zu beobachten. Da die Varianzanalyse (ANOVA) homogene Varianzen innerhalb der Gruppen voraussetzt, konnte dieser Test hier nicht angewendet werden [8]. Durch den vorliegenden Datensatz sind mehrere unabhängige Stichproben gegeben, die Proben der jeweiligen Herkunftsländer, wobei angenommen wird, dass diese sich nur in einem Merkmal, der Herkunft, unterscheiden. Aufgrund der genannten Voraussetzungen wurde der Kruskal-Wallis-Test als geeignetes Verfahren ausgewählt. Mit diesem kann überprüft werden, ob sich für mehrere unabhängige Stichproben zentrale Tendenzen unterscheiden. Die abhängige Variable muss dazu mindestens ordinalskaliert sein, was hier zutreffend ist, da die Elementkonzentrationen

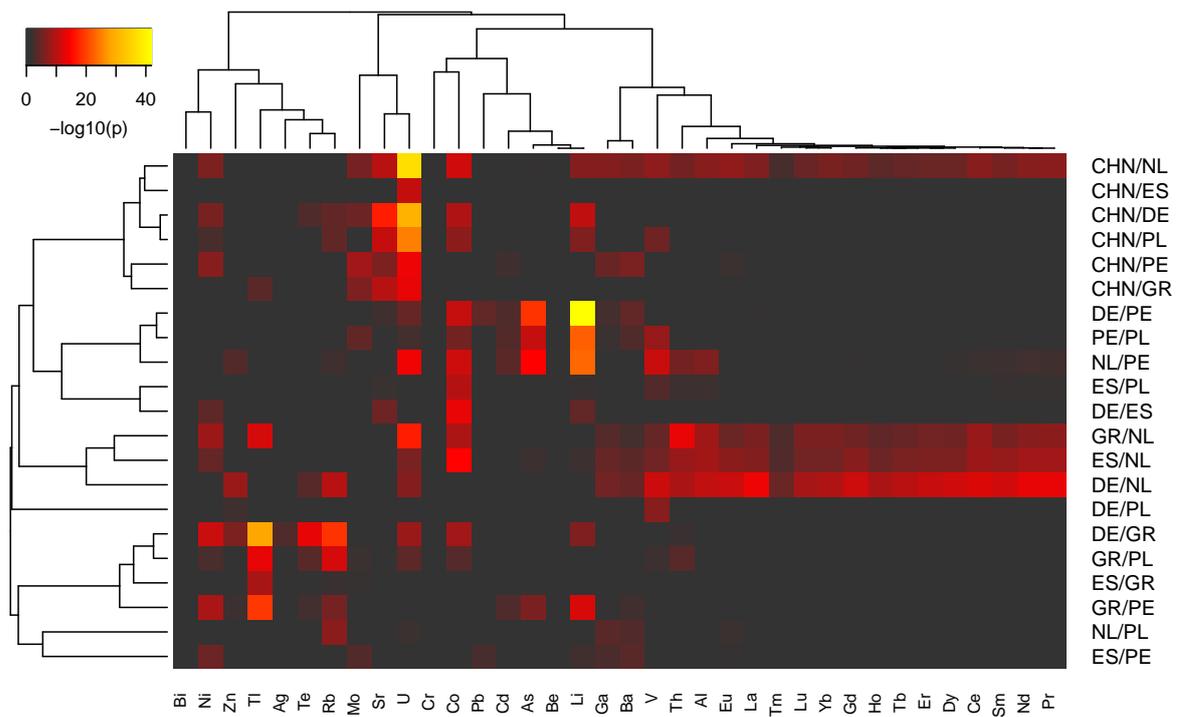
trationen metrisch skalierte Variablen darstellen. Der Kruskal-Wallis-Test ist zudem gut geeignet wenn kleine Stichproben vorliegen, wie es bei den Ländern Spanien und China mit nur 8 bzw. 7 Proben der Fall ist [7]. Es wurde daher mit dem Kruskal-Wallis-Test auf signifikante Konzentrationsunterschiede zwischen den verschiedenen Ländern geprüft. Der errechnete Betrag der Teststatistik ist mit dem kritischen Wert zu vergleichen und überschreitet diesen beim Vorliegen von Unterschieden zwischen den Gruppen. Wenn der Test für die Untersuchung von mehr als fünf Gruppen herangezogen wird, ist von einer asymptotischen Chi-Quadrat-Verteilung auszugehen [7]. Anhand der Teststatistik wird zudem der p-Wert berechnet, der eine Aussage über die Signifikanz erlaubt. Das Signifikanzlevel wurde dazu auf 0.05 festgesetzt.

Beim Testen einer Hypothese, in diesem Fall ob ein Element signifikante Unterschiede zwischen mindestens zwei Ländern aufweist, liegt die Wahrscheinlichkeit eines signifikanten Ergebnisses bei dem gewählten Signifikanzniveau von 0.05 auch bei zufälliger Verteilung der Daten bei 5%, was auch als  $\alpha$ -Fehler bezeichnet wird. Dagegen wird beim multiplen Testen, d.h. beim Testen von mehr als einer Hypothese auf demselben Datensatz, in diesem Fall das Testen mehrerer Elemente, eine ansteigende Fehlerwahrscheinlichkeit beobachtet, je mehr Hypothesen getestet werden. Man spricht daher auch von der  $\alpha$ -Fehler-Kumulierung. Dies führt dazu, dass auch bei nicht signifikanten Unterschieden die Wahrscheinlichkeit steigt, dass der Test fälschlicherweise eine Signifikanz anzeigt und somit zu einem falsch-positiven Ergebnis führt, was den globalen  $\alpha$ -Fehler deutlich erhöht [307, 308]. Um dies zu verhindern, wurde die Bonferroni-Korrektur angewendet. Diese stellt eine vergleichsweise strenge Korrekturmethode für multiples Testen dar, bei der der im statistischen Test berechnete p-Wert angepasst wird, indem dieser mit der Anzahl an Tests, in diesem Fall der Zahl an untersuchten Elementen, multipliziert wird. Ein signifikanter Unterschied liegt somit nur dann vor, wenn auch der angepasste p-Wert das hier gewählte Signifikanzniveau von 0.05 unterschreitet [3, 307, 309]. Eine Übersicht der korrigierten p-Werte im Kruskal-Wallis-Test für alle Elemente ist Tabelle 3 zu entnehmen. Hoch signifikante Konzentrationsunterschiede sind demnach für die Elemente Cobalt, Uran, Lithium, Rubidium, Vanadium, Nickel, Thallium und Arsen zu verzeichnen, in jeweils absteigender Reihenfolge, mit p-Werten in der Größenordnung von  $10^{-23}$  bis  $10^{-14}$ . Viele weitere Elemente weisen p-Werte im Bereich von  $10^{-10}$  bis  $10^{-12}$  auf. Insgesamt wurde für alle Elemente, mit Ausnahme von Beryllium, ein p-Wert unterhalb der Signifikanzschwelle festgestellt, sodass für jedes dieser Elemente signifikante Unterschiede zwischen mindestens zwei der untersuchten Länder vorliegen. Dies bedeutet zudem, dass fast alle Elemente einen bedeutsamen Einfluss auf die Unterscheidung der Herkunftsgebiete zeigen. Der Kruskal-Wallis-Test sagt zunächst nur aus, dass Unterschiede zwischen den Gruppen vorliegen, nicht jedoch zwischen welchen Gruppen. Um zu prüfen, welche Länder signifikant unterschiedliche Elementkonzentrationen aufweisen, wurde zusätzlich ein

**Tabelle 3:** Übersicht der ermittelten p-Werte im Kruskal-Wallis-Test nach Bonferroni-Korrektur.

Element	p-Wert	Element	p-Wert
Li	$2.34 \cdot 10^{-21}$	Ce	$3.53 \cdot 10^{-11}$
Be	0.072	Pr	$2.12 \cdot 10^{-12}$
Al	$2.77 \cdot 10^{-11}$	Nd	$3.35 \cdot 10^{-12}$
V	$3.24 \cdot 10^{-15}$	Sm	$1.3 \cdot 10^{-11}$
Cr	0.021	Eu	$5.04 \cdot 10^{-13}$
Co	$5.4 \cdot 10^{-23}$	Gd	$9 \cdot 10^{-12}$
Ni	$1.66 \cdot 10^{-14}$	Tb	$6.48 \cdot 10^{-11}$
Zn	$9 \cdot 10^{-14}$	Dy	$1.8 \cdot 10^{-11}$
Ga	$3.6 \cdot 10^{-12}$	Ho	$2.99 \cdot 10^{-10}$
As	$3.6 \cdot 10^{-14}$	Er	$3.56 \cdot 10^{-11}$
Rb	$6.84 \cdot 10^{-17}$	Tm	$1.55 \cdot 10^{-9}$
Sr	$6.12 \cdot 10^{-11}$	Yb	$1.69 \cdot 10^{-10}$
Mo	$2.16 \cdot 10^{-10}$	Lu	$5.4 \cdot 10^{-10}$
Ag	$1.87 \cdot 10^{-5}$	Tl	$4.32 \cdot 10^{-15}$
Cd	$4.68 \cdot 10^{-5}$	Pb	$2.38 \cdot 10^{-8}$
Te	$1.91 \cdot 10^{-13}$	Bi	0.006
Ba	$1.33 \cdot 10^{-12}$	Th	$1.94 \cdot 10^{-13}$
La	$1.98 \cdot 10^{-12}$	U	$3.28 \cdot 10^{-22}$

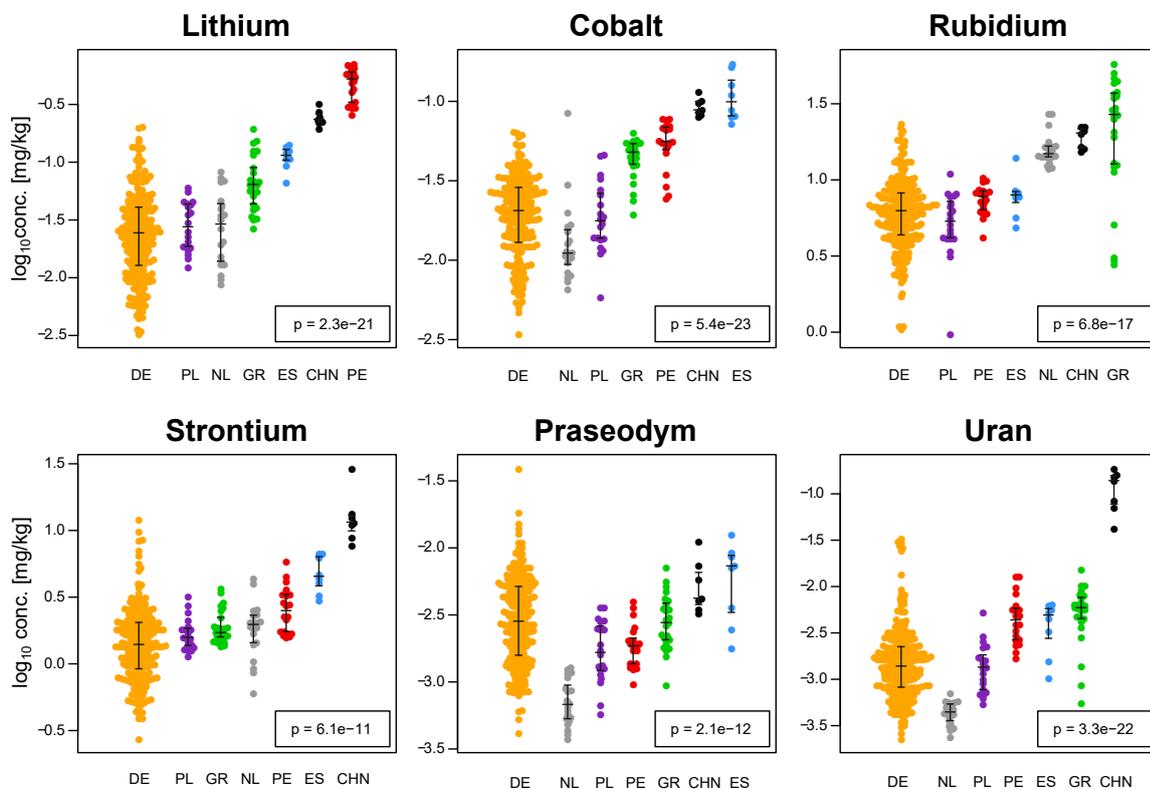
paarweiser Vergleich der Konzentrationslevel für alle Länderkombinationen mithilfe eines Zweistichproben-Tests vorgenommen. Der klassische Zweistichproben-t-Test sollte hier nicht angewendet werden, da dieser wiederum Varianzhomogenität voraussetzt. Daher wurde der Welch-Test herangezogen, welcher eine Abwandlung des t-Tests darstellt und auch bei ungleichen Varianzen einsetzbar ist [9]. Es wurde für jedes Element ein zweiseitiger Welch-Test durchgeführt, da Konzentrationsunterschiede in beide Richtungen von Interesse sind. Auch hier wurde die Bonferroni-Korrektur angewandt. Die ermittelten p-Werte sind als Länderpaar-Element-Matrix in Form einer *Heatmap* in Abbildung 27 dargestellt. Bei den grau-schwarzen Kästchen handelt es sich um Kombinationen aus Element und Länderpaar, bei denen kein signifikanter Unterschied vorliegt. Für alle weiteren Kombinationen wurden dagegen signifikante Abweichungen zwischen den untersuchten Ländern im Konzentrationsprofil der jeweiligen Elemente beobachtet, wobei die p-Werte hin zu helleren Farben abnehmen. Die kleinsten p-Werte und somit hochsignifikante Konzentrationsunterschiede sind für die Elemente Lithium, Uran und Thallium zu verzeichnen. Lithium zeigt eine hohe Relevanz bei der Unterscheidung von Proben aus Deutschland und Peru sowie aus Polen und Peru



**Abbildung 27:** Mittels Welch-Test ermittelte p-Werte nach Bonferroni-Korrektur; grau/schwarz: p-Wert  $\geq 0.05$ , dunkelrot bis gelb: p-Wert  $< 0.05$  in abnehmender Reihenfolge.

bzw. den Niederlanden und Peru. Somit lässt sich insgesamt ein hoher Stellenwert der Lithiumkonzentration bei der Charakterisierung der peruanischen Herkunft ableiten. Des Weiteren zeigen die Urangelhalte signifikante Unterschiede, besonders zwischen den Länderpaaren China/Niederlande sowie China/Deutschland und China/Polen. Da alle Länderpaare mit China für das Element Uran signifikant niedrige p-Werte aufweisen, lässt sich auch hier auf eine große Bedeutung des Urangelhaltes für die Herkunft China schließen. Zudem wurde bei Uran die größte Anzahl an signifikanten p-Werten beobachtet, mit Unterschieden im Konzentrationslevel bei 13 der insgesamt 21 Länderpaare. Für das Element Thallium sind signifikante Abweichungen zwischen den Ländern Deutschland und Griechenland sowie allen weiteren Länderpaaren unter Beteiligung von Griechenland zu verzeichnen, sodass für diese Herkunft ein Zusammenhang mit der Thalliumkonzentration vorliegt. Der *Heatmap* in Abbildung 27 können viele weitere signifikante Konzentrationsunterschiede entnommen werden, wie etwa bei den seltenen Erden für die meisten Länderpaare mit den Niederlanden. Für kein Länderpaar relevant sind dagegen die Elemente Bismut, Cobalt und Beryllium, die auch im zuvor gezeigten Kruskal-Wallis-Test die höchsten p-Werte aufwiesen. Bedingt durch die in diesem Fall deutlich größere Zahl an Tests, da alle Länderpaare separat betrachtet wurden, und die strenge Bonferroni-Korrektur, liegt der p-Wert im Welch-Test insgesamt etwas höher, sodass für die genannten Elemente keine Signifikanz zu beobachten ist.

Auf der Grundlage der Erkenntnisse aus den statistischen Tests wurden die Konzentrationsverteilungen der einzelnen Länder für alle Elemente separat betrachtet. Zur bildlichen Darstellung wurde für jedes Element ein Konzentrationsprofil in Form eines *Beeswarm*-Plots erstellt. Dabei ist jede Probe als ein Punkt im Plot abgebildet, was zu einem Bienenschwarm-ähnlichen Aussehen führt. Zusätzlich sind auch der Median sowie das obere und untere Quartil angegeben, welche ein Maß für die Streuung der Daten darstellen. In Abbildung 28 sind die Konzentrationsprofile für die Elemente Lithium, Cobalt, Rubidium, Strontium, Praseodym und Uran gezeigt, welche eine Auswahl der Elemente mit größter statistischer Signifikanz sowie Relevanz für die Unterscheidung der Herkunftsländer darstellen. Die Konzentrationsübersicht für Lithium zeigt sehr hohe Gehalte in den peruanischen Proben, die deutlich höher liegen als die der übrigen Proben. Die Konzentrationen an Lithium sind in den chinesischen Proben ebenfalls etwas erhöht, im Mittel aber deutlich niedriger als in den peruanischen. Für das Element Cobalt sind hohe Konzentrationen bei der Herkunft Spanien und China zu verzeichnen, mit jeweils sehr vergleichbaren Mittelwerten. Diese liegen deutlich höher als bei den übrigen Herkunftsgebieten und zeigen im Vergleich dazu eine geringere Varianz der Konzentrationsverteilung. Das Element Rubidium weist besonders hohe Gehalte für Proben aus Griechenland, China und den Niederlanden auf. Insbesondere tritt hier



**Abbildung 28:** *Beeswarm*-Plots der Konzentrationsverteilungen relevanter Elemente für die Herkunftsunterscheidung; p-Werte aus dem Kruskal-Wallis-Test nach Bonferroni-Korrektur.

der hohe Gehalt in den griechischen Proben hervor, wobei diese eine recht große Streuung in der Konzentrationsverteilung zeigen. Beim Konzentrationsprofil von Strontium sind sehr hohe Konzentrationen für die chinesischen Proben zu verzeichnen, wie sie sonst nur vereinzelt in einigen deutschen Proben zu finden sind. Die Proben niederländischer Herkunft demonstrieren, wie zuvor in der *Heatmap* gezeigt (Abb. 26), sehr niedrige Gehalte an seltenen Erden, was beispielhaft im *Beeswarm*-Plot für das Element Praseodym zu erkennen ist. Alle anderen Herkunftsländer zeigen dagegen höhere und untereinander sehr vergleichbare Konzentrationen an seltenen Erden mit etwas größerer Varianz. Ähnlich wie beim Element Strontium weisen die chinesischen Proben auch beim Uran sehr hohe Gehalte auf, wobei hier jedoch eine vollständige Abgrenzung aller weiteren Proben möglich ist. Bei den niederländischen Proben sind dagegen sehr niedrige Urangehalte zu erkennen. Neben den hier gezeigten wurden einige weitere Elemente identifiziert, die für eine Herkunftsregion besonders charakteristisch sind, darunter Vanadium, Arsen und Thallium. Die Konzentrationsprofile aller weiteren Elemente sind im Anhang im Abschnitt 8.7 dargestellt.

Insgesamt konnten anhand der gezeigten statistischen Methoden für alle untersuchten Herkunftsländer, mit Ausnahme von Polen, Elemente identifiziert werden, die für das jeweilige Land charakteristisch sind. Zusätzlich ist jedoch festzuhalten, dass die Untersuchung eines charakteristischen Elementes nicht ausreichend ist, um präzise und vertrauenswürdige Aussagen über die Herkunft einer Spargelprobe treffen zu können, da keine scharfen Grenzen zwischen den Konzentrationsverteilungen der einzelnen Länder zu beobachten waren. Erst die Hinzunahme weiterer Elemente lässt spezifische Muster entstehen, anhand derer die verschiedenen Herkunftsländer eindeutig identifiziert werden können und die eine Unterscheidung der geographischen Herkunft ermöglichen. Ein einfaches Beispiel bietet das Element Cobalt, dessen hohe Konzentration für die Herkunft Spanien als charakteristisch anzusehen ist. Allein durch eine Messung des Cobaltgehaltes kann jedoch nicht zwischen den Herkunftsländern Spanien und China unterschieden werden, erst der Vergleich mit den Konzentrationen an Strontium und Uran ermöglicht eine deutliche Unterscheidung der beiden Länder. Diese Erkenntnis wird auch von den vorherigen Beobachtungen der statistischen Auswertung gestützt. Zum einen haben die allgemein sehr niedrigen p-Werte aus dem Kruskal-Wallis-Test gezeigt, dass fast alle Elemente signifikante Unterschiede zwischen den Ländern aufweisen und damit einen wichtigen Beitrag zu deren Unterscheidung leisten, auch wenn dies in der Regel nicht bedeutet, dass die Konzentration eines Elementes bei allen Ländern differiert. Die Bedeutung einer größeren Anzahl an Elementen bei der Herkunftsunterscheidung wird zum anderen untermauert durch die vergleichsweise geringe Varianz (66.5%), welche mit den ersten drei Hauptkomponenten der PCA beschrieben wird (vgl. Abschnitt 4.2.3). Dies deutet darauf hin, dass der Datensatz vielschichtig aufgebaut ist und wenig Möglichkeiten zur Vereinfachung bietet, ohne dass dies mit einem bedeutenden Informationsverlust

bezüglich der Herkunftsdifferenzierung einhergeht. Die Ergebnisse suggerieren somit ein komplexes Zusammenspiel vieler Elemente zur Charakterisierung einer geographischen Herkunft, was eine Einbeziehung aller zur Verfügung stehenden Elementdaten für die Klassifikation der Herkunft nahelegt.

Die einzige Ausnahme bei der Konzentrationsabgrenzung stellt das Element Uran dar, bei dem Proben oberhalb eines definierten Konzentrationslevels exklusiv mit der Herkunft China assoziiert werden können, ohne dass Proben aus anderen Ländern vereinzelt ähnlich hohe Werte aufweisen. Die beobachteten hohen Urankonzentrationen im chinesischen Spargel stehen möglicherweise im Zusammenhang mit der Verwendung von weniger hochwertigen Phosphat-Düngemitteln. Phosphathaltige Gesteine enthalten etwa 50 - 200 ppm Uran, welches in Spuren auch im daraus hergestellten Dünger zu finden ist [310–312]. In diesem Fall wären die Uragehalte im Spargel auf anthropogene Einflüsse zurückzuführen, sodass man nicht von einem zuverlässigen Markerelement für die Herkunft China sprechen könnte. Um dies zu prüfen und auch aufgrund des kleinen Stichprobenumfangs ist die Untersuchung weiterer Proben notwendig, um sicherzustellen, dass die Varianz innerhalb dieses Landes statistisch abgedeckt wird.

Zusätzlich wurden die Ergebnisse dieser Arbeit mit den von Zannella *et al.* veröffentlichten mittleren Elementgehalten für die Spargelproben aus der italienischen Region *Bassano del Grappa* in Norditalien verglichen, die auf Proben von 5 Produktionsstandorten im Umkreis von 8 km basieren [107]. Beide Untersuchungen weisen von allen quantifizierten Analyten eine gemeinsame Schnittmenge von 15 Elementen auf. Bei Molybdän, Zink, Nickel, Cobalt, Strontium, Barium, Rubidium und Lithium sind im italienischen Spargel Gehalte zu verzeichnen, die den in dieser Arbeit gezeigten Konzentrationen stark ähneln. Dagegen zeigten sich Unterschiede der Konzentrationsbereiche von Blei, Silber, Cadmium, Vanadium, Lanthan, Cer und Chrom zwischen dem italienischen Spargel und den hier untersuchten Proben aus allen anderen Ländern. Chrom gehört zudem als einziges der letztgenannten Elemente zu den von Zannella *et al.* auf der Grundlage der Variabilität ermittelten potenziellen Schlüsselementen, die die untersuchte italienische Herkunft charakterisieren sollen. In den italienischen Proben wurde ein Chromgehalt von  $2.03 \pm 0.20$  mg/kg ermittelt, wohingegen die hier analysierten Proben aus den verschiedenen Herkunftsländern deutlich niedrigere Gehalte im Bereich von 0.07 - 0.67 mg/kg aufweisen. Der Vergleich liefert somit ein Indiz dafür, dass es sich bei Chrom tatsächlich um ein für die Herkunft Italien relevantes Element handeln könnte. Bei einer Erweiterung der Herkunftsuntersuchung um den Ursprung Italien wären die erwähnten abweichenden Elemente und im Speziellen das Element Chrom potentielle Ansatzpunkte zur Abgrenzung italienischer Proben.

### 4.2.5 Klassifikation zur Unterscheidung des Herkunftslandes

Die Multielementdaten der Spargelproben lassen wie zuvor gezeigt einige Unterschiede in Bezug auf die geographische Herkunft erkennen. Auf dieser Grundlage wurden Vorhersagemodelle erstellt und validiert, mit dem Ziel, diese auch auf unbekannte Proben anwenden zu können. Dabei wurden verschiedene Möglichkeiten der Herkunftsunterscheidung thematisiert, darunter die Differenzierung des Ursprungslandes, eine Abtrennung von deutschen und ausländischen Proben sowie eine innerdeutsche Herkunftsunterscheidung.

Die Erstellung der Klassifikationsmodelle erfolgte mithilfe der *Support Vector Machine* (SVM), welche eine aktuelle und vielfach erfolgreich angewandte Methode des maschinellen Lernens darstellt [313–315]. Im Gegensatz zu anderen häufig in diesem Zusammenhang eingesetzten Methoden wie der PLS oder der *Linear Discriminant Analysis* (LDA) erlaubt die SVM zusätzlich zur linearen auch eine nicht-lineare Trennung der Daten. Dies ist von großem Vorteil, da die zuvor gezeigten PCA-Plots vermuten lassen, dass trotz erkennbarer Klassenunterschiede eine rein lineare Trennung kein zufriedenstellendes Ergebnis liefern wird, da die Cluster der verschiedenen Herkunftsgebiete häufig Überlappungen aufweisen (vgl. Abb. 21). Ein weiteres Argument für die Anwendung der SVM ist ihre Eignung für kleine Klassengrößen, wie sie im Fall der spanischen und chinesischen Herkunft gegeben ist ( $n = 8$  bzw.  $7$  Proben) [316].

Als kritischer Punkt bei der Erstellung des jeweiligen Klassifikationsmodells ist die Validierung zu nennen. Diese soll einerseits gewährleisten, dass die für die Klassentrennung erforderlichen optimalen Parameter des Modells gefunden werden, sodass die Proben mit hoher Wahrscheinlichkeit richtig zugeordnet werden können. Auf der anderen Seite soll die Validierung sicherstellen, dass das Modell kein *Overfitting* aufweist. Dies würde bedeuten, dass mithilfe des Modells die vorliegenden Proben zwar mit hoher Genauigkeit vorhergesagt werden können, bei unbekanntem und unabhängigen Proben wäre jedoch eine deutliche Abnahme der Vorhersagegenauigkeit zu erwarten, da das Modell in diesem Fall zu stark an den vorhandenen Datensatz angepasst wurde. In der vorliegenden Arbeit wurden daher alle Vorhersagemodelle mit einer *nested cross-validation* optimiert und validiert (vgl. Abschnitte 2.5.8 und 6.3.6).

Das erste Vorhersagemodell zielte auf die Unterscheidung des Herkunftslandes der Spargelproben ab. Daraus ergeben sich 7 zu unterscheidende Klassen für die Proben der Länder Deutschland, Polen, den Niederlanden, Griechenland, Spanien, China und Peru. Zunächst wurden die verschiedenen Kern-Typen der SVM auf ihre Eignung hinsichtlich der genannten Fragestellung untersucht. Im Vergleich von linearem, polynomialen und RBF-Kern lieferte letzterer die höchsten Klassifikationsgenauigkeiten, der lineare Kern schnitt wie erwartet deutlich schlechter ab. Mithilfe des RBF-basierten SVM-Modells

konnte das Herkunftsland mit einer Genauigkeit von 91.2% richtig vorhergesagt werden. Detaillierte Ergebnisse der Klassifikation sind in Tabelle 4 dargestellt. Die Konfusionsmatrix zeigt mit der Beschriftung am linken Rand die tatsächliche Herkunft laut Ernteort bzw. Deklaration an. Die Kennzeichnung in horizontaler Richtung spiegelt dagegen die in der Klassifikation vorhergesagte Herkunft wider. Stimmen beide Angaben überein, so liegt eine richtige Zuordnung vor, welche auf der diagonalen Achse in dick gedruckten Zahlen hervorgehoben ist. Aus der Matrix wird deutlich, dass sowohl peruanische als auch chinesische Proben jeweils mit einer Genauigkeit von 100% vorhergesagt werden können, da weder falsch-positive noch falsch-negative Zuordnungen für diese Proben vorliegen. Auch die deutschen Proben konnten mit einer sehr hohen Genauigkeit vorhergesagt werden, da 211 der insgesamt 219 Proben richtig eingeordnet wurden (96.3%). Bei der Klassifikation der niederländischen und griechischen Proben wurden 80% bzw. 76% richtig zugeordnet, wohingegen die Herkunft der spanischen Proben mit einer etwas höheren Genauigkeit von 87.5% richtig vorhergesagt wurde. Die meisten falschen Klassifikationen sind für die polnischen Proben zu verzeichnen, von denen nur 55% korrekt eingeordnet wurden. Alle falsch-negativen Ergebnisse führten zu einer Klassifizierung als deutsche Proben. Eine Erklärung dafür könnte die geographische Nähe von Polen zum angrenzenden deutschen Gebiet liefern. Auch die Niederlande, deren falsch klassifizierte Proben ausschließlich eine Einordnung nach Deutschland erhielten, weist eine geographische Nähe zu Deutschland auf. Diese Vermutung wird dadurch gestützt, dass auch die meisten falsch-negativen Vorhersagen von deutschen Proben als Klassifikationsergebnis eine niederländische oder polnische Herkunft aufweisen. Die polnischen und niederländischen Proben zeigen daher vermutlich eine große Ähnlichkeit zu den deutschen Proben aus dem jeweils angrenzenden Gebiet im Osten bzw. Westen Deutschlands, die zur Überlappung der Proben führt, wie sie in den zuvor gezeigten PCA- und t-SNE-Plots zu beobachten war (Abb. 21 und 22). Da keine Verwechslungen

**Tabelle 4:** Konfusionsmatrix zur Klassifizierung des Herkunftslandes basierend auf Multielementdaten (Gesamtgenauigkeit: 91.2%).

	<b>DE</b>	<b>NL</b>	<b>PL</b>	<b>ES</b>	<b>GR</b>	<b>PE</b>	<b>CHN</b>
<b>DE</b>	<b>211</b>	3	4	0	1	0	0
<b>NL</b>	4	<b>16</b>	0	0	0	0	0
<b>PL</b>	9	0	<b>11</b>	0	0	0	0
<b>ES</b>	1	0	0	<b>7</b>	0	0	0
<b>GR</b>	3	1	1	1	<b>19</b>	0	0
<b>PE</b>	0	0	0	0	0	<b>20</b>	0
<b>CHN</b>	0	0	0	0	0	0	<b>7</b>

zwischen Polen und den Niederlanden auftraten, kann geschlussfolgert werden, dass beide Länder gut voneinander abgetrennt werden können und Überschneidungen mit deutschen Proben sich daher auf unterschiedliche deutsche Herkunftsgebiete beziehen. Spanien und Griechenland, die nicht in Nachbarschaft zu einem anderen Herkunftsland gelegen sind, konnten mit guten Ergebnissen richtig zugeordnet werden und zeigen keine Systematik bei vereinzelt Fehlklassifikationen. Dass griechische Proben sowohl als deutsch, niederländisch, polnisch und spanisch falsch eingeordnet wurden, zeigt, dass diese Klasse nicht in allen Fällen präzise abgetrennt werden kann. Im Gegensatz dazu weisen die Herkunftsländer Peru und China die größte geographische Entfernung sowohl untereinander als auch zu den untersuchten europäischen Ländern auf, was mit der höchstmöglichen Genauigkeit bei der Klassifikation einhergeht. Insgesamt wurde mit dem gezeigten SVM-Modell in den meisten Fällen eine sehr gute Unterscheidung und eine präzise Zuordnung des geographischen Herkunftslandes der Spargelproben erreicht.

Des Weiteren wurde eine Klassifikation des Herkunftslandes auf der Grundlage der identifizierten Schlüsselemente für alle Herkunftsländer vorgenommen. Dabei wurde bei jeder Reduktion der Anzahl in Betracht gezogener Elemente eine Verringerung der Klassifikationsgenauigkeit beobachtet. Dies steht im Einklang mit der zuvor diskutierten Vielschichtigkeit des Datensatzes und der großen Bedeutung einer hohen Zahl an Dimensionen. Daher wurden die Klassifikationsmodelle im weiteren Verlauf ausschließlich anhand aller Elementdaten erstellt.

Schwierigkeiten bei der Klassifikation zeigten sich insbesondere für die Trennung einiger polnischer und niederländischer Proben von denen aus Deutschland aufgrund einer vermuteten großen Ähnlichkeit. Dies führt dazu, dass Klassengrenzen nicht immer präzise gezogen werden können, da ein *Overfitting* aufgrund eines zu stark an den aktuellen Datensatz angepassten Modells unbedingt zu vermeiden ist. Um daraus resultierende Fehlklassifikationen zu umgehen, wurde auf der Grundlage des gezeigten Vorhersagemodells eine Konfidenzabschätzung der Klassifikationsergebnisse vorgenommen, um die Güte der Klassifikationsentscheidungen differenziert beurteilen zu können. Dazu wurde für jede Probe und jede Klasse die Wahrscheinlichkeit berechnet, mit der eine Probe in die entsprechende Klasse eingeordnet wird. Die Klasse mit der höchsten Wahrscheinlichkeit ist dabei immer jene, die als Klassifikationsergebnis vom Modell ausgegeben wird. Präzise Vorhersagen können erzielt werden, wenn die Klassen für eine Probe jeweils eine sehr hohe oder eine sehr niedrige Wahrscheinlichkeit aufweisen. Im Gegensatz dazu können ähnlich große Wahrscheinlichkeitswerte bei mehreren Klassen leicht zu Fehlklassifikationen führen. Im theoretischen Idealfall liegt die Wahrscheinlichkeit für die korrekte Klasse somit bei 1, die aller weiteren Klassen bei 0. Da die meisten fehlerhaften Klassifikationen unter der Beteiligung der deutschen Proben entstanden und die deutsche Herkunft einen wesentlichen Schwerpunkt dieser Untersuchungen bildete,

wurde die folgende Konfidenzabschätzung auf der Grundlage der Wahrscheinlichkeitswerte für die Klasse der deutschen Proben vorgenommen. Es wurden verschiedene Intervalle von Wahrscheinlichkeiten abgetrennt, in denen unpräzise Ergebnisse beobachtet und die daher als „nicht klassifizierbar“ definiert wurden. Die Klassifikationsergebnisse wurden anschließend neu ausgewertet, indem nur die übrigen Proben oberhalb und unterhalb des Bereichs der nicht klassifizierbaren Proben herangezogen wurden. Der so gewonnene Konfidenzbereich, der von verschiedenen Wahrscheinlichkeitswerten ( $p$ ) begrenzt wird, führte zu deutlich verbesserten Klassifikationsergebnissen, was in Tabelle 5 dargestellt ist. Die Klassifikationsgenauigkeit konnte so auf bis zu 98 % angehoben werden. Auf der anderen Seite geht die Eingrenzung des Konfidenzbereiches mit einer zunehmend reduzierten Anzahl an klassifizierten Proben einher, je strikter der Bereich abgegrenzt wird. Im Falle des Konfidenzbereichs von  $p < 0.1$  oder  $p > 0.9$  wurden daher nur 63 % der Proben in die Auswertung miteinbezogen. Die Anwendung der Konfidenzabschätzung ist in Tabelle 6 beispielhaft für den Konfidenzbereich mit  $p < 0.2$  oder  $p > 0.8$  dargestellt. Die Konfusionsmatrix zeigt nur noch vereinzelt Fehlklassifikationen in der Klasse der deutschen, griechischen und niederländischen Proben, während alle Proben der Herkunftsländer Spanien, Peru und China richtig eingeordnet wurden. Bei der Untersuchung des polnischen Ursprungs konnten jedoch noch immer nur etwa die Hälfte der Proben richtig klassifiziert werden, die übrigen wurden wiederum als Proben deutscher Herkunft eingeordnet. Diese Beobachtung zieht sich auch durch die Matrices aller weiteren untersuchten Konfidenzbereiche, was erneut auf eine große Ähnlichkeit zwischen den polnischen und einigen deutschen Proben hindeutet, sodass die Herkunft Polen mit der gewählten Methodik nicht präzise abgetrennt werden kann. Da die Parameter Klassifikationsgenauigkeit und Anzahl klassifizierter Proben in einem entgegengesetzten Verhältnis stehen, ist bei der Wahl des geeigneten Klassifikations-

**Tabelle 5:** Klassifikationsergebnisse für die Betrachtung verschiedener Konfidenzbereiche begrenzt durch unterschiedliche Wahrscheinlichkeitswerte  $p$  der Klasse der deutschen Proben basierend auf Multielementdaten.

Konfidenzbereich	Klassifikationsgenauigkeit	klassifizierte Proben
$p < 0.25$ oder $p > 0.75$	94.7 %	88.1 %
$p < 0.2$ oder $p > 0.8$	95.6 %	83.1 %
$p < 0.15$ oder $p > 0.85$	96.7 %	76.5 %
$p < 0.1$ oder $p > 0.9$	98.0 %	63.0 %

**Tabelle 6:** Konfusionsmatrix zur Klassifizierung des Herkunftslandes anhand von Proben mit hoher Konfidenz ( $p < 0.2$  oder  $p > 0.8$ ) basierend auf Multielementdaten (Gesamtgenauigkeit: 95.6 %).

	DE	NL	PL	ES	GR	PE	CHN
DE	<b>192</b>	2	0	0	1	0	0
NL	1	<b>11</b>	0	0	0	0	0
PL	6	0	<b>7</b>	0	0	0	0
ES	0	0	0	<b>5</b>	0	0	0
GR	1	0	0	1	<b>17</b>	0	0
PE	0	0	0	0	0	<b>20</b>	0
CHN	0	0	0	0	0	0	<b>7</b>

modells je nach Situation ein Abwägen zwischen beiden Parametern erforderlich. Im Hinblick auf die Fragestellung der Verfälschung deutscher Produkte mit importiertem Spargel bzw. einer gezielten Umetikettierung der Herkunft könnten höhere Genauigkeiten zu Lasten der Anzahl klassifizierbarer Proben bevorzugt sein, da eine höhere Belastbarkeit der Ergebnisse gegeben ist. So wird vermutlich eher in Kauf genommen, dass eine Probe nicht klassifiziert werden kann, als dass sie als ‘nicht deutsch’ eingeordnet wird mit womöglich strafrechtlichen Konsequenzen, obwohl die Probe tatsächlich eine deutsche Herkunft aufweist.

Eine Kombination aus dem gezeigten SVM-Modell und der Konfidenzabschätzung ermöglicht insgesamt eine sehr präzise Vorhersage der geographischen Herkunft einer Spargelprobe und kann daher einen wichtigen Beitrag zur Aufdeckung einer verfälschten Herkunftsangabe leisten.

#### 4.2.6 Klassifikation deutscher und importierter Produkte

Neben der eindeutigen Vorhersage des Herkunftslandes ist auch die Abgrenzung ausländischer Proben vom deutschen Spargel von großem Interesse. Dies soll die Möglichkeit bieten, Verfälschungen nachzuweisen, bei denen importierter Spargel mit der Deklaration einer deutschen Herkunft versehen wird, um so einen höheren Verkaufspreis zu erzielen. Die genaue Herkunft ist dabei zweitrangig, von größerem Interesse ist hier, ob eine fälschlicherweise als deutsch deklarierte Herkunft vorliegt und mit welcher Wahrscheinlichkeit diese erkannt werden kann. Dazu wurde ein zweites vereinfachtes Vorhersagemodell entwickelt, bei dem die deutschen Proben allen anderen gegenübergestellt wurden, die im Folgenden in der zweiten Klasse der ausländischen

Proben zusammengefasst werden.

Das resultierende SVM-Klassifikationsmodell (RBF-Kern) lieferte eine richtige Vorhersage für 91.8 % der Proben und schnitt somit geringfügig besser ab als das zuvor gezeigte Modell zur Länderklassifikation. Tabelle 7 zeigt im Detail die Konfusionsmatrix der Klassifikation. Es wurde bei 95.4 % der deutschen sowie 84 % der ausländischen Proben die Herkunft richtig vorhergesagt. Im Vergleich zur Klassifikation des Herkunftslandes, bei der 8 deutsche Proben falsch eingeordnet wurden, ist bei diesem Klassifikationsmodell mit 10 falschen Zuordnungen keine Verbesserung zu verzeichnen. Dagegen sind auf der Seite der ausländischen Proben 16 nicht korrekte Klassifikationen festzustellen, während diese Anzahl bei der Länderklassifikation in der Summe bei 20 lag.

**Tabelle 7:** Konfusionsmatrix zur Klassifizierung einer deutschen oder ausländischen Herkunft basierend auf Multielementdaten (Gesamtgenauigkeit: 91.8 %).

	Deutsch	Ausländisch
Deutsch	209	10
Ausländisch	16	84

Insgesamt stellt dieses Modell für die Abgrenzung der deutschen Proben von jenen aus allen anderen Herkunftsgebieten keine signifikante Verbesserung im Vergleich zu dem Modell der Ländervorhersage dar. Ein Grund dafür könnte die Varianz der Gruppe der ausländischen Proben sein, die bei der Kombination zu einer Klasse im Vergleich zu den einzelnen Ländern stark zunimmt. Diese Klasse umfasst sehr heterogene Proben mit zum Teil weit entfernten Ursprungsgebieten von China über Europa bis hin zu Peru, was die Charakterisierung und Abgrenzung dieser Klasse erschwert und sich unvorteilhaft auf die Klassifikation auswirkt. Da die Elementdaten, wie zuvor gezeigt, für fast alle untersuchten Länder spezifische Muster aufweisen, hat sich somit eine Zuordnung zu kleineren und homogeneren Herkunftsgebieten als sehr präzise und zuverlässig erwiesen, weshalb diese der Einordnung in einen deutschen oder ausländischen Herkunftsort vorzuziehen ist.

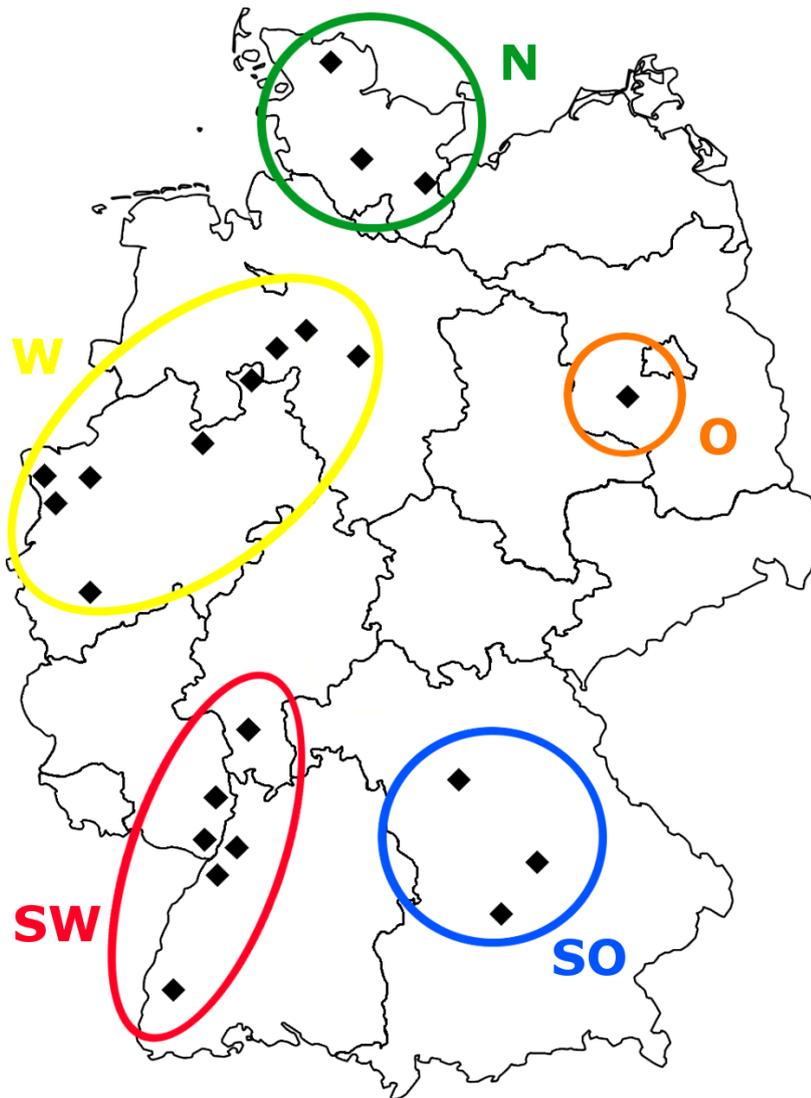
#### 4.2.7 Innerdeutsche Probenklassifikation

Da die zuvor gezeigten Ergebnisse bereits eine hohe Leistungsfähigkeit der elementbasierten Vorhersagemodelle gezeigt haben, wurde das Herkunftsgebiet im nächsten Schritt auf die Proben deutschen Ursprungs begrenzt, um den Datensatz noch genauer

zu untersuchen und Belastbarkeit und Präzision der resultierenden Modelle zu testen. Dafür musste eine geeignete Klasseneinteilung vorgenommen werden. Als Unterteilungskriterium wurde zunächst die geographische Lage der Spargelhöfe, von denen die Proben bezogen wurden, in Erwägung gezogen. In Abbildung 29 sind die entsprechenden Standorte der insgesamt 43 Spargelhöfe in einer Deutschlandkarte eingetragen. Da einige Höfe sehr nah beieinander liegen, wurden diese in einem Punkt zusammengefasst. Aus der Abbildung wird deutlich, dass eine Probeneinteilung nach dem Bundesland der Herkunft wenig sinnvoll ist, da einige Höfe trotz unterschiedlicher Bundesländer in direkter Nachbarschaft liegen, wie dies etwa bei den Proben aus Rheinland-Pfalz, Hessen und dem Norden von Baden-Württemberg der Fall ist. Stattdessen wurden die Herkunftsgebiete einiger Bundesländer zusammengefasst und eine Probeneinteilung in fünf deutsche Regionen vorgenommen, welche in Abbildung 29 farblich markiert sind. Die Proben aus Schleswig-Holstein wurden dafür als „Nord“ eingeordnet, jene aus Niedersachsen und Nordrhein-Westfalen als „Westen“ zusammengefasst und die Proben aus Baden-Württemberg, Hessen und Rheinland-Pfalz in der Klasse „Südwesten“ zusammengeschlossen. Die Proben aus Bayern und Brandenburg bilden jeweils die Klassen „Südost“ bzw. „Ost“. Die Erstellung und Optimierung eines Klassifikationsmodells für die Unterscheidung dieser fünf Regionen erfolgte mithilfe einer SVM mit RBF-Kern (vgl. Abschnitt 6.3.6).

Die Ergebnisse dieses Modells sind in Tabelle 8 dargestellt. Die Herkunft wurde bei insgesamt 183 der 219 deutschen Proben richtig vorhergesagt, was einer Genauigkeit von 83.6% entspricht. Ein besonders gutes Klassifikationsergebnis wurde mit 91.8% richtiger Vorhersage für die Proben aus dem Südwesten erzielt. Dagegen wurde der Spargel mit westdeutscher Herkunft nur in 60.7% der Fälle richtig eingeordnet, wobei falsche Zuordnungen dieser Proben zu etwa gleichen Teilen nach Norden, Südwesten und Südosten auftraten. Eine Erklärung für diese vergleichsweise schlechte Vorhersageleistung könnte eine Schwäche bei der gewählten Klasseneinteilung sein. Die Probenstandorte innerhalb der Klasse Westdeutschland weisen eine maximale geographische Entfernung auf, die deutlich größer ist als die Distanz der nächstgelegenen Standorte zwischen den Klassen Westen und Norden bzw. Westen und Südwesten. Eine Überlappung der westdeutschen Proben mit den Regionen Norden und Südwesten erscheint daher plausibel. Eine Möglichkeit zur Verbesserung von Intra- und Inter-Gruppen-Distanz wäre eine Unterteilung der Klassen in kleinere und homogenere Gruppen. Ein solcher Versuch der Klasseneinteilung, z. B. die Unterteilung der Klasse Westen in zwei separate Klassen, führte jedoch zu keiner Erhöhung der Präzision der Ergebnisse.

Mit der gezeigten Klassifikation konnten insgesamt gute Ergebnisse erzielt werden, sodass das Modell für eine grobe Einordnung der innerdeutschen Herkunft als geeignet zu bewerten ist. Das Klassifikationsmodell schnitt mit einer Genauigkeit von 83.6%



**Abbildung 29:** Klasseneinteilung der deutschen Spargelproben in die 5 Regionen Nord (N, grün), Ost (O, orange), West (W, gelb), Südost (SO, blau) und Südwest (SW, rot) sowie Standorte der bei der Probenakquise beteiligten Spargelhöfe (schwarze Rauten).

**Tabelle 8:** Konfusionsmatrix zur Klassifizierung der deutschen Herkunftsregionen (Gesamtgenauigkeit: 83.6%).

	O	N	W	SW	SO
O	12	1	1	3	0
N	1	48	1	1	5
W	0	4	17	3	4
SW	2	1	1	67	2
SO	0	1	0	5	39

schlechter ab als die zuvor gezeigte Unterscheidung der Herkunftsländer mit 91.2% korrekter Klassifikationen (vgl. Abschnitt 4.2.5). Angesichts der viel kleineren Unterschiede zwischen den einzelnen Klassen bei der Betrachtung von ausschließlich deutschen Proben und der zufällig richtigen Klassifikationsrate, die bedingt durch die Probenverteilung dieses 5-Klassen-Trennproblems bei etwa 24% liegt, stellt dies dennoch ein sehr gutes Ergebnis dar. Trotz allem bleibt zu hinterfragen, ob die hier gewählte Klasseneinteilung bzw. der Ansatz zur Gruppeneinteilung der am besten geeignetste ist. Bei einem genaueren Blick auf die Bodenverhältnisse in Deutschland mithilfe von Bodenübersichtskarten fällt auf, dass die Bodenart und die Bodenzusammensetzung sehr kleinteiligen Änderungen unterworfen sind, sodass bereits nah benachbarte Gebiete große Unterschiede aufweisen können [317]. Eine variierende Bodenzusammensetzung bedeutet auch, dass die im Boden enthaltenen Elemente differieren können, was benachbarte Probenstandorte allein auf der Grundlage ihrer geographischen Lage bzw. Nähe zueinander nur bedingt vergleichbar erscheinen lässt. Dies führte zu der Überlegung, anstelle von Probengebieten einzelne Probenstandorte gegenüberzustellen.

#### 4.2.8 Klassifikation zur Standortunterscheidung

Wie in der vorangegangenen Diskussion erläutert, wurde die Klassifikation der Spargelproben nach ihrem Produktionsstandort in Erwägung gezogen. So sollte der Bezug zum Boden des Erzeugerfeldes stärker in den Mittelpunkt gestellt werden, da die Bodenzusammensetzung bekanntermaßen einen großen Einfluss auf das Elementprofil einer darauf wachsenden Pflanze hat [107, 318, 319]. Zusätzlich wurde neben der Ähnlichkeit vieler Proben mit gleicher Herkunftsregion bereits im Clustering der Spargelproben in der zuvor gezeigten *Heatmap* (Abb. 26) beobachtet, dass die Proben eines Bauern im Dendrogramm oftmals sehr dicht beieinander liegen und somit trotz vorliegender Unterschiede bei Sorte und Erntejahr große Ähnlichkeiten aufweisen. Um statistisch belastbare und aussagekräftige Ergebnisse zu erhalten, wurden für die Modellerstellung nur die Produktionsstandorte einbezogen, von denen mindestens sechs Proben vorlagen. Dies traf auf insgesamt 13 Standorte zu, deren Proben sich auf die Erntezeiträume 2014 - 2016 erstreckten. Darauf basierend wurde ein SVM-Vorhersagemodell (RBF-Kern) für die Unterscheidung des Standortes analog zu den drei zuvor gezeigten Modellen anhand einer *nested cross-validation* trainiert und optimiert.

Der Produktionsstandort einer Spargelprobe konnte auf diese Weise mit einer Genauigkeit von 82.6% richtig vorhergesagt werden, wie in Tabelle 9 im Detail dargestellt ist. Unter Berücksichtigung der großen Anzahl an Klassen stellt dies ein extrem gutes Ergebnis dar. Im Vergleich dazu ist bei einem Umfang von 13 Klassen und der hier

**Tabelle 9:** Konfusionsmatrix zur Klassifizierung des Produzenten (P) mittels SVM basierend auf Multielementdaten (Gesamtgenauigkeit: 82.6%).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13
P1	3	0	0	0	0	1	0	1	1	0	0	0	0
P2	0	6	0	0	0	0	0	0	0	0	0	0	1
P3	0	0	6	0	0	0	0	0	0	0	0	0	0
P4	0	0	0	6	0	0	0	0	0	1	0	0	0
P5	0	0	0	0	6	0	0	0	0	0	0	0	0
P6	0	0	0	0	0	4	0	0	2	0	0	0	0
P7	1	0	0	0	0	0	3	0	3	0	1	0	0
P8	0	0	0	0	0	0	0	7	0	0	0	1	1
P9	0	0	0	0	0	0	0	0	12	0	1	0	0
P10	0	0	0	0	0	0	0	0	0	20	0	0	0
P11	0	0	0	0	0	0	1	0	1	0	6	0	0
P12	0	0	0	0	0	0	0	0	0	0	0	7	0
P13	0	0	0	0	0	0	0	2	0	0	0	0	4

vorliegenden Probenverteilung eine zufällig richtige Vorhersage mit einer Wahrscheinlichkeit von nur 9.3% zu erwarten. Ein Blick auf die Konfusionsmatrix zeigt bei 4 Produzenten eine Genauigkeit von 100% und damit die richtige Zuordnung aller Spargelproben. Bei 3 Produzenten wurde nur eine, bei weiteren 4 Produzenten wurden zwei Spargelproben falsch eingeordnet. Zwei der 13 Standorte schnitten mit drei bzw. fünf nicht korrekt klassifizierten Proben etwas schlechter ab.

Das Modell lieferte eine Genauigkeit, die fast identisch zu der zuvor gezeigten Klassifikation der deutschen Herkunftsregionen ausfiel. Unter Berücksichtigung des enormen Unterschieds bei der Klassenanzahl (5 Klassen vs. 13 Klassen) konnte somit gezeigt werden, dass der Produktionsstandort ein deutlich aussagekräftigeres und zuverlässigeres Kriterium für die Klasseneinteilung bei der innerdeutschen Herkunftsunterscheidung darstellt. Trotz eines recht kleinen Stichprobenumfangs von 6-20 Proben pro Klasse erlaubt das gezeigte Klassifikationsmodell eine sehr präzise Zuordnung einer Spargelprobe zum Standort ihrer Produktion. Daher ist anzunehmen, dass der Spargel durch den zugrundeliegenden Produktionsstandort wesentlich charakterisiert wird, während Unterschiede bei Erntejahr und Sorte in den Hintergrund treten. Das untersuchte Elementprofil des Spargels von einem Standort zeigte sich als über mehrere Jahre hinweg stabil, was eine Vorhersage des Standortes von Spargelproben aus verschiedenen Erntejahren ermöglichte. All diese Proben haben als gemeinsames Merkmal den Boden, auf dem sie gewachsen sind. Dies lässt die Schlussfolgerung zu, dass ein starker Zusammenhang zwischen dem untersuchten Elementprofil des Spargels und dem des Bodens vom jeweils zugehörigen

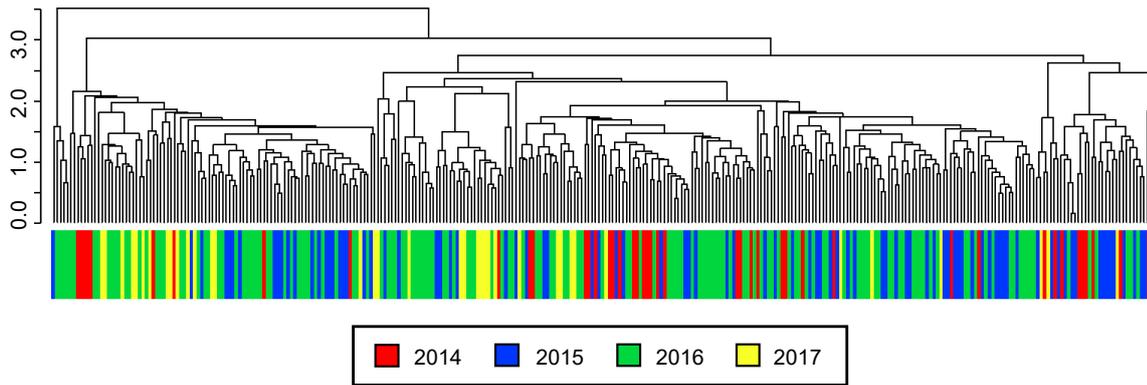
Standort besteht und der Spargel somit das Elementprofil des Bodens widerspiegelt. Daraus folgt, dass die Spargelpflanze offenbar über funktionierende Aufnahmewege für die untersuchten Elemente verfügt, was zu einem charakteristischen und über Jahre reproduzierbaren Elementprofil im Spargel führt. Dies bedeutet im Umkehrschluss, dass auch die Elementzusammensetzung des jeweiligen Bodens über Jahre hinweg sehr stabil ist, was plausibel erscheint, sofern der Boden keinen anthropogenen oder sonstigen Einflüssen unterliegt, die das Elementprofil bzw. die Eigenschaften, die die Elementaufnahme bedingen, verändern und die hier nicht berücksichtigt wurden.

Mit dem Klassifikationsmodell zum Produktionsstandort konnten insgesamt sehr präzise Ergebnisse erzielt werden. Die hohen Genauigkeiten der Klassifikation trotz großer Klassen- und niedriger Probenzahl unterstreichen zudem die Robustheit des elementbasierten Vorhersagemodells. Die dargestellten Ergebnisse zeigen somit einen neuen Ansatz zur Herkunftsbestimmung von Spargel und damit verbunden die Möglichkeit zur Überprüfung der angegebenen Produktdeklaration und insbesondere der geschützten Herkunftsangaben. Voraussetzung hierfür ist der Aufbau einer Datenbank, die eine ausreichend große Anzahl an Proben jedes zu klassifizierenden Standortes enthält, um diese zuverlässig charakterisieren zu können. Die hohe Vergleichbarkeit der Proben aus verschiedenen Erntejahren legt zudem nahe, dass die Datenbank fortlaufend um neue Erntejahre erweitert werden kann. Vorstellbar wäre des Weiteren, dass die Genauigkeit der Vorhersage weiter angehoben werden kann, wenn die zugrunde liegende Datenlage weiter verbessert wird. Bei dennoch auftretenden Problemen bei der Abgrenzbarkeit mehrerer Klassen stellt auch hier die Anwendung einer Konfidenzabschätzung eine gute Möglichkeit zur Verbesserung von Präzision und Belastbarkeit des Modells dar. Aufgrund der kleinen Probenanzahl pro Klasse für das gezeigte Modell wurde hier jedoch auf eine Konfidenzabschätzung verzichtet.

#### **4.2.9 Untersuchung von Einflussfaktoren auf die Herkunftsbestimmung**

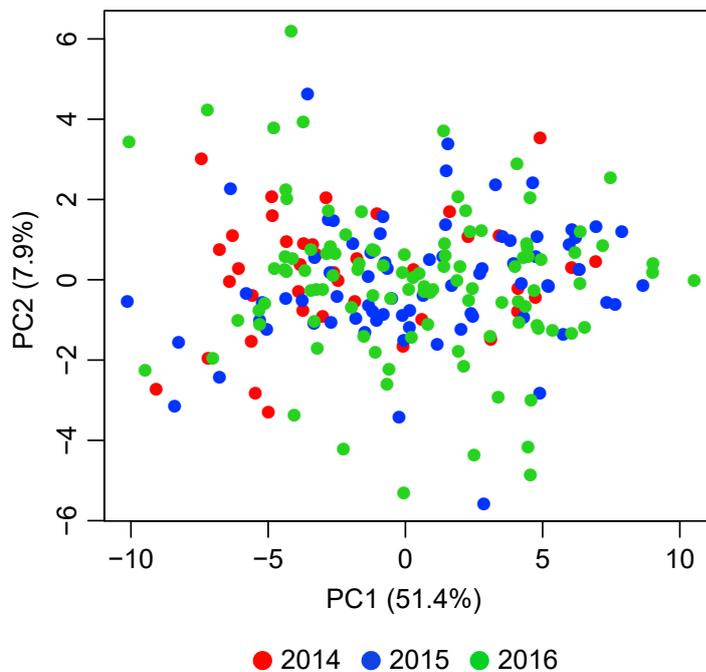
Bei der zuvor gezeigten Klassifikation des Produktionsstandortes fiel bereits auf, dass die Sorte und das Erntejahr einer Spargelprobe im Vergleich zum Herkunftsort eine untergeordnete Rolle beim Vergleich des Elementprofils spielen. Als mögliche Einflussfaktoren auf die Herkunftsbestimmung wurden Spargelsorte, Erntejahr sowie das Messdatum noch einmal genauer anhand einer hierarchischen Clusteranalyse untersucht, basierend auf dem euklidischen Abstand als Distanzmaß.

Abbildung 30 zeigt das Clustering für die Untersuchung des Erntejahreinflusses. Bei den untersuchten Spargelproben ist keine Clusterbildung für die einzelnen Erntejahre zu



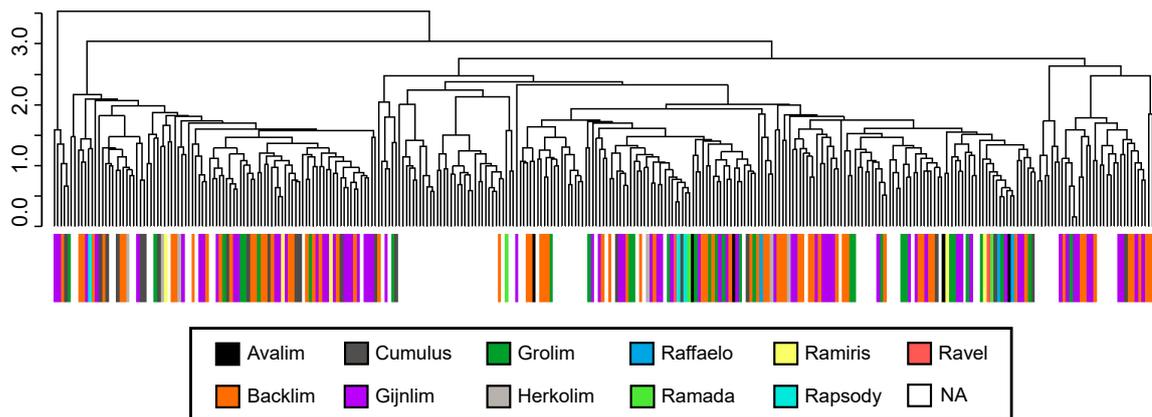
**Abbildung 30:** Hierarchisches Clustering mit Markierung des Erntejahres basierend auf Multielementdaten.

beobachten, vielmehr zeigt sich eine gleichmäßige Aufteilung der vier Jahre auf alle Cluster. Die einzige Einschränkung bilden die Proben des Erntejahres 2017, die sich fast ausschließlich in der linken Hälfte des Clusterings befinden. Diese Beobachtung ist möglicherweise auf die Struktur des Datensatzes zurückzuführen, da für das Jahr 2017 keine deutschen Proben vorlagen. Somit ist nicht auszuschließen, dass ein möglicher Einfluss des Erntejahres von der Herkunft überlagert wird. Aus diesem Grund wurde zusätzlich eine PCA angefertigt, bei der nur die deutschen Proben aller Erntejahre berücksichtigt wurden, sodass der Faktor Herkunft konstant bleibt. In der PCA in Abbildung 31 ist eine sehr homogene Verteilung der Proben aus den drei Erntejahren zu verzeichnen, aus der kein Ansatz einer Clusterbildung hervorgeht. Aufgrund einer fehlenden



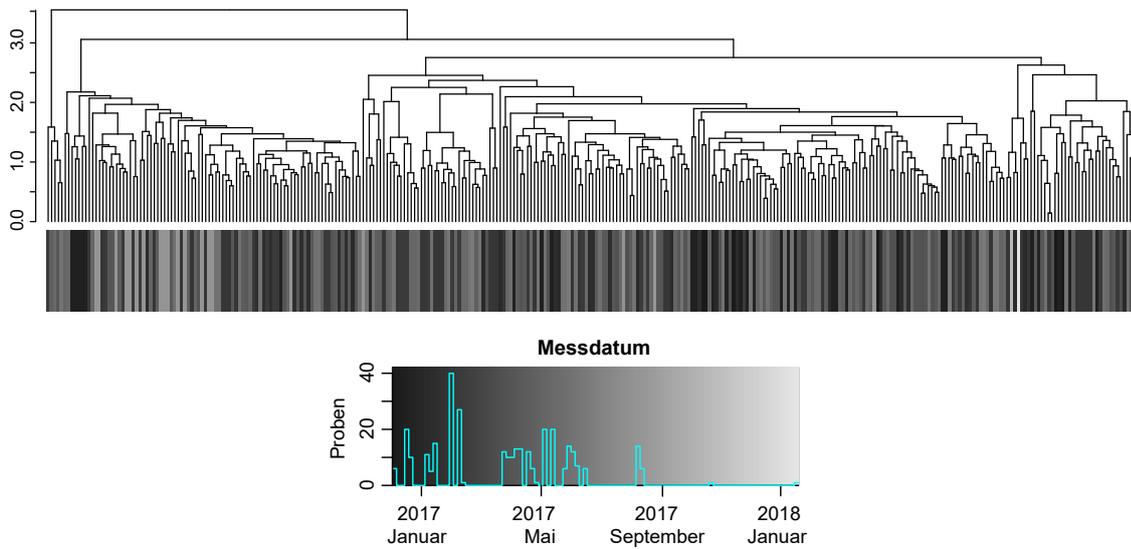
**Abbildung 31:** PCA der deutschen Proben aus den Erntejahren 2014 - 2016 basierend auf Multielementdaten.

Systematik in Bezug auf das Erntejahr ist dieses nicht als bedeutender Einflussfaktor für das Elementprofil zu bewerten. Dies steht zudem im Einklang mit den Ergebnissen der Klassifikation des Produktionsstandortes, welche in Abschnitt 4.2.8 gezeigt wurden. Zudem wurde keine systematische Abweichung der Proben des Erntejahres 2014 beobachtet, was die eingangs angestellte Vermutung bekräftigt, dass die Lagerungstemperatur während Probenakquise und -transport keinen bedeutsamen Einfluss auf die Elementgehalte hat. Des Weiteren wurde auch der Einfluss der Spargelsorte genauer beleuchtet, da die zugrundeliegende Genetik der Pflanze ebenfalls Auswirkungen auf die Elementaufnahme haben kann. Dazu wurde im hierarchischen Clustering die jeweilige Spargelsorte aller Proben farblich gekennzeichnet, wie Abbildung 32 zu entnehmen ist. Diese zeigt keine Gruppierung der Proben in Abhängigkeit einer Sorte. Aus der gleichmäßigen Verteilung aller Spargelsorten über das gesamte Dendrogramm kann abgeleitet werden, dass die Sorte ebenfalls keinen bedeutenden Einfluss auf das Elementprofil und somit auf die darauf basierende Herkunftsvorhersage ausübt.



**Abbildung 32:** Hierarchisches Clustering zur Unterscheidung der Spargelsorte basierend auf Multielementdaten.

Als weiterer Faktor wurde das Messdatum hinsichtlich eines Einflusses auf die Elementdaten untersucht. Die Aufarbeitung und Messung der Spargelproben erfolgte soweit wie möglich in randomisierter Reihenfolge, um messbedingte Einflüsse zu vermeiden. Bedingt durch die zeitliche Abfolge der Probenakquise in unterschiedlichen Erntejahren war dies jedoch nur mit Einschränkungen möglich. Da die Methodvalidierung sowohl über kurze als auch längere Zeiträume eine hohe Stabilität der Messung zeigte, wird jedoch kein Einfluss durch die Messabfolge vermutet. Abbildung 33 zeigt ein hierarchisches Clustering aller Spargelproben, bei dem das jeweilige Messdatum mithilfe eines Farbverlaufes von dunkelgrau (erster Messtag) bis hin zu hellgrau (letzter Messtag) dargestellt ist. Dem Histogramm im unteren Teil der Abbildung ist zudem die Verteilung der gemessenen Probenanzahl über den gesamten Messzeitraum von 13 Monaten zu



**Abbildung 33:** Hierarchisches Clustering mit Markierung des Messdatums basierend auf Multielementdaten; unten: Darstellung der gemessenen Probenzahlen als Histogramm und Farbverlauf über die gesamte Messzeit.

entnehmen. Anhand der Färbung des Messtages sind auch hier keine systematischen Zusammenhänge in Form einer Clusterbildung zu erkennen. Insgesamt konnte für keinen der genannten Faktoren ein Einfluss auf die Elementzusammensetzung festgestellt werden. Somit ist eine sehr gute Vergleichbarkeit der Proben hinsichtlich ihrer Herkunft gegeben, trotz Unterschieden bei Erntejahr und Sorte. Demnach stellt der elementbasierte Ansatz zur Herkunftsbestimmung von Spargel ein sehr robustes Verfahren dar, welches wenig anfällig gegenüber anderen Einflüssen als der Herkunft ist. Dies ermöglicht eine Kombination aller zur Verfügung stehenden Proben mit dem Ziel des Aufbaus einer Datenbank zur präzisen Bestimmung des geographischen Ursprungs.

---

## 4.3 Vergleich des Elementprofils von Spargel und Boden

Bei der Klassifikation der Spargelproben nach ihrem Produktionsstandort im Abschnitt 4.2.8 war eine große Ähnlichkeit der Proben eines Standortes trotz unterschiedlicher Sorten und Erntejahre zu verzeichnen. Da diese Proben jeweils denselben Boden gemeinsam haben, liegt die Schlussfolgerung nahe, dass auch der Boden ein über die Jahre hinweg stabiles Elementprofil besitzt. Die Vermutung, dass das Elementprofil der Spargelproben die Elementzusammensetzung des Bodens widerspiegelt bzw. dieser stark ähnelt, führt des Weiteren zu der Hypothese, dass eine systematische Aufnahme der Elemente durch den Spargel aus dem umgebenden Boden stattfindet. Die Bodenzusammensetzung zählt bekanntermaßen zu den wichtigsten Einflussfaktoren für das Elementprofil einer Pflanze [318, 319]. In der Literatur wurden jedoch mitunter sehr unterschiedliche Elementaufnahmen beschrieben, die von hoher Korrelation bishin zu keinem erkennbaren Zusammenhang zwischen den Konzentrationen in Pflanze und Boden reichen, abhängig von der untersuchten Kultur [130, 320–323]. Daher soll im Folgenden untersucht werden, ob und welche Zusammenhänge sich aus dem Elementprofil von Spargel und Boden ableiten lassen. Dazu wurden die im Jahr 2014 akquirierten Bodenproben hinzugezogen, welche jeweils vom selben Feld wie die Spargelproben aus dem Erntejahr 2014 stammen. Es wurden Bodenproben von 36 verschiedenen Spargelfeldern akquiriert, wobei an jedem Standort 3 Proben aus einer Bodentiefe von jeweils 0 - 30 cm, 30 - 60 cm sowie 60 - 90 cm entnommen wurden. Vor der Analyse der Bodenproben wurde eine Extraktion vorgenommen, um den bioverfügbaren Anteil der Elemente im Boden zu ermitteln.

### 4.3.1 Korrelationen zwischen den Elementen

Zunächst wurden alle Elementkonzentrationen auf Korrelationen untersucht. Dafür wurde der Korrelationskoeffizient nach Pearson ( $r_P$ ) aller Elemente für die im Jahr 2014 geernteten Spargelproben berechnet und mit denen der Bodenproben verglichen. Um eine ähnliche Anzahl an Wertepaaren von Spargel (41 Proben) und Boden (108 Proben) gegenüberstellen zu können, wurde zunächst nur auf die Bodenproben einer Tiefe zurückgegriffen. Die Hauptwurzelmasse des Spargels liegt in einer Tiefe von 20 - 70 cm, was am besten durch die Proben der Bodentiefe von 30 - 60 cm abgebildet wird [53]. Aufgrund der Annahme einer Elementaufnahme durch das Wurzelgeflecht der Spargelpflanze aus dem umgebenden Erdreich wurden diese Proben für den Korrelationsvergleich herangezogen. Die Berechnung der Korrelationen erfolgte anhand der im Abschnitt 8.11





Die Berechnungen wurden kombiniert mit einem t-Test zur Prüfung auf signifikante Zusammenhänge. Dazu wurde eine Prüfgröße ermittelt (s. Abschnitt 8.11), welche einer t-Verteilung folgt, und diese mit dem kritischen t-Wert verglichen [324–326]. Weist die berechnete Prüfgröße  $t$  einen Wert oberhalb des kritischen Werts auf, so liegt eine signifikante Korrelation vor. Die dabei berechneten t-Werte sind den Abbildungen 61 und 62 im Anhang zu entnehmen.

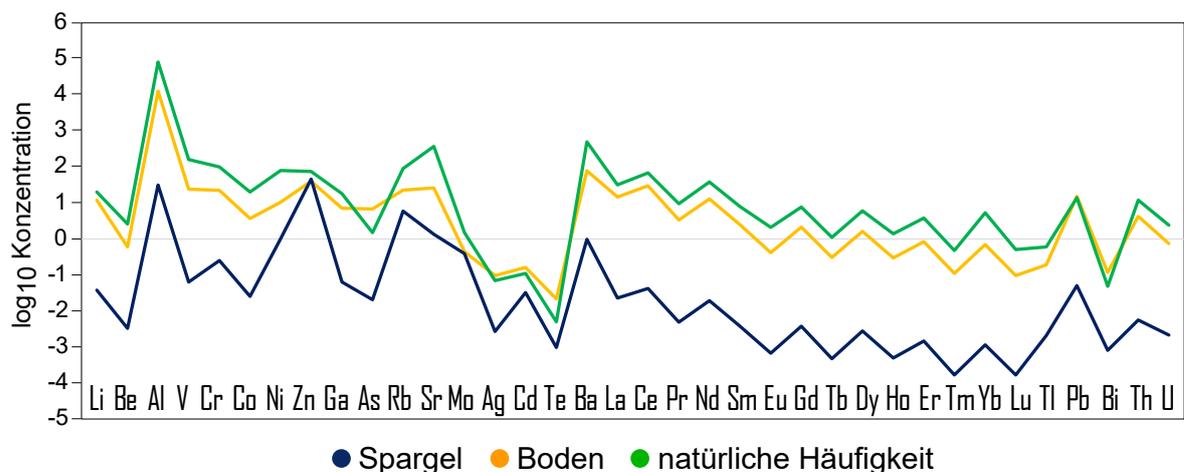
Analog zur zuvor gezeigten *Heatmap* (Abbildung 25) wurden in Abbildung 34 hochsignifikante Korrelationen zwischen den Metallen der seltenen Erden untereinander beobachtet (Lanthan bis Lutetium) mit Korrelationskoeffizienten zwischen 0.657 und 0.996, welche in grüner Farbe dargestellt sind. Zudem liegen ebenfalls signifikante Korrelationen zwischen den Elementen Thorium, Uran, Aluminium, Cobalt, Arsen, Vanadium bzw. Barium einerseits und den seltenen Erden andererseits vor, die in der genannten Reihenfolge schwächer werden mit Korrelationskoeffizienten von 0.957 bis 0.406. Des Weiteren konnten vereinzelt Elementpaare mit einer hohen Korrelation identifiziert werden, darunter Gallium-Barium und Beryllium-Tellur ( $r_P = 0.983$  und  $0.908$ ). Dagegen sind keine Korrelationen der Elemente Chrom, Nickel, Zink, Rubidium, Strontium, Molybdän, Cadmium, Thallium und Bismut mit fast allen anderen Elementen erkennbar. Bei den letztgenannten Elementen treten zum Teil auch negative Korrelationskoeffizienten auf, die jedoch einen niedrigen Betrag haben und keine signifikante Korrelation zeigen. Einzige Ausnahme stellt Zink dar, welches eine signifikant negative Korrelation zu den seltenen Erden Holmium bis Lutetium aufweist ( $|r_P| = 0.40 - 0.46$ ). Alle übrigen Elemente weisen Koeffizienten von etwa 0.5 auf, sodass man hier von schwachen Korrelationen sprechen kann.

Im Gegensatz dazu sind bei den in Abbildung 35 dargestellten Korrelationskoeffizienten der im Boden gemessenen Elemente keine negativen Korrelationen zu verzeichnen. Im Vergleich zu den Spargelproben sind im Boden insgesamt stärkere Korrelationen und größere t-Werte für alle Elemente zu beobachten, was für eine Vielzahl an Korrelationen zwischen den untersuchten Elementen spricht. Die stärksten Korrelationen sind auch hier zwischen den Lanthanoiden zu verzeichnen, welche Korrelationskoeffizienten zwischen 0.860 und 0.999 aufweisen. Insgesamt zeigen alle Elemente signifikante Korrelationen zueinander, mit Ausnahme von Zink, Molybdän und Tellur mit deutlich niedrigeren Korrelationskoeffizienten ( $r_P$  im Mittel: 0.23).

In beiden Matrices, Spargel und Boden, war bei den Elementen Zink und Molybdän keine Korrelation mit anderen Elementen bzw. im Spargel eine negative Korrelation zwischen Zink und einigen seltenen Erden zu beobachten. Während im Spargel auch weitere Elemente ohne signifikante Korrelationen vorliegen, ist dies im Boden nur für das Element Thallium zu verzeichnen. Recherchen haben gezeigt, dass zusätzlich zu den eingangs erwähnten Elementen, welche beim Anbau der hier verwendeten Spargelproben

in Form von Dünge- oder Pflanzenschutzmitteln verwendet wurden, im Spargelanbau auch der Einsatz von Zink und Molybdän in Düngemitteln verbreitet ist [327–329]. Eine solche nicht natürliche Beeinflussung der Elementkonzentration könnte daher eine plausible Erklärung für das abweichende Verhalten dieser beiden Elemente bei der Untersuchung der Korrelationen sein.

Die insgesamt schwächeren Korrelationen im Spargel sind vermutlich auf die Aufnahmeprozesse der Elemente vom Boden in die Pflanze zurückzuführen, da eine Elementaufnahme nur dann ermöglicht wird, wenn die Pflanze über Transportmechanismen für die jeweiligen Elemente verfügt. Dagegen wird vermutet, dass die sehr starken Korrelationen im Boden im Zusammenhang mit der natürlichen Häufigkeit der Elemente in der Erdkruste stehen. Aus diesem Grund wurde die natürliche Häufigkeit mit den Konzentrationen in Spargel und Boden gegenübergestellt. Zuvor wurde eine Logarithmierung der Konzentrationen vorgenommen, um die unterschiedlichen Größenordnungen vergleichen zu können. Die daraus resultierenden Konzentrationskurven sind in Abbildung 36 dargestellt. Für die drei dargestellten Parameter sind insgesamt sehr ähnliche Kurvenverläufe zu erkennen. Die Konzentrationskurven von Boden und natürlicher Häufigkeit verlaufen für die meisten Elemente nahezu parallel, wobei die Kurve für die natürliche Häufigkeit etwas oberhalb liegt. Eine Abweichung von diesem Verlauf ist besonders bei Arsen zu beobachten, aber auch bei den Elementen Silber, Cadmium, Tellur, Blei und Bismut, bei denen die logarithmierte Konzentration des Bodens die natürliche Häufigkeit überschreitet. Dies deutet darauf hin, dass die erwähnten Elemente im bioverfügbaren Anteil des hier untersuchten Bodens in etwas höheren Konzentrationen als im Mittel in der Erdkruste vorkommen. Ein ähnlicher Verlauf ist auch bei der Konzentrationskurve des Spargels zu beobachten, die jedoch deutlich unterhalb der beiden anderen Kurven liegt und somit insgesamt



**Abbildung 36:** Mittlere logarithmierte Konzentrationen aller Elemente in Spargel und Boden sowie die natürliche Häufigkeit in der Erdkruste.

geringere Konzentrationen anzeigt. Auch hier sind einzelne Abweichungen erkennbar, insbesondere bei den Elementen Zink, Rubidium, Molybdän, Cadmium und Tellur, deren Konzentrationen deutlich höher liegen und im Fall von Zink und Molybdän sogar jene im Boden erreichen. Diese Beobachtungen legen eine besonders gute Aufnahme dieser Elemente durch die Spargelpflanze nahe, sodass im Verhältnis zu den übrigen Elementen höhere Konzentrationen erreicht werden. Dies unterstreicht zudem die Annahme einer höheren Verfügbarkeit von Zink und Molybdän aufgrund von Düngemitteln zur Versorgung der Pflanze mit Mikronährstoffen. Bei allen genannten Elementen mit Abweichungen im Verhältnis zur natürlichen Häufigkeit handelt es sich des Weiteren um solche Elemente, bei denen mit wenigen Ausnahmen keine signifikanten Korrelationen im Spargel beobachtet wurden. Dies verweist auf eine insgesamt unspezifische und ungerichtete Aufnahme dieser Elemente.

#### 4.3.2 Korrelationen der seltenen Erden

Da zuvor an verschiedenen Stellen eine starke Korrelation zwischen den Metallen der seltenen Erden auffiel, wurden die Daten dieser Elemente noch einmal genauer untersucht. Dazu wurden die Korrelationskoeffizienten nach Pearson isoliert betrachtet, was in Abbildung 37 gezeigt ist. Dabei fällt auf, dass bei allen seltenen Erdelementen die stärkste Korrelation jeweils zu den direkten Nachbarn mit ähnlichster Massenzahl besteht. Mit zunehmender Massendifferenz zwischen zwei Elementen nimmt dagegen sowohl im Spargel als auch im Boden die Korrelation ab, sodass Lanthan/Cer und Lutetium als leichteste und schwerste Elemente dieser Gruppe die schwächste Korrelation zueinander aufweisen mit einem Koeffizienten von 0.657 im Spargel und 0.842 im Boden.

Ein von diesem Trend abweichendes Verhalten ist für das Element Europium in den Spargelproben und in abgeschwächter Form auch in den Bodenproben zu beobachten. Europium zeigt merklich schwächere Korrelationen zu allen anderen Elementen, auch den unmittelbaren Nachbarn, eine starke Korrelation wurde dagegen zuvor bereits mit dem Element Barium beobachtet (vgl. Abb. 34 und 35). Eine Erklärung für diese Abweichung könnte das Vorliegen einer massenspezifischen Interferenz liefern, ausgelöst durch polyatomische Bariumoxid-Ionen, wie sie bereits in der Literatur in diesem Zusammenhang beschrieben wurde [330]. Zur Bestimmung der Europiumkonzentrationen wurde das Isotop  $^{153}\text{Eu}$  herangezogen. Ein Bariumoxid-Teilchen mit der Isotopenverteilung  $^{137}\text{Ba}^{16}\text{O}$  weist ganzzahlig betrachtet dieselbe Masse wie  $^{153}\text{Eu}$  auf, sodass bei gleichzeitigem Vorliegen beider Ionen eine Massentrennung im Quadrupol nicht möglich wäre [214]. Die Messung des Isotops  $^{153}\text{Eu}$  erfolgte daher im Helium-Modus bei aktivierter Kollisionszelle. Aufgrund der hier gezeigten Ergebnisse ist anzunehmen, dass unter den gewählten Messbedingungen keine vollständige Auslöschung der

<b>a</b>	La	Ce	Pr	Nd	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
La	1.000	0.964	0.960	0.942	0.881	0.691	0.891	0.866	0.826	0.800	0.774	0.722	0.716	0.679
Ce	0.964	1.000	0.977	0.971	0.933	0.733	0.934	0.896	0.844	0.801	0.767	0.696	0.696	0.657
Pr	0.960	0.977	1.000	0.996	0.970	0.784	0.963	0.925	0.879	0.831	0.804	0.714	0.744	0.694
Nd	0.942	0.971	0.996	1.000	0.985	0.810	0.979	0.942	0.897	0.849	0.821	0.729	0.759	0.712
Sm	0.881	0.933	0.970	0.985	1.000	0.857	0.994	0.961	0.924	0.871	0.848	0.741	0.790	0.743
Eu	0.691	0.733	0.784	0.810	0.857	1.000	0.869	0.848	0.825	0.794	0.785	0.717	0.771	0.756
Gd	0.891	0.934	0.963	0.979	0.994	0.869	1.000	0.981	0.954	0.912	0.891	0.796	0.837	0.796
Tb	0.866	0.896	0.925	0.942	0.961	0.848	0.981	1.000	0.984	0.967	0.947	0.887	0.903	0.859
Dy	0.826	0.844	0.879	0.897	0.924	0.825	0.954	0.984	1.000	0.986	0.982	0.909	0.950	0.918
Ho	0.800	0.801	0.831	0.849	0.871	0.794	0.912	0.967	0.986	1.000	0.993	0.961	0.972	0.944
Er	0.774	0.767	0.804	0.821	0.848	0.785	0.891	0.947	0.982	0.993	1.000	0.956	0.987	0.967
Tm	0.722	0.696	0.714	0.729	0.741	0.717	0.796	0.887	0.909	0.961	0.956	1.000	0.955	0.943
Yb	0.716	0.696	0.744	0.759	0.790	0.771	0.837	0.903	0.950	0.972	0.987	0.955	1.000	0.978
Lu	0.679	0.657	0.694	0.712	0.743	0.756	0.796	0.859	0.918	0.944	0.967	0.943	0.978	1.000

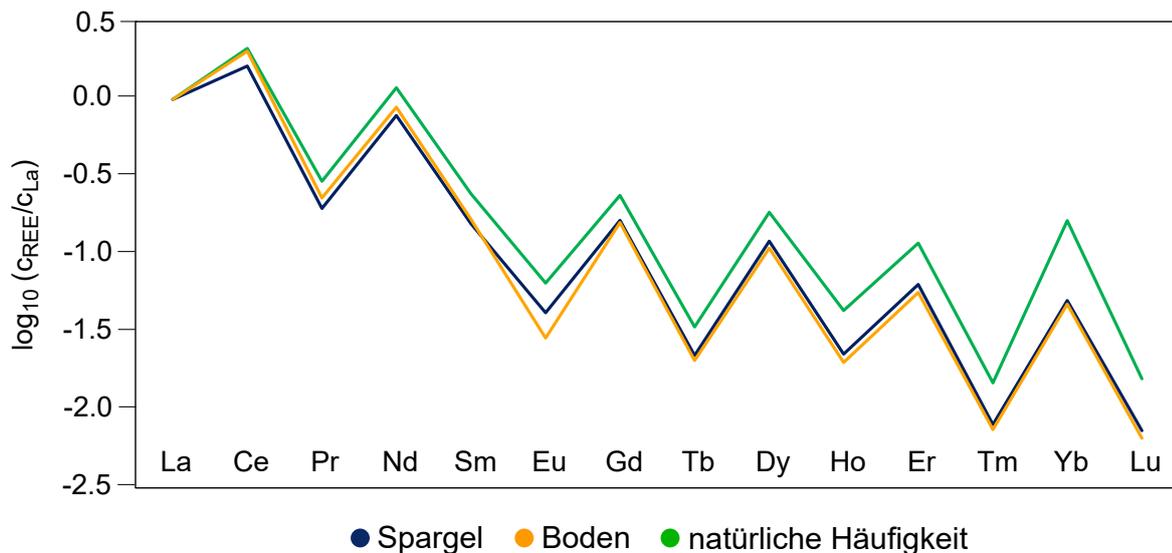
<b>b</b>	La	Ce	Pr	Nd	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
La	1.000	0.997	0.998	0.997	0.990	0.970	0.987	0.976	0.954	0.927	0.907	0.873	0.856	0.842
Ce	0.997	1.000	0.995	0.996	0.990	0.968	0.987	0.976	0.955	0.929	0.909	0.875	0.860	0.846
Pr	0.998	0.995	1.000	0.999	0.993	0.974	0.990	0.980	0.958	0.932	0.911	0.877	0.859	0.847
Nd	0.997	0.996	0.999	1.000	0.997	0.981	0.995	0.985	0.964	0.937	0.917	0.882	0.865	0.852
Sm	0.990	0.990	0.993	0.997	1.000	0.989	0.999	0.991	0.971	0.946	0.925	0.890	0.872	0.860
Eu	0.970	0.968	0.974	0.981	0.989	1.000	0.991	0.992	0.980	0.961	0.943	0.913	0.895	0.885
Gd	0.987	0.987	0.990	0.995	0.999	0.991	1.000	0.995	0.980	0.957	0.938	0.906	0.889	0.877
Tb	0.976	0.976	0.980	0.985	0.991	0.992	0.995	1.000	0.994	0.979	0.964	0.939	0.924	0.914
Dy	0.954	0.955	0.958	0.964	0.971	0.980	0.980	0.994	1.000	0.995	0.987	0.970	0.959	0.952
Ho	0.927	0.929	0.932	0.937	0.946	0.961	0.957	0.979	0.995	1.000	0.997	0.988	0.980	0.976
Er	0.907	0.909	0.911	0.917	0.925	0.943	0.938	0.964	0.987	0.997	1.000	0.996	0.992	0.987
Tm	0.873	0.875	0.877	0.882	0.890	0.913	0.906	0.939	0.970	0.988	0.996	1.000	0.998	0.997
Yb	0.856	0.860	0.859	0.865	0.872	0.895	0.889	0.924	0.959	0.980	0.992	0.998	1.000	0.999
Lu	0.842	0.846	0.847	0.852	0.860	0.885	0.877	0.914	0.952	0.976	0.987	0.997	0.999	1.000

**Abbildung 37:** Korrelationen nach Pearson für die seltenen Erden in  
 (a) Spargel- und (b) Bodenproben des Erntejahres 2014.

Bariumoxid-Interferenz möglich war. Da die Bariumkonzentrationen in beiden Matrices deutlich oberhalb der Europiumgehalte liegen, erscheint diese Vermutung plausibel. Hinzu kommt, dass das Verhältnis der Gehalte an Barium und Europium im Spargel etwa 10 Mal höher liegt als im Boden, was erklären könnte, warum die Abweichung im Spargel besonders ausgeprägt ist. Des Weiteren könnte der Beobachtung eine veränderte Aufnahme des Elementes Europium in die Spargelpflanzen zugrunde liegen. Die seltenen Erden kommen natürlicherweise bevorzugt als trivalente Kationen vor, sodass grundsätzlich eine Aufnahme in die Spargelpflanze in dieser Form angenommen wird. Eine ähnliche Abweichung der Europiumkonzentrationen wie in den Spargelproben wurde auch von Tyler bei seiner Untersuchung von Wurzeln der Rotbuche im Vergleich zu Proben aus dem umgebenden Boden beobachtet [331]. Aufgrund von höheren Aufnahmeraten von Europium im Verhältnis zu den übrigen seltenen Erden, was sich auch in einem größeren Verhältnis der Konzentrationen von Pflanze und Boden äußerte, wurde in der erwähnten Arbeit auf eine erleichterte Europiumaufnahme in Form eines divalenten Kations geschlossen [331]. Auch in der vorliegenden Arbeit weist Europium

ein erhöhtes Verhältnis der Konzentration in Spargel und Boden auf. Dieses liegt bei 0.0023, während die übrigen seltenen Erden Verhältnisse zwischen 0.0015 und 0.0019 zeigen. Die aus dem Boden aufgenommene Menge ist damit insgesamt extrem niedrig.

Für einen präziseren Vergleich der Konzentrationsverhältnisse wurden die Gehalte der seltenen Erden einer Lanthannormalisierung unterzogen, bei der jedes Lanthanoid als Verhältnis zur Lanthankonzentration ausgedrückt wird. Nach zusätzlicher Logarithmierung ergeben sich Konzentrationsverhältnisse in derselben Größenordnung, die eine sehr gute Vergleichbarkeit ermöglichen. Die mittleren lanthannormalisierten Konzentrationen der seltenen Erden in den deutschen Spargel- und Bodenproben aus dem Erntejahr 2014 kombiniert mit der Verteilung der natürlichen Häufigkeit sind in Abbildung 38 dargestellt. Die Kurven der Konzentrationsverhältnisse von Spargel und Boden verlaufen fast deckungsgleich, sodass die Verhältnisse der seltenen Erden untereinander in beiden Matrices fast identisch sind. Im Detail betrachtet liegt das Konzentrationsverhältnis im Boden bei den Elementen Cer, Praseodym und Neodym etwas höher als im Spargel, während dieses bei den übrigen Elementen etwas unterhalb des Spargels verläuft. Eine Ausnahme bildet wie zuvor angedeutet das Element Europium, welches in den Spargelproben ein deutlich höheres Verhältnis aufweist. Die Kurve der natürlichen Häufigkeit verläuft oberhalb und beinahe parallel zu den Kurven von Spargel- und Bodenproben. Die gezeigten relativen Konzentrationen lassen somit den Schluss zu, dass für die meisten Lanthanoide eine geringe Selektivität bei der Aufnahme vorliegt, weshalb die Verteilung in weiten Teilen die Bodenverhältnisse und die natürliche Häufigkeit widerspiegelt. Eine spezifische Aufnahme könnte dagegen für das Element Europium vorliegen.



**Abbildung 38:** Logarithmierte mittlere Konzentrationen der seltenen Erden bezogen auf die Lanthan-Konzentration für die deutschen Spargel- und Bodenproben des Erntejahres 2014 im Vergleich zur natürlichen Häufigkeit.

In der Literatur werden Aufnahmewege für seltene Erden diskutiert, die die Mechanismen der Calciumaufnahme nutzen. Grundlage dieser Annahme ist die chemische Ähnlichkeit von Calcium- und Lanthanoidionen, insbesondere in Bezug auf den Ionenradius [332–336]. Daher ist eine Aufnahme der seltenen Erden beispielsweise über Calcium-Ionenkanäle auch im Spargel denkbar.

### 4.3.3 Korrelationen zwischen Spargel und Boden

Der Zusammenhang des Elementprofils von Spargel und Boden wurde schließlich anhand der Korrelation zwischen beiden Matrices untersucht. Dafür wurde der Korrelationskoeffizient nach Pearson herangezogen, zusätzlich erfolgte die Berechnung des jeweils zugehörigen t-Wertes, um die Signifikanz der Korrelation beurteilen zu können. Die Ergebnisse sind in Tabelle 10 dargestellt (Berechnung s. Abschnitt 8.11).

Demnach weisen 25 der insgesamt 36 Elemente hochsignifikante positive Korrelationen zwischen den beiden Matrices auf, die in Tabelle 10 in grauer Farbe hervorgehoben sind. Die stärkste Korrelation wurde für das Element Barium beobachtet mit einem Korrelationskoeffizienten von 0.81, gefolgt von den Elementen Gallium, Lithium und Cobalt mit Koeffizienten von 0.78 - 0.74. Zuvor wurde bereits die starke Korrelation zwischen den Elementen Barium und Gallium diskutiert, die sowohl im Spargel als auch im Boden auftritt und die sich auch beim Vergleich beider Matrices durch starke Korrelationen dieser Elemente widerspiegelt. Des Weiteren zeigte sich auch für die seltenen Erden ein starker Zusammenhang der Konzentration in Spargel und Boden ( $r_P = 0.63 - 0.75$ ). Dagegen weisen die Elemente Beryllium, Vanadium, Chrom, Zink, Rubidium, Molybdän, Silber, Cadmium, Tellur, Thallium und Bismut keine signifikanten Korrelationen zwischen Spargel- und Bodenproben auf. Die fehlende Korrelation für Zink und Molybdän könnte, wie zuvor erwähnt, im Zusammenhang mit einer möglichen Verwendung in Düngemitteln stehen.

Insgesamt konnte für die Mehrzahl der untersuchten Elemente ein positiver linearer Zusammenhang der Konzentration im Spargel zu der im Boden nachgewiesen werden. Dies bedeutet, dass ein erhöhter Gehalt dieser Elemente im Boden auch zu höheren Konzentrationen im Spargel führt. Dies wiederum setzt eine Aufnahme der Elemente durch die Spargelpflanze voraus, sodass davon auszugehen ist, dass Aufnahmewege für die verschiedenen Elemente bestehen. Während einige Elemente in Abhängigkeit von der Verfügbarkeit im Boden aufgenommen werden, ist die Aufnahme der Elemente ohne Korrelation zwischen Boden und Spargel jedoch begrenzt.

**Tabelle 10:** Elementkorrelationen in Spargel und Boden: Korrelationskoeffizienten nach Pearson ( $r_P$ ) sowie t-Werte (krit. t-Wert = 2.728, 34 Freiheitsgrade, Signifikanzniveau = 0.01, zweiseitige Verteilung [326]).

Element	$r_P$	t-Wert
Li	0.75	6.664
Be	0.10	0.571
Al	0.64	4.899
V	0.36	2.259
Cr	0.12	0.733
Co	0.74	6.509
Ni	0.69	5.600
Zn	0.24	1.446
Ga	0.78	7.311
As	0.56	3.949
Rb	0.05	0.289
Sr	0.60	4.331
Mo	0.00	0.009
Ag	-0.02	0.090
Cd	0.03	0.181
Te	-0.07	0.407
Ba	0.81	8.012
La	0.64	4.830
Ce	0.65	4.957
Pr	0.72	6.007
Nd	0.72	6.061
Sm	0.73	6.306
Eu	0.69	5.523
Gd	0.75	6.573
Tb	0.74	6.398
Dy	0.74	6.339
Ho	0.73	6.187
Er	0.73	6.164
Tm	0.59	4.241
Yb	0.72	5.998
Lu	0.63	4.759
Tl	0.24	1.468
Pb	0.56	3.961
Bi	-0.03	0.159
Th	0.74	6.384
U	0.55	3.846

Prinzipiell ist die Überwindung der Zellmembran und somit die Aufnahme von Elementen in die Pflanze über verschiedene Transportproteine möglich, beispielsweise über Carrier-Proteine, Ionenkanäle oder Ionenpumpen. Der Transport kann grundsätzlich aktiv oder passiv erfolgen, je nachdem ob dieser mit oder ohne Energieverbrauch stattfindet. Ein Übergang durch einfache Diffusion ist dagegen unwahrscheinlich, da die Elemente für den Transport im Boden gelöst vorliegen müssen und als Ion in geladener Form die unpolare Membran nicht überwinden können [337, 338]. Eine Reihe von Elementen werden in der Pflanze als Makro- und Mikronährstoffe benötigt, darunter Natrium, Kalium, Calcium, Eisen, Kupfer und Zink, da sie essentielle Funktionen beispielsweise als Signalmolekül, zur Membranstabilisierung oder als Cofaktor für diverse Proteine haben [339]. Für diese Elemente existieren bekannte Transportsysteme, die für eine Regulierung der Konzentration auf ein physiologisch notwendiges Level sorgen, da höhere Mengen oftmals toxische Wirkungen in der Pflanze aufweisen [339–341].

Bei den in der vorliegenden Arbeit untersuchten Elementen handelt es sich dagegen mit großer Mehrheit nicht um essentielle Elemente, die die Spargelpflanze für physiologische Prozesse benötigt. Daher ist es unwahrscheinlich, dass die Pflanze über spezifische Transportsysteme für Schwermetalle wie Cadmium, Blei oder Uran verfügt, dennoch lässt der analytische Befund auf eine Aufnahme dieser Elemente schließen. Ein möglicher Aufnahmeweg besteht in den P-Typ ATPasen, den sogenannten Schwermetall-ATPasen, die in pflanzlichen Zellen vorkommen und beispielsweise für die Regulierung von Kupfer- und Zinkkonzentrationen zuständig sind, aber auch eine Aufnahme von Cadmium und Blei ermöglichen [342]. Des Weiteren haben Untersuchungen gezeigt, dass die meisten Kationen-Transporter in den Zellmembranen ein breites Substratspektrum aufweisen und somit auch nicht-essentiellen Elementen zugänglich sind, sodass beispielsweise Cadmium und Blei mithilfe dieser Transporterproteine in die Zelle gelangen können. Für weitere Elemente wie etwa Nickel wird ein solcher Transportmechanismus ebenfalls vermutet [140, 343–345]. Zudem verfügen pflanzliche Zellen über Mechanismen der Komplexierung von Metallen, beispielsweise mithilfe von Salzen organischer Säuren wie Citrat oder Malat, die zur Mobilisierung beitragen und den Transport von Metallen durch die Membran ermöglichen. Zudem wurde eine Maskierung toxikologischer Effekte durch die Komplexierung der Metalle beobachtet, weshalb Pflanzen mitunter deutlich erhöhte und sonst schädliche Metallkonzentrationen tolerieren können [140, 346]. Andere Arbeiten zeigen eine erhöhte Toleranz gegenüber hohen Konzentrationen an Schwermetallen beim Vorliegen einer symbiotischen Beziehung zwischen Pflanze und Rhizobien [347].

Es existiert somit eine Vielzahl an Ansatzpunkten zur Erklärung der beobachteten Elementaufnahmen durch die Spargelpflanze und die vielfach starken Korrelationen zwischen den Matrices Spargel und Boden. Die Ergebnisse haben insgesamt gezeigt, dass ein sehr enger Zusammenhang des Elementprofils im Spargel mit

dem im bioverfügbaren Anteil des Bodens existiert, der zu einer über Jahre reproduzierbaren Elementzusammensetzung im Spargel führt. Diese Erkenntnisse sind zudem konsistent zu den Arbeiten von Swoboda *et al.* und Zannella *et al.*, die von nahezu übereinstimmenden Strontium-Isotopenverhältnissen ( $^{87}\text{Sr}/^{86}\text{Sr}$ ) im Spargel und im bioverfügbaren Anteil des zugehörigen Bodens berichtet haben [105, 107].

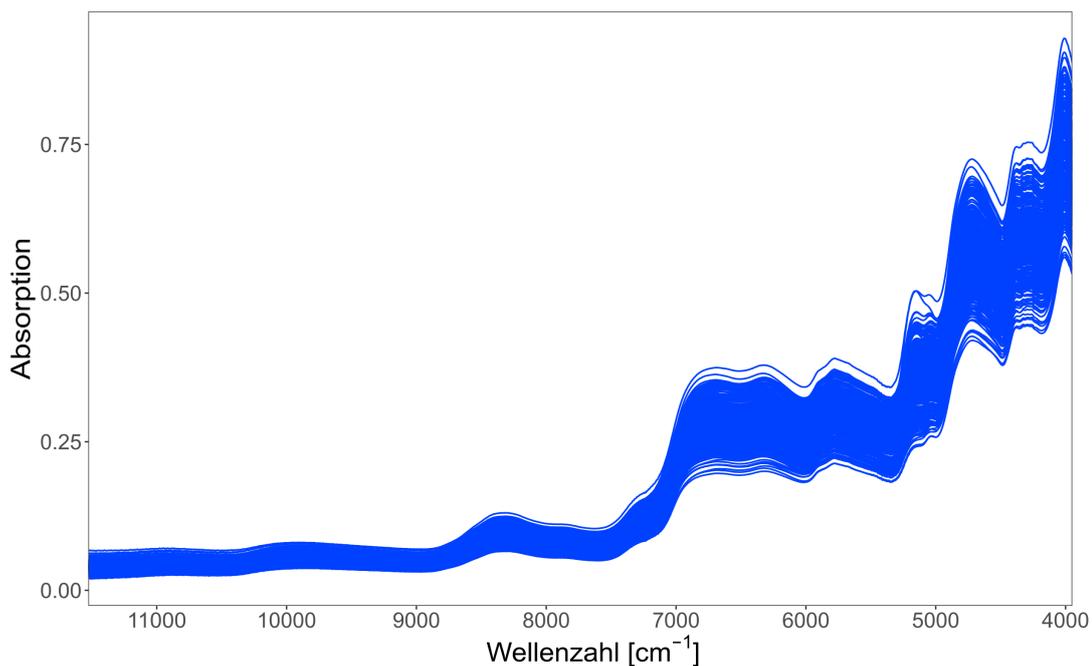
---

## 4.4 Herkunftsbestimmung mittels FT-NIR

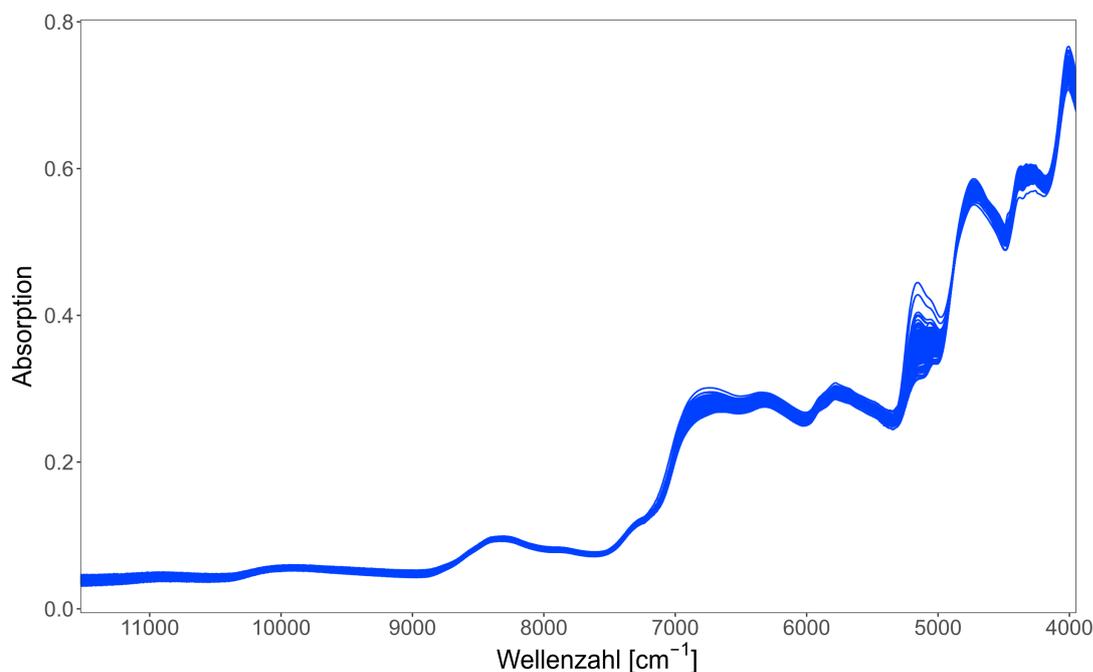
Als zweiter Ansatzpunkt für die geographische Herkunftsunterscheidung wurde die Nahinfrarotspektroskopie herangezogen, die überwiegend Metabolite des Spargels anspricht. Dafür wurden FT-NIR-Messungen der Proben durchgeführt und die Daten anschließend einer statistischen Auswertung unterzogen.

### 4.4.1 Datenpräprozessierung

Bei NIR-spektroskopischen Messungen werden eine Reihe unerwünschter Effekte beobachtet, die eine Abweichung des linearen Zusammenhangs zwischen Absorption und Konzentration der Probe hervorrufen, weshalb eine sorgfältige Datenvorbehandlung vor der eigentlichen Auswertung von großer Bedeutung ist (vgl. Abschnitt 2.4.3). Da die Spargelproben für die vorliegende Arbeit in pulverisierter Form vermessen wurden, waren insbesondere eine Streuung an den Probenpartikeln und Streueffekte durch nicht vollständig homogene Partikelgrößen zu erwarten. Abbildung 39 zeigt die Rohspektren aller Proben vor der Prozessierung. Diese weisen deutliche Offset-Effekte auf, da der Abstand zwischen den einzelnen Spektren im Verlauf hin zu niedrigeren Wellenzahlen immer weiter zunimmt. Da dieses Phänomen mit einer Lichtstreuung an den Probenpartikeln zu erklären ist, wurde eine Streulichtkorrektur (MSC) durchgeführt. Die MSC-korrigierten Spektren, welche Abbildung 40 zu entnehmen sind, liegen insgesamt deutlich näher und über den gesamten Spektrenverlauf vergleichbar beieinander, sodass die Offset-Effekte auf diese Weise erfolgreich ausgeglichen werden konnten.

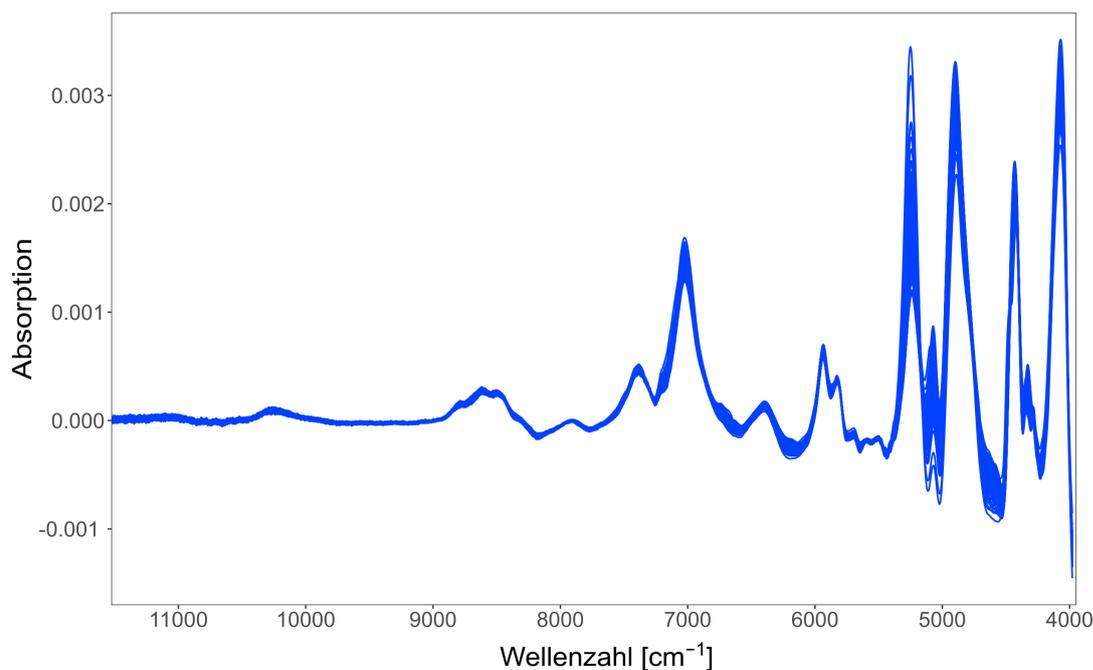


**Abbildung 39:** NIR-Rohdaten der Spargelproben ohne Datenpräprozessierung.



**Abbildung 40:** MSC-korrigierte NIR-Spektren der Spargelproben.

Zudem werden einzelne Regionen sichtbar, in denen deutliche Unterschiede zwischen den Spektren erkennbar sind. Im Gegensatz dazu wurde das Ansteigen der Basislinie mithilfe der MSC nicht verringert. Aus diesem Grund wurden die Spektren zusätzlich einer Ableitung ersten Grades unterzogen, wie in Abbildung 41 gezeigt ist. Mithilfe der Ableitung konnten sowohl Basislinienunterschiede als auch der Trend der linear ansteigenden Basislinie korrigiert werden.



**Abbildung 41:** NIR-Spektren der Spargelproben nach Anwendung von MSC-Korrektur und 1. Ableitung.

Da die Spektren mit einer recht hohen Auflösung von  $2\text{ cm}^{-1}$  aufgenommen wurden, lagen insgesamt 3'725 Variablen pro Spektrum vor. Im Zuge der Auswertung zeigte sich, dass die benachbarten Wellenzahlen sehr häufig fast identische Absorptionswerte aufweisen. Um redundante Daten zu vermeiden, wurde daher ein *Binning* vorgenommen, bei dem jeweils der Mittelwert der Absorption für 10 benachbarte Wellenzahlen gebildet und die Messpunkte somit zu 373 Variablen zusammengefasst wurden.

#### 4.4.2 Probenauswahl

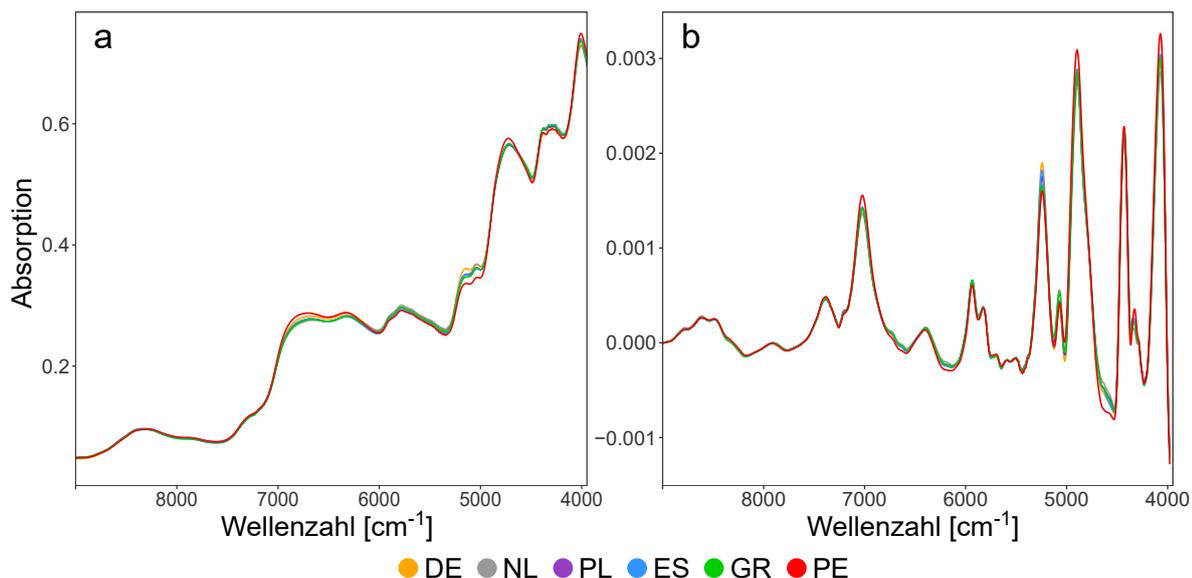
Im Zuge der Probenakquise im Jahr 2014 wurden die Spargelproben bei einer Temperatur von  $-20^\circ\text{C}$  transportiert, abweichend zu den Folgejahren, in denen der Transport bei  $+3^\circ\text{C}$  erfolgte (vgl. Abschnitt 6.1.1). Da die Temperatur einen bekannten Einflussfaktor für das Metabolom darstellt, ist eine Vergleichbarkeit der Proben aus den unterschiedlichen Erntejahren hier nicht mehr gegeben [156, 158]. Aus diesem Grund wurden die Proben aus dem Jahr 2014 nicht in die weitere Datenauswertung einbezogen. Hinzu kommt, dass aus diesem Erntejahr nur eine ausländische Probe vorlag. Bei allen weiteren handelte es sich um Proben deutscher Herkunft, sodass deren Berücksichtigung bei der statistischen Auswertung eher noch zu einer Verstärkung des Klassenungleichgewichtes führen würde, was bei der Klassifikation jedoch möglichst vermieden werden sollte.

Des Weiteren wurden die chinesischen Proben aus den Messungen mittels NIR ausgeschlossen. Diese wurden als einzige Proben tiefgefroren und geschält erworben, was ebenfalls als bedeutsamen Eingriff in das Metabolom des Spargels zu bewerten ist. Wie bereits beschrieben, ist beim Spargel rasch nach der Ernte eine hohe metabolische Aktivität zu beobachten, die zu einer Reihe von Veränderungen von Struktur und Zusammensetzung führt [93, 348]. Das Entfernen der Schale stellt jedoch eine weit massivere Verletzung der Pflanzenstruktur dar, als dies beim Abtrennen der Spargelstange bei der Ernte der Fall ist. Da die Schale Schutzfunktionen für die darunterliegenden Zellen ausübt, ist somit davon auszugehen, dass dies eine gravierende Veränderung der Stoffwechselaktivitäten mit sich bringt, beispielsweise durch das Freisetzen von Enzymen und daraus resultierenden Reaktionswegen [349, 350]. Aus den genannten Gründen ist keine Vergleichbarkeit mit den übrigen Proben gegeben, bei denen die Schale zu keinem Zeitpunkt nach der Ernte oder während der Probenaufarbeitung entfernt wurde.

Insgesamt standen somit 275 Proben aus den Ländern Deutschland (180), Polen (21), den Niederlanden (20), Griechenland (25), Spanien (8) und Peru (21) sowie den Erntejahren 2015 - 2017 für die Methodenentwicklung zur Herkunftsbestimmung mittels NIR-Spektroskopie zur Verfügung.

### 4.4.3 Datenvisualisierung

Um einen Überblick über den Datensatz und mögliche Unterschiede zwischen den Proben der verschiedenen Länder zu erhalten, wurden zu Beginn die Mittelwertspektren der einzelnen Länder miteinander verglichen. Dabei wurde im Hinblick auf die Peakinterpretation zunächst auf die Anwendung der 1. Ableitung verzichtet (Abb. 42 a). Wie in Abbildung 40 bereits zu erkennen war, sind die Proben in der ersten Hälfte des Spektrums sehr ähnlich und überlagern fast vollständig. Erst bei kleineren Wellenzahlen zeigen sich Unterschiede, daher wurde der Bereich  $> 9000 \text{ cm}^{-1}$  in dieser Darstellung zunächst nicht berücksichtigt. Die mittleren Spektren weisen einige Maxima auf, bei denen sich einzelne Länder abgrenzen lassen. Am deutlichsten weichen die peruanischen Proben ab, deren Mittelwertspektrum an einigen Stellen deutlich ober- oder unterhalb der übrigen Spektren liegt. Auch das mittlere Spektrum der deutschen Proben hebt sich an einigen Stellen merklich von den anderen ab. Die vollständigen Spektren über den gesamten Wellenzahlenbereich sind in Abbildung 64 im Anhang im Abschnitt 8.9 dargestellt.

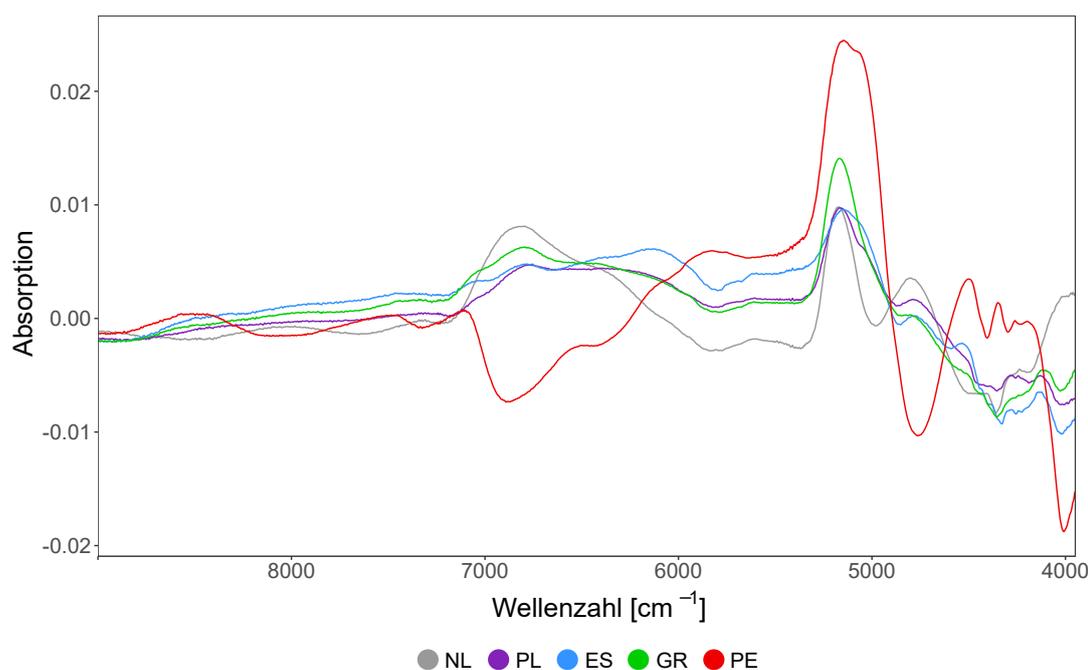


**Abbildung 42:** Mittlere MSC-korrigierte FT-NIR-Spektren der Spargelproben aller Herkunftsländer (a) vor und (b) nach Anwendung der 1. Ableitung.

NIR-Spektren sind prinzipiell verhältnismäßig schwer zu interpretieren, da die Absorptionspeaks in hohem Maße überlappen, dennoch lassen sich einige strukturelle Merkmale ableiten [222]. Da der Spargel im Zuge der Probenvorbereitung lyophilisiert wurde, konnte der hohe Wasseranteil von mehr als 90 % im Spargel effektiv entfernt werden. Auf diese Weise wurde die stark ausgeprägte Absorption des Wassers vermieden, sodass die markanten Absorptionsbanden bei  $6900 \text{ cm}^{-1}$  und  $5150 \text{ cm}^{-1}$  in den gemessenen Spektren nicht eindeutig in Erscheinung treten [351]. Das gefrierge-

trocknete Spargelpulver enthält statt Wasser als Hauptkomponenten Kohlenhydrate, Proteine und Aminosäuren. Fette sind dagegen auch in der Trockenmasse nur ein Minorbestandteil [40]. Diese Substanzklassen können auch im Spektrum wiedererkannt werden. Der breite Absorptionspeak bei  $8333\text{ cm}^{-1}$  wurde einer CH-Streckschwingung (2. Oberschwingung) zugeordnet, wobei diese von jedem der erwähnten Inhaltsstoffe ausgehen kann. Eine breite Absorption bei  $6666\text{ cm}^{-1}$  spricht für eine NH-Streckschwingung (1. Oberschwingung), die auf Proteine und Aminosäuren zurückzuführen ist. Zusätzlich konnten NH-Kombinationsschwingungen auch bei einer Wellenzahl von  $4687\text{ cm}^{-1}$  beobachtet werden. Eine weitere CH-Streckschwingung (1. Oberschwingung) wurde aus der Absorption bei  $5760\text{ cm}^{-1}$  abgeleitet. Die beiden kleineren Peaks bei  $5020\text{ cm}^{-1}$  und  $5120\text{ cm}^{-1}$  wurden einer Absorption von OH- und C=O-Gruppen zugeordnet, wie sie bei Proteinen und Kohlenhydraten beobachtet werden [222, 352, 353]. Nach Anwendung der 1. Ableitung treten in Abbildung 42 b einige Peaks in den Spektren durch eine schmalere Peakform deutlich prägnanter in Erscheinung und insbesondere Intensitätsunterschiede bei den einzelnen Ländern werden deutlich.

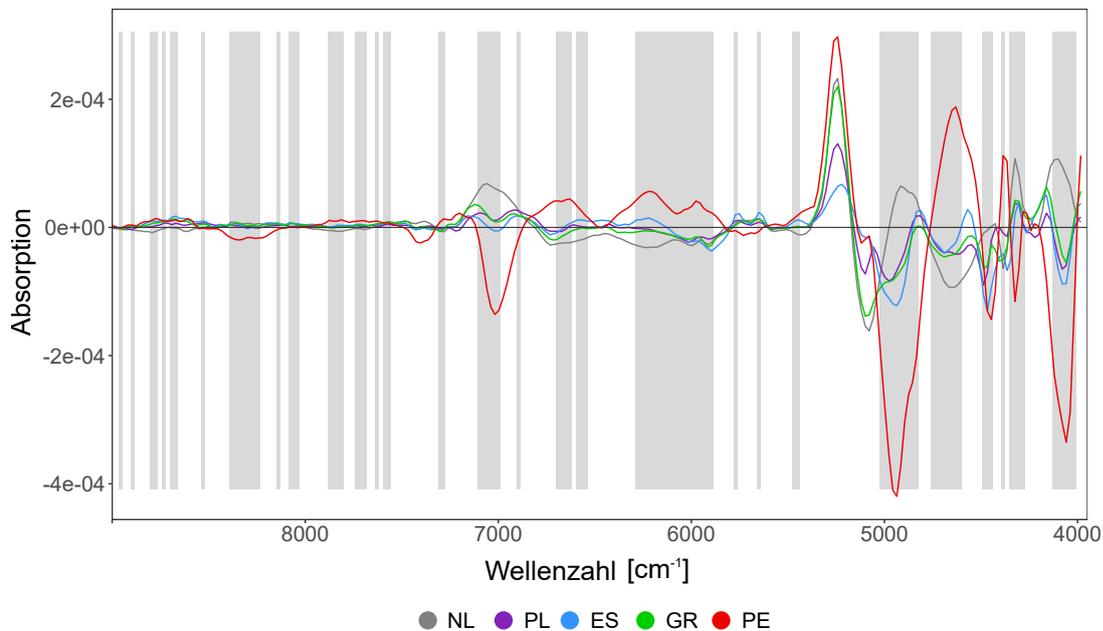
Insgesamt wurden einige Strukturmerkmale des Probenmaterials im Spektrum identifiziert, aus dem Vergleich der Mittelwertspektren konnten jedoch noch keine herausstechenden Unterschiede zwischen dem Spargel der verschiedenen Herkunftsgebiete abgeleitet werden. Daher wurde im nächsten Schritt ein Differenzspektrum erstellt, bei dem vom Mittelwertspektrum jeden Landes das mittlere Spektrum der deutschen Proben abgezogen wurde, was in Abbildung 43 gezeigt ist. Somit beleuchtet die Darstellung der Differenzspektren die Unterschiede aller Herkunftsgebiete zum deutschen Probenursprung.



**Abbildung 43:** Differenzspektrum der MSC-korrigierten Spektren nach Abzug des Spektrums der deutschen Proben vom mittleren NIR-Spektrum des jeweiligen Landes.

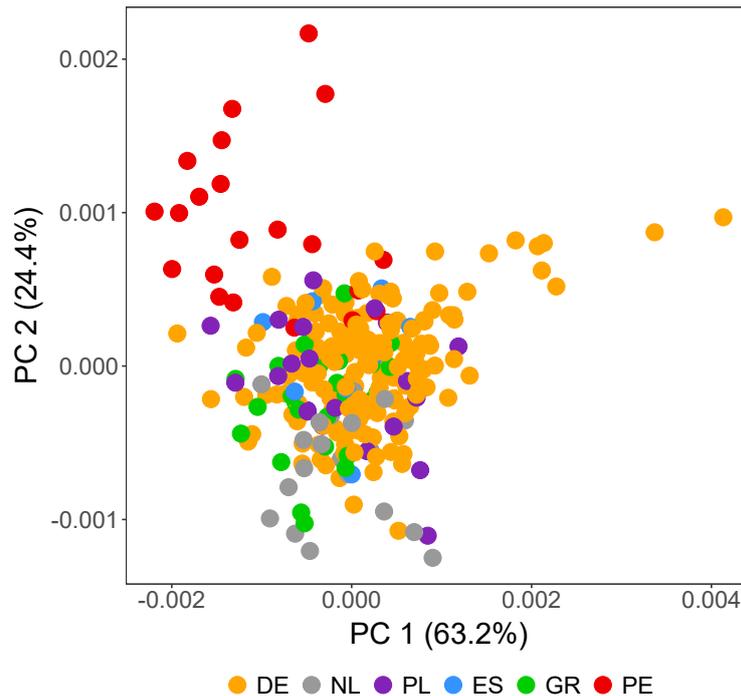
Durch die Bildung der Differenz können nun auch kleinere Unterschiede wahrgenommen werden. Die deutlichste Abweichung zeigt sich, wie erwartet, bei den peruanischen Proben, bei denen über weite Teile des Spektrums ein gegensätzlicher Verlauf zu den Spektren der anderen Länder erkennbar ist. Besonders deutlich kommt dies beim Vergleich der Differenzspektren von Peru und den Niederlanden zum Ausdruck, die oberhalb einer Wellenzahl von etwa  $5500\text{ cm}^{-1}$  und unterhalb von  $5000\text{ cm}^{-1}$  beinahe spiegelgleich verlaufen. Daneben weisen auch alle anderen Länder erkennbare Abweichungen auf, die sich besonders auf den Bereich zwischen  $5000\text{ cm}^{-1}$  und  $7000\text{ cm}^{-1}$  konzentrieren.

Um das Erkennen relevanter Abschnitte für die Herkunftsunterscheidung im Spektrum nicht nur optisch, sondern objektiv an die Fragestellung angepasst, zu ermöglichen, wurde eine *Feature Selection* mithilfe des Boruta-Algorithmus durchgeführt. Die dabei ermittelten Wellenzahlen sind im Differenzspektrum in Abbildung 44 dargestellt. Insgesamt führte der Algorithmus zu einer Anzahl von 122 Wellenzahlen, die für die Unterscheidung der Herkunftsländer als relevant betrachtet werden. In Abbildung 44 fällt auf, dass sich diese oftmals in enger Nachbarschaft zu anderen relevanten Wellenzahlen befinden, insbesondere die Bereiche um  $6000\text{ cm}^{-1}$  und  $4000 - 5000\text{ cm}^{-1}$ , sodass breitere Abschnitte des Spektrums eine zentrale Rolle bei der Herkunftsunterscheidung spielen. Des Weiteren geht die Position der relevanten Wellenzahlen mit den beobachteten Unterschieden der einzelnen Länder im Differenzspektrum einher, da jedem relevanten Bereich ein Peak aus dem Spektrum zugeordnet werden kann, bei dem sich mindestens eines der Herkunftsländer deutlich von den anderen abhebt.



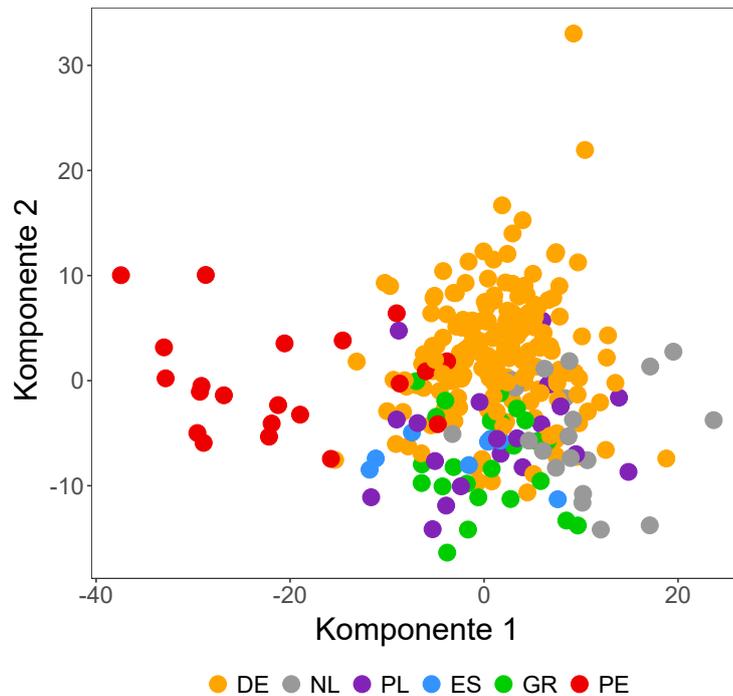
**Abbildung 44:** Differenzspektrum nach MSC-Korrektur und 1. Ableitung als Mittelwertspektrum eines Landes nach Abzug des deutschen Spektrums; relevante Wellenzahlen (*selected Features* mittels Boruta) sind in grau hervorgehoben.

Die Möglichkeit zur Differenzierung der Proben unterschiedlicher Herkunft wurde des Weiteren anhand einer PCA aller Spargelproben untersucht. Abbildung 45 zeigt einen *Scores Plot* für die Unterscheidung der Proben nach ihrem Herkunftsland. Die PCA verdeutlicht eine gute Separierung der peruanischen Proben, alle weiteren Länder sind dagegen schwer voneinander abzugrenzen. Die Proben aus Griechenland, Spanien, den Niederlanden und Polen überlagern sowohl untereinander als auch mit einem großen Teil der deutschen Proben.



**Abbildung 45:** *Scores Plot* der PCA für die Herkunftsunterscheidung basierend auf NIR-Daten.

Da die PCA keine ausreichende Länderunterscheidung der Spargelproben ermöglichte, wurde des Weiteren eine PLS hinzugezogen. Im Gegensatz zur PCA kann bei der PLS als Methode des überwachten Lernens durch die Vorgabe der Klassenzuordnung gezielt eine Auftrennung erreicht werden. In Abbildung 46 ist die PLS für diese Fragestellung gezeigt. Diese lässt wiederum eine gute Abtrennung der Proben peruanischer Herkunft erkennen. Auch hier sind insgesamt einige Überlappungen zu beobachten, im Vergleich zur PCA liegt jedoch ein viel größerer Teil der deutschen Proben isoliert vor. Die griechischen Proben liegen im unteren mittleren Part der Abbildung, die niederländischen sind eher am unteren rechten Rand angeordnet. Dennoch erscheint eine Auftrennung aller einbezogenen Herkunftsländer auch auf der Grundlage der dargestellten PLS eher schwierig.



**Abbildung 46:** PLS für die Herkunftsunterscheidung basierend auf NIR-Daten.

#### 4.4.4 Klassifikation zur Unterscheidung des Herkunftslandes

Auf der Grundlage der FT-NIR-Daten der Spargelproben wurden verschiedene Modelle für die Herkunftsvorhersage erstellt. Es wurden zwei Methoden des maschinellen Lernens, *Support Vector Machine* (SVM) und *Random Forest* (RF), auf ihre Eignung für die genannte Fragestellung untersucht, wobei im Vorfeld verschiedene SVM-Kerne getestet wurden. Da die zuvor gezeigten PCA- und PLS-Plots stellenweise eine größere Probenüberlappung für mehrere Länder aufwiesen, wurde angenommen, dass lineare Trennmodelle für diesen Zweck nicht ausreichend sein werden. Um dies zu prüfen, wurde die SVM mit linearem bzw. RBF-Kern eingesetzt, sodass lineare und nicht-lineare Klassifikatoren gegenübergestellt werden konnten. Für alle genannten Methoden wurde eine Parameteroptimierung und eine Validierung des Modells mithilfe der *nested cross-validation* vorgenommen (vgl. Abschnitt 6.4.2).

Zu Beginn wurde ein statistisches Modell für die Vorhersage des Herkunftslandes einer Spargelprobe separat mit jeder der drei genannten Klassifizierungsmethoden erstellt. Aus den sechs verschiedenen Ursprungsländern ergibt sich somit eine Probeneinteilung in sechs Klassen. Die Klassifizierungsergebnisse sind beispielhaft in Tabelle 11 anhand des SVM-Modells mit linearem Kern dargestellt. Mit diesem Modell wurde eine Klassifizierungsgenauigkeit von 86.9% erreicht. Mit einer besonders hohen Genauigkeit konnten dabei die deutschen und peruanischen Proben eingeordnet werden, mit einer richtigen Zuordnung

**Tabelle 11:** Konfusionsmatrix zur Klassifizierung des Herkunftslandes mittels linearer SVM basierend auf NIR-Daten (Gesamtgenauigkeit: 86.9%).

	<b>DE</b>	<b>NL</b>	<b>PL</b>	<b>ES</b>	<b>GR</b>	<b>PE</b>
<b>DE</b>	<b>176</b>	1	2	0	1	0
<b>NL</b>	6	<b>14</b>	0	0	0	0
<b>PL</b>	10	0	<b>9</b>	0	2	0
<b>ES</b>	3	0	0	<b>0</b>	5	0
<b>GR</b>	2	0	1	1	<b>21</b>	0
<b>PE</b>	2	0	0	0	0	<b>19</b>

von 176 der insgesamt 180 deutschen und 19 der 21 peruanischen Proben (98% und 90% Genauigkeit). Gute Ergebnisse wurden ebenfalls für die Proben aus Griechenland und den Niederlanden erzielt, die in jeweils 84% und 70% der Fälle richtig klassifiziert wurden. Dagegen traten vermehrt Fehler bei der Einordnung der polnischen Proben auf, von denen weniger als die Hälfte richtig zugeordnet wurden. Stattdessen erfolgte eine falsch-positive Einordnung zumeist in die Gruppe deutscher Herkunft. Die Proben aus dem Herkunftsland Spanien konnten in diesem Modell nicht ausreichend gut beschrieben werden, da keine einzige der spanischen Proben richtig klassifiziert wurde. Diese Proben wurden dagegen größtenteils in die Klasse der griechischen Proben eingeordnet. Ähnliches war auch in den beiden weiteren Modellen basierend auf der SVM mit RBF-Kern und dem RF zu beobachten (siehe Tabelle 35 im Anhang, Abschnitt 8.10). Als Ursache dafür ist die sehr geringe Klassengröße mit nur acht spanischen Proben zu nennen. Problematisch sind kleine Klassengrößen auch im Hinblick auf die Kreuzvalidierung. Bei der Unterteilung des Datensatzes in fünf Segmente zur Definition von innerer und äußerer Schleife enthält jedes Segment bei stratifizierter Probeneinteilung nur ein bis zwei spanische Proben. Die Klassifikationsergebnisse legen nahe, dass diese kleine Probenanzahl nicht genügt, um die Herkunft Spaniens zu charakterisieren und die Varianz für dieses Land ausreichend zu beschreiben. Zusätzlich wurde beobachtet, dass sich die einzelnen spanischen Proben verhältnismäßig stark voneinander unterscheiden und somit eine etwas größere Varianz aufweisen. Eine so geringe Anzahl an Proben reicht daher nicht aus, um stellvertretend für die übrigen Proben zu stehen. Des Weiteren war eine große Ähnlichkeit zwischen den griechischen und spanischen Proben festzustellen, da die Proben aus Spanien vorwiegend einer griechischen Herkunft zugeordnet wurden. Ebenso erfolgte eine Klassifizierung von griechischen Proben als Proben spanischer Herkunft, was suggeriert, dass mitunter eine größere Ähnlichkeit der griechischen zu einzelnen spanischen Proben vorliegt als dies bei den spanischen Proben untereinander der Fall ist. Als möglicher Grund können die ähnlichen mediterranen Klimabedingungen in beiden Ländern genannt werden. Daraus kann geschlussfolgert werden, dass Spargelproben

aus diesen beiden Herkunftsländern, auf der Grundlage der vorhandenen Datenlage, anhand von NIR-Daten nicht unterschieden werden können. Aus diesem Grund wurden die Länder im Folgenden in einer Klasse zusammengefasst und neue Vorhersagemodelle für die Unterscheidung der verbleibenden fünf Klassen entwickelt.

Die Ergebnisse der 5-Klassen-Modelle sind in Tabelle 12 dargestellt. Im Vergleich zum zuvor gezeigten Modell konnten hier deutliche Verbesserungen der Klassifizierungsgenauigkeit für die Klasse der griechischen und spanischen Proben erreicht werden. Wo vorher nur 18 der insgesamt 33 Proben aus Griechenland und Spanien richtig vorhergesagt wurden, konnte diese Anzahl in den SVM-Modellen auf 28 bzw. 27 angehoben werden, was einer Verbesserung der Genauigkeit von 55 % auf 85 % bzw. 82 % entspricht (Tabelle 12 a und b). Insgesamt lieferten beide SVM-Modelle gute Ergebnisse für die Vorhersage der Herkunft mit sehr ähnlichen Klassifizierungsergebnissen und Genauigkeiten von 88.4 % und 89.5 %, wobei die SVM mit RBF-Kern leicht besser abschnitt. Die besten Vorhersageergebnisse wurden dabei für die Klasse der deutschen sowie der peruanischen Proben erzielt (jeweils 97 % und 90 % Genauigkeit). Etwas schlechter schnitten die spanischen/griechischen sowie die niederländischen Proben ab, von denen 82 % bzw. 75 % richtig vorhergesagt wurden. Von den untersuchten polnischen Proben konnte dagegen nicht einmal die Hälfte der Proben (48 %) ihrer korrekten Herkunft zugeordnet werden. Dies ist wiederum sehr wahrscheinlich auf eine große Ähnlichkeit zwischen den polnischen und einigen deutschen Proben zurückzuführen. Analog wird dies auch für die niederländischen Proben vermutet, von denen 25 % falsch klassifiziert und ausschließlich in die Klasse der deutschen Proben eingeordnet wurden.

Im Gegensatz zur SVM schnitt das *Random Forest*-Modell mit einer Gesamtgenauigkeit von 77.8 % merklich schlechter ab (Tabelle 12 c). Insgesamt konnte nur die Klasse der deutschen Proben mit einer zur SVM vergleichbaren Genauigkeit richtig vorhergesagt werden. Alle anderen Klassen konnten dagegen nicht zuverlässig beschrieben und von den übrigen Klassen abgetrennt werden. Am gravierendsten wird dies bei den polnischen Proben deutlich, von denen nur 2 der insgesamt 21 Proben richtig klassifiziert wurden. Des Weiteren wird beim RF-Modell ein starker *Bias* der deutschen Proben deutlich, eine Verzerrung des Vorhersagemodells hin zur Klasse mit den meisten Proben. In der Praxis bedeutet dies, dass der Klassifikator zu häufig eine Einordnung in die Klasse der deutschen Proben vornimmt, da aufgrund der großen Probenanzahl eine hohe Wahrscheinlichkeit für ein richtiges Ergebnis gegeben ist. Diese Gefahr besteht grundsätzlich beim Erstellen von Vorhersagemodellen, die auf einem Datensatz mit einem ausgeprägten Klassenungleichgewicht beruhen, wie es in der vorliegenden Arbeit der Fall ist. Der *Bias* hin zu den deutschen Proben beeinträchtigt wie zuvor beschrieben besonders die Genauigkeit der kleineren Klassen. In deutlich abgeschwächter Form ist dieser Effekt auch bei den SVM-Modellen zu beobachten.

**Tabelle 12:** Konfusionsmatrices und Gesamtgenauigkeit zur Klassifizierung des Herkunftslandes im 5-Ländermodell mittels (a) SVM mit linearem Kern, (b) SVM mit RBF-Kern und (c) RF basierend auf NIR-Daten.

(a) SVM, linearer Kern: 88.4 %

	DE	NL	PL	ES/GR	PE
DE	174	2	3	1	0
NL	5	14	1	0	0
PL	12	0	9	0	0
ES/GR	5	0	0	28	0
PE	3	0	0	0	18

(b) SVM, RBF-Kern: 89.5 %

	DE	NL	PL	ES/GR	PE
DE	175	3	1	1	0
NL	5	15	0	0	0
PL	11	0	10	0	0
ES/GR	6	0	0	27	0
PE	2	0	0	0	19

(c) RF: 77.8 %

	DE	NL	PL	ES/GR	PE
DE	174	4	0	2	0
NL	10	10	0	0	0
PL	18	0	2	1	0
ES/GR	19	1	0	13	0
PE	6	0	0	0	15

Dennoch wurden dort für alle Klassen akzeptable Ergebnisse erzielt, mit Ausnahme der Herkunft Polen, bei der die falsch als deutsch eingeordneten Proben jedoch nicht allein durch den *Bias* zu erklären sind. Da die kleineren Klassen insgesamt gut charakterisiert werden können, muss aufgrund des leichten *Bias* somit nicht grundsätzlich die Eignung der SVM-Modelle zur Herkunftsvorhersage infrage gestellt werden.

Aufgrund der gezeigten Ergebnisse kann geschlossen werden, dass der *Random Forest* für die untersuchte Fragestellung nicht geeignet ist, da das Klassenungleichgewicht einen zu großen Einfluss auf das Modell hat und keine zuverlässigen Vorhersagen erlaubt. Aus diesem Grund wird der RF nicht für weitere Klassifikationen herangezogen.

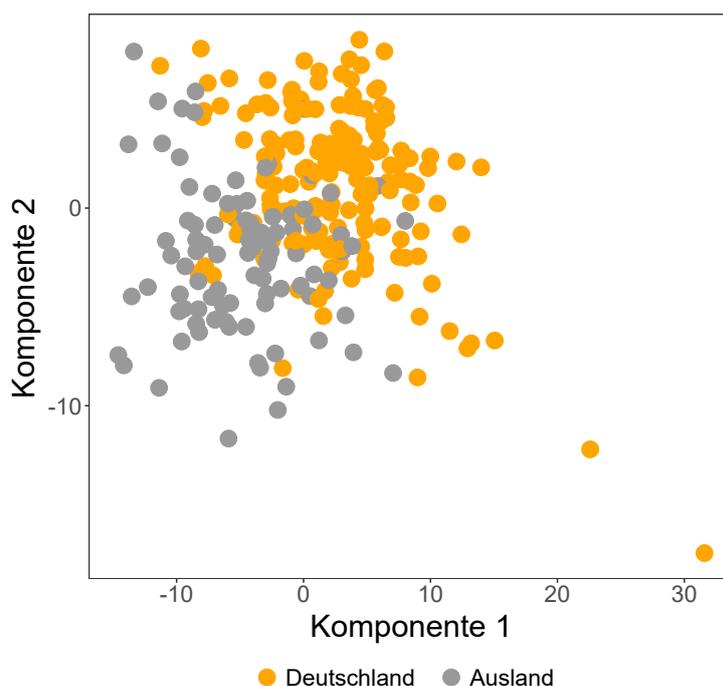
Entgegen der ursprünglichen Erwartung, dass für dieses Trennproblem die nicht linearen Methoden besser abschneiden, hat sich gezeigt, dass die SVM mit linearem Kern fast identische Ergebnisse wie das Modell basierend auf dem RBF-Kern liefert. Daraus wird deutlich, dass die PCA- bzw. PLS-Plots oft nur eine begrenzte Aussagekraft haben, da im zweidimensionalen Raum häufig nur bedingt erkennbar ist, ob sich die untersuchten Klassen unterscheiden.

#### 4.4.5 Klassifikation deutscher und importierter Produkte

Bei der Überprüfung der Herkunft einer Spargelprobe steht in einigen Fällen nicht die genaue Herkunftsregion im Vordergrund, sondern vielmehr die Frage, ob eine fragliche als deutsch deklarierte Probe einer deutschen Herkunft zugeordnet werden kann oder ob möglicherweise eine Verfälschung mit importierten Produkten vorliegt. Daher wurde ein weiteres Vorhersagemodell entwickelt, bei dem zwischen der Herkunft „Deutschland“ und „Ausland“ unterschieden werden sollte. Dafür wurde eine SVM mit linearem bzw. RBF-Kern angewendet. Die von Deutschland abweichenden Länder wurden in der Klasse der ausländischen Proben zusammengefasst, sodass ein 2-Klassen-Trennproblem mit 180 deutschen und 95 ausländischen Proben vorlag. Im Vergleich zu den vorher gezeigten 5- bzw. 6-Klassen-Modellen sind die Klassengrößen hier somit deutlich ausgewogener.

Die Klassentrennung wurde zunächst bildlich dargestellt, indem mithilfe einer PLS zwischen deutscher und ausländischer Herkunft unterschieden wurde, was in Abbildung 47 gezeigt ist. Da die PLS zuvor bereits einen leichten Vorteil bei der Klassentrennung gegenüber der PCA gezeigt hat, wurde auf Letztere verzichtet. Der PLS-Plot zeigt einen klar erkennbaren Trend für die Auftrennung in die beiden Klassen Deutschland und Ausland. In dieser Darstellung sind drei Bereiche erkennbar: jeweils ein Part, in dem sich fast ausschließlich deutsche bzw. ausländische Proben befinden, und ein Bereich dazwischen, in dem Proben aus beiden Klassen überlagern.

Zur Unterscheidung dieser beiden Gruppen wurden SVM-Klassifikationsmodelle erstellt und in Tabelle 13 gegenübergestellt. Im Vergleich zum zuvor gezeigten 5-Klassen-Klassifikationsmodell konnten etwas höhere Genauigkeiten bei der Unterscheidung von deutschen und ausländischen Proben erzielt werden. Dies ist vor allem auf die verbesserte Vorhersage der ausländischen Proben zurückzuführen. Beim linearen Modell wurden Fehlklassifikationen für 17 ausländische Proben beobachtet, bei der Klassifikation des Herkunftslandes lagen hingegen 26 falsche Zuordnungen vor. Etwas besser schnitt die SVM mit RBF-Kern ab, bei der die Anzahl falscher Klassifikationen im Vergleich zum 5-Klassenmodell von 24 auf 13 gesenkt werden konnte. Für beide Modelle wurde dagegen eine Zunahme der deutschen Proben beobachtet, die nicht richtig



**Abbildung 47:** PLS für die Unterscheidung einer deutschen oder ausländischen Herkunft basierend auf NIR-Daten.

klassifiziert wurden. Beim Vergleich der Gesamtgenauigkeiten zeigen beide SVM-Modelle fast identische Ergebnisse, die sich nur in der Nachkommastelle unterscheiden (90.9 % und 90.5 %), mit einem minimalen Vorteil für die SVM mit linearem Kern.

**Tabelle 13:** Konfusionsmatrices und Gesamtgenauigkeit zur Klassifizierung einer deutschen oder ausländischen Herkunft mittels (a) SVM mit linearem Kern und (b) SVM mit RBF-Kern basierend auf NIR-Daten.

(a) SVM, linearer Kern: 90.9 %

	Deutsch	Ausländisch
Deutsch	172	8
Ausländisch	17	78

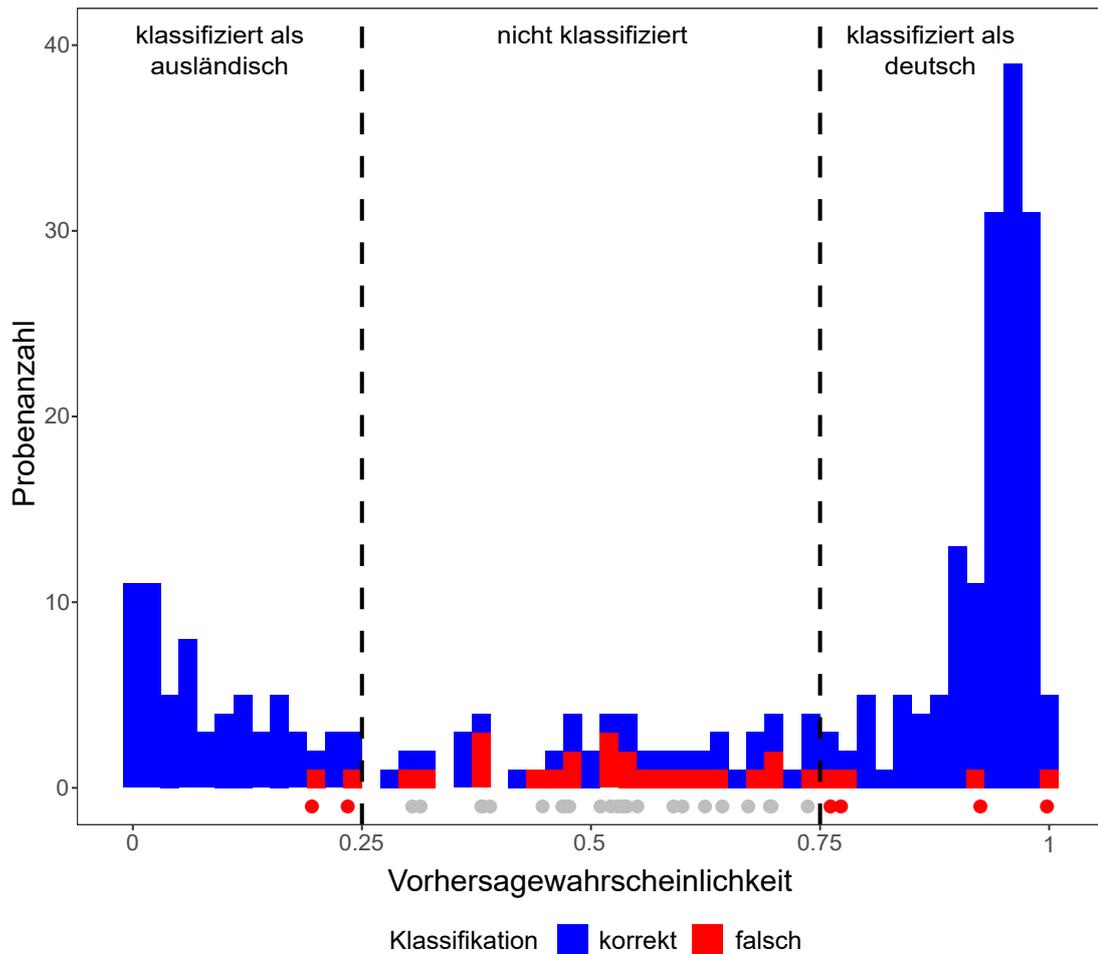
(b) SVM, RBF-Kern: 90.5 %

	Deutsch	Ausländisch
Deutsch	167	13
Ausländisch	13	82

Zur besseren Beurteilung der Leistungsfähigkeit der beiden Klassifikatoren wurden zudem jeweils die Sensitivität sowie die Spezifität berechnet. Die Sensitivität wurde definiert als der Anteil der ausländischen Proben, die richtig als solche erkannt wurden (richtig-positiv), da das Aufdecken von verfälschten Herkunftsangaben im Mittelpunkt dieser Untersuchungen steht. Die Spezifität spiegelt dagegen den Anteil der deutschen Proben wider, die als deutsch klassifiziert wurden (richtig-negativ). Auf diese Weise wurde für das SVM-Modell mit linearem Kern eine Sensitivität von 82.1 % und eine Spezifität von 95.6 % berechnet. Die SVM mit RBF-Kern erreichte eine Sensitivität von 86.3 % sowie eine Spezifität von 92.8 %. Bei beiden Modellen ist somit die Spezifität höher als die Sensitivität, wobei die lineare SVM die höchste Spezifität aufweist. Bei der SVM mit RBF-Kern ist diese etwas niedriger, dafür liegt die Sensitivität im Vergleich zum linearen Modell höher und beide Werte sind insgesamt ausgeglichener. Bei der Frage, welches der Modelle für die Herkunftsbestimmung besser geeignet ist, wird auch hier angenommen, dass bei der praktischen Anwendung eine nicht erkannte ausländische Probe eher in Kauf genommen wird, als dass eine deutsche Probe als ausländisch eingeordnet wird, woraus sich unbegründete Anschuldigungen ableiten könnten. Daraus folgt, dass der Spezifität eine größere Bedeutung beigemessen wird, wonach das lineare SVM-Modell vorzuziehen wäre. Für die routinemäßige Anwendung könnte des Weiteren die Tatsache ins Gewicht fallen, dass der lineare Klassifikator das einfachere Modell in Bezug auf die Berechnung darstellt. Andererseits ist der Anteil der falschen Klassifikationen beim SVM-Modell mit RBF-Kern deutlich ausgeglichener, sodass der leichte *Bias* aufgrund der größeren Zahl an deutschen Proben hier am geringsten ausfällt. In der Praxis muss daher im Einzelfall je nach Anwendungsbereich entschieden werden, welchen Faktoren die höchste Priorität zugeordnet wird.

#### 4.4.6 Eingrenzung des Konfidenzbereiches der Klassifikation

Bei der zuvor gezeigten Datenvisualisierung sowie den Klassifikationsmodellen für die Vorhersage des Herkunftslandes bzw. einer deutschen oder ausländischen Herkunft zeigten sich an verschiedenen Stellen Einbußen bei der Genauigkeit aufgrund der Überlagerung von Proben aus verschiedenen Klassen und daraus resultierenden falschen Klasseneinordnungen. Um die Klassifikationsgenauigkeit zu erhöhen und präzisere Aussagen bezüglich der Probenherkunft ableiten zu können, wurde eine Untersuchung der Vertrauensbereiche vorgenommen. Über die Abschätzung der Konfidenz sollten auf diese Weise jene Bereiche identifiziert werden, in denen Proben mit sehr hoher Präzision richtig zugeordnet werden können. Proben in Überlappungsbereichen, bei denen die Herkunft nicht zuverlässig vorhersagbar ist, wurden dagegen aus der Klassifikation ausgeschlossen. Ziel dabei war es, eine hohe Genauigkeit der Klassifikation zu erreichen,



**Abbildung 48:** Vorgehensweise bei der Abschätzung des Vertrauensbereiches: Verteilung der Klassifikationsentscheidungen anhand der Vorhersagewahrscheinlichkeiten der Klasse der deutschen Proben dargestellt als Histogramm für richtige (blau) und falsche (rot) Klassifikationen; die Mehrzahl der Fehlklassifikationen wird eliminiert (graue Punkte), indem Proben mit geringer Konfidenz ausgeschlossen werden (Intervall innerhalb der gestrichelten Linien).

aber gleichzeitig eine möglichst große Zahl an Proben in das Modell einzubeziehen. Das Vorgehen erfolgte analog zur Konfidenzabschätzung auf der Grundlage der Multi-elementdaten in Abschnitt 4.2.5. Um die Konfidenz der Klassifikation abschätzen zu können, wurde für jede einzelne Probe die Wahrscheinlichkeit der Vorhersage separat für alle Klassen berechnet. Die Beurteilung wurde auch hier anhand der Klasse der deutschen Proben vorgenommen. Die Intervallgrenzen wurden symmetrisch um den Mittelpunkt bei einer Wahrscheinlichkeit von 0.5 nach oben und unten gesetzt, da sowohl hohe als auch niedrige Klassenwahrscheinlichkeiten gleichermaßen zu einer präzisen Klassifikation beitragen. Es wurden systematisch verschiedene Intervallgrenzen mit einer Wahrscheinlichkeit von 0.4 und 0.6 bis hin zu 0.1 und 0.9 getestet, die Proben innerhalb des Intervalls wurden bei der anschließenden Beurteilung der Klassifikation nicht berücksichtigt. Anhand der übrigen Proben wurde die Klassifikationsgenauigkeit und der Anteil an klassifizierbaren Proben bestimmt. Diese Vorgehensweise ist beispielhaft

in Abbildung 48 für die Klassifikation von deutschen und ausländischen Produkten dargestellt. Innerhalb der gestrichelten Linien befindet sich das Intervall der Proben, die aufgrund einer Vorhersagewahrscheinlichkeit im mittleren Bereich (hier 0.25 - 0.75) besonders anfällig für Fehlklassifikationen sind und daher ausgeschlossen werden. Auf diese Weise können die meisten falschen Klassifikationsergebnisse vermieden werden (graue Punkte), sodass nur noch wenige Proben falsch eingeordnet werden (rote Punkte). Die auf diese Weise erhaltenen Ergebnisse der Konfidenzabschätzung sind in Tabelle 14 für das 5-Klassenmodell der Länderzuordnung sowie in Tabelle 15 für das 2-Klassenmodell der Unterscheidung von deutschem und ausländischem Ursprung dargestellt.

Bei der Vorhersage des Herkunftslandes führte die Konfidenzabschätzung zu einer deutlichen Steigerung der Genauigkeit auf 97.6 % (linear) bzw. 98.8 % (RBF), wobei mit einem klassifizierbarem Anteil von jeweils etwa 60 % auch eine deutliche Abnahme der in das Modell einbezogenen Proben zu verzeichnen ist (Tabelle 14). Innerhalb der betrachteten Konfidenzbereiche ist die mit Abstand größte Zunahme der Genauigkeit

**Tabelle 14:** Klassifikationsergebnisse der Ländervorhersage für die Betrachtung verschiedener Konfidenzbereiche begrenzt durch unterschiedliche Wahrscheinlichkeitswerte  $p$  der Klasse der deutschen Proben im SVM-Modell mit linearem und RBF-Kern basierend auf NIR-Daten.

(a) SVM (linear)

Konfidenzbereich	Klassifikationsgenauigkeit	klassifizierte Proben
$p < 0.4$ oder $p > 0.6$	92.1 %	91.6 %
$p < 0.3$ oder $p > 0.7$	95.7 %	85.1 %
$p < 0.25$ oder $p > 0.75$	96.4 %	81.8 %
$p < 0.2$ oder $p > 0.8$	97.1 %	74.5 %
$p < 0.1$ oder $p > 0.9$	97.6 %	59.6 %

(b) SVM (RBF)

Konfidenzbereich	Klassifikationsgenauigkeit	klassifizierte Proben
$p < 0.4$ oder $p > 0.6$	93.2 %	91.3 %
$p < 0.3$ oder $p > 0.7$	96.9 %	83.3 %
$p < 0.25$ oder $p > 0.75$	97.3 %	80.4 %
$p < 0.2$ oder $p > 0.8$	98.6 %	75.6 %
$p < 0.1$ oder $p > 0.9$	98.8 %	60.0 %

zwischen dem ersten und zweiten Bereich zu verzeichnen, mit Zunahmen von 3.6 % (linear) und 3.7 % (RBF). Dies deutet darauf hin, dass der erste untersuchte Bereich noch nicht ausreichte, um die Mehrzahl der Fehlklassifikationen zu eliminieren. Bei allen weiteren betrachteten Konfidenzbereichen ist jeweils ein geringerer Anstieg der Genauigkeit zu erkennen.

Ähnliche Beobachtungen sind der Konfidenzabschätzung für das 2-Klassenmodell zu entnehmen (Tabelle 15). Die Gesamtgenauigkeit des Modells konnte für beide Kern-Typen bis etwa 97 % gesteigert werden, mit einem Anteil an klassifizierten Proben von 63.3 % (linear) bzw. 70.5 % (RBF). Auch bei diesem Klassifikationsansatz wurde die größte Steigerung der Genauigkeit zwischen dem ersten und zweiten Konfidenzbereich erzielt (+2.7 % und +2.1 %), die folgenden Übergänge führen zu einer geringeren Zunahme. Im Falle des linearen Modells ist beim vierten im Vergleich zum dritten betrachteten Konfidenzbereich sogar eine Abnahme zu verzeichnen. Dies ist darauf zurückzuführen, dass durch die weitere Eingrenzung hauptsächlich richtig klassifizierte

**Tabelle 15:** Klassifikationsergebnisse der Vorhersage einer deutschen oder ausländischen Herkunft für die Betrachtung verschiedener Konfidenzbereiche begrenzt durch Wahrscheinlichkeitswerte  $p$  der Klasse der deutschen Proben im SVM-Modell mit linearem und RBF-Kern basierend auf NIR-Daten.

(a) SVM (linear)

Konfidenzbereich	Klassifikationsgenauigkeit	klassifizierte Proben
$p < 0.4$ oder $p > 0.6$	91.6 %	95.3 %
$p < 0.3$ oder $p > 0.7$	94.3 %	89.5 %
$p < 0.25$ oder $p > 0.75$	95.7 %	84.7 %
$p < 0.2$ oder $p > 0.8$	95.4 %	78.5 %
$p < 0.1$ oder $p > 0.9$	97.1 %	63.3 %

(b) SVM (RBF)

Konfidenzbereich	Klassifikationsgenauigkeit	klassifizierte Proben
$p < 0.4$ oder $p > 0.6$	93.1 %	95.3 %
$p < 0.3$ oder $p > 0.7$	95.2 %	90.5 %
$p < 0.25$ oder $p > 0.75$	96.3 %	87.3 %
$p < 0.2$ oder $p > 0.8$	97.3 %	81.1 %
$p < 0.1$ oder $p > 0.9$	97.4 %	70.5 %

Proben entfernt wurden, sodass der prozentuale Anteil der verbleibenden fehlerhaften Klassifikationen an der berücksichtigten Gesamtprobenzahl ansteigt. Dies verdeutlicht, dass eine noch strengere Eingrenzung des Vertrauensbereiches nicht sinnvoll ist.

Da sich die Parameter Gesamtgenauigkeit und Anteil klassifizierbarer Proben grundsätzlich gegensätzlich zueinander verhalten, ist abzuwägen, welchem Kennwert die höhere Priorität beigemessen wird. Im vorliegenden Fall wurden beide gegeneinander aufgewogen und der Bereich einer Klassenwahrscheinlichkeit der deutschen Proben von  $p < 0.25$  oder  $p > 0.75$  als der geeignetste bewertet. Bei einer weiteren Eingrenzung des Vertrauensbereiches steigt die Genauigkeit nur noch sehr geringfügig an, dagegen nimmt der Anteil der klassifizierten Proben stark ab. Bei Anwendung des genannten Konfidenzbereiches wird eine zufriedenstellende Annäherung an beide Parameter erreicht, da sowohl eine sehr gute Genauigkeit als auch die Berücksichtigung einer großen Mehrheit der Proben gegeben ist, sodass ein aussagekräftiges und sehr präzises Modell vorliegt. Beim Vergleich der 5- bzw. 2-Klassenmodelle fällt auf, dass bei jedem der untersuchten Konfidenzbereiche die Genauigkeit des 5-Klassenmodells höher ist, während beim 2-Klassenmodell ein größerer Anteil an klassifizierten Proben vorliegt. Anhand dieser Parameter ist somit kein klarer Vorteil für eines der Modelle festzumachen. In Anbetracht der Tatsache, dass das 5-Klassenmodell mit dem Herkunftsland eine detailliertere Vorhersage ermöglicht und das 2-Klassenmodell bei gleichzeitig geringerem Informationsgehalt keine signifikant besseren Vorhersageergebnisse liefert, ist das Modell zur Klassifikation des Herkunftslandes vorzuziehen.

Die Anwendung des Vertrauensbereiches von  $p < 0.25$  oder  $p > 0.75$  auf die Klassifikation des Herkunftslandes ist im Detail anhand der Konfusionsmatrices in Tabelle 16 dargestellt. Durch die ausschließliche Berücksichtigung der Proben mit hoher Konfidenz konnte die Zahl der Fehlklassifikationen deutlich reduziert werden. Die Proben der Klassen Spanien/Griechenland und Peru wurden zu 100 % richtig eingeordnet, während bei den deutschen und niederländischen nur jeweils eine bzw. zwei Proben falsch klassifiziert wurden. Dagegen wurde auch hier bei 25 - 50 % der polnischen Proben die Herkunft nicht richtig vorhergesagt. Durch die Eingrenzung des Vertrauensbereiches ist somit keine ausreichend gute Vorhersage der polnischen Proben möglich, was sehr wahrscheinlich darauf zurückzuführen ist, dass mithilfe der zugrunde liegenden Datenlage die polnische Herkunft nicht vollständig charakterisiert werden kann.

Zur Verbesserung der Vorhersagegenauigkeit für die polnischen Proben wäre daher eine entsprechende Erweiterung des Datensatzes notwendig. Zudem weisen wenige polnische Proben eine hohe Vorhersagewahrscheinlichkeit für die Klasse der deutschen Proben auf, was auch aus Abbildung 48 ersichtlich wird, sodass eine Abgrenzung dieser Proben prinzipiell sehr schwierig ist und auch durch die Konfidenzabschätzung nicht zufriedenstellend erreicht wurde.

**Tabelle 16:** Konfusionsmatrices und Gesamtgenauigkeit zur Klassifikation des Herkunftslandes im 5-Ländermodell der Proben mit hoher Konfidenz mittels (a) SVM mit linearem Kern und (b) SVM mit RBF-Kern basierend auf NIR-Daten.

(a) SVM, linearer Kern: 96.4 %

	DE	NL	PL	ES/GR	PE
DE	<b>156</b>	0	0	1	0
NL	1	<b>13</b>	0	0	0
PL	4	1	<b>6</b>	1	0
ES/GR	0	0	0	<b>25</b>	0
PE	0	0	0	0	<b>17</b>

(b) SVM, RBF-Kern: 97.3 %

	DE	NL	PL	ES/GR	PE
DE	<b>151</b>	1	1	0	0
NL	1	<b>13</b>	0	0	0
PL	3	0	<b>9</b>	0	0
ES/GR	0	0	0	<b>24</b>	0
PE	0	0	0	0	<b>18</b>

Insgesamt schneidet bei dieser Klassifikation die SVM mit RBF-Kern etwas besser ab. In beiden Modellen ist zudem kein *Bias* hin zu deutschen Proben mehr erkennbar, da die falschen Einordnungen sehr gleichmäßig verteilt sind, mit Ausnahme einiger polnischer Proben aus zuvor diskutierten Gründen.

Bei allen gezeigten Klassifikationen zur Bestimmung der Herkunft basierend auf NIR-Daten, sowohl auf der Grundlage aller Proben als auch der Proben mit hoher Konfidenz, waren für die SVM-Modelle mit linearem bzw. RBF-Kern sehr ähnliche Ergebnisse zu verzeichnen. Bei der SVM mit RBF-Kern war eine etwas bessere Leistungsfähigkeit in Bezug auf die Klassifikationsgenauigkeit und den *Bias* hin zur deutschen Herkunft zu beobachten. Insgesamt ist jedoch kein klarer Vorzug für einen der beiden Kerne abzuleiten.

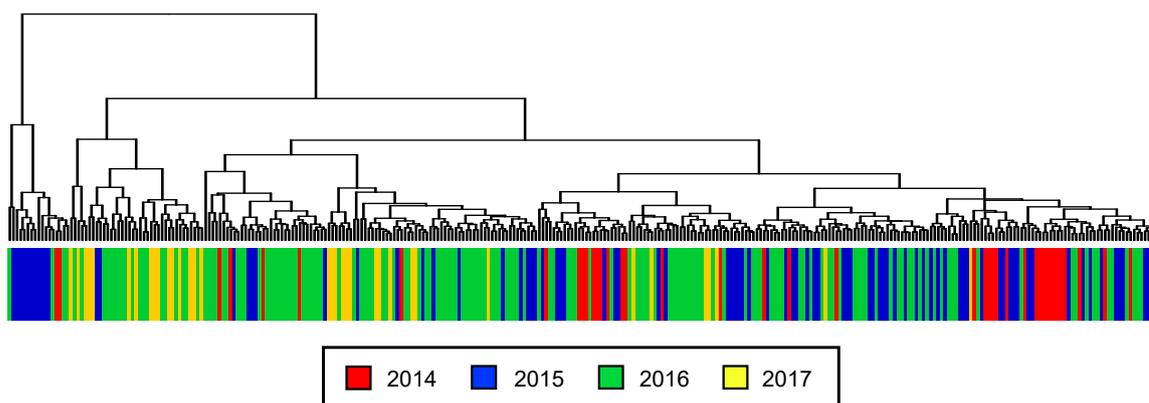
Neben der Konfidenzabschätzung wurden weitere Ansätze zur Verbesserung der Klassifikationsergebnisse in Erwägung gezogen. Zum einen wurde eine *Feature Selection* mithilfe des Boruta-Algorithmus durchgeführt und anstelle aller Wellenzahlen ein Klassifikationsmodell nur basierend auf den ausgewählten *Features* berechnet. Der Vergleich der SVM-Modelle zeigte jedoch keine deutliche Verbesserung der

Klassifikationsgüte, die Genauigkeit der Klassifikation stieg nur bei den Modellen mit RBF-Kern leicht an, bei der linearen SVM sank sie sogar leicht ab. Des Weiteren kann eine Wiederholung der Erstellung und Validierung des Modells genutzt werden, um das Auftreten von zufälligen Effekten zu vermeiden, beispielsweise aufgrund der im Einzelfall gewählten Teilung des Datensatzes während der *nested cross-validation* [354]. Daher wurde eine 10-fache sowie eine 20-fache Wiederholung der Modellerstellung durchgeführt. Dabei waren sehr geringe Standardabweichungen für die Klassifikationsgenauigkeiten von 0.5 % bis 1.2 % zu beobachten, was die hohe Stabilität der Modelle unterstreicht [291]. Sowohl *Feature Selection* als auch eine mehrfache Wiederholung der Modellerstellung haben bei der gewählten Methodik und dem zugrunde liegenden Datensatz somit keine relevanten Auswirkungen auf das Klassifikationsergebnis.

#### 4.4.7 Untersuchung von Einflussfaktoren auf die Herkunftsbestimmung

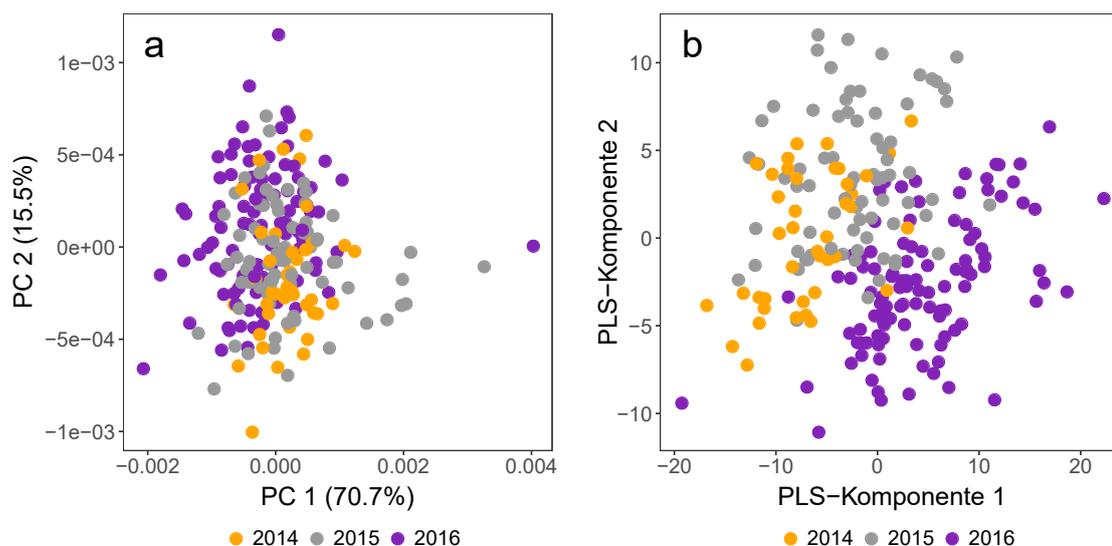
Bei der Methodenentwicklung zur Bestimmung der geographischen Herkunft ist es aus statistischer Sicht wünschenswert, die weiteren Parameter außer der Herkunft bei allen Proben konstant zu halten, um präzise allein die Unterschiede des Faktors Herkunft untersuchen zu können. In der Praxis ist es jedoch mitunter schwierig, alle anderen Faktoren konstant zu halten. Des Weiteren ist dies für einige Merkmale auch nicht erwünscht, da die isolierte Betrachtung beispielsweise jedes Erntejahres oder jeder Sorte für den praktischen Einsatz des statistischen Modells nicht zielführend wäre. Für valide Ergebnisse wäre eine weitaus größere Probenzahl erforderlich und ein jahresspezifisches Modell wäre erst am Ende der Saison einsatzfähig, nach der Ernte und Untersuchung der Proben mit anschließender Datenauswertung, was in der Praxis wenig Nutzen mit sich bringen würde. Daher wurden Proben unterschiedlicher Sorte aus verschiedenen Erntejahren im statistischen Modell für die Herkunftsbestimmung zusammengefasst. Zur Beurteilung der Validität und Robustheit des Modells werden die erwähnten Faktoren im Folgenden gezielt hinsichtlich eines möglichen Einflusses untersucht.

Im Hinblick auf den Einfluss des Erntejahres wurde zunächst ein hierarchisches Clustering aller Proben erstellt, bei dem das jeweilige Erntejahr der Proben markiert ist, gezeigt in Abbildung 49. Hier wurden ebenfalls die Proben des Erntejahres 2014 miteinbezogen. Die Abbildung zeigt die Aufteilung der meisten Proben aus dem Jahr 2014 in zwei Clusterstrukturen im mittleren und rechten Bereich. Die Proben des Erntejahres 2017 sind dagegen mehrheitlich in der linken Hälfte des Clusterings angeordnet. Problematisch ist jedoch, dass bedingt durch die Zusammensetzung des Datensatzes eine Überlagerung durch die Herkunftsinformation als wahrscheinlich zu bewerten ist,



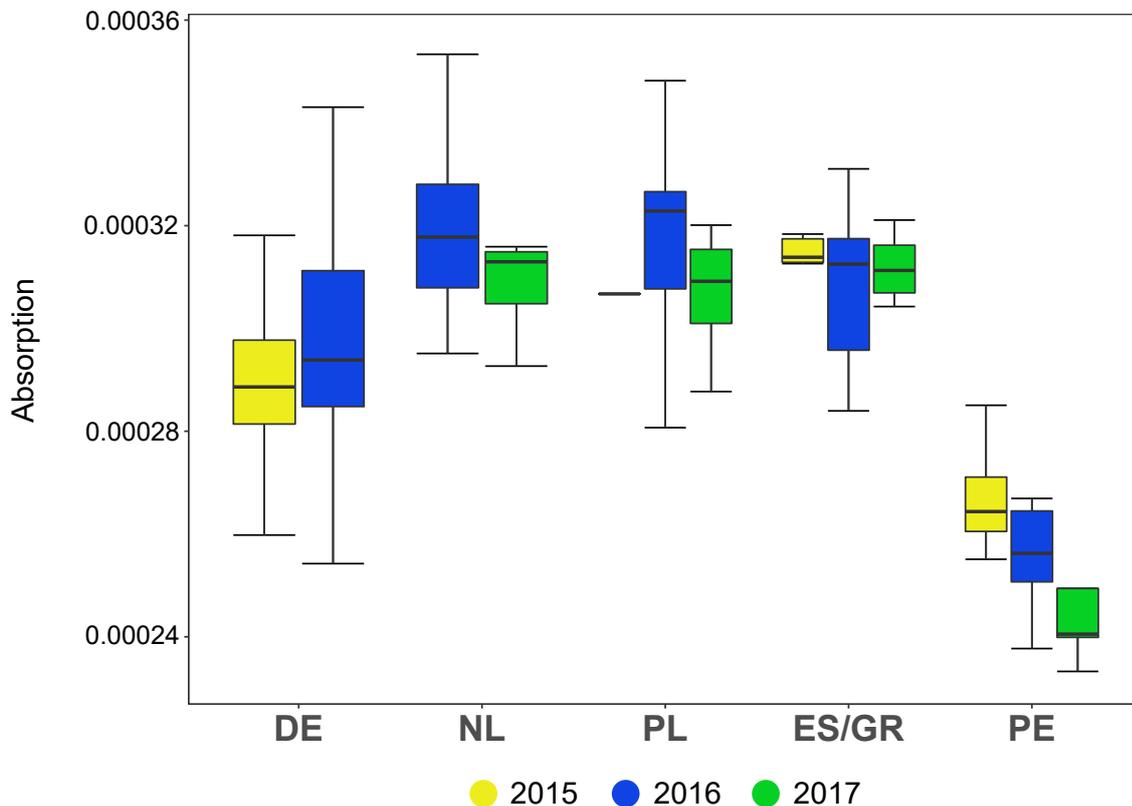
**Abbildung 49:** Hierarchisches Clustering der Spargelproben mit Markierung der Erntejahre basierend auf NIR-Daten.

da für das Jahr 2014 keine ausländischen, für 2017 dagegen keine deutschen Proben vorlagen. Daher erfolgte die Betrachtung der deutschen Proben isoliert mithilfe von PCA- und PLS-Plots, die in Abbildung 50 dargestellt sind. In der PCA ist eine großflächige Überlagerung der Proben aus allen Erntejahren zu verzeichnen, wobei die Proben aus 2016 etwas weiter links, jene aus dem Jahr 2014 tendenziell weiter rechts in der Abbildung angeordnet sind. Eine Gruppierung der Proben eines einzelnen Erntejahres ist jedoch nicht zu erkennen (Abb. 50 a). Dagegen kann im PLS-Plot in Abbildung 50 b eine eindeutige Clusterbildung beobachtet werden, wenn auch mit größeren Überlappungsbereichen. Dies belegt die Vermutung, dass das Erntejahr aufgrund des variierenden Metaboloms einen Einfluss auf die mittels FT-NIR untersuchten Proben hat.



**Abbildung 50:** (a) PCA und (b) PLS der deutschen Spargelproben zur Untersuchung des Erntejahreinflusses basierend auf NIR-Daten.

Zur Beantwortung der Frage, in welcher Weise dieser Jahreseinfluss auch die Klassifikation der Herkunft beeinflusst, wurde eine Untersuchung der jahresspezifischen Verteilung der Intensitäten der einzelnen Wellenzahlen vorgenommen. Dazu wurden Boxplots erstellt, die für jedes Land und jedes Erntejahr jeweils die Verteilung des Absorptionslevels widerspiegeln. Exemplarisch ist dies in Abbildung 51 für die Wellenzahl  $5976.8 \text{ cm}^{-1}$  gezeigt. Dabei sind deutliche Unterschiede bei der Intensität der Proben eines Landes je nach Erntejahr zu beobachten, mit recht unterschiedlicher Größe der Boxen und somit der Spannweite der Intensität. Letztere ist auch von der Probenanzahl abhängig, daher ist für die Herkunft Deutschland die größte Varianz zu verzeichnen. Eine große Spannweite des Absorptionslevels ist ebenfalls bei den polnischen Proben gegeben, was sich nachteilig auf die Abgrenzung von Proben anderer Herkunft auswirkt und konsistent mit den zuvor gezeigten Klassifikationsergebnissen ist.

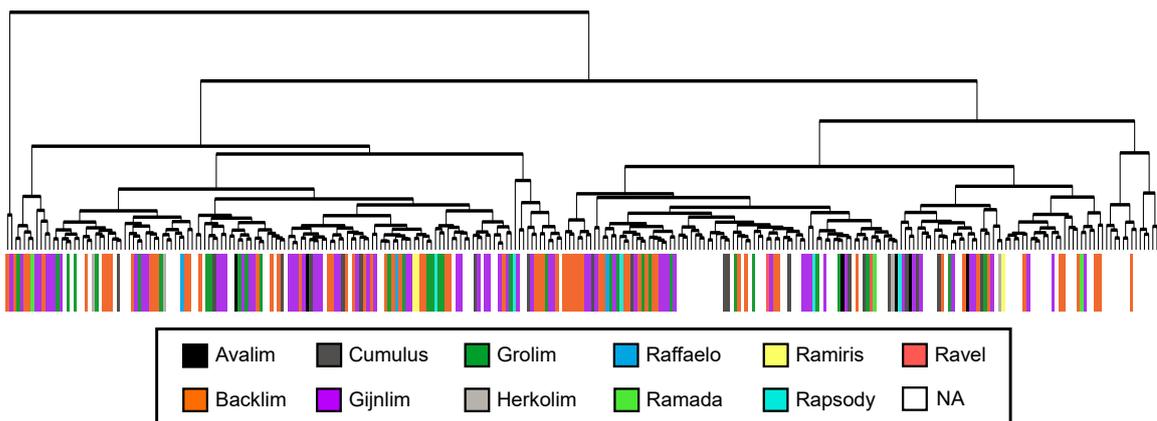


**Abbildung 51:** Boxplot für die Verteilung des Absorptionslevels der Proben aus unterschiedlichen Herkunftsländern und Erntejahren bei der Wellenzahl  $5976.8 \text{ cm}^{-1}$ .

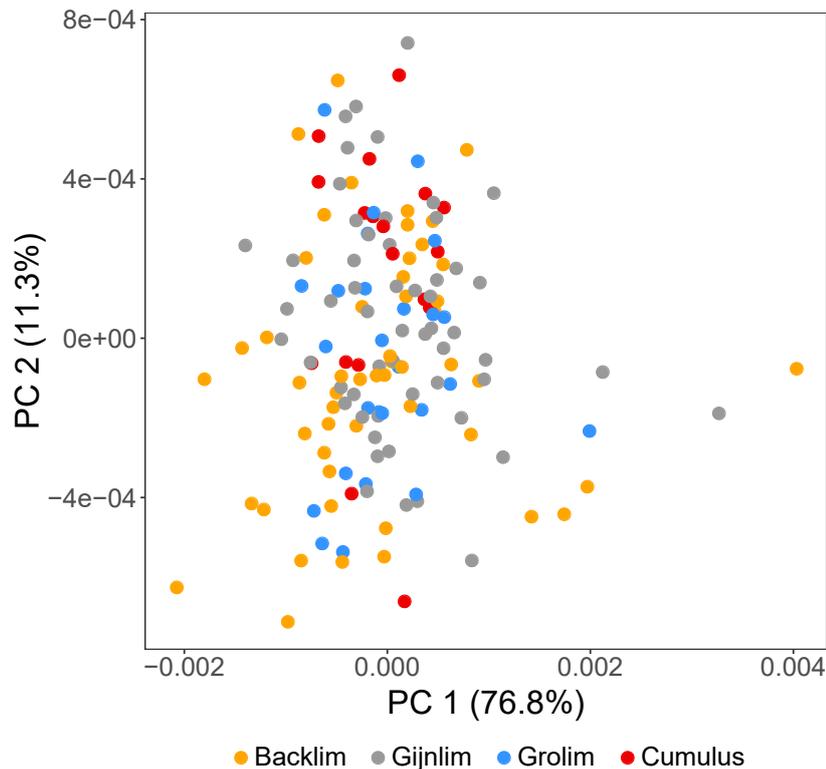
Anhand des in Abbildung 51 dargestellten Boxplots ist eine Abgrenzung der peruanischen Proben nachvollziehbar. Auch diese Proben weisen deutliche Schwankungen des Intensitätslevels für die unterschiedlichen Erntejahre auf, dennoch ist der Abstand

zu den Boxen aller übrigen Proben groß genug, um eine Abtrennung vornehmen zu können. Anhand dieses Beispiels wird deutlich, dass die jährlichen Schwankungen der Intensität die Klassifikation der Herkunft zwar beeinflussen, diese aber dennoch möglich ist, da der Einfluss des geographischen Ursprungs deutlich überwiegt. Eine plausible Erklärung liefern die Wetterbedingungen, die im Allgemeinen zwischen den einzelnen Ländern deutlich stärker variieren als an einem Standort in unterschiedlichen Jahren. Daher ist auch die Kombination mehrerer Erntejahre in einem Klassifikationsmodell als legitim und sinnvoll zu bewerten, sodass trotz eines leichten Jahreseinflusses Herkunftsvorhersagen mit hohen Genauigkeiten erzielt werden können. Dies geht auch mit dem Ziel einher, eine Methode zu entwickeln, in der jahresübergreifend verschiedene Proben einbezogen werden können, um eine jährliche Neuberechnung des Modells zu verhindern und den Aufbau einer Datenbank zu ermöglichen.

Des Weiteren wurde untersucht, ob die Spargelsorte bei der Fragestellung der Herkunftsbestimmung einen relevanten Einfluss hat. Daher wurde ebenfalls zunächst ein hierarchisches Clustering erstellt, welches in Abbildung 52 gezeigt ist. Das Clustering zeigt über die gesamte Breite eine gleichmäßige Verteilung aller Spargelsorten, bei der keine systematische Clusterbildung aufgrund der Sorte erkennbar ist. Zusätzlich wurde eine PCA mit den Proben der vier relevantesten Sorten erstellt, die auch zahlenmäßig am stärksten vertreten waren. Abbildung 53 zeigt wiederum, dass die untersuchten Sorten homogen über den Plot verteilt sind und keine Clusterbildung stattfindet. Daraus geht hervor, dass die Spargelsorte keinen relevanten Einfluss auf die Herkunftsbestimmung mittels FT-NIR hat.

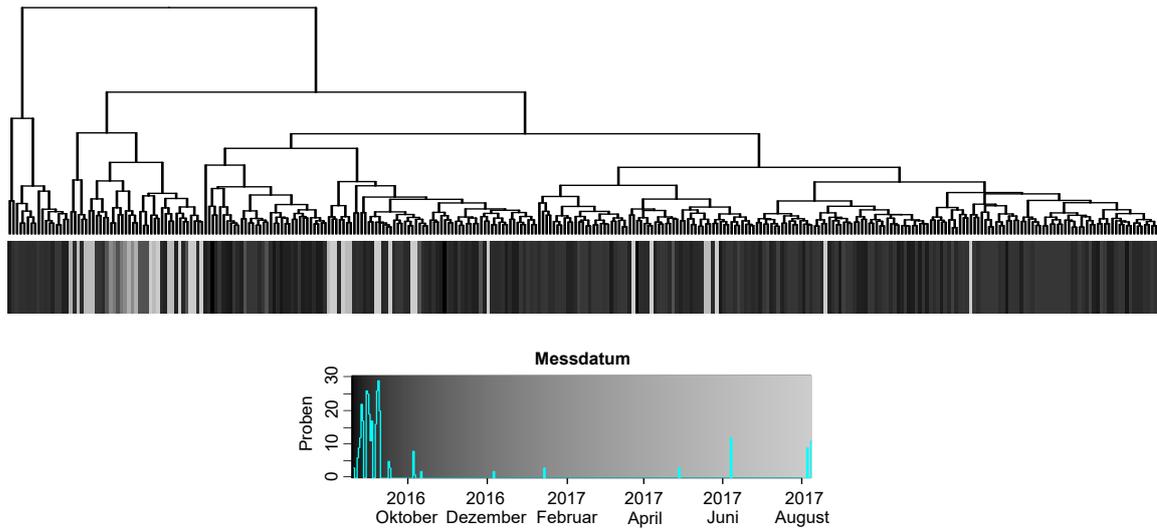


**Abbildung 52:** Hierarchisches Clustering aller Spargelproben mit Markierung der Spargelsorte basierend auf NIR-Daten.



**Abbildung 53:** PCA der deutschen Spargelproben der Sorten Backlim, Gijnlim, Grolim und Cumulus basierend auf NIR-Daten.

Schließlich erfolgte eine Untersuchung des Messdatums ebenfalls anhand eines hierarchischen Clusterings, welches in Abbildung 54 dargestellt ist. Dieses zeigt eine Färbung der meisten Proben in dunkelgrau, da die überwiegende Mehrheit der Proben zu Beginn des Messzeitraums analysiert wurden. Beim Ablauf der Messungen wurde eine randomisierte Abfolge angestrebt, die jedoch nur begrenzt umgesetzt werden konnte. Die Ursache hierfür lag zum einen in der Verfügbarkeit der Proben aufgrund der verschiedenen Erntejahre, sodass dieser Parameter nicht vollständig unter Einbeziehung aller Erntejahre randomisiert werden konnte. Zum anderen wurde dasselbe Material nach Abschluss der NIR-Messungen für den Aufschluss als Vorbereitung der ICP-MS-Messungen verwendet, sodass eine zeitnahe NIR-spektroskopische Untersuchung notwendig war. Dies hat zur Folge, dass die beiden Parameter Erntejahr und Messdatum überlagern und daher nicht getrennt voneinander dargestellt werden können. So steht beispielsweise die deutlich hellere Färbung des zweiten Clusters am linken Rand, die eine Messung zum Ende des Messzeitraums symbolisiert, in deutlichem Zusammenhang mit dem Erntejahr, da dieses Cluster fast ausschließlich Proben der Erntejahre 2016 und 2017 enthält (vgl. Abbildung 49). Aufgrund der Häufung der Anzahl gemessener Proben zu Beginn des Messzeitraums und der somit nicht vollständig randomisierten Messreihenfolge kann der Einfluss des Messdatums daher nicht abschließend beurteilt werden.



**Abbildung 54:** Hierarchisches Clustering aller Spargelproben mit Markierung des Messdatums anhand eines Farbverlaufs basierend auf NIR-Daten; unten: Histogramm mit Anzahl gemessener Proben im gesamten Messzeitraum.

---

## 4.5 Vergleich von ICP-MS- und FT-NIR-basierter Herkunftsbestimmung

Mit der Analyse der Spargelproben mittels FT-NIR und ICP-MS wurden zwei sehr unterschiedliche analytische Ansätze herangezogen. Im Hinblick auf ihre Eignung zur Herkunftsbestimmung von Spargel müssen beide Disziplinen daher differenziert bewertet werden.

Die Messung mittels FT-NIR stellt eine apparativ wenig aufwendige Herangehensweise dar. Das NIR-Spektrometer ist ein vergleichsweise günstiges Gerät, leicht zu bedienen und liefert schnelle Analysenergebnisse mit einer Messzeit von etwa 1 min/Probe. Das Probenmaterial erfordert nach der Lyophilisation keine weitere Aufarbeitung, was zu einer schnellen Datenakquirierung beiträgt. Zudem handelt es sich um eine nicht-destruktive Methode, sodass das Probenmaterial ggf. weiterverwendet werden kann. Hinzu kommt, dass NIR-Geräte und entsprechende Methoden in vielen Laboren zur Standardausrüstung gehören.

Im Vergleich zur FT-NIR handelt es sich bei der ICP-MS-Analyse um eine komplexere Methodik, die mit deutlich höheren Anschaffungs- und Unterhaltskosten verbunden ist. Des Weiteren erfolgt die Analyse in der Regel in flüssiger Form, was einen vorherigen Aufschluss des Probenmaterials erforderlich macht. Bei einer reinen Messzeit von unter 7 min/Probe, wie sie in der vorliegenden Arbeit erzielt wurde, kann man auch hier von einer kurzen Analysenzeit sprechen. Zudem verfügen mittlerweile diverse Labore über Aufschlussstechnik, die einen Aufschluss von mindestens 40 bis über 100 Proben gleichzeitig erlaubt, sodass ein hoher Probendurchsatz ermöglicht wird. Auch die Elementanalytik mittels ICP-MS stellt daher eine wirtschaftlich relevante und effiziente Technik dar, die bereits vielfach in Handels- und amtlichen Laboren eingesetzt wird.

Basierend auf NIR-spektroskopischen Messungen wurde eine Klassifikation des Herkunftslandes unter Berücksichtigung von 5 Klassen mit einer Genauigkeit von 89.5 % erreicht. Bei der Vorhersage einer deutschen oder nicht deutschen Herkunft lag die Genauigkeit mit 90.9 % geringfügig höher. Durch die Anwendung einer Konfidenzabschätzung wurde eine deutlich verbesserte Präzision der Klassifikation erzielt, indem Vorhersagen mit einer geringen Konfidenz ausgeschlossen wurden. Unter Einbeziehung von mindestens 80 % der Proben konnte die Genauigkeit somit auf bis zu 97 % angehoben werden.

Mithilfe eines 7-Klassen-Klassifikationsmodells auf der Grundlage der Elementdaten war eine Unterscheidung des Herkunftslandes mit einer höheren Genauigkeit von 91.2 % möglich. Eine Konfidenzabschätzung analog zum NIR-Modell führte zur Erhöhung der Genauigkeit auf > 95 % für die Klassifikation von mehr als 80 % der Proben. Des Weiteren kann anhand der ICP-MS-Daten eine Zuordnung zu kleineren Herkunftsgebieten

bzw. zum Produktionsstandort vorgenommen werden, wobei in beiden Fällen eine Genauigkeit von etwa 83 % erreicht wurde.

Beide analytischen Ansätze ermöglichen eine Vorhersage des Herkunftslandes mit einer ähnlich hohen Genauigkeit, die mithilfe einer Konfidenzabschätzung noch weiter gesteigert werden kann. Die NIR-Modelle unterliegen jedoch der Einschränkung, dass auf der Grundlage des vorhandenen Datensatzes keine getrennte Vorhersage der Herkunftsländer Spanien und Griechenland möglich war. Im Gegensatz dazu wird die Zuordnung zu beiden Ländern anhand der ICP-MS-Daten mit hoher Genauigkeit ermöglicht, ebenso wie die zusätzliche Klassifikation der Herkunft China. Größere Schwierigkeiten bereitete bei beiden analytischen Herangehensweisen die Abtrennung des polnischen Spargels insbesondere von den deutschen Proben. Dies liefert einen Hinweis auf eine unzureichende Datenlage für die Charakterisierung des Ursprungs Polen mit dem Probenmaterial, welches für diese Arbeit zur Verfügung stand. Der Ansatz der Reduzierung der Ländervorhersage auf die Unterscheidung einer deutschen oder nicht-deutschen Herkunft hat für beide Analysetechniken keine grundlegende Verbesserung der Vorhersagegenauigkeit mit sich gebracht. Aufgrund der präziseren Herkunftseinordnung ist daher die Vorhersage einzelner Ursprungsländer in beiden Fällen vorzuziehen.

Das Ziel eines jahresübergreifenden Modells zur Herkunftsbestimmung von Spargel konnte mit beiden gewählten Ansätzen, FT-NIR und ICP-MS, erreicht werden. Unterschiede der beiden Disziplinen sind jedoch im Hinblick auf den Einfluss des Parameters Erntejahr zu verzeichnen. Während bei den ICP-MS-Messungen kein Jahreseinfluss nachgewiesen wurde, war dieser bei den NIR-Daten deutlich zu erkennen. Da der Einfluss der Herkunft auf das zugrundeliegende Metabolom deutlich stärker ausgeprägt ist, ist eine jahresübergreifende Nutzung der zur Verfügung stehenden Daten dennoch als legitim zu bewerten. Etwaige Einbußen bei der Präzision der Vorhersagen im Vergleich zur isolierten Betrachtung jedes Erntejahres werden dabei im Hinblick auf den viel größeren praktischen Nutzen eines jahresübergreifenden Klassifikators in Kauf genommen. Mit beiden analytischen Strategien konnte daher, unter Hinzunahme der Konfidenzabschätzung, eine Zuordnung des Herkunftslandes unabhängig vom Erntejahr vorgenommen werden. Das Fehlen eines messbaren Jahreseinflusses bei der Elementanalytik ist als besonderer Vorteil zu unterstreichen, da dies sehr präzise Ergebnisse bei der Herkunftsbestimmung ermöglicht. Im Gegensatz dazu sind bei den verschiedenen Metabolom-basierten Methoden durch die Jahresschwankungen in unterschiedlichem Maße immer Einbußen bei Präzision und Genauigkeit der Vorhersage zu verzeichnen bzw. die jährliche Berechnung eines neuen Modells erforderlich.

Eine äußerst präzise Bestimmung des geographischen Ursprungs einer Spargelprobe anhand ihres Elementprofils wurde mit der Klassifikation des Produktionsstandortes erzielt. Eine solche Zuordnung zu kleineren Regionen war dagegen auf der Grundlage der

NIR-Daten nicht möglich. Auch bei der geforderten Robustheit des Vorhersagemodells schneidet die ICP-MS-basierte Methode deutlich besser ab. Dies wurde insbesondere anhand des hohen Maßes an Stabilität im Elementprofil belegt, welches an einem Standort über Jahre hinweg reproduzierbar im Spargel vorgefunden wurde. Aus der Perspektive der Wirtschaftlichkeit können beide Methoden mit einer kurzen Analysenzeit überzeugen, mit einem Vorteil für die NIR-Messungen, insbesondere wegen der einfacheren Probenvorbereitung. Dem gegenüber steht jedoch eine verhältnismäßig lange Dauer der Probenaufarbeitung bis zum Vorliegen des Probenmaterials in lyophilisierter Form, sodass diese Phase des Analysenprozesses bislang als zu aufwändig einzustufen ist. Ein weiteres Ziel dieser Arbeit war die Entwicklung einer Analysenmethode, die zum Schutz der Produkte mit einer geographisch geschützten Angabe (ggA) eingesetzt werden kann. Mithilfe des ICP-MS-basierten Analysenansatzes wird die Zuordnung einer Spargelprobe zum jeweiligen Produzenten mit einer hohen Genauigkeit ermöglicht, sodass auf dieser Grundlage leicht beurteilt werden kann, ob der Spargel aus der angegebenen Region stammt und eine ggA gerechtfertigterweise in der Deklaration aufgeführt wird. Auf diese Weise konnte der ggA-Schutz erfolgreich adressiert werden.

In der vorliegenden Arbeit wurden verschiedene multivariate Methoden zur Datenauswertung herangezogen. Für beide Analysenansätze wurden unter Einsatz von Verfahren des maschinellen Lernens, insbesondere der *Support Vector Machine*, Klassifikationsmodelle mit hoher Genauigkeit für die Herkunftsvorhersage entwickelt. Somit konnte gezeigt werden, dass die Anwendung aktueller statistischer Methoden zu einer effizienten Datenauswertung und zum Erreichen präziser Herkunftsvorhersagen entscheidend beiträgt.

Zusammenfassend ist festzuhalten, dass anhand von FT-NIR-Messungen, verknüpft mit einer multivariaten Datenauswertung, eine grobe Zuordnung des Herkunftslandes von weißem Spargel für eine große Mehrzahl der untersuchten Proben mit einer hohen Genauigkeit möglich ist. Diese Herangehensweise besticht besonders durch einen geringen analytischen Aufwand, geringe Kosten und die Möglichkeit zur Weiterverwendung des Probenmaterials und eignet sich daher besonders gut als *Screening*-Ansatz. Darüber hinaus erzielt eine Untersuchung des Elementprofils im Spargel in Kombination mit einem SVM-gestützten Klassifikationsmodell eine sehr genaue Vorhersage des Herkunftslandes unter Einbeziehung einer überwiegenden Mehrheit der Proben. Zudem ist mithilfe dieser Methode eine präzisere regionale Einordnung des geographischen Ursprungs möglich, bis hin zur Bestimmung des Produktionsstandortes. Bei etwas größerem analytischen Aufwand kann man hier von einem extrem leistungsfähigen und robusten Verfahren sprechen, welches zuverlässige und belastbare Aussagen zur geographischen Herkunft des Spargels liefert und daher das Potential zur Generierung gerichtsfester Daten unter Beweis gestellt hat.

## 5 Ausblick

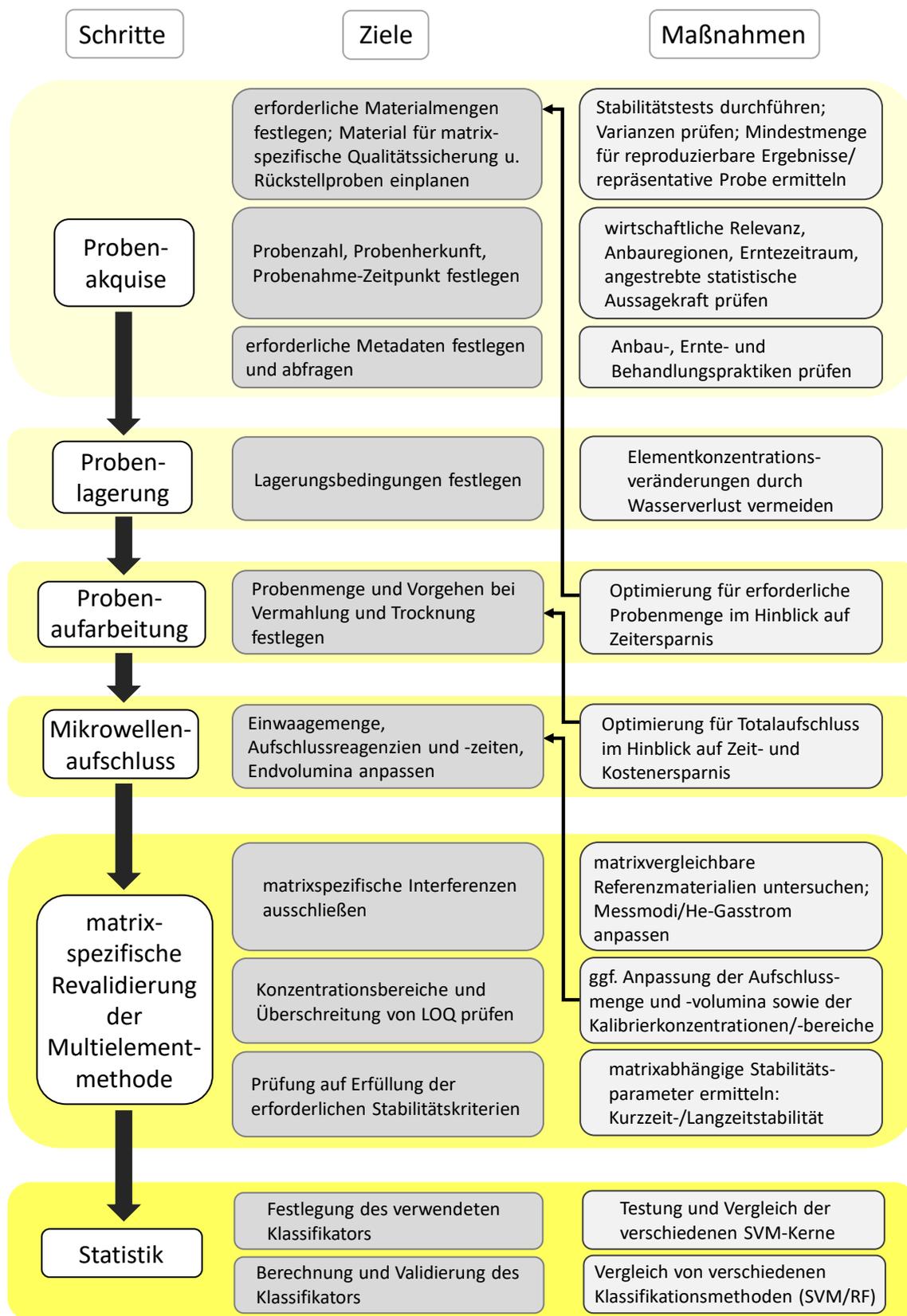
Anhand der zuvor diskutierten Ergebnisse wurde eine hohe Leistungsfähigkeit der entwickelten Methoden zur Herkunftsbestimmung von Spargel gezeigt, die eine perspektivische Anwendung in der routinemäßigen Deklarationsprüfung ermöglicht. Der Einsatz dieser Verfahren ist insbesondere für die amtliche Lebensmittelüberwachung bzw. in privaten Handelslaboren denkbar, in denen mitunter bereits eine Herkunftsüberprüfung des Spargels vorgenommen wird. Da verarbeitete Spargelprodukte selten mit einer Herkunftsangabe versehen sind, spielt deren Kontrolle in der lebensmittelverarbeitenden Industrie dagegen eine untergeordnete Rolle. Für die Lebensmittelüberwachung hat die ICP-MS-basierte Methode großes Potential zur Verbesserung der bislang erzielten Ergebnisse, besonders aufgrund der hohen Präzision mit der Vorhersage des Produktionsortes und der enormen Stabilität des Elementprofils über mehrere Jahre. Dies ermöglicht das Erstellen und die stetige Erweiterung einer Datenbank, da die Klassifikationsmodelle nicht für jedes Erntejahr separat trainiert werden müssen, sodass eine umfassende Datengrundlage erreicht werden kann. Denkbar wäre des Weiteren eine Kombination beider hier gezeigter Verfahren mit der Nutzung der NIR-Methode als *Screening*-Ansatz und der Erweiterung auf die Untersuchung des Elementprofils bei erkennbaren Auffälligkeiten oder Ungenauigkeiten der NIR-Ergebnisse bzw. zu deren Absicherung. Diese Herangehensweise ist aufgrund der Limitationen des NIR-Ansatzes nicht für jede Fragestellung ausreichend, dafür spricht jedoch der geringe Aufwand der NIR-Messungen und dass für beide Methoden dieselbe Probenaufarbeitung zur Gewinnung von lyophilisiertem Material genutzt werden kann. Für das Erstellen gerichtsfester Analysenergebnisse ist aufgrund von höherer Präzision und Genauigkeit der Vorhersage die Analytik mittels ICP-MS jedoch klar vorzuziehen. Vor diesem Hintergrund sollte auch der Ansatz der Konfidenzabschätzung weiter verfolgt werden. In der vorliegenden Arbeit wurde gezeigt, dass dieses Vorgehen zu einer deutlichen Steigerung der Klassifikationsgenauigkeit führt, sodass die resultierenden Herkunftszuordnungen extrem zuverlässig sind. Da dies das übergeordnete Ziel bei der Anwendung in der Routineanalytik darstellt, birgt dieser Ansatz großes Potential für die stetige Methodenverbesserung. Des Weiteren wäre zu prüfen, ob eine Kombination der Datensätze aus der Analytik mittels ICP-MS und FT-NIR, was auch als *data fusion* bezeichnet wird, als Grundlage eines Klassifikationsmodells zu

verbesserten Vorhersageergebnissen führt.

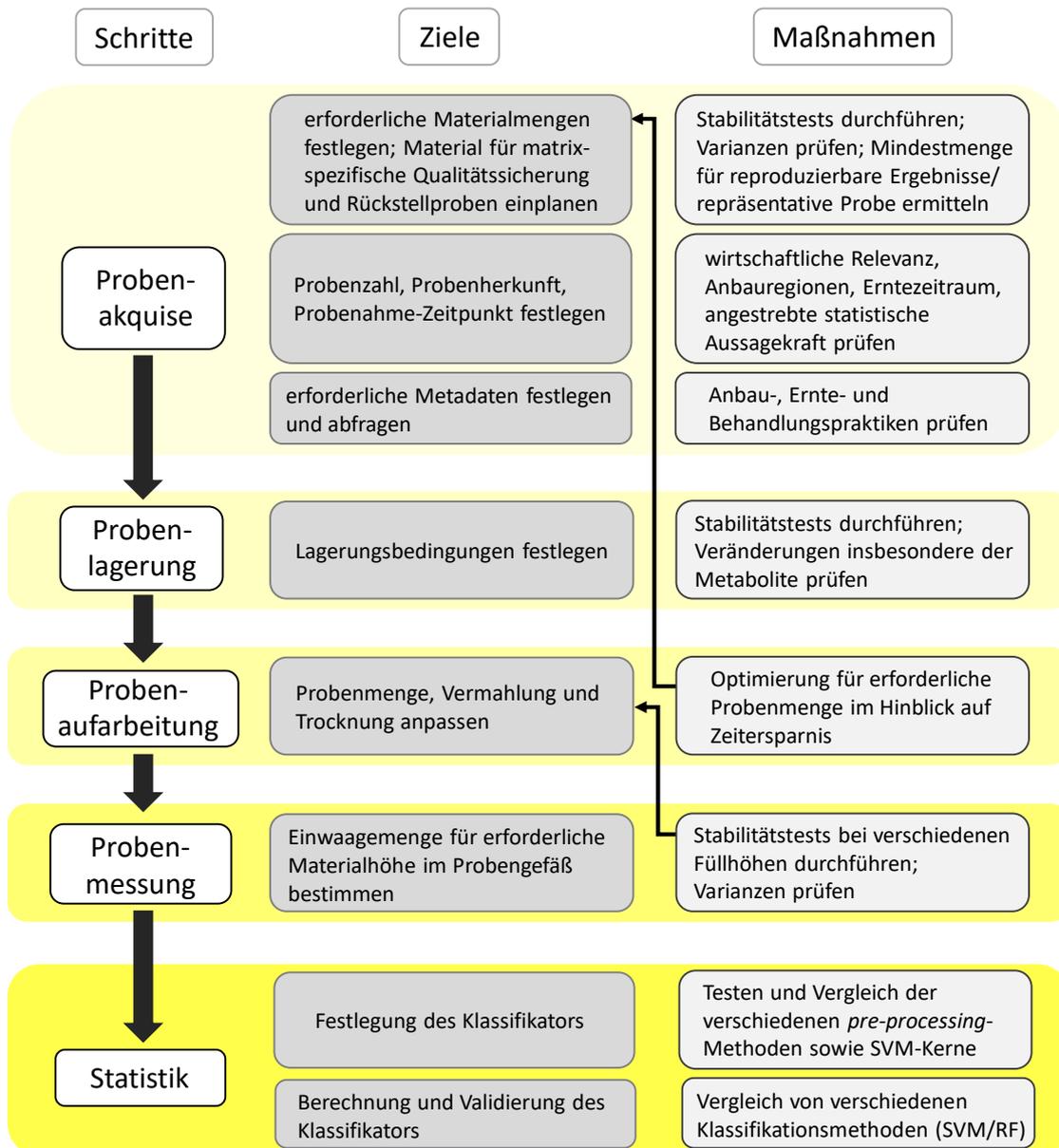
In dieser Arbeit wurde anhand eines umfangreichen Datensatzes exemplarisch eine Möglichkeit zur sehr genauen Rückverfolgung der Herkunft aufgezeigt. Für die Etablierung der Methodik in der Routineanalytik sind jedoch weitere Anpassungen notwendig. Dazu gehört das Erstellen einer Datenbank, welche Proben aus weiteren Erntejahren und eine große Probenanzahl für alle relevanten Herkunftsgebiete enthält. Dabei ist ein Schwerpunkt auf die Akquise von authentischem Probenmaterial am Produktionsstandort zu legen. Dies ist mit einem erhöhten finanziellen Aufwand verbunden, trägt jedoch essentiell zu einer vollständigen und zuverlässigen Charakterisierung eines Herkunftsgebietes bei und sichert somit die Richtigkeit der Klassifikationsmodelle. Im Zuge des Schutzes der ggA sollte zudem auch besonderen Wert gelegt werden auf eine systematische Probenakquise von allen beteiligten Produzenten mit jeweils ausreichender Probenzahl, etwa durch das Entnehmen von Proben an verschiedenen Stellen der in Frage kommenden Spargelfelder. Die Datenbank muss zudem kontinuierlich erweitert werden, um die Datengrundlage stetig zu verbessern und etwaige Veränderungen im Metabolom bzw. im Isotopolom abbilden zu können. Mit einer verbesserten Probenanzahl im Vergleich zum hier verwendeten Datensatz ist außerdem eine höhere Genauigkeit bei der Vorhersage der polnischen Herkunft sowie im Falle der NIR-Methode eine Unterscheidung der Herkunftsländer Spanien und Griechenland denkbar.

Des Weiteren ist für die Routineanwendung eine Verkürzung der Lyophilisationszeit erforderlich, da diese mit einer Dauer von bislang 72 Stunden als nicht wirtschaftlich zu bewerten ist. Die Lyophilisation des Probenmaterials ist jedoch insbesondere für die Analytik mittels ICP-MS unabdingbar, um für alle Analyten Konzentrationen oberhalb der erreichbaren Nachweisgrenzen zu erzielen. Im Rahmen dieser Arbeit wurden recht große Probenmengen für die Lyophilisation eingesetzt, um sicherzugehen, dass ausreichend Material, auch für etwaige Nachmessungen, zur Verfügung steht und da dies die einzig praktikable Möglichkeit der längerfristigen Lagerung des Probenmaterials darstellt. Zudem wurde eine Dreifachbestimmung jeder Probe durchgeführt, wobei in der Routineanwendung zu prüfen wäre, ob eine Einfachbestimmung ausreichen könnte. Somit besteht in verschiedenen Teilprozessen Potential zur Verringerung der benötigten Probenmenge und somit auch zur Optimierung der Lyophilisationszeit.

Für die hier gezeigten Methoden zur Herkunftsbestimmung ist darüber hinaus eine Anwendbarkeit auch auf andere Lebensmittel gegeben. Insbesondere die erreichte Präzision und die hohe Stabilität des Elementprofils unterstreichen, dass diese Herangehensweise auch auf andere Produkte übertragbar ist. Die Vorgehensweise zur Anwendung und Anpassung der gezeigten Methoden auf ein anderes Lebensmittel ist nachfolgend in den Abbildungen 55 und 56 dargestellt.



**Abbildung 55:** Schematische Darstellung eines Handlungsstrangs zur Übertragung der in dieser Arbeit gezeigten ICP-MS-basierten Methodik auf ein anderes Lebensmittel.



**Abbildung 56:** Schematische Darstellung eines Handlungsstrangs zur Übertragung der in dieser Arbeit gezeigten FT-NIR-basierten Methodik auf ein anderes Lebensmittel.

# 6 Experimenteller Teil

## 6.1 Probenmaterial

### 6.1.1 Akquise der Spargelproben

Zu Beginn des Projektes erfolgte eine umfangreiche Akquise von Probenmaterial. Im Vordergrund standen dabei Spargelproben aus allen relevanten Anbaugebieten Deutschlands und insbesondere aus den deutschen Regionen mit einer geschützten geographischen Herkunftsangabe (ggA). Die deutschen Spargelproben wurden von 43 Bauern aus 13 Regionen bezogen. Die Authentizität dieser Proben konnte sichergestellt werden, da der Spargel vor Ort beim Bauern und direkt nach der Ernte abgeholt wurde. Auf diese Weise sollte zudem gewährleistet werden, dass alle Proben nach der Ernte möglichst gleich behandelt werden, um außer der geographischen Herkunft so wenig wie möglich Varianz innerhalb der Proben entstehen zu lassen. Aus diesem Grund wurde darauf geachtet, dass alle Proben nach dem Stechen vorgewaschen und schockgekühlt wurden, bevor Lagerung und Transport in einem Kühltransporter bei ca. 3 °C erfolgten. Die Probenakquise erstreckte sich über einen Zeitraum von einer Woche, sodass eine geringe Varianz bei der Transportdauer nicht vermieden werden konnte. Auf diese Weise wurden während der Erntesaison in den Jahren 2015 und 2016 insgesamt 181 Proben akquiriert. Davon abweichend erfolgte der Probentransport der Spargelproben im Jahr 2014 in einem Kühltransporter bei – 20 °C. Im Jahr 2014 wurden zusätzlich 41 Proben erhalten. Kombiniert mit dem Abholen der Spargelproben direkt beim Erzeuger wurden zudem wichtige Metadaten generiert, wie z. B. Spargelsorte, Feldbezeichnung, Jahr der wiederholten Spargelernte sowie die verwendeten Dünge- und Pflanzenschutzmittel.

In den Jahren 2016 und 2017 wurde zusätzlich Probenmaterial aus den beiden Nachbarländern Polen und den Niederlanden bezogen, welches ebenfalls direkt beim Bauern abgeholt wurde. Spargelproben aus weiteren Ländern wurden aus dem Einzel- und Großhandel bezogen. Dabei wurde, soweit möglich, eine Überprüfung der Herkunftsdeklaration auf der Grundlage vorhandener Dokumente vorgenommen. Die Authentizität dieser Proben konnte somit nicht zweifelsfrei belegt werden. Auf diese Weise wurden

99 Proben nicht deutschen Ursprungs generiert.

Um eine repräsentative Stichprobe eines Standortes zu erreichen, umfasste jede Probe mindestens 2 kg Spargel. Alle Spargelproben wurden nach dem Transport zunächst in flüssigem Stickstoff schockgefroren, um Stoffwechselfvorgänge zu stoppen und Veränderungen des Metaboloms zu verhindern. Anschließend erfolgte die Lagerung der Proben bis zur weiteren Verarbeitung bei  $-20^{\circ}\text{C}$ . Eine Übersicht der akquirierten Spargelproben ist den Tabellen 20, 21 und 22 im Anhang im Abschnitt 8.1 zu entnehmen.

### 6.1.2 Akquise der Bodenproben

Zusätzlich zu dem im Jahr 2014 akquirierten Spargel wurden von jedem Standort Bodenproben entnommen, welche vom selben Feld stammten. Die Probenahme erfolgte mit einem Hohlmeißelbohrer. Dabei handelt es sich um ein Metallrohr mit einer Länge von einem Meter, wobei der Rohrquerschnitt zu einer Seite hin offen ist, sodass das Instrument einer Schaufel ähnelt. Der Hohlmeißelbohrer wurde mithilfe eines Schlaghammers in den Boden eingebracht, sodass in kurzer Zeit eine Tiefe von bis zu einem Meter erreicht werden konnte. Der Bohrer wurde anschließend aus dem Boden herausgezogen und aus dem im Rohr haftenden Bodenmaterial konnten nun Proben entnommen werden. An jedem Standort wurden drei Proben unterschiedlicher Tiefe akquiriert, indem jeweils etwa 100 g Material aus einer Tiefe von 0 - 30 cm, 30 - 60 cm und 60 - 90 cm entnommen wurden. Das Probenmaterial wurde in Plastikröhren gefüllt und diese bis zur weiteren Behandlung bei  $-20^{\circ}\text{C}$  gelagert. Eine Übersicht der akquirierten Bodenproben ist Tabelle 23 im Anhang im Abschnitt 8.2 zu entnehmen.

### 6.1.3 Vorbereitung der Spargelproben

Im Zuge der Probenvorbereitung für die spätere Analyse wurde der Spargel im gefrorenen Zustand zunächst mit Reinstwasser (Direct-Q 3 UV-R, Millipore, Bedford, MA, USA) gereinigt, um verbleibende Erdanhaftungen und weiteren Schmutz zu entfernen, da dies zu Verfälschungen, insbesondere bei der Elementanalytik, führen kann. Anschließend wurde jede Spargelstange auf eine Länge von 15 cm zugeschnitten und der obere Teil nochmals in drei Stücke geschnitten, um eine adäquate Größe für den folgenden Vermahlungsprozess zu erreichen. Die Stücke wurden nach dem Schneiden in flüssigem Stickstoff gelagert, um ein Auftauen und damit verbunden einer Veränderung des Metaboloms entgegenzuwirken. Nur der obere Teil der Spargelstange wurde für die Analyse verwendet, um für jede Spargelprobe eine vergleichbare und reproduzierbare

Probenzusammensetzung zu erreichen. Es folgte das Vermahlen der Probe mithilfe einer Messermühle (Grindomix GM 300, Retsch, Haan, Deutschland) unter Hinzunahme von Trockeneis, um das Auftauen der Probe durch die entstehende Reibungswärme zu vermeiden. Dazu wurden Spargel und Trockeneis zu gleichen Teilen (je 400 g) in den Mahlbecher eingewogen, was mehrere Mahlgänge erforderte. Das Mahlprogramm umfasste drei Zyklen à 30 s bei einer Drehzahl von 4000 rpm, die jeweils durch eine Pause von 5 s unterbrochen wurden. Dem ersten Zyklus ging eine Vorlaufzeit von 20 s bei 1000 rpm voraus, um zunächst eine grobe Zerkleinerung des Probenmaterials zu erreichen. Da es sich bei Mahlbecher und Messer um Edelstahlteile handelte, konnte ein geringer Elementeintrag, beispielsweise von Eisen, an dieser Stelle nicht vermieden werden. Das erhaltene Pulver wurde mit dem der übrigen Mahlgänge vereint und nach ausreichender Durchmischung wurde ein Aliquot von 60 g in einen Rundkolben eingewogen. Das gemahlene Probenmaterial wurde anschließend gefriergetrocknet (Beta 1-8 LDplus, Martin Christ Gefrier Trocknungsanlagen GmbH, Osterode am Harz, Deutschland). Die Trocknung erforderte eine Lyophilisationszeit von 72 h, wobei nach 48 h eine Lockerung und gleichmäßige Durchmischung des Materials erfolgte. Die trockenen Proben wurden schließlich im Mörser in ein feines Pulver überführt, in Plastikröhren gefüllt und bis zur weiteren Behandlung vor der Messung bei  $-20\text{ }^{\circ}\text{C}$  gelagert.

#### **6.1.4 Vorbereitung der Bodenproben**

Die bei  $-20\text{ }^{\circ}\text{C}$  gelagerten Bodenproben mussten vor dem Aufschluss zunächst getrocknet und von Verunreinigungen wie Steinen und Wurzelfragmenten befreit werden. Die Trocknung der Proben erfolgte in Porzellanschalen im Trockenschrank bei  $105\text{ }^{\circ}\text{C}$  für 24 h. Anschließend wurden zusammenhaftende Erdklumpen mit Mörser und Stößel zerkleinert und Fremdkörper mithilfe eines Siebes der Maschenweite 2 mm entfernt. Die Proben wurden anschließend in Plastikröhren überführt und bis zur weiteren Verwendung luftdicht bei Raumtemperatur gelagert.

---

## 6.2 Probenaufschlüsse

### 6.2.1 Initialer Aufschluss für pflanzliche Proben

Die Messung der Spargelproben in flüssiger Form erforderte einen vorherigen Probenaufschluss. Es wurde ein Aufschluss unter Druck in einer Mikrowelle durchgeführt (Ethos.lab, MLS, Leutkirch, Deutschland), der sich gegenüber dem offenen Aufschluss besonders durch höhere Temperaturen und stark verkürzte Aufschlusszeiten auszeichnet. Dazu wurde initial mit der Methode für den Aufschluss pflanzlicher Materialien gearbeitet, die in der Software des Mikrowellengerätes der Firma MLS hinterlegt ist. Diese sieht eine Einwaage von bis zu 700 mg Trockenmasse vor, die mithilfe von 8 mL konzentrierter Salpetersäure, 3 mL konzentriertem Wasserstoffperoxid und 10 mL Wasser aufgeschlossen werden sollen. Menge und Art der Aufschlussreagenzien wurden wie in Abschnitt 4.1.1 beschrieben für die vorliegende Fragestellung optimiert. Nach dem Einwiegen der Probe direkt in die Teflonbehälter und der Zugabe der Reagenzien wurden die Gefäße mit einem Deckel versehen und in die Segmente des Rotors fest eingeschraubt, um ein geschlossenes System zu erhalten. Das Temperaturprogramm sieht ein Erhitzen der Proben auf 200 °C innerhalb von 16 min und ein Halten dieser Temperatur für weitere 14 min vor. Es folgt eine Abkühlphase mindestens bis zur Unterschreitung einer Temperatur von 80 °C. Nach dem Abkühlen der Proben wurden diese unter mehrmaligem Spülen mit Reinstwasser in vorgereinigte Plastikröhren überführt und ebenfalls mit Reinstwasser auf ein Volumen von 20 mL aufgefüllt. Die so erhaltene klare Lösung wurde direkt zur Messung mittels ICP-MS eingesetzt. Es erfolgte für jede Probe eine Dreifachbestimmung, indem das Probenmaterial drei Mal aufgeschlossen wurde.

### 6.2.2 Aufschluss der Bodenproben

Beim Aufschluss der Bodenproben wurde auf einen Totalaufschluss mit Flusssäure verzichtet, stattdessen erfolgte ein partieller Aufschluss nach folgendem Procedere: Es wurden jeweils 250 mg der gemäß Abschnitt 6.1.4 vorbereiteten Bodenproben in einen Aufschlussbehälter eingewogen. Dazu wurden 5 mL Salpetersäure (69 %, ROTIPURAN<sup>®</sup> Supra), 3 mL Salzsäure (30 %, Suprapur<sup>®</sup>) und 2 mL Reinstwasser gegeben, die Behälter verschlossen und in den Mikrowellenrotor eingesetzt. Die Proben wurden innerhalb von 10 min auf 210 °C erhitzt und die Temperatur wurde anschließend für weitere 20 min gehalten. Nach dem Abkühlen der Proben wurden diese in vorgereinigte Plastikröhren überführt und mit Reinstwasser auf ein Volumen von 40 mL aufgefüllt. Im Anschluss

erfolgte eine Zentrifugation der Proben für 20 min bei 3000 *g*, um unlösliche Rückstände der Probe abzutrennen. Zusätzlich wurde eine 1:40-Verdünnung vorgenommen, indem 17.5 mL Reinstwasser mit 2 mL Salpetersäure (69 %, ROTIPURAN® Supra) und 0.5 mL des Zentrifugationsüberstandes versetzt wurden. Diese Verdünnung wurde zur Messung mittels ICP-MS eingesetzt.

### 6.2.3 Reinigungsaufschluss und Gefäßreinigung

Zur Vermeidung von Kontaminationen wurde im Anschluss an jeden Aufschluss ein Reinigungsaufschluss durchgeführt. Dazu wurden 5 mL konzentrierte Salpetersäure (65 %, p.a.) und 5 mL Reinstwasser in die Aufschlussbehälter pipettiert, diese verschlossen und in den Mikrowellenrotor eingesetzt. Es folgte ein Aufschluss, bei dem die Lösung innerhalb von 8 min auf 170 °C erhitzt und diese Temperatur für weitere 6 min gehalten wurde. Nach dem Abkühlen der Behälter wurde die Aufschlusslösung entsorgt und die Behälter mit destilliertem Wasser und Reinstwasser gespült und über Nacht getrocknet.

Des Weiteren wurden die Plastikröhrchen, in die die Proben nach dem Aufschluss überführt wurden, vor der Benutzung gereinigt, um einer Migration von Elementen aus dem Gefäß in die Probenlösung entgegenzuwirken. Dazu wurden die Röhrchen für 24 h mit verdünnter Salpetersäure (5 %, p.a.) befüllt. Anschließend erfolgte die Reinigung der Röhrchen mit Reinstwasser sowie die Trocknung über Nacht.

---

## 6.3 Messung mittels ICP-MS

### 6.3.1 Gerätekalibration

Vor jeder Messung erfolgte eine mindestens 30-minütige Aufwärmphase, die der Stabilisierung des Plasmas diente. Daran schloss sich die Geräteoptimierung an, um adäquate Countzahlen für den gesamten Massenbereich zu erzielen. Dazu wurde eine *Tuning*-Lösung der Konzentration 1 ppb verwendet, die die Elemente Lithium, Cobalt, Yttrium und Thallium mit den jeweiligen Massen 7, 59, 89 und 205 enthielt. Diese Elemente stehen stellvertretend für die Elemente im niedrigen, mittleren und hohen Massenbereich, womit sichergestellt werden soll, dass die Geräteparameter für alle Elemente gleichermaßen geeignet eingestellt werden. Die Massen 7, 89 und 205 fanden Anwendung im *NoGas*-Modus, wohingegen der Helium-Modus mithilfe der Massen 59, 89 und 205 kalibriert wurde. Die verwendeten Massen spiegeln dabei den für den jeweiligen Messmodus typischen Massenbereich wider. Das Einstellen geeigneter Parameter erfolgte separat für die beiden Messmodi *NoGas*- und Helium-Modus. Neben dem Erreichen ausreichend hoher Countzahlen und damit verbunden einer hohen Sensitivität, ist beim *Tuning* auf eine niedrige Oxidrate sowie auf ein geringes Maß an doppelt geladenen Ionen zu achten, da diese Störfaktoren darstellen, die das  $m/z$ -Verhältnis der Analytionen verfälschen können. Die genannten Effekte werden anhand von Cer (Masse: 140) untersucht, welches ebenfalls in der *Tuning*-Lösung enthalten ist. Anhand der Verhältnisse der Massen 156/140 kann die Oxidrate und mit dem Verhältnis 70/140 die Rate an doppelt geladenen Ionen bestimmt werden, die beide möglichst niedrig gehalten werden sollen. Die Oxidrate lag im Mittel bei 1.38 % im *NoGas*-Modus und bei 0.26 % im Helium-Modus. Die mittlere Rate an doppelt geladenen Ionen betrug 1.44 % (*NoGas*-Modus) bzw. 1.57 % (Helium-Modus). Zusätzlich wurde ein *P/A-Tuning* durchgeführt, bei dem die beiden Modi des Detektors, Pulse- und Analog-Modus, aufeinander abgestimmt werden. So wird sichergestellt, dass aus den Messwerten beider Detektormodi eine durchgehend lineare Kalibrierkurve erstellt wird, ohne dass eine Änderung der Steigung zwischen den beiden Messbereichen auftritt. Dafür wurde eine Kalibrierlösung der Konzentration 50 µg/L verwendet, die alle Zielanalyten enthielt. Während des *P/A-Tunings* wurden für alle Massen spezifische Faktoren ermittelt, mit denen kleine Abweichungen der Steigung beider Teilgeraden ausgeglichen wurden.

Wichtige Parameter, die nicht in der Messmethodik festgelegt sind, sondern an die momentane Geräteperformance angepasst werden müssen, sind die vertikale und horizontale Position der Plasmafackel, die Spannungen verschiedener Linsen am *Interface*-Linsensystem sowie die Spannungen an Oktupol und Quadrupol und die Parameter des Detektors. Ein beispielhafter Bericht für das *Tuning* der Sensitivität

ist Abbildung 57 im Anhang im Abschnitt 8.5 zu entnehmen. Der jeweils ausgewählte Messmodus für jedes Isotop ist in Tabelle 24 dargestellt.

### 6.3.2 Messparameter der Multielementmethode

Die in Tabelle 17 aufgeführten Parameter mit Spannbreiten wurden im *Tuning* vor jeder Messung auf der Grundlage der aktuellen Geräteperformance innerhalb des angegebenen Rahmens optimiert.

**Tabelle 17:** Geräteparameter der finalen Multielementmethode.

<b>Analyten</b>	<b>Parameterwahl</b>
Elemente	Ag, Al, As, Ba, Be, Bi, Ce, Cd, Co, Cr, Dy, Er, Eu, Ga, Gd, Ho, La, Li, Lu, Ni, Mo, Nb, Pb, Pr, Rb, Sm, Sr, Tb, Te, Th, Tl, Tm, U, V, Yb, Zn
Massenbereich	7 - 238
<b>Plasma Conditions</b>	<b>Parameterwahl</b>
Radio Frequency Power	1550 W
Radio Frequency Matching	1.9 V
Sample Depth	8 mm
Carrier Gas	0.75 L/min
Dilution Mode	OFF
Makeup Gas	0.35 L/min
Nebulizer Pump	0.15 rps
S/C Temperature	2 °C
<b>Ion Lenses</b>	<b>Parameterwahl</b>
Extract 1	0 V
Extract 2	-(140 - 190 V)
Omega Bias	- 70 V
Omega Lens	6.8 - 8.6 V
Cell Entrance	- 30 V
Cell Exit	- 50 V
Deflect	11.8 - 14.6 V
Plate Bias	- 40 V

**Tabelle 17:** Geräteparameter der finalen Multielementmethode - Fortsetzung.

<b>Octopole Parameters</b>	<b>Parameterwahl</b>
Octopole Radio Frequency	150 - 180 V
Octopole Bias	- 8 V
<b>Quadrupole Parameters</b>	<b>Parameterwahl</b>
AMU Gain	126 - 132
AMU Offset	127 - 128
Axis Gain	0.9997 - 1.003
Axis Offset	0.01 - 0.08
Quadrupole Bias	- 3 V
<b>Detector Parameters</b>	<b>Parameterwahl</b>
Discriminator	4.8 mV
Analog High Voltage	1828 V
Pulse High Voltage	1034 V
<b>Collision Cell</b>	<b>Parameterwahl</b>
He Gas Flow	5.5 mL/min

Während der Aufwärmphase des Gerätes sowie nach der Messung jeder Probe erfolgte ein Spülvorgang, um Probenrückstände und andere Verunreinigungen aus den Schläuchen und allen sonstigen Teilen des Proben transports zu entfernen. Dazu wurde eine verdünnte Salpetersäure-Lösung der Konzentration 3 % (p.a.) verwendet. Auch die Probennadel am Autosampler wurde durch das Umspülen für mindestens 5 s mit einer Salpetersäure-Lösung (2 %, p.a.) nach dem Eintauchen in eine Probenlösung von Rückständen gereinigt.

### 6.3.3 Externe Kalibrierung

Die Kalibriergerade, die zur Quantifizierung der Analyten herangezogen wurde, umfasste 20 Kalibrierpunkte, deren Konzentrationen in Tabelle 18 aufgelistet sind. Die Kalibriergerade wurde unter Hinzunahme des Signals für den internen Standard generiert. Dazu wurde die Countzahl des jeweiligen Kalibrierpunktes mit der Countzahl des internen Standards ins Verhältnis gesetzt und dieses Verhältnis gegen die Standardkonzentration aufgetragen.

**Tabelle 18:** Konzentrationen der Kalibrierpunkte und abweichende Konzentrationen für die Elemente Arsen, Beryllium, Zink sowie die seltenen Erden; Angaben in  $\mu\text{g/L}$ .

Kalibrierpunkt	Konzentration	Konz. As/Be/Zn	Konz. REE
1	0	0	0
2	0	0	0
3	0.001	0.01	0.001
4	0.0025	0.025	0.0025
5	0.005	0.05	0.005
6	0.01	0.1	0.01
7	0.025	0.25	0.025
8	0.05	0.5	0.05
9	0.1	1	0.1
10	0.25	2.5	0.25
11	0.5	5	0.5
12	1	10	1
13	2.5	25	2.5
14	5	50	5
15	10	100	0
16	25	250	0
17	50	500	0
18	100	1000	0
19	500	5'000	0
20	1'000	10'000	0

### 6.3.4 Qualitätskontrolle

Um eine ausreichende Stabilität der Messungen zu gewährleisten und gravierende Abweichungen ausschließen zu können, erfolgte die Messung einer QC-Lösung nach jeder 10. Probe in jeder Sequenz. Diese bestand aus einer Verdünnung der Multi-elementstandards und enthielt alle gemessenen Elemente mit einer Konzentration von  $5 \mu\text{g/L}$  bzw. die Elemente Arsen, Beryllium und Zink mit  $50 \mu\text{g/L}$ . Zusätzlich zu einer konstanten Leistungsfähigkeit des Gerätes wurde auch die gleichbleibende Qualität des Probenaufschlusses überprüft. Dazu wurde eine zufällig ausgewählte Spargelprobe in größerer Menge aufgearbeitet und in jedem Probenaufschluss erneut mit aufgeschlossen.

### 6.3.5 Blindwert

Um Verfälschungen durch mögliche Verunreinigungen der Aufschlussbehälter oder der Aufschlussreagenzien zu vermeiden, wurde jedem Aufschlussdurchgang ein Blindwert hinzugefügt, der jeweils vom Elementgehalt der Proben abgezogen wurde.

### 6.3.6 Statistische Datenauswertung

Es wurde für jede Probe der Mittelwert der Dreifachbestimmung herangezogen.

Das t-SNE wurde mit einer *perplexity* von 15 erstellt, alle weiteren Parameter waren identisch mit den Standardeinstellungen, welche im R-Paket 'Rtsne' hinterlegt sind [355].

Das hierarchische Clustering wurde mithilfe des agglomerativen Verfahrens erstellt, welches nach der Methode des *Weighted Average Linkage* mit dem Korrelationskoeffizienten nach Pearson als Distanzmaß durchgeführt wurde.

Die Erstellung und Validierung der SVM-Modelle erfolgte mithilfe einer *nested cross-validation*. Der Datensatz wurde dafür in 5 Teile ( $m$ ) getrennt, wobei ein Teil zur Validierung diente, während die übrigen Teile für eine Optimierung des Modells mittels 4-facher ( $k$ ) Kreuzvalidierung herangezogen wurden. Für die SVM-Modelle wurde der RBF-Kern verwendet. Die beiden dafür notwendigen Parameter  $C$  und  $\gamma$  wurden optimiert und validiert mit  $C = 2^p$  und  $\gamma = 0.01 \cdot 2^p$ , wobei  $p$  jeweils aus der Menge  $\{-5, -4, \dots, 5\}$  entnommen wurde.

---

## 6.4 Messung mittels FT-NIR

### 6.4.1 Messparameter

Das bei der Probenaufarbeitung erhaltene Pulver der lyophilisierten Spargelproben (vgl. Abschnitt 6.1.3) wurde vor der NIR-Messung gesiebt. Mithilfe eines Siebes der Maschenweite 500  $\mu\text{m}$  wurden größere Partikel abgetrennt, um so die Homogenität der Partikelgrößen zu erhöhen und Streulichteffekte zu reduzieren. Die Proben wurden direkt in ein Probengläschen gesiebt, welches anschließend zur Messung auf dem NIR-Spektrometer platziert wurde. Die Füllhöhe der Probe im Gläschen betrug dabei mindestens 2 cm. Die Messung der Spargelproben erfolgte über den Wellenzahlenbereich von 3'940 - 11'550  $\text{cm}^{-1}$  mit einer Auflösung von 2  $\text{cm}^{-1}$  und einer Scanzahl von 50. Es wurde eine Dreifachmessung jeder Probe vorgenommen, wobei das Probengläschen zwischen jeder Einzelmessung um 120° rotiert wurde. Aufgrund des Vorliegens von Festkörperproben wurden Messungen basierend auf diffuser Reflexion durchgeführt, für die das FT-NIR-Spektrometer TANGO der Firma Bruker (Bruker Optik GmbH, Ettlingen, Deutschland) verwendet wurde. Eine Übersicht der Geräte- und Messparameter ist in Tabelle 19 dargestellt.

**Tabelle 19:** Geräte- und Messparameter des FT-NIR-Spektrometers TANGO.

Parameter	Wert
Messbereich	3948 - 11547 $\text{cm}^{-1}$
Auflösung	2 $\text{cm}^{-1}$
Detektor	thermoelektrisch gekühlter Indium-Gallium-Arsenid-Detektor
Durchmesser des Messbereichs	10 mm
Messmethodik	diffuse Reflexion (mit Integrationskugel)
Scanzahl Probe	50 Scans
Scanzahl Hintergrund	50 Scans

## 6.4.2 Statistische Datenauswertung

Bei jeder Probe wurde der Mittelwert der Dreifachbestimmung für die Datenauswertung herangezogen. Mithilfe der präprozessierten Daten (vgl. Abschnitt 4.4.1) wurden Klassifikationsmodelle anhand von *Support Vector Machine* und *Random Forest* erstellt. Die Validierung der Modelle erfolgte mithilfe einer *nested cross-validation*. Dafür wurden die Daten in 5 Teile unterteilt, von denen ein Teil als Testdatensatz abgetrennt wurde (*outer loop*). Mit den übrigen Daten wurde mittels 10-facher Kreuzvalidierung ein Klassifikationsmodell trainiert, dessen Leistungsfähigkeit anschließend auf dem Testdatensatz geprüft wurde. Dieser Vorgang wurde fünf Mal wiederholt, sodass jeder Teil des Datensatzes einmal als Testdatensatz fungierte. Die für die SVM notwendigen Parameter  $C$  und  $\gamma$  wurden während der Kreuzvalidierung in der *inner loop* optimiert und validiert, wobei für  $C$  der Wertebereich  $\{10^{-6}, 10^{-4}, 10^{-2}, 1, 10^2, 10^4, 10^6\}$  berücksichtigt und  $\gamma$  aus der Menge  $\{0.0001, 0.001, 0.01, 0.1\}$  entnommen wurde. Für die Erstellung der RF-Modelle wurde eine Anzahl von 1000 Bäumen verwendet. Bei dem Parameter ‘mtry’, der Zahl der Variablen, die in einem Baum bei jedem Split betrachtet werden, wurde auf die Wurzel der Zahl der *Features* zurückgegriffen, die bei 370 *Features* (nach *Binning*) 19 betrug.

Für das hierarchische Clustering wurde ein agglomeratives Verfahren nach der Methode des *Complete Linkage* mit dem euklidischen Abstand als Distanzmaß herangezogen.

## 6.5 Software

Die Messung und Auswertung der ICP-MS-Daten erfolgte mithilfe der *MassHunter*-Software der Firma Agilent (Agilent Technologies, Inc., Santa Clara, CA, USA) in der Version B.01.01 [356, 357].

Die NIR-Messungen wurden mithilfe der Software *OPUS* der Firma Bruker (Bruker Optik GmbH, Ettlingen, Deutschland) in der Version 7.5 vorgenommen [358].

Berechnungen, Datenformatierungen und Datenauswertungen wurden mittels *Excel* (Microsoft Corporation, Redmond, WA, USA) durchgeführt [359]. Für die Datenauswertung wurde zudem der *Unscrambler*<sup>®</sup> X 10.3 (Camo Analytics, Oslo, Norwegen) herangezogen [360]. Multivariate Datenanalyse, Modellerstellung für die Klassifikation sowie das Erstellen von Abbildungen erfolgte des Weiteren mit der Software *R* [361]. Es wurden die *R*-Pakete *e1071*, *gplots*, *beeswarm*, *caret*, *prospectr*, *Rtsne* und *Boruta* hinzugezogen [2, 355, 362–366].

## 7 Literatur

- [1] IUPAC, Periodic Table of Elements. <https://iupac.org/what-we-do/periodic-table-of-elements/> (besucht am 15. 11. 2019).
- [2] Eklund, A. beeswarm: The Bee Swarm Plot, an Alternative to Stripchart., Version 0.2.3, 2016, <https://CRAN.R-project.org/package=beeswarm> (besucht am 20. 11. 2019).
- [3] Abdi, H., *Bonferroni and Šidák corrections for multiple comparisons in Encyclopedia of measurement and statistics*; Sage Thousand Oaks, CA: 2007, S. 103–107.
- [4] Otto, M., *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, 2. vollst. überarb. Aufl.; Wiley-VCH: Weinheim, 2007.
- [5] Pflanzenforschung.de c/o Genius GmbH, Hybride; Hybridsorte. <https://www.pflanzenforschung.de/de/themen/lexikon/hybride-hybridsorte-1980> (besucht am 10. 03. 2020).
- [6] Institut Technik-Theologie-Naturwissenschaften, Hybridzüchtung – Mehr Ertrag bei den Nachkommen. <https://www.pflanzen-forschung-ethik.de/verfahren/hybridzuechtung.html> (besucht am 10. 03. 2020).
- [7] Universität Zürich, Kruskal-Wallis-Test. [https://www.methodenberatung.uzh.ch/de/datenanalyse\\_spss/unterschiede/zentral/kruskal.html](https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/unterschiede/zentral/kruskal.html) (besucht am 20. 11. 2019).
- [8] Universität Zürich, Einfaktorielle Varianzanalyse. [https://www.methodenberatung.uzh.ch/de/datenanalyse\\_spss/unterschiede/zentral/evarianz.html](https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/unterschiede/zentral/evarianz.html) (besucht am 10. 03. 2020).
- [9] Klenke, A. (Johannes Gutenberg-Universität Mainz), Biostatistik, Vergleich zweier Stichproben: Gepaarter, Ungepaarter t-Test, Welch Test. [https://www.staff.uni-mainz.de/klenke/vorlesungen/vorl\\_ss17-3/biostatistik-SS2017-V12-online.pdf](https://www.staff.uni-mainz.de/klenke/vorlesungen/vorl_ss17-3/biostatistik-SS2017-V12-online.pdf) (besucht am 20. 11. 2019).
- [10] European Union: Knowledge Centre for Food Fraud and Quality, Food fraud and quality. [https://ec.europa.eu/knowledge4policy/sites/know4pol/files/a0infographic\\_kc\\_food\\_fraud\\_final\\_0.pdf](https://ec.europa.eu/knowledge4policy/sites/know4pol/files/a0infographic_kc_food_fraud_final_0.pdf) (besucht am 15. 11. 2019).

- [11] Bundesministerium für Ernährung und Landwirtschaft, Lebensmittel mit geschützter geografischer Herkunftsangabe. [https://www.bmel.de/DE/Landwirtschaft/Agrarpolitik/1\\_EU-Marktregelungen/\\_Texte/GeschuetzteBezeichnungen.html](https://www.bmel.de/DE/Landwirtschaft/Agrarpolitik/1_EU-Marktregelungen/_Texte/GeschuetzteBezeichnungen.html) (besucht am 29.10.2019).
- [12] Bundeszentrum für Ernährung, Wissenswertes zu Spargel: Genuss aus der Tiefe. <https://www.bzfe.de/inhalt/wissenswertes-zu-spargel-30027.html> (besucht am 15.11.2019).
- [13] Berlin-Brandenburgische Akademie der Wissenschaften, Spargel – Digitales Wörterbuch der deutschen Sprache. <https://www.dwds.de/wb/Spargel> (besucht am 15.11.2019).
- [14] *Duden - Wörterbuch medizinischer Fachbegriffe*, 9. Aufl.; Dudenverlag: Mannheim, 2012.
- [15] Iqbal, M. *et al.* Review on therapeutic and pharmaceutically important medicinal plant *Asparagus officinalis* L. *Journal of Plant Biochemistry & Physiology* **2017**, 5, 2.
- [16] Hondelmann, W., *Die Kulturpflanzen der griechisch-römischen Welt - Pflanzliche Ressourcen der Antike*; Gebrüder Borntraeger: Berlin, 2002.
- [17] Löbell, G.; Löbell, E., *Spargel*; Umschau Buchverlag: Neustadt/Weinstraße, 2000.
- [18] Thompson, H. C.; Kelly, W. C., *Vegetable Crops*, 5. Aufl.; McGraw-Hill Book Co.: New York, 1957.
- [19] *Asparagus: How to Grow, Harvest and Preserve Asparagus*; U.S. Department of Agriculture: 2017.
- [20] Thompson, H. C., *Asparagus Production*; Read Books Ltd: 2013.
- [21] Lenhardt, A., Spargel. <https://boellenhof.de/produkte/spargel/> (besucht am 13.11.2019).
- [22] Schwarz, J. *et al.* Sustainability of Global and Local Food Value Chains: An Empirical Comparison of Peruvian and Belgian Asparagus. *Sustainability* **2016**, 8, 344.
- [23] Zafiriou, P. *et al.* Analysis of energy flow and greenhouse gas emissions in organic, integrated and conventional cultivation of white asparagus by PCA and HCA: cases in Greece. *Journal of Cleaner Production* **2012**, 29-30, 20–27.
- [24] Benson, B. 2009 Update of the world's asparagus production areas, spear utilization and production periods. *XII International Asparagus Symposium 950* **2009**, 87–100.
- [25] Unilever Deutschland GmbH, Die drei wichtigsten Spargelsorten. [https://asset-eu.unileversolutions.com/content/dam/unilever/knorr\\_world/global/unilever\\_corporate\\_images/unilever\\_brand/other\\_foods/all/article-detail-alleszuspargel-spargelsorten-2880x1620-961003.jpg.ulenscale.768x576.jpg](https://asset-eu.unileversolutions.com/content/dam/unilever/knorr_world/global/unilever_corporate_images/unilever_brand/other_foods/all/article-detail-alleszuspargel-spargelsorten-2880x1620-961003.jpg.ulenscale.768x576.jpg) (besucht am 15.11.2019).

- [26] Nestlé Professional, Spargel: Staude mit ihren Sprossen. [https://www.nestleprofessional.de/sites/g/files/gfb201/f/styles/wysiwyg\\_half/public/media/keyword-spargel-staude.jpg?itok=zMIqeQ7v](https://www.nestleprofessional.de/sites/g/files/gfb201/f/styles/wysiwyg_half/public/media/keyword-spargel-staude.jpg?itok=zMIqeQ7v) (besucht am 15. 11. 2019).
- [27] Spektrum Lexikon der Biologie: Spargel. <https://www.spektrum.de/lexikon/biologie/spargel/62343> (besucht am 15. 11. 2019).
- [28] Park, J.-H. *et al.* Expression of AODEF, a B-functional MADS-box gene, in stamens and inner tepals of the dioecious species *Asparagus officinalis* L. *Plant Molecular Biology* **2003**, *51*, 867–875.
- [29] Harkess, A. *et al.* Sex-biased gene expression in dioecious garden asparagus (*Asparagus officinalis*). *New Phytologist* **2015**, *207*, 883–892.
- [30] Bundeszentrum für Ernährung, Spargel: Erzeugung – Wie wird Spargel angebaut?, <https://www.bzfe.de/inhalt/spargel-erzeugung-5862.html> (besucht am 15. 11. 2019).
- [31] Gao, W. *et al.* Identification of two markers linked to the sex locus in dioecious *Asparagus officinalis* plants. *Russian Journal of Plant Physiology* **2007**, *54*, 816–821.
- [32] Fuchs, R.; Fuchs, D. Spargel. <https://www.wildfind.com/pflanzen/spargel> (besucht am 18. 11. 2019).
- [33] Universitätsklinikum Bonn, Informationszentrale gegen Vergiftungen: Garten-Spargel. <http://www.gizbonn.de/119.0.html> (besucht am 13. 11. 2019).
- [34] Statista, Pro-Kopf-Konsum von Spargel in Deutschland in den Jahren 2005/06 bis 2017/18. <https://de.statista.com/statistik/daten/studie/290815/umfrage/pro-kopf-konsum-von-spargel-in-deutschland/> (besucht am 18. 11. 2019).
- [35] Statistisches Bundesamt, Pressemitteilung vom 19. Juli 2017 – 248/17. [https://www.destatis.de/DE/Presse/Pressemitteilungen/2017/07/PD17\\_248\\_412pdf.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Presse/Pressemitteilungen/2017/07/PD17_248_412pdf.pdf?__blob=publicationFile) (besucht am 15. 11. 2019).
- [36] Food and Agriculture Organization of the United Nations, Crop data, asparagus production. <http://www.fao.org/faostat/en/#data/QC> (besucht am 18. 11. 2019).
- [37] European Commission, Food and agricultural products – DOOR database. [http://ec.europa.eu/agriculture/quality/door/list.html?locale=de&recordStart=0&filter.dossierNumber=&filter.comboName=spargel&filterMin.milestone\\_\\_mask=&filterMin.milestone=&filterMax.milestone\\_\\_mask=&filterMax.milestone=&filter.country=&filter.category=&filter.type=&filter.status=](http://ec.europa.eu/agriculture/quality/door/list.html?locale=de&recordStart=0&filter.dossierNumber=&filter.comboName=spargel&filterMin.milestone__mask=&filterMin.milestone=&filterMax.milestone__mask=&filterMax.milestone=&filter.country=&filter.category=&filter.type=&filter.status=) (besucht am 04. 02. 2019).
- [38] Regulation (EU) No 1151/2012 of the European Parliament and of the Council of 21 November 2012 on quality schemes for agricultural products and foodstuffs. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02012R1151-20130103&from=EN> (besucht am 18. 11. 2019).

- [39] Bundesanstalt für Landwirtschaft und Ernährung, geschützte geografische Angabe. [https://www.ble.de/SharedDocs/Bilder/DE/Ernaehrung-Lebensmittel/EU-Qualitaetskennzeichen/Logo\\_ggA.jpg;jsessionid=E197B93F02C7487231E60BF41FA13EC9.2\\_cid335?\\_\\_blob=normal&v=3](https://www.ble.de/SharedDocs/Bilder/DE/Ernaehrung-Lebensmittel/EU-Qualitaetskennzeichen/Logo_ggA.jpg;jsessionid=E197B93F02C7487231E60BF41FA13EC9.2_cid335?__blob=normal&v=3) (besucht am 29.10.2019).
- [40] Souci, S. W.; Fachmann, W.; Kraut, H. in *Food Composition and Nutrition Tables*. 6. Aufl.; medpharm GmbH Scientific Publishers, Stuttgart, Germany: 2000, S. 739–740.
- [41] Verordnung (EU) Nr. 1169/2011 des Europäischen Parlaments und des Rates vom 25. Oktober 2011 betreffend die Information der Verbraucher über Lebensmittel. <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:02011R1169-20180101&from=EN> (besucht am 18.11.2019).
- [42] Hafizur, R. M.; Kabir, N.; Chishti, S. Asparagus officinalis extract controls blood glucose by improving insulin secretion and  $\beta$ -cell function in streptozotocin-induced type 2 diabetic rats. *British Journal of Nutrition* **2012**, *108*, 1586–1595.
- [43] Negi, J. S. *et al.* Chemical constituents of *Asparagus*. *Pharmacognosy Reviews* **2010**, *4*, 215–220.
- [44] Mitchell, S. C., *Biological Interactions Of Sulfur Compounds*; CRC Press: 2003.
- [45] Waring, R. H.; Mitchell, S. C.; Fenwick, G. R. The chemical nature of the urinary odour produced by man after asparagus ingestion. *Xenobiotica* **1987**, *17*, 1363–1371.
- [46] Pelchat, M. L.; Bykowski, C.; Duke, F. F.; Reed, D. R. Excretion and Perception of a Characteristic Odor in Urine after Asparagus Ingestion: a Psychophysical and Genetic Study. *Chemical Senses* **2011**, *36*, 9–17.
- [47] RÖMPP, Thieme; Kolter, T. Asparagusinsäure. <https://roemp.thieme.de/roemp4.0/do/data/RD-01-05571> (besucht am 18.11.2019).
- [48] Shimoyamada, M. *et al.* Antifungal activity of the saponin fraction obtained from *Asparagus officinalis* L. and its active principle. *Agricultural and Biological Chemistry* **1990**, *54*, 2553–2557.
- [49] Edenharter, R.; John, K.; Ivo-Boor, H. Antimutagenic activity of vegetable and fruit extracts against in-vitro benzo(a)pyrene. *Zeitschrift für die gesamte Hygiene und ihre Grenzgebiete* **1990**, *36*, 144–147.
- [50] Shao, Y. *et al.* Anti-tumor activity of the crude saponins obtained from asparagus. *Cancer letters* **1996**, *104*, 31–36.
- [51] Zhao, Q. *et al.* *In vitro* antioxidant and antitumor activities of polysaccharides extracted from *Asparagus officinalis*. *Carbohydrate Polymers* **2012**, *87*, 392–396.
- [52] Bousserouel, S. *et al.* Methanolic extract of white asparagus shoots activates TRAIL apoptotic death pathway in human cancer cells and inhibits colon carcinogenesis in a preclinical model. *International Journal of Oncology* **2013**, *43*, 394–404.

- [53] Eva Morgenstern, Dienstleistungszentren Ländlicher Raum, Rheinland-Pfalz, Spargel richtig anbauen. [http://www.dlr.rlp.de/Internet/global/themen.nsf/se\\_quick/BOCCD1501F53E650C12577E70032137A?OpenDocument](http://www.dlr.rlp.de/Internet/global/themen.nsf/se_quick/BOCCD1501F53E650C12577E70032137A?OpenDocument) (besucht am 18. 11. 2019).
- [54] Bayerische Landesanstalt für Landwirtschaft, Erosionsschutz beim Anbau von Spargel. [https://www.lfl.bayern.de/mam/cms07/publikationen/daten/informationen/erosionsschutz-anbau-spargel\\_lfl-information.pdf](https://www.lfl.bayern.de/mam/cms07/publikationen/daten/informationen/erosionsschutz-anbau-spargel_lfl-information.pdf) (besucht am 18. 11. 2019).
- [55] Landwirtschaftskammer Nordrhein-Westfalen, Spargel selbst angebaut. <https://www.landwirtschaftskammer.de/verbraucher/rezepte/spargelrezepte.pdf> (besucht am 13. 08. 2019).
- [56] Rheinischer Landwirtschafts-Verlag GmbH, Bodenvorbereitung in Spargelkulturen. <https://www.spargel-erdbeerprofi.de/nachricht-spargelerdbeer/detail/bodenvorbereitung-in-spargelkulturen/> (besucht am 18. 11. 2019).
- [57] Monokrousos, N.; Papatheodorou, E. M.; Stamou, G. P. The response of soil biochemical variables to organic and conventional cultivation of *Asparagus* sp. *Soil Biology and Biochemistry* **2008**, *40*, 198–206.
- [58] Goldhofer, H.; Dunst, A. Feldgemüseanbau in Bayern. [https://www.lfl.bayern.de/mam/cms07/publikationen/daten/informationen/p\\_19984.pdf](https://www.lfl.bayern.de/mam/cms07/publikationen/daten/informationen/p_19984.pdf) (besucht am 18. 11. 2019).
- [59] May, H. Infos zum Frühjahrsgemüse Spargel., NABU - Naturschutzbund Deutschland e.V. <https://www.nabu.de/umwelt-und-ressourcen/oekologisch-leben/essen-und-trinken/06608.html> (besucht am 18. 11. 2019).
- [60] Peter Lederer, Spargelfeld in der Frühsonne. <https://peterlederer.de/wp-content/uploads/2017/04/spargelfeld-gr-700x467.jpg> (besucht am 18. 11. 2019).
- [61] Bosch, Aparagus field 2.0. [https://www.bosch-presse.de/pressportal/de/media/migrated\\_media/pi8947/1-rb-21311-e-2.jpg](https://www.bosch-presse.de/pressportal/de/media/migrated_media/pi8947/1-rb-21311-e-2.jpg) (besucht am 29. 10. 2019).
- [62] Landwirtschaftskammer Niedersachsen, Hinweise zur Düngung im Spargelanbau. <https://www.lwk-niedersachsen.de/index.cfm/portal/96/nav/2207/article/33974.html> (besucht am 18. 11. 2019).
- [63] Ländliches Fortbildungsinstitut, Pflanzenschutzmittel Liste der Kultur Spargel (Freiland). <https://warndienst.lko.at/pages/pflanzenschutz/Spargel.pdf> (besucht am 18. 11. 2019).
- [64] Barnes, H. The asparagus miner (*Melanagromyza simplex* H. Loew) (Agromyzidae; Diptera). *Ann. Appl. Biol* **1937**, *24*, 574–588.
- [65] Gilbertson, R.; Manning, W.; Ferro, D. Association of the asparagus miner with stem rot caused in asparagus by *Fusarium* species. *Phytopathology* **1985**, *75*, 1188–1191.

- [66] Voigt, D.; Gorb, S. Egg attachment of the asparagus beetle *Crioceris asparagi* to the crystalline waxy surface of *Asparagus officinalis*. *Proceedings of the Royal Society B: Biological Sciences* **2009**, *277*, 895–903.
- [67] Capinera, J. L. Asparagus beetle defense behavior: adaptations for survival in dispersing and non-dispersing species. *Annals of the Entomological Society of America* **1976**, *69*, 269–272.
- [68] Williamson, B.; Tudzynski, B.; Tudzynski, P.; van Kan, J. A. Botrytis cinerea: the cause of grey mould disease. *Molecular Plant Pathology* **2007**, *8*, 561–580.
- [69] Lord, F. Untersuchungen zum Einfluss spezieller Rhizosphärenbakterien auf Pflanzenwachstum und Fusarium spp.-Toleranz bei Spargel (*Asparagus officinalis* L.), Diss., 2002, <https://edoc.hu-berlin.de/bitstream/handle/18452/15631/Lord.pdf?sequence=1&isAllowed=y> (besucht am 18.11.2019).
- [70] Blok, W. J.; Coenen, T. C.; Termorshuizen, A. J.; Lamers, J. G. The potential of biological soil disinfestation to manage Fusarium foot and root rot in asparagus. *XI International Asparagus Symposium 776* **2005**, 135–144.
- [71] Yergeau, E. *et al.* Relationships between Fusarium population structure, soil nutrient status and disease incidence in field-grown asparagus. *FEMS Microbiology Ecology* **2006**, *58*, 394–403.
- [72] Hein, A. Über Viruskrankheiten des Spargels (*Asparagus officinalis* L.): Spargelvirus I. *Zeitschrift für Pflanzenkrankheiten (Pflanzenpathologie) und Pflanzenschutz* **1969**, *76*, 395–406.
- [73] Blok, W. J.; Bollen, G. J. Fungi on roots and stem bases of asparagus in the Netherlands: species and pathogenicity. *European Journal of Plant Pathology* **1995**, *101*, 15–24.
- [74] Elmer, W. H.; Johnson, D. A.; Mink, G. I. Epidemiology and management of the diseases causal to asparagus decline. *Plant Disease* **1996**, *80*, 117–125.
- [75] Benson, B. L. Second international asparagus cultivar trial final report. *X International Asparagus Symposium 589* **2001**, 159–166.
- [76] Wilson, D.; Cloughley, C.; Sinton, S. Model of the Influence of Temperature on the Elongation Rate of Asparagus Spears. *IX International Asparagus Symposium 479* **1999**, 297–304.
- [77] Velten, K.; Paschold, P.-J.; Stahel, A. Optimization of cultivation measures affecting soil temperature. *Scientia Horticulturae* **2003**, *97*, 163–184.
- [78] Frenz, D. F.; Munz, R. The influence of plastic covering of asparagus ridges on soil temperature and yield. *Symposium Plastics on Crops, Crops under Plastics 9* **1968**, 139–146.
- [79] Dienstleistungszentren Ländlicher Raum Rheinland-Pfalz, Spargeldammtemperaturen. <https://www.am.rlp.de/Internet/global/inetcntr.nsf/dlrsearesult.xsp?src=TU917E3NDM&p3=76G85T2XQW&p4=XJPZBV4849&p1=title%5C%3DVersuchswesen+online+Spargeldammtemperaturen%5C%7E%5C%7Eur1%5C%3DCN%5C%3DDL-R-WEB-SRV1%5C%2FOU%5C%3DWEB%5C%2FO%5C%3DAV-RPL%5C%7E%5C%7EInternet%5C%2Fglobal%5C%2Fthemen.nsf%5C%7E%5C%7EB87000EEFC010587C12575B000267B80%5C%7E%5C%7E1%5C%7E%5C%7EExpcontent.xsp%5C%7E%5C%7Eundefined> (besucht am 18.11.2019).

- [80] Spargelmesser Firmenich GmbH & Co. KG, Folien & Vlies. <http://www.firmenich.de/folieninfo.htm> (besucht am 18.11.2019).
- [81] Böckenhoff Folien GmbH, Folien - Überblick. <https://www.ehmo-tec-folien.de/folien/ueberblick/> (besucht am 18.11.2019).
- [82] Verband Süddeutscher Spargel- und Erdbeeranbauer e.V., Folieneinsatz in der Spargelproduktion. [https://www.vsse.de/fileadmin/media/Dokumente/Werbemittel/Folie\\_in\\_Spargelproduktion\\_Argumentation20042017.pdf](https://www.vsse.de/fileadmin/media/Dokumente/Werbemittel/Folie_in_Spargelproduktion_Argumentation20042017.pdf) (besucht am 18.11.2019).
- [83] Van den Elzen Plants, Gijnlim - Spargelpflanzen. <https://www.vandelzenplants.com/de/spargelpflanzen/gijnlim/> (besucht am 18.11.2019).
- [84] Südwestdeutsche Saatzucht GmbH & Co. KG, Ramires. <https://www.suedwestsaat.de/spargel/ramires.html> (besucht am 18.11.2019).
- [85] Van den Elzen Plants, Avalim - Spargelpflanzen. <https://www.vandelzenplants.com/de/spargelpflanzen/avalim/> (besucht am 18.11.2019).
- [86] Südwestdeutsche Saatzucht GmbH & Co. KG, Ramada. <https://www.suedwestsaat.de/spargel/ramada.html> (besucht am 18.11.2019).
- [87] Van den Elzen Plants, Grolim - Spargelpflanzen. <https://www.vandelzenplants.com/de/spargelpflanzen/grolim/> (besucht am 18.11.2019).
- [88] Van den Elzen Plants, Cumulus - Spargelpflanzen. <https://www.vandelzenplants.com/de/spargelpflanzen/cumulus/> (besucht am 18.11.2019).
- [89] Südwestdeutsche Saatzucht GmbH & Co. KG, Rapsody. <https://www.suedwestsaat.de/spargel/rapsody.html> (besucht am 18.11.2019).
- [90] Vissers Aardbeiplanten B.V., Herkolim - Spargelpflanzen. <http://www.vissers.com/de/spargelpflanzen/herkolim> (besucht am 18.11.2019).
- [91] Vissers Aardbeiplanten B.V., Backlim - Spargelpflanzen. <http://www.vissers.com/de/spargelpflanzen/backlim> (besucht am 18.11.2019).
- [92] Kling, S.; Kling, A. Erntefrischer Spargel direkt vom Bauern. <https://www.lehmgrubenhof.de/erntefrischer-spargel-direkt-vom-bauern> (besucht am 18.11.2019).
- [93] Sánchez, M.-T. *et al.* Use of near-infrared reflectance spectroscopy for shelf-life discrimination of green asparagus stored in a cool room under controlled atmosphere. *Talanta* **2009**, *78*, 530–536.
- [94] Henning, R.; Paschold, P.-J. Spargel-Qualität- Nacherntebehandlung. <https://www.hortigate.de/bericht?nr=40053> (besucht am 18.11.2019).

- [95] „Weißes Gold“: Griechische Bauern setzen auf deutsche Liebe zum Spargel. *Berliner Zeitung* **2018**, <https://www.berliner-zeitung.de/wirtschaft/-weisses-gold---griechische-bauern-setzen-auf-deutsche-liebe-zum-spargel-29896298> (besucht am 18.11.2019).
- [96] Prechtel, B.; Kayser, M.; Theuvsen, L. Organisation von Wertschöpfungsketten in der Gemüseproduktion: Das Beispiel Spargel. *DARE Diskussionsbeiträge* **2013**.
- [97] Shimizu, Tatsuya, Institute of Developing Economies, Expansion of asparagus production and exports in Peru. [https://www.researchgate.net/profile/Tatsuya\\_Shimizu3/publication/5141241\\_Expansion\\_of\\_Asparagus\\_Production\\_and\\_Exports\\_in\\_Peru/links/55d7377d08ae9d65948d851a/Expansion-of-Asparagus-Production-and-Exports-in-Peru.pdf](https://www.researchgate.net/profile/Tatsuya_Shimizu3/publication/5141241_Expansion_of_Asparagus_Production_and_Exports_in_Peru/links/55d7377d08ae9d65948d851a/Expansion-of-Asparagus-Production-and-Exports-in-Peru.pdf) (besucht am 18.11.2019).
- [98] Toledo, J. Asparagus Production in Peru. *Acta Horticulturae* **1990**, 203–210, [https://www.actahort.org/books/271/271\\_26.htm](https://www.actahort.org/books/271/271_26.htm).
- [99] Lewis, M. W. Asparagus Land: Coastal Peru’s Fruit and Vegetable Export Boom. <http://www.geocurrents.info/economic-geography/asparagus-land-coastal-perus-fruit-and-vegetable-export-boom> (besucht am 18.11.2019).
- [100] Bundesanstalt für Landwirtschaft und Ernährung, Marktbericht - Obst - Gemüse - Südfrüchte - KW 12/19. <https://www.fruchtportal.de/download/ble/b9683333> (besucht am 18.11.2019).
- [101] Nordwest-Zeitung, Erntebeginn in Niedersachsen: Spargel ist 40 Prozent teurer als im letzten Jahr. [https://www.nwzonline.de/wirtschaft/weser-ems/oldenburg-hannover-erntebeginn-in-niedersachsen-spargel-ist-40-prozent-teurer-als-im-letzten-jahr\\_a\\_50,1,1891731210.html](https://www.nwzonline.de/wirtschaft/weser-ems/oldenburg-hannover-erntebeginn-in-niedersachsen-spargel-ist-40-prozent-teurer-als-im-letzten-jahr_a_50,1,1891731210.html) (besucht am 13.08.2019).
- [102] Mitteldeutscher Rundfunk, Polen entdeckt den Spargel. <https://www.mdr.de/nachrichten/osteuropa/land-leute/spargel-polen-trend-100.html> (besucht am 18.11.2019).
- [103] Fruitnet Media International GmbH, Polen: Erster Spargel erhältlich. <https://www.fruchthandel.de/newsnet/aktuelle-news/einzelmeldung-newsnet/polen-erster-spargel-erhaeltlich/> (besucht am 18.11.2019).
- [104] Märkisches Medienhaus GmbH & Co. KG, Spargel für 1,20 Euro je Kilo. <https://www.moz.de/landkreise/maerkisch-oderland/bad-freienwalde/artikel1/dg/0/1/14003/> (besucht am 18.11.2019).
- [105] Swoboda, S. *et al.* Identification of Marchfeld asparagus using Sr isotope ratio measurements by MC-ICP-MS. *Analytical and Bioanalytical Chemistry* **2008**, *390*, 487–494.
- [106] Menzel, M. I. *et al.* Comparing <sup>1</sup>H-NMR imaging and relaxation mapping of German white asparagus from five different cultivation sites. *Journal of Plant Nutrition and Soil Science* **2007**, *170*, 24–38.

- [107] Zannella, C. *et al.* Genetic and geochemical signatures to prevent frauds and counterfeit of high-quality asparagus and pistachio. *Food Chemistry* **2017**, *237*, 545–552.
- [108] Creydt, M. *et al.* Food Authentication: Small-Molecule Profiling as a Tool for the Geographic Discrimination of German White Asparagus. *Journal of Agricultural and Food Chemistry* **2018**, *66*, 13328–13339.
- [109] Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit, Isotopen-Herkunftsanalyse am Beispiel Spargel. [https://www.lgl.bayern.de/lebensmittel/kennzeichnung/echtheit\\_herkunft/pflanzliche\\_lebensmittel/et\\_isotopen\\_herkunftsanalyse\\_spargel.htm](https://www.lgl.bayern.de/lebensmittel/kennzeichnung/echtheit_herkunft/pflanzliche_lebensmittel/et_isotopen_herkunftsanalyse_spargel.htm) (besucht am 18. 11. 2019).
- [110] Christoph, N.; Schellenberg, A. Stabilisotopenanalytik – Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit. [https://www.lgl.bayern.de/lebensmittel/kennzeichnung/echtheit\\_herkunft/stabilisotopenanalytik.htm](https://www.lgl.bayern.de/lebensmittel/kennzeichnung/echtheit_herkunft/stabilisotopenanalytik.htm) (besucht am 18. 11. 2019).
- [111] Niedersächsisches Landesamt für Verbraucherschutz u. Lebensmittelsicherheit, Stabilisotopenanalyse – Authentizitätsanalyse im Lebensmittel- und Veterinärinstitut Oldenburg. <http://www.laves.niedersachsen.de/lebensmittel/authentizitaetsanalyse/stabilisotopenanalyse-133613.html> (besucht am 18. 11. 2019).
- [112] Schlicht, C.; Roßmann, A.; Brunner, E. Anwendung der Multielement-Multikomponenten Isotopenverhältnismassenspektrometrie (IRMS) zur Prüfung der geographischen Herkunft von Spargel. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **2006**, *1*, 97–105.
- [113] Arbeitsgruppe Stabilisotopenanalytik, LChG, Stellungnahme der Lebensmittelchemischen Gesellschaft – Möglichkeiten und Grenzen der Sauerstoff Isotopenanalyse von Wasser in der Herkunftsbestimmung von Lebensmitteln. [https://www.gdch.de/fileadmin/downloads/Netzwerk\\_und\\_Strukturen/Fachgruppen/Lebensmittelchemiker/Arbeitsgruppen/stabilisotopen/stabil\\_wasser.pdf](https://www.gdch.de/fileadmin/downloads/Netzwerk_und_Strukturen/Fachgruppen/Lebensmittelchemiker/Arbeitsgruppen/stabilisotopen/stabil_wasser.pdf) (besucht am 18. 11. 2019).
- [114] Annweiler, E.; Erich, S. Chemisches und Veterinäruntersuchungsamt Freiburg – Nachweis der Herkunft. [http://www.ua-bw.de/pub/beitrag.asp?ID=2231&subid=0&Thema\\_ID=2&lang=DE](http://www.ua-bw.de/pub/beitrag.asp?ID=2231&subid=0&Thema_ID=2&lang=DE) (besucht am 19. 11. 2019).
- [115] Rossier, J. S.; Maury, V.; de Voogd, B.; Pfammatter, E. Use of isotope ratio mass spectrometry (IRMS) determination ( $^{18}\text{O}/^{16}\text{O}$ ) to assess the local origin of fish and asparagus in western Switzerland. *CHIMIA International Journal for Chemistry* **2014**, *68*, 696–700.
- [116] Joyce, A. R.; Palsson, B. Ø. The model organism as a system: integrating 'omics' data sets. *Nature Reviews Molecular Cell Biology* **2006**, *7*, 198.
- [117] Horgan, R. P.; Kenny, L. C. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist* **2011**, *13*, 189–195.

- [118] Bujak, R.; Struck-Lewicka, W.; Markuszewski, M. J.; Kaliszan, R. Metabolomics for laboratory diagnostics. *Journal of Pharmaceutical and Biomedical analysis* **2015**, *113*, 108–120.
- [119] Brunetti, A. E. *et al.* An integrative omics perspective for the analysis of chemical signals in ecological interactions. *Chemical Society Reviews* **2018**, *47*, 1574–1591.
- [120] Hasin, Y.; Seldin, M.; Lusic, A. Multi-omics approaches to disease. *Genome Biology* **2017**, *18*, 83.
- [121] Zhang, W.; Li, F.; Nie, L. Integrating multiple ‘omics’ analysis for microbial biology: application and methodologies. *Microbiology* **2010**, *156*, 287–301.
- [122] Raamsdonk, L. M. *et al.* A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology* **2001**, *19*, 45–50.
- [123] Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **2009**, *10*, 57–63.
- [124] James, P. Protein identification in the post-genome era: the rapid rise of proteomics. *Quarterly Reviews of Biophysics* **1997**, *30*, 279–331.
- [125] Ryan, D.; Robards, K. Metabolomics: the greatest omics of them all? *Analytical Chemistry* **2006**, *78*, 7954–7958.
- [126] Rochfort, S. Metabolomics reviewed: a new “omics” platform technology for systems biology and implications for natural products research. *Journal of Natural Products* **2005**, *68*, 1813–1820.
- [127] Fröhlich, E. Role of omics techniques in the toxicity testing of nanoparticles. *Journal of Nanobiotechnology* **2017**, *15*, 84.
- [128] Felbinger, C.; Schelm, S.; Fischer, M. Food Fraud - Hindernisse und Lösungswege bei der Authentizitätsbestimmung von Lebensmitteln. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **2015**, *10*, 25–30.
- [129] Baumann, F. *et al.* Lebensmittelsicherheit & Verpackung. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **2014**, *9*, 391–420.
- [130] Kabata-Pendias, A. Soil-plant transfer of trace elements – an environmental issue. *Geoderma* **2004**, *122*, 143–149.
- [131] Dobran, S.; Zagury, G. J. Arsenic speciation and mobilization in CCA-contaminated soils: Influence of organic matter content. *Science of The Total Environment* **2006**, *364*, 239–250.
- [132] Reid, R.; Horvath, D. Soil chemistry and mineral problems in farm livestock. A review. *Animal Feed Science and Technology* **1980**, *5*, 95–167.
- [133] Nye, P. H.; Tinker, P. B., *Solute movement in the soil-root system*; University of California Press: 1977; Bd. 4.
- [134] Tyler, G.; Olsson, T. Plant uptake of major and minor mineral elements as influenced by soil acidity and liming. *Plant and Soil* **2001**, *230*, 307–321.
- [135] Graham, R. D.; Stangoulis, J. C. Trace element uptake and distribution in plants. *The Journal of Nutrition* **2003**, *133*, 1502S–1505S.

- [136] Máthé-Gáspár, G.; Anton, A. Phytoremediation study: Factors influencing heavy metal uptake of plants. *Acta Biologica Szegediensis* **2005**, *49*, 69–70.
- [137] Ure, A. Single extraction schemes for soil analysis and related applications. *Science of The Total Environment* **1996**, *178*, 3–10.
- [138] Hodgson, J. Chemistry of the micronutrient elements in soils. *Advances in Agronomy* **1963**, *15*, 119–159.
- [139] Li, X.; Thornton, I. Chemical partitioning of trace and major elements in soils contaminated by mining and smelting activities. *Applied Geochemistry* **2001**, *16*, 1693–1706.
- [140] Clemens, S.; Palmgren, M. G.; Krämer, U. A long way ahead: understanding and engineering plant metal accumulation. *Trends in Plant Science* **2002**, *7*, 309–315.
- [141] Jamal, A.; Ayub, N.; Usman, M.; Khan, A. G. Arbuscular mycorrhizal fungi enhance zinc and nickel uptake from contaminated soil by soybean and lentil. *International Journal of Phytoremediation* **2002**, *4*, 205–221.
- [142] Clark, R. B., *Genetic Aspects of Plant Nutrition*; Plant genotype differences in the uptake, translocation, accumulation, and use of mineral elements required for plant growth; Springer: 1983, S. 49–70.
- [143] Ogra, Y.; Hirata, T., *Metallomics: Recent Analytical Techniques and Applications*; Springer: Japan, 2017.
- [144] Michalke, B., *Metallomics: Analytical Techniques and Speciation Methods*; John Wiley & Sons: Weinheim, Germany, 2016.
- [145] Capozzi, F.; Trimigno, A. 11 - Using metabolomics to describe food in detail. *Metabolomics as a Tool in Nutrition Research* **2015**, 203–229.
- [146] Dunn, W. B.; Ellis, D. I. Metabolomics: current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry* **2005**, *24*, 285–294.
- [147] Weckwerth, W. Metabolomics in systems biology. *Annual Review of Plant Biology* **2003**, *54*, 669–689.
- [148] Creydt, M.; Fischer, M. Omics approaches for food authentication. *Electrophoresis* **2018**, *39*, 1569–1581.
- [149] Cevallos-Cevallos, J. M. *et al.* Metabolomic analysis in food science: a review. *Trends in Food Science & Technology* **2009**, *20*, 557–566.
- [150] Cubero-Leon, E.; Peñalver, R.; Maquet, A. Review on metabolomics for food authentication. *Food Research International* **2014**, *60*, 95–107.
- [151] Hollywood, K.; Brison, D. R.; Goodacre, R. Metabolomics: Current technologies and future trends. *Proteomics* **2006**, *6*, 4716–4723.
- [152] Fernie, A. R.; Trethewey, R. N.; Krotzky, A. J.; Willmitzer, L. Metabolite profiling: from diagnostics to systems biology. *Nature Reviews Molecular Cell Biology* **2004**, *5*, 763–769.
- [153] Lu, W.; Bennett, B. D.; Rabinowitz, J. D. Analytical strategies for LC–MS-based targeted metabolomics. *Journal of Chromatography B* **2008**, *871*, 236–242.
- [154] Wei, R.; Li, G.; Seymour, A. B. High-throughput and multiplexed LC/MS/MRM method for targeted metabolomics. *Analytical Chemistry* **2010**, *82*, 5527–5533.

- [155] Weljie, A. M. *et al.* Targeted profiling: quantitative analysis of  $^1\text{H}$  NMR metabolomics data. *Analytical Chemistry* **2006**, *78*, 4430–4442.
- [156] Lee, J.-E. *et al.* Geographical and climatic dependencies of green tea (*Camellia sinensis*) metabolites: a  $^1\text{H}$  NMR-based metabolomics study. *Journal of Agricultural and Food Chemistry* **2010**, *58*, 10582–10589.
- [157] Vinale, F. *et al.* Major secondary metabolites produced by two commercial *Trichoderma* strains active against different phytopathogens. *Letters in Applied Microbiology* **2006**, *43*, 143–148.
- [158] Verma, N.; Shukla, S. Impact of various factors responsible for fluctuation in plant secondary metabolites. *Journal of Applied Research on Medicinal and Aromatic Plants* **2015**, *2*, 105–113.
- [159] Hovell, A.; Pereira, E.; Arruda, N.; Rezende, C. Evaluation of alignment methods and data pretreatments on the determination of the most important peaks for the discrimination of coffee varieties Arabica and Robusta using gas chromatography–mass spectroscopy. *Analytica Chimica Acta* **2010**, *678*, 160–168.
- [160] Rochfort, S.; Ezernieks, V.; Bastian, S. E.; Downey, M. O. Sensory attributes of wine influenced by variety and berry shading discriminated by NMR metabolomics. *Food Chemistry* **2010**, *121*, 1296–1304.
- [161] Beckmann, M.; Enot, D. P.; Overy, D. P.; Draper, J. Representation, comparison, and interpretation of metabolome fingerprint data for total composition analysis and quality trait investigation in potato cultivars. *Journal of Agricultural and Food Chemistry* **2007**, *55*, 3444–3451.
- [162] Thomas, R., *Practical Guide to ICP-MS: A Tutorial for Beginners*, 3rd New Edition; CRC press: 2013.
- [163] PerkinElmer Instruments, The 30-Minute Guide to ICP-MS. [https://www.csub.edu/~dbaron/TCH\\_ICPMSThirtyMinuteGuide.pdf](https://www.csub.edu/~dbaron/TCH_ICPMSThirtyMinuteGuide.pdf) (besucht am 19. 11. 2019).
- [164] Agilent Technologies, Inc., *Hardware Maintenance Manual, Agilent 7700 Series ICP-MS*, Printed in USA, 2011.
- [165] Todoli, J.-L.; Mermet, J., *Liquid Sample Introduction in ICP Spectrometry: A Practical Guide*; Elsevier Science: 2008.
- [166] Fairclough, C. Exploring the Venturi Effect. <https://www.comsol.de/blogs/exploring-the-venturi-effect/> (besucht am 19. 11. 2019).
- [167] Canals, A.; Wagner, J.; Browner, R. F.; Hernandis, V. Empirical model for estimating drop size distributions of aerosols generated by inductively coupled plasma nebulizers. *Spectrochimica Acta Part B: Atomic Spectroscopy* **1988**, *43*, 1321–1335.
- [168] Browner, R. F.; Boorn, A. W.; Smith, D. D. Aerosol transport model for atomic spectrometry. *Analytical Chemistry* **1982**, *54*, 1411–1419.
- [169] Wilbur, S. M. Factors Determining Sensitivity in ICP-MS. <http://www.spectroscopyonline.com/factors-determining-sensitivity-icp-ms> (besucht am 19. 11. 2019).

- [170] Cornelis, R.; Caruso, J.; Crews, H.; Heumann, K., *Handbook of Elemental Speciation: Techniques and Methodology*; John Wiley & Sons: 2003.
- [171] Leclercq, A. *et al.* Introduction of organic/hydro-organic matrices in inductively coupled plasma optical emission spectrometry and mass spectrometry: a tutorial review. Part I. Theoretical considerations. *Analytica Chimica Acta* **2015**, *885*, 33–56.
- [172] Wilbur, S.; Jones, C. Simple, reliable analysis of high matrix samples according to US EPA Method 6020A using the Agilent 7700x/7800 ICP-MS. <https://www.agilent.com/cs/library/applications/5990-5514EN.pdf> (besucht am 19.11.2019).
- [173] Agilent Technologies, Inc., ICP-MS High Matrix Introduction (HMI) Accessory. [https://www.agilent.com/en/products/icp-ms/icp-ms-accessories/icp-ms-high-matrix-introduction-\(hmi\)-accessory#literature](https://www.agilent.com/en/products/icp-ms/icp-ms-accessories/icp-ms-high-matrix-introduction-(hmi)-accessory#literature) (besucht am 19.11.2019).
- [174] McCurdy, E.; Proper, W. Improving ICP-MS analysis of samples containing high levels of total dissolved solids. *Spectroscopy* **2014**, *29*.
- [175] Michalke, B. The coupling of LC to ICP-MS in element speciation: I. General aspects. *TrAC Trends in Analytical Chemistry* **2002**, *21*, 142–153.
- [176] Brüchert, W.; Bettmer, J. On-Line Coupling of Gel Electrophoresis and Inductively Coupled Plasma-Sector Field-Mass Spectrometry for the Determination of dsDNA Fragments. *Analytical Chemistry* **2005**, *77*, 5072–5075.
- [177] Wuilloud, J. C.; Wuilloud, R. G.; Vonderheide, A. P.; Caruso, J. A. Gas chromatography/plasma spectrometry—an important analytical tool for elemental speciation studies. *Spectrochimica Acta Part B: Atomic Spectroscopy* **2004**, *59*, 755–792.
- [178] Grote, J.-P.; Zeradjanin, A. R.; Cherevko, S.; Mayrhofer, K. J. Coupling of a scanning flow cell with online electrochemical mass spectrometry for screening of reaction selectivity. *Review of Scientific Instruments* **2014**, *85*, 104101.
- [179] Michalke, B.; Schramel, P. Coupling of capillary electrophoresis with ICP-MS for speciation investigations. *Fresenius' Journal of Analytical Chemistry* **1997**, *357*, 594–599.
- [180] Vickers, G. H.; Ross, B. S.; Hieftje, G. M. Reduction of Mass-Dependent Interferences in Inductively Coupled Plasma-Mass Spectrometry by using Flow-Injection Analysis. *Applied Spectroscopy* **1989**, *43*, 1330–1333.
- [181] Günther, D.; Horn, I.; Hattendorf, B. Recent trends and developments in laser ablation-ICP-mass spectrometry. *Fresenius' Journal of Analytical Chemistry* **2000**, *368*, 4–14.
- [182] Russo, R. E. *et al.* Laser ablation in analytical chemistry—a review. *Talanta* **2002**, *57*, 425–451.
- [183] Pozebon, D.; Scheffler, G. L.; Dressler, V. L.; Nunes, M. A. Review of the applications of laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) to the analysis of biological samples. *Journal of Analytical Atomic Spectrometry* **2014**, *29*, 2204–2228.

- [184] Matsubara, A.; Kojima, H.; Itoga, T.; Kanehori, K. A ceramic plasma torch for determining silicon content in HF solutions by inductively coupled plasma atomic emission spectrometry. *Japanese Journal of Applied Physics* **1996**, *35*, 4541.
- [185] Ammann, A. A. Inductively coupled plasma mass spectrometry (ICP MS): a versatile tool. *Journal of Mass Spectrometry* **2007**, *42*, 419–427.
- [186] Bazzano, A. Isotopic analysis of environmental samples using inductively coupled plasma mass spectrometry., Diss., Ghent University, 2016.
- [187] Bazilio, A.; Weinrich, J. Inductively Coupled Plasma- Mass Spectrometry (ICP-MS). <http://www.ecs.umass.edu/eve/facilities/equipment/ICPMS/ICPMS%5C%20quick%5C%20guide.pdf> (besucht am 19.11.2019).
- [188] Batsala, M. *et al.* Inductively Coupled Plasma Mass Spectrometry (ICP-MS). *International Journal of Research in Pharmacy and Chemistry* **2012**, *2*, 671–680.
- [189] Sakata, K.; Kawabata, K. Reduction of fundamental polyatomic ions in inductively coupled plasma mass spectrometry. *Spectrochimica Acta Part B: Atomic Spectroscopy* **1994**, *49*, 1027–1038.
- [190] Koirtiyohann, S.; Jones, J. S.; Yates, D. A. Nomenclature system for the low-power argon inductively coupled plasma. *Analytical Chemistry* **1980**, *52*, 1965–1966.
- [191] Baralkiewicz, D.; Gramowska, H.; Hanc, A.; Krzyzaniak, I. A Comparison of ICP-OES and ICP-MS in the Determination of Elements in Lake Water. *Atomic Spectroscopy -Norwalk Connecticut-* **2007**, *28*, 164–170.
- [192] Thermo Fisher Scientific, Comparison of ICP-OES and ICP-MS for Trace Element Analysis. <https://www.thermofisher.com/de/de/home/industrial/environmental/environmental-learning-center/contaminant-analysis-information/metal-analysis/comparison-icp-oes-icp-ms-trace-element-analysis.html> (besucht am 19.11.2019).
- [193] Pearson, G. This is How you ICP-MS: Mastering the Art of Cone Performance – Agilent Technologies., <https://www.agilent.com/cs/library/troubleshootingguide/public/technicaloverview-ICP-MS-guide-5994-0860EN-agilent.pdf> (besucht am 19.11.2019).
- [194] Schmidt, K. H.; Gebel, A. ICP-MS Technik: Interface und Ionenoptik. <http://www.icp-ms.de/tech/inter.html> (besucht am 19.11.2019).
- [195] Košler, J.; Sylvester, P. J. Present Trends and the Future of Zircon in Geochronology: Laser Ablation ICPMS. *Reviews in Mineralogy & Geochemistry* **2003**, *53*, 243–275.
- [196] Hu, K.; Clemons, P. S.; Houk, R. Inductively coupled plasma mass spectrometry with an enlarged sampling orifice and offset ion lens. I. Ion trajectories and detector performance. *Journal of the American Society for Mass Spectrometry* **1993**, *4*, 16–27.
- [197] Becker, S., *Inorganic Mass Spectrometry: Principles and Applications*; John Wiley & Sons: Chichester, UK, 2007.

- [198] College, W. Chapter 5 - Mass Spectrometry. [https://www.whitman.edu/chemistry/edusolns\\_software/FAAS\\_ICP\\_2017/CH5\\_MS\\_Part1\\_2017.pdf](https://www.whitman.edu/chemistry/edusolns_software/FAAS_ICP_2017/CH5_MS_Part1_2017.pdf) (besucht am 19. 11. 2019).
- [199] Tanner, S. D.; Baranov, V. I.; Bandura, D. R. Reaction cells and collision cells for ICP-MS: a tutorial review. *Spectrochimica Acta Part B: Atomic Spectroscopy* **2002**, *57*, 1361–1452.
- [200] Agilent Technologies, Inc., Collision/Reaction Cells in ICP-MS. [https://www.agilent.com/cs/library/articlereprints/public/5990\\_5955EN%5C%20HR.pdf](https://www.agilent.com/cs/library/articlereprints/public/5990_5955EN%5C%20HR.pdf) (besucht am 19. 11. 2019).
- [201] Yamada, N. Kinetic energy discrimination in collision/reaction cell ICP-MS: Theoretical review of principles and limitations. *Spectrochimica Acta Part B: Atomic Spectroscopy* **2015**, *110*, 31–44.
- [202] McCurdy, E.; Woods, G.; Potter, D. Unmatched Removal of Spectral Interferences in ICP-MS Using the Agilent Octopole Reaction System with Helium Collision Mode., <https://pdfs.semanticscholar.org/4b67/506215c3def0ae4505853b4df114ef9fbef3.pdf> (besucht am 19. 11. 2019).
- [203] Agilent Technologies, Inc., First ICP-QQQ in US Delivers More Accurate Results for Complex Sample Analysis. *Agilent ICP-MS Journal* **2012**, *51*, 4.
- [204] Moldovan, M.; Krupp, E. M.; Holliday, A. E.; Donard, O. F. High resolution sector field ICP-MS and multicollector ICP-MS as tools for trace metal speciation in environmental studies: a review. *Journal of Analytical Atomic Spectrometry* **2004**, *19*, 815–822.
- [205] Thermo Fisher Scientific, ICP-MS Systems and Technologies. <https://www.thermofisher.com/de/de/home/industrial/spectroscopy-elemental-isotope-analysis/spectroscopy-elemental-isotope-analysis-learning-center/trace-elemental-analysis-tea-information/inductively-coupled-plasma-mass-spectrometry-icp-ms-information/icp-ms-systems-technologies.html> (besucht am 19. 11. 2019).
- [206] Hendriks, L.; Gundlach-Graham, A.; Hattendorf, B.; Günther, D. Characterization of a new ICP-TOFMS instrument with continuous and discrete introduction of solutions. *Journal of Analytical Atomic Spectrometry* **2017**, *32*, 548–561.
- [207] Restek, Electron Multipliers for Mass Spectrometry. <https://chromspec.com/pdf/e/rk123.pdf> (besucht am 19. 11. 2019).
- [208] Kitson, F. G.; Larsen, B. S.; McEwen, C. N., *Gas Chromatography and Mass Spectrometry: A Practical Guide*; Academic Press: 1996.
- [209] Medhe, S. Mass Spectrometry: Detectors Review. *Chemical and Biomolecular Engineering* **2018**, *3*, 51–58.
- [210] Vanhaecke, F.; Degryse, P., *Isotopic Analysis: Fundamentals and Applications Using ICP-MS*; John Wiley & Sons: Weinheim, 2012.
- [211] Wiedenbeck, M. *et al.* GGR Biennial critical review: analytical developments since 2010. *Geostandards and Geoanalytical Research* **2012**, *36*, 337–398.

- [212] Hamester, M.; Toms, A.; Chemnitzer, R. High Sensitivity ICP-MS: Overcome the problem of complex samples – Bruker Daltonics. [http://chemistry.unt.edu/~verbeck/LIMS/Manuals/820\\_tutorial.pdf](http://chemistry.unt.edu/~verbeck/LIMS/Manuals/820_tutorial.pdf) (besucht am 19. 11. 2019).
- [213] Bulska, E.; Wagner, B. Quantitative aspects of inductively coupled plasma mass spectrometry. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2016**, *374*, 20150369.
- [214] May, T. W.; Wiedmeyer, R. H. A table of polyatomic interferences in ICP-MS. *Atomic Spectroscopy -Norwalk Connecticut-* **1998**, *19*, 150–155.
- [215] Rüdell, D. H.; Kösters, D. J. Guidelines for Chemical Analysis: Determination of the Elemental Content of Environmental Samples using ICP-MS. [https://www.ime.fraunhofer.de/content/dam/ime/en/documents/AE/SOP\\_ICP-MS\\_en.pdf](https://www.ime.fraunhofer.de/content/dam/ime/en/documents/AE/SOP_ICP-MS_en.pdf) (besucht am 19. 11. 2019).
- [216] Evans, E. H.; Giglio, J. J. Interferences in inductively coupled plasma mass spectrometry. A review. *Journal of Analytical Atomic Spectrometry* **1993**, *8*, 1–18.
- [217] Vanhaecke, F.; Dams, R.; Vandecasteele, C. ‘Zone model’ as an explanation for signal behaviour and non-spectral interferences in inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* **1993**, *8*, 433–438.
- [218] Halicz, L.; Galy, A.; Belshaw, N.; O’Nions, R. High-precision measurement of calcium isotopes in carbonates and related materials by multiple collector inductively coupled plasma mass spectrometry (MC-ICP-MS). *Journal of Analytical Atomic Spectrometry* **1999**, *14*, 1835–1838.
- [219] Wiedenbeck, M. *et al.* GGR biennial critical review: analytical developments since 2012. *Geostandards and Geoanalytical Research* **2014**, *38*, 467–512.
- [220] Gremlich, H.-U., *Ullmann’s Encyclopedia of Industrial Chemistry*; John Wiley & Sons: 2000.
- [221] Gey, M. H., *Instrumentelle Analytik und Bioanalytik: Biosubstanzen, Trennmethode, Strukturanalytik, Applikationen*, 2. Aufl.; Springer-Lehrbuch; Springer-Verlag: Berlin Heidelberg, 2008.
- [222] Workman, J.; Weyer, L., *Practical Guide to Interpretive Near-Infrared Spectroscopy*; CRC Press: 2008.
- [223] Sun, D.-W., *Infrared Spectroscopy for Food Quality Analysis and Control*; Academic Press: 2009.
- [224] Reich, G. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. *Advanced Drug Delivery Reviews* **2005**, *57*, 1109–1143.
- [225] Siesler, H. W.; Ozaki, Y.; Kawata, S.; Heise, H. M., *Near-Infrared Spectroscopy: Principles, Instruments, Applications*; John Wiley & Sons: Weinheim, 2001.
- [226] Bienz, S.; Bigler, L.; Fox, T.; Meier, H., *Spektroskopische Methoden in der organischen Chemie*, 9. Aufl.; Thieme: Stuttgart New York, 2016.
- [227] Bruker, Die FT-IR-Spektroskopie — Prinzipieller Aufbau eines FT-IR-Spektrometers – Bruker OPUS 7.5 Handbuch.

- [228] Bruker Optics, The Bruker ROCKSOLID™ Interferometer.  
[http://www.orgchm.bas.bg/~irlab/Site\\_BG/Laboratory/rocksolid.pdf](http://www.orgchm.bas.bg/~irlab/Site_BG/Laboratory/rocksolid.pdf)  
(besucht am 19. 11. 2019).
- [229] Kok, André (Bruker Optics), From benchtop to process NIR control.  
Andr%C3%A9%20Kok%20Int.%20Key%20Account%20Manager%20Food,%20feed%  
20&%20Agri (besucht am 19. 11. 2019).
- [230] Thermo Fisher Scientific Inc., Smart NIR Ulbrichtkugel.  
<https://www.thermofisher.com/order/catalog/product/840-172400#/840-172400> (besucht am 19. 11. 2019).
- [231] Pike Technologies, Inc., Integrating spheres.,  
<https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=7&ved=2ahUKEwiw46G5rflAhWI6qQKHUR8AGgQFjAGegQIARAC&url=https%5C%3A%5C%2F%5C%2Fwww.norlab.com%5C%2Ffile%5C%2F3358%5C%2Fdownload%5C%3Ftoken%5C%3DMVg09x5u&usg=A0vVaw228TEy0e6lBxqqBhOX3IMO> (besucht am 19. 11. 2019).
- [232] Bruker Optics, TANGO. The next generation FT-NIR spectrometer.  
[http://bruker.poznan.pl/wp-content/uploads/2018/12/TANGO\\_Brochure\\_EN.pdf](http://bruker.poznan.pl/wp-content/uploads/2018/12/TANGO_Brochure_EN.pdf) (besucht am 19. 11. 2019).
- [233] Bruker, Messtechniken — Diffuse Reflexion – Bruker OPUS 7.5 Handbuch.
- [234] Bruker, Messtechniken — Diffuse Reflexion - die Integrationskugel – Bruker OPUS 7.5 Handbuch.
- [235] Rinnan, Å.; Van Den Berg, F.; Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry* **2009**, *28*, 1201–1222.
- [236] Rinnan, Å. *et al.*, *Infrared spectroscopy for food quality analysis and control*; Academic Press: 2008.
- [237] Geladi, P.; MacDougall, D.; Martens, H. Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy* **1985**, *39*, 491–500.
- [238] Savitzky, A.; Golay, M. J. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* **1964**, *36*, 1627–1639.
- [239] Windig, W.; Shaver, J.; Bro, R. Loopy MSC: a simple way to improve multiplicative scatter correction. *Applied Spectroscopy* **2008**, *62*, 1153–1159.
- [240] Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy* **1989**, *43*, 772–777.
- [241] Backhaus, K.; Erichson, B.; Plinke, W.; Weiber, R., *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*, 15., vollst. überarb. Aufl. 2018; Springer-Verlag: Berlin Heidelberg, 2018.
- [242] Day, W. H.; Edelsbrunner, H. Efficient algorithms for agglomerative hierarchical clustering methods. *Journal of Classification* **1984**, *1*, 7–24.
- [243] Zhao, Y.; Karypis, G.; Fayyad, U. Hierarchical clustering algorithms for document datasets. *Data Mining and Knowledge Discovery* **2005**, *10*, 141–168.

- [244] Stein, P.; Vollnhals, S. Grundlagen clusteranalytischer Verfahren - Institut für Soziologie - Universität Duisburg-Essen. [https://www.uni-due.de/imperia/md/content/soziologie/stein/skript\\_clusteranalyse\\_sose2011.pdf](https://www.uni-due.de/imperia/md/content/soziologie/stein/skript_clusteranalyse_sose2011.pdf) (besucht am 19. 11. 2019).
- [245] Universität Zürich, Clusteranalyse. [https://www.methodenberatung.uzh.ch/de/datenanalyse\\_spss/interdependenz/gruppierung/cluster.html](https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/interdependenz/gruppierung/cluster.html) (besucht am 19. 11. 2019).
- [246] Sneath, P. H. A. The application of computers to taxonomy. *Journal of General Microbiology* **1957**, *17*, 201–226.
- [247] Wilkinson, L.; Friendly, M. The History of the Cluster Heat Map. *The American Statistician* **2009**, *63*, 179–184.
- [248] Pryke, A.; Mostaghim, S.; Nazemi, A. Heatmap Visualization of Population Based Multi Objective Algorithms. *Evolutionary Multi-Criterion Optimization* **2007**, hrsg. von Obayashi, S. *et al.*, 361–375.
- [249] Jolliffe, I. T., *Principal Component Analysis*, 2. Aufl.; Springer-Verlag: New York, 2002.
- [250] Bro, R.; Smilde, A. K. Principal component analysis. *Analytical Methods* **2014**, *6*, 2812–2831.
- [251] Abdi, H.; Williams, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* **2010**, *2*, 433–459.
- [252] Wentura, D.; Pospeschill, M., *Multivariate Datenanalyse: Eine kompakte Einführung*; Springer: Wiesbaden, 2015.
- [253] Massart, D.; Heyden, Y. From tables to visuals: principal component analysis, Part 2. *LC GC Europe* **2005**, *17*, 586–591.
- [254] Lee, L. C.; Liong, C.-Y.; Jemain, A. A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* **2018**, *143*, 3526–3539.
- [255] Wold, H. Soft Modelling by Latent Variables: The Non-Linear Iterative Partial Least Squares (NIPALS) Approach. *Journal of Applied Probability* **1975**, *12*, 117–142.
- [256] Haenlein, M.; Kaplan, A. M. A Beginner's Guide to Partial Least Squares Analysis. *Understanding Statistics* **2004**, *3*, 283–297.
- [257] Brereton, R. G.; Lloyd, G. R. Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics* **2014**, *28*, 213–225.
- [258] Barker, M.; Rayens, W. Partial least squares for discrimination. *Journal of Chemometrics* **2003**, *17*, 166–173.
- [259] Höskuldsson, A. A combined theory for PCA and PLS. *Journal of Chemometrics* **1995**, *9*, 91–123.
- [260] Van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **2008**, *9*, 2579–2605.

- [261] Van Der Maaten, L. Fast optimization for t-SNE. In *Neural Information Processing Systems (NIPS) 2010 Workshop on Challenges in Data Visualization 2010*, 100.
- [262] Wattenberg, M.; Viégas, F.; Johnson, I. How to Use t-SNE Effectively. <http://distill.pub/2016/misread-tsne> (besucht am 19.11.2019).
- [263] Bioinformatics Group, Babraham Institute, The TSNE Plot. <https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/Help/3%5C%20Visualisation/3.2%5C%20Figures%5C%20and%5C%20Graphs/3.2.26%5C%20The%5C%20TSNE%5C%20Plot.html> (besucht am 19.11.2019).
- [264] Pizarro, C.; Esteban-Díez, I.; González-Sáiz, J.-M.; Forina, M. Use of near-infrared spectroscopy and feature selection techniques for predicting the caffeine content and roasting color in roasted coffees. *Journal of Agricultural and Food Chemistry* **2007**, *55*, 7477–7488.
- [265] Mladeníć, D. Feature selection for dimensionality reduction. *International Statistical and Optimization Perspectives Workshop: Subspace, Latent Structure and Feature Selection* **2005**, 84–102.
- [266] Kursa, M. B.; Rudnicki, W. R. Feature Selection with the Boruta Package. *Journal of Statistical Software* **2010**, *36*, 1–13.
- [267] Cortes, C.; Vapnik, V. Support-Vector Networks. *Machine Learning* **1995**, *20*, 273–297.
- [268] Weston, J.; Watkins, C. Support vector machines for multi-class pattern recognition. *Esann* **1999**, *99*, 219–224.
- [269] Duan, K.-B.; Keerthi, S. S. Which is the best multiclass SVM method? An empirical study. *International workshop on multiple classifier systems* **2005**, 278–285.
- [270] Hsu, C.-W.; Lin, C.-J. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks* **2002**, *13*, 415–425.
- [271] Fischer, J. (Universität Ulm), Support Vector Machines (SVM), Seminar Statistische Lerntheorie und ihre Anwendungen. [http://www.mathematik.uni-ulm.de/stochastik/lehre/ss07/seminar\\_sl/fischer.pdf](http://www.mathematik.uni-ulm.de/stochastik/lehre/ss07/seminar_sl/fischer.pdf) (besucht am 19.11.2019).
- [272] Grieser, G.; Technische Universität Darmstadt, Support-Vektor-Maschinen: Übersicht. <https://www.ke.tu-darmstadt.de/lehre/archiv/ws0405/mldm/svm.pdf> (besucht am 19.11.2019).
- [273] Heinert, M. (Technische Universität Braunschweig), Support Vector Machines Teil 1: Ein theoretischer Überblick. [https://www.tu-braunschweig.de/Medien-DB/igp/heinert\\_2010a.pdf](https://www.tu-braunschweig.de/Medien-DB/igp/heinert_2010a.pdf) (besucht am 19.11.2019).
- [274] Breiman, L. Random forests. *Machine Learning* **2001**, *45*, 5–32.
- [275] Biau, G. Analysis of a random forests model. *Journal of Machine Learning Research* **2012**, *13*, 1063–1095.
- [276] Robnik-Šikonja, M. Improving random forests. *European conference on machine learning - Machine Learning: ECML 2004* **2004**, 359–370.

- [277] Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Machine Learning* **2006**, *63*, 3–42.
- [278] Hartshorn, S., *Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners*; Kindle edition: 2016.
- [279] Kessler, W., *Multivariate Datenanalyse für die Pharma, Bio- und Prozessanalytik*; John Wiley & Sons: 2006.
- [280] Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence* **1995**, *14*, 1137–1145.
- [281] Fushiki, T. Estimation of prediction error by using  $K$ -fold cross-validation. *Statistics and Computing* **2011**, *21*, 137–146.
- [282] Kim, J.-H. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics & Data Analysis* **2009**, *53*, 3735–3745.
- [283] Rodriguez, J. D.; Perez, A.; Lozano, J. A. Sensitivity analysis of  $k$ -fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2010**, *32*, 569–575.
- [284] Reitermanová, Z. Data splitting. *WDS'10 Proceedings of Contributed Papers* **2010**, *10*, 31–36.
- [285] Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Statistics Surveys* **2010**, *4*, 40–79.
- [286] Krstajic, D.; Buturovic, L.; Leahy, D.; Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics* **2014**, *6*, 10.
- [287] Varma, S.; Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* **2006**, *7*, 91.
- [288] Scheffer, T. Error estimation and model selection., Diss., Technische Universität Berlin, 1999, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1.1.5846&rep=rep1&type=pdf> (besucht am 19. 11. 2019).
- [289] Raschka, S. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808* **2018**.
- [290] Richter, B. *et al.* Food authentication: Multi-elemental analysis of white asparagus for provenance discrimination. *Food Chemistry* **2019**, *286*, 475–482.
- [291] Richter, B. *et al.* Food monitoring: Screening of the geographical origin of white asparagus using FT-NIR and machine learning. *Food Control* **2019**, *104*, 318–325.
- [292] Cuprozin<sup>®</sup> progress.  
<https://www.agrarinfo.de/certis/deutschland/56.htm> (besucht am 19. 11. 2019).
- [293] Lebosol<sup>®</sup> Dünger GmbH, Blattdüngung im Spargel.  
<https://www.lebosol.de/kultur/spargel/> (besucht am 19. 11. 2019).

- [294] Düngemittel NPK 18-8-16., Yara Deutschland, <https://www.yara.de/pflanzenernaehrung/duengemittel/yaramila/yaramila-npk-18-8-16/> (besucht am 19. 11. 2019).
- [295] Deutsches Institut für Normung e.V., DIN 32645 - Chemische Analytik: Nachweis-, Erfassungs- und Bestimmungsgrenze. **1994**.
- [296] Food and Drug Administration, Guidance for Industry: Bioanalytical Method Validation.  
<http://academy.gmp-compliance.org/guidemgr/files/UCM070107.PDF> (besucht am 18. 11. 2019).
- [297] Food and Drug Administration, Analytical procedures and methods validation for drugs and biologics - Guidance for industry.  
<http://www.gmp-compliance.org/guidemgr/files/UCM386366.PDF> (besucht am 19. 11. 2019).
- [298] Wellmitz, J.; Gluschke, M., *Leitlinie zur Methodenvalidierung*; Umweltbundesamt: 2005.
- [299] European Commission: Joint Research Centre, BCR-670 Duck-Weed (Lemna Minor/aquatic plant); rare elements, Th, U.  
<https://crm.jrc.ec.europa.eu/p/40455/By-material-matrix/BCR-670-DUCK-WEED-Lemna-Minor-aquatic-plant-rare-elements-Th-U/BCR-670> (besucht am 19. 11. 2019).
- [300] Dienstleistung Lebensmittel Analytik GbR, Ahrensburg.  
Laborvergleichsuntersuchung 41/2015: Schwermetalle und Spurenelemente in pflanzlichem Nahrungsergänzungsmittel.
- [301] Braune, C. Nanoindentation an Phytolithen ausgewählter Nahrungspflanzenherbivorer Säugetiere und der potentielle Einfluss der Härte auf die Usur der Dentition., Diss., 2014, <https://ediss.sub.uni-hamburg.de/volltexte/2015/7219/pdf/Dissertation.pdf> (besucht am 19. 11. 2019).
- [302] Zaremba, C. M.; Stucky, G. D. Biosilicates and biomimetic silicate synthesis. *Current Opinion in Solid State and Materials Science* **1996**, *1*, 425–429.
- [303] European Commission: Joint Research Centre, BCR-667 Estuarine Sediment; rare elements, Th, U. <https://crm.jrc.ec.europa.eu/p/40454/40470/By-application-field/Environment/BCR-667-ESTUARINE-SEDIMENT-rare-elements-Th-U/BCR-667> (besucht am 19. 11. 2019).
- [304] Deutsches Institut für Normung e.V., DIN 38402-51:2017-05, Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung- Allgemeine Angaben (Gruppe A) - Teil 51: Kalibrierung von Analysenverfahren - Lineare Kalibrierfunktion (A 51). **2017**.
- [305] Koch, M. Die neue DIN 38402 - A51 zur linearen Kalibrierung von Analysenverfahren.  
[https://aqs.iswa.uni-stuttgart.de/pdf/JT16\\_Koch\\_A51.pdf](https://aqs.iswa.uni-stuttgart.de/pdf/JT16_Koch_A51.pdf) (besucht am 19. 11. 2019).

- [306] Eurachem's Proficiency Testing Working Group: How can proficiency testing help my laboratory? [https://www.eurachem.org/images/stories/leaflets/pt/labhelp/how\\_can\\_PT\\_help\\_my\\_lab\\_2013\\_EN.pdf](https://www.eurachem.org/images/stories/leaflets/pt/labhelp/how_can_PT_help_my_lab_2013_EN.pdf) (besucht am 20. 11. 2019).
- [307] Shaffer, J. P. Multiple hypothesis testing. *Annual Review of Psychology* **1995**, *46*, 561–584.
- [308] Feise, R. J. Do multiple outcome measures require p-value adjustment? *BMC Medical Research Methodology* **2002**, *2*, 8.
- [309] Streiner, D. L.; Norman, G. R. Correction for multiple testing: is there a resolution? *Chest* **2011**, *140*, 16–18.
- [310] World Information Service on Energy: Uranium project Uranium in Fertilizers. <https://www.wise-uranium.org/ufert.html> (besucht am 20. 11. 2019).
- [311] Al-Eshaikh, M. A.; Kadachi, A. N.; Sarfraz, M. Determination of uranium content in phosphate ores using different measurement techniques. *Journal of King Saud University - Engineering Sciences* **2016**, *28*, 41–46.
- [312] Yamazaki, I. M.; Geraldo, L. P. Uranium content in phosphate fertilizers commercially produced in Brazil. *Applied Radiation and Isotopes* **2003**, *59*, 133–136.
- [313] Maione, C. *et al.* Classification of geographic origin of rice by data mining and inductively coupled plasma mass spectrometry. *Computers and Electronics in Agriculture* **2016**, *121*, 101–107.
- [314] Li, C. *et al.* Determining the geographical origin of the medicinal plant *Marsdenia tenacissima* with multi-element analysis and data mining techniques. *Chemometrics and Intelligent Laboratory Systems* **2014**, *136*, 115–120.
- [315] Alcázar, Á. *et al.* Recognition of the geographical origin of beer based on support vector machines applied to chemical descriptors. *Food Control* **2012**, *23*, 258–262.
- [316] Mubaris, N. Support Vector Machines for Classification. <https://mubaris.com/posts/svm/> (besucht am 20. 11. 2019).
- [317] Bundesanstalt für Geowissenschaften und Rohstoffe, Bodenübersichtskarte 1:200.000. [https://fisbo.bgr.de/app/fisbobgr\\_produktauswahl/IMap/BUEK200\\_Arbeitsstand/map/m10000.html](https://fisbo.bgr.de/app/fisbobgr_produktauswahl/IMap/BUEK200_Arbeitsstand/map/m10000.html) (besucht am 20. 11. 2019).
- [318] Chesworth, W., *Encyclopedia of Soil Science*; Springer Science & Business Media: 2007.
- [319] Kabata-Pendias, A., *Trace Elements in Soils and Plants*, 3. Aufl.; CRC Press: 2000.
- [320] Wyttenbach, A.; Furrer, V.; Schleppei, P.; Tobler, L. Rare earth elements in soil and in soil-grown plants. *Plant and Soil* **1998**, *199*, 267–273.
- [321] Tyler, G. Rare earth elements in soil and plant systems - A review. *Plant and Soil* **2004**, *267*, 191–206.
- [322] Chen, S.; Zhu, Y.; Hu, Q. Soil to plant transfer of  $^{238}\text{U}$ ,  $^{226}\text{Ra}$  and  $^{232}\text{Th}$  on a uranium mining-impacted soil from southeastern China. *Journal of Environmental Radioactivity* **2005**, *82*, 223–236.

- [323] Nan, Z.; Li, J.; Zhang, J.; Cheng, G. Cadmium and zinc interactions and their transfer in soil-crop system under actual field conditions. *Science of the Total Environment* **2002**, *285*, 187–195.
- [324] Tobinski, D. (Universität Duisburg Essen), Deskriptive Statistik - Merkmalszusammenhänge. [https://www.uni-due.de/imperia/md/content/dokforum/skript\\_teil\\_\\_3](https://www.uni-due.de/imperia/md/content/dokforum/skript_teil__3) (besucht am 20. 11. 2019).
- [325] Schäfer, T. (Technische Universität Chemnitz), Methodenlehre I, WS 2013/14. <https://www.tu-chemnitz.de/hsw/psychologie/professuren/method/homepages/ts/methodenlehre/Meth13Web.pdf> (besucht am 20. 11. 2019).
- [326] Fakultät für Statistik, Technische Universität Dortmund, Tabelle der t-Verteilung. [https://www.statistik.tu-dortmund.de/fileadmin/user\\_upload/Lehrstuehle/Oekonometrie/Lehre/WiSoOekoSS17/tabelleV.pdf](https://www.statistik.tu-dortmund.de/fileadmin/user_upload/Lehrstuehle/Oekonometrie/Lehre/WiSoOekoSS17/tabelleV.pdf) (besucht am 20. 11. 2019).
- [327] Fischer-Klüver, D. G. Spargel braucht Aufmerksamkeit auch nach der Ernte. *Gemüse* **2016**, 47–49, [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&cad=rja&uact=8&ved=2ahUKEwjrk\\_ZitPjAhWwvQBHdcMDjcQFjAJegQIBhAC&url=https%5C%3A%5C%2F%5C%2Fwww.syngenta.de%5C%2Ffile%5C%2F35211%5C%2Fdownload%5C%3Ftoken%5C%3DH3VnHdpW&usg=AOvVaw2v\\_zaqVRB6F0FB2Pf0Mq1v](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&cad=rja&uact=8&ved=2ahUKEwjrk_ZitPjAhWwvQBHdcMDjcQFjAJegQIBhAC&url=https%5C%3A%5C%2F%5C%2Fwww.syngenta.de%5C%2Ffile%5C%2F35211%5C%2Fdownload%5C%3Ftoken%5C%3DH3VnHdpW&usg=AOvVaw2v_zaqVRB6F0FB2Pf0Mq1v) (besucht am 20. 11. 2019).
- [328] JOST GmbH, Informationen für unsere Kunden zum Spargelanbau. [https://www.jost-group.com/wissenswertes/downloads-pdf?eID=tx\\_rtgfiles\\_download&tx\\_rtgfiles\\_pi1%5C%5Buid%5C%5D=112&cHash=463ef7ccfe0242c8191efd4d1656fc11](https://www.jost-group.com/wissenswertes/downloads-pdf?eID=tx_rtgfiles_download&tx_rtgfiles_pi1%5C%5Buid%5C%5D=112&cHash=463ef7ccfe0242c8191efd4d1656fc11) (besucht am 20. 11. 2019).
- [329] Liang, L.; He, L.; Hu, N. Effects of B, Mn, Zn and Mo on yield and quality of asparagus. *Journal of Jilin Agricultural University* **2009**, *31*, 297–300.
- [330] Oddone, M. *et al.* Authentication and traceability study of hazelnuts from Piedmont, Italy. *Journal of Agricultural and Food Chemistry* **2009**, *57*, 3404–3408.
- [331] Tyler, G. Ionic charge, radius, and potential control root/soil concentration ratios of fifty cationic elements in the organic horizon of a beech (*Fagus sylvatica*) forest podzol. *Science of the Total Environment* **2004**, *329*, 231–239.
- [332] Evans, C. H., *Biochemistry of the Lanthanides*; Springer Science & Business Media: 2013.
- [333] Li, Z. *et al.* Direct measurement of lanthanum uptake and distribution in internodal cells of *Chara*. *Plant Science* **2008**, *174*, 496–501.
- [334] Shtangeeva, I.; Ayrault, S. Effects of Eu and Ca on yield and mineral nutrition of wheat (*Triticum aestivum*) seedlings. *Environmental and Experimental Botany* **2007**, *59*, 49–58.
- [335] Kindler & Glielch GbR Germany, Rutherford: Ionenradien. [http://www.uniterra.de/rutherford/tab\\_ior.htm](http://www.uniterra.de/rutherford/tab_ior.htm) (besucht am 20. 11. 2019).
- [336] Homer, R.; Mortimer, B. Europium II as a replacement for calcium II in concanavalin A. *FEBS Letters* **1978**, *87*, 69–72.

- [337] Schäfer, U. *et al.*, *Molekularbiologie der Zelle*, 5. Aufl.; Wiley-Blackwell: Weinheim, 2011.
- [338] Clarke, R. J.; Khalid, M. A. A., *Pumps, Channels and Transporters: Methods of Functional Analysis*; John Wiley & Sons: 2015.
- [339] Fox, T. C.; Guerinot, M. L. Molecular biology of cation transport in plants. *Annual Review of Plant Biology* **1998**, *49*, 669–696.
- [340] Silva, P. *et al.* Mechanism of active chloride secretion by shark rectal gland: role of Na-K-ATPase in chloride transport. *American Journal of Physiology-Renal Physiology* **1977**, *233*, F298–F306.
- [341] Carraretto, L. *et al.* Calcium flux across plant mitochondrial membranes: possible molecular players. *Frontiers in Plant Science* **2016**, *7*, 354.
- [342] Mills, R. F. *et al.* Functional expression of AtHMA4, a P<sub>1B</sub>-type ATPase of the Zn/Co/Cd/Pb subclass. *The Plant Journal* **2003**, *35*, 164–176.
- [343] Korshunova, Y. O. *et al.* The IRT1 protein from *Arabidopsis thaliana* is a metal transporter with a broad substrate range. *Plant Molecular Biology* **1999**, *40*, 37–44.
- [344] Thomine, S. *et al.* Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to *Nramp* genes. *Proceedings of the National Academy of Sciences* **2000**, *97*, 4991–4996.
- [345] Hall, J.; Williams, L. E. Transition metal transporters in plants. *Journal of Experimental Botany* **2003**, *54*, 2601–2613.
- [346] De la Fuente, J. M.; Ramírez-Rodríguez, V.; Cabrera-Ponce, J. L.; Herrera-Estrella, L. Aluminum tolerance in transgenic plants by alteration of citrate synthesis. *Science* **1997**, *276*, 1566–1568.
- [347] Sobariu, D. L. *et al.* Rhizobacteria and plant symbiosis in heavy metal uptake and its implications for soil bioremediation. *New Biotechnology* **2017**, *39*, 125–134.
- [348] Scheer, C. Spargelqualität im Produktions- und Aufbereitungsprozess., Diss., Technische Universität Berlin, 2002, [https://depositonce.tu-berlin.de/bitstream/11303/803/1/Dokument\\_13.pdf](https://depositonce.tu-berlin.de/bitstream/11303/803/1/Dokument_13.pdf) (besucht am 20.11.2019).
- [349] Sedlacek, K.-D.; Wrobel, N., *Die Lebenskraft: Wie Enzyme, Bewusstsein und quantenbiologische Effekte das Leben regulieren*; BoD – Books on Demand: 2019.
- [350] Vilgis, T.; Vierich, T., *Aroma Gemüse: Der Weg zum perfekten Geschmack*; Stiftung Warentest: 2017.
- [351] Alander, J. T. *et al.* A review of optical nondestructive visual and near-infrared methods for food quality and safety. *International Journal of Spectroscopy* **2013**, *2013*.
- [352] Socrates, G., *Infrared and Raman Characteristic Group Frequencies: Tables and Charts*; John Wiley & Sons: 2004.
- [353] Barbin, D. F. *et al.* Application of infrared spectral techniques on quality and compositional attributes of coffee: An overview. *Food Research International* **2014**, *61*, 23–32.

- 
- [354] Krell, M. M. Generalizing, Decoding, and Optimizing Support Vector Machine Classification., Diss., Universität Bremen,  
<https://elib.suub.uni-bremen.de/edocs/00104380-1.pdf> (besucht am 20. 11. 2019).
- [355] Krijthe, J.; van der Maaten, L. Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation.  
<https://cran.r-project.org/web/packages/Rtsne/index.html> (besucht am 20. 11. 2019).
- [356] Agilent Technologies, Inc., MassHunter G7200B B.01.01.  
<https://www.agilent.com/en/support/software-informatics/atomic-spectroscopy-data-systems/icp-ms-masshunter-software/7700-7500-b0101-update> (besucht am 20. 11. 2019).
- [357] Agilent Technologies, Inc., ICP-MS MassHunter Software.  
<https://www.agilent.com/en/products/software-informatics/atomic-spectroscopy-data-systems/icp-ms-masshunter-software> (besucht am 20. 11. 2019).
- [358] Bruker, OPUS Spektroskopiesoftware.  
<https://www.bruker.com/de/products/infrared-near-infrared-and-raman-spectroscopy/opus-spectroscopy-software.html> (besucht am 20. 11. 2019).
- [359] Microsoft Corporation, Excel: Tabellenkalkulation, XLS und XLSX.  
<https://products.office.com/de-de/excel> (besucht am 20. 11. 2019).
- [360] Camo Analytics, Unscrambler. <https://www.camo.com/unscrambler/>  
(besucht am 20. 11. 2019).
- [361] The R Foundation, R: A language and environment for statistical computing.  
<https://www.r-project.org/> (besucht am 20. 11. 2019).
- [362] Meyer, D. *et al.* e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien., Version 1.7-0, 2018,  
<https://CRAN.R-project.org/package=e1071> (besucht am 20. 11. 2019).
- [363] Warnes, G. R. *et al.* gplots: Various R Programming Tools for Plotting Data., Version 3.0.1, 2016, <https://CRAN.R-project.org/package=gplots> (besucht am 20. 11. 2019).
- [364] Kuhn, M. caret: Classification and Regression Training., Version 6.0-80, 2018,  
<https://cran.r-project.org/web/packages/caret/index.html> (besucht am 20. 11. 2019).
- [365] Stevens, A.; Ramirez-Lopez, L. prospectr: Miscellaneous functions for processing and sample selection of vis-NIR diffuse reflectance data., Version 0.1.3, 2014,  
<https://cran.r-project.org/web/packages/prospectr/index.html>  
(besucht am 20. 11. 2019).
- [366] Kursa, M. B.; Rudnicki, W. R. Boruta: Wrapper Algorithm for All Relevant Feature Selection., Version 6.0.0, 2018,  
<https://CRAN.R-project.org/package=Boruta> (besucht am 20. 11. 2018).



# 8 Anhang

## 8.1 Akquirierte Spargelproben

Tabelle 20: Übersicht der akquirierten deutschen Spargelproben.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
14-DE-BB-01	2014	BB	-	Beelitz	NA
14-DE-BB-02	2014	BB	-	Beelitz	NA
14-DE-BB-03	2014	BB	-	Beelitz	NA
14-DE-BWN-01	2014	BW	Nord	Forst	NA
14-DE-BWN-02	2014	BW	Nord	Wiesental	Gi
14-DE-BWN-03	2014	BW	Nord	Rastatt	NA
14-DE-BWS-01	2014	BW	Süd	Hartheim- Feldkirch	Gi
14-DE-BWS-02	2014	BW	Süd	Freiburg	Gi
14-DE-BWS-03	2014	BW	Süd	Schlatt	Gi
14-DE-BYAb-01	2014	Bayern	Abensberg	Offenstetten	Ba
14-DE-BYAb-02	2014	Bayern	Abensberg	Abensberg	Gr
14-DE-BYAb-03	2014	Bayern	Abensberg	Train	Av
14-DE-BYFr-01	2014	Bayern	Franken	Nürnberg	Ba
14-DE-BYFr-03	2014	Bayern	Franken	Prichsenstadt	Gr
14-DE-BYFr-04	2014	Bayern	Franken	Hausen	Ba
14-DE-BYFr-02	2014	Bayern	Franken	Nürnberg	Gi
14-DE-BYSr-01	2014	Bayern	Schrobenhausen	Inchenhofen	Gi
14-DE-BYSr-02	2014	Bayern	Schrobenhausen	Hohenwart	Vi
14-DE-BYSr-03	2014	Bayern	Schrobenhausen	Schrobenhausen	Gr
14-DE-HE-01	2014	Hessen	-	Griesheim	Gi
14-DE-HE-02	2014	Hessen	-	Weiterstadt	Gr
14-DE-HE-03	2014	Hessen	-	Darmstadt	He
14-DE-NI-01	2014	NI	-	Kirchdorf	Rps

Tabelle 20: Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
14-DE-NI-02	2014	NI	-	Fuhrberg	Ba
14-DE-NI-03	2014	NI	-	Hoyerhagen	Ba
14-DE-NWBo-01	2014	NRW	Bornheim	Bornheim	Ba
14-DE-NWBo-02	2014	NRW	Bornheim	Zülpich	Gr
14-DE-NWBo-03	2014	NRW	Bornheim	Dorsten	Gi
14-DE-NW-01	2014	NRW	Westfalen	Sassenberg	Ba
14-DE-NW-02	2014	NRW	Westfalen	Sassenberg	Gi
14-DE-NWWa-01	2014	NRW	Walbeck	Walbeck	Ba
14-DE-NWWa-02	2014	NRW	Walbeck	Goch	Gi
14-DE-NWWa-03	2014	NRW	Walbeck	Walbeck	Ba
14-DE-NW-03	2014	NRW	Westfalen	Rahden	Ba
14-DE-RP-01	2014	RP	-	Neustadt an der Weinstraße	Ba
14-DE-RP-02	2014	RP	-	Weisenheim/Sand	Ba
14-DE-RP-03	2014	RP	-	Kandel	NA
14-DE-SH-01	2014	SH	-	Freienwill	Ba
14-DE-SH-02	2014	SH	-	Buchholz am Ratzeburger See	Ba
14-DE-SH-03	2014	SH	-	Wiemersdorf	Ba
14-DE-SH-04	2014	SH	-	Wiemersdorf	Gi
15-DE-BB-01	2015	BB	-	Beelitz	Gi
15-DE-BB-02	2015	BB	-	Beelitz	Ba
15-DE-BB-03	2015	BB	-	Beelitz	Gi
15-DE-BB-04	2015	BB	-	Beelitz	Gr
15-DE-BB-05	2015	BB	-	Beelitz	Gi
15-DE-BWS-01	2015	BW	Süd	Freiburg	Gi
15-DE-BWS-02	2015	BW	Süd	Hartheim- Feldkirch	Cu
15-DE-BWS-03	2015	BW	Süd	Hartheim- Feldkirch	Ba
15-DE-BWS-04	2015	BW	Süd	Hartheim- Feldkirch	Gi
15-DE-BWS-05	2015	BW	Süd	Schlatt	Gi
15-DE-BWS-06	2015	BW	Süd	Schlatt	Ba
15-DE-BWN-01	2015	BW	Nord	Rastatt	Rvl

**Tabelle 20:** Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
15-DE-BWN-02	2015	BW	Nord	Rastatt	Rfl
15-DE-BWN-03	2015	BW	Nord	Rastatt	Gr
15-DE-BWN-04	2015	BW	Nord	Rastatt	Gi
15-DE-BWN-05	2015	BW	Nord	Rastatt	Rps
15-DE-BWN-06	2015	BW	Nord	Rastatt	Rmr
15-DE-BWN-07	2015	BW	Nord	Rastatt	Ba
15-DE-BWN-08	2015	BW	Nord	Rastatt	Av
15-DE-BWN-09	2015	BW	Nord	Rastatt	Cu
15-DE-BWN-10	2015	BW	Nord	Rastatt	Rmd
15-DE-BWN-11	2015	BW	Nord	Wiesental	Gi
15-DE-BYAb-01	2015	Bayern	Abensberg	Abensberg	Gr
15-DE-BYAb-02	2015	Bayern	Abensberg	Offenstetten	Ba
15-DE-BYAb-03	2015	Bayern	Abensberg	Train	Gi
15-DE-BYFr-01	2015	Bayern	Franken	Hausen	Rmd
15-DE-BYFr-02	2015	Bayern	Franken	Nürnberg	Gi
15-DE-BYFr-03	2015	Bayern	Franken	Nürnberg	Ba
15-DE-BYFr-04	2015	Bayern	Franken	Prichsenstadt	Gr
15-DE-BYFr-05	2015	Bayern	Franken	Prichsenstadt	Ba
15-DE-BYSr-01	2015	Bayern	Schrobenhausen	Hohenwart	Gi
15-DE-BYSr-02	2015	Bayern	Schrobenhausen	Inchenhofen	Ba
15-DE-BYSr-03	2015	Bayern	Schrobenhausen	Inchenhofen	Gi
15-DE-BYSr-04	2015	Bayern	Schrobenhausen	Inchenhofen	Cu
15-DE-BYSr-05	2015	Bayern	Schrobenhausen	Schrobenhausen	Gr
15-DE-HE-01	2015	Hessen	-	Darmstadt	Ba
15-DE-HE-02	2015	Hessen	-	Darmstadt	Gr
15-DE-HE-03	2015	Hessen	-	Griesheim	Ba
15-DE-HE-04	2015	Hessen	-	Griesheim	Gr
15-DE-HE-05	2015	Hessen	-	Griesheim	Ba
15-DE-HE-06	2015	Hessen	-	Griesheim	Gi
15-DE-NI-01	2015	NI	-	Hoyerhagen	Gi
15-DE-NW-01	2015	NRW	Westfalen	Dorsten	Gi
15-DE-NW-02	2015	NRW	Westfalen	Dorsten	Ba
15-DE-NW-03	2015	NRW	Westfalen	Rahden	Cu
15-DE-NW-04	2015	NRW	Westfalen	Rahden	Gi
15-DE-NW-05	2015	NRW	Westfalen	Rahden	Ba

**Tabelle 20:** Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
15-DE-NW-06	2015	NRW	Westfalen	Sassenberg	Cu
15-DE-NW-07	2015	NRW	Westfalen	Sassenberg	Ba
15-DE-NW-08	2015	NRW	Westfalen	Sassenberg	Gi
15-DE-NW-09	2015	NRW	Westfalen	Sassenberg	Gi
15-DE-NWBo-01	2015	NRW	Bornheim	Bornheim	Ba
15-DE-NWBo-02	2015	NRW	Bornheim	Bornheim	Gr
15-DE-NWBo-03	2015	NRW	Bornheim	Zülpich	Gr
15-DE-NWBo-04	2015	NRW	Bornheim	Zülpich	Av
15-DE-NWWa-01	2015	NRW	Walbeck	Goch	Gi
15-DE-NWWa-02	2015	NRW	Walbeck	Walbeck	Gi
15-DE-NWWa-03	2015	NRW	Walbeck	Walbeck	Ba
15-DE-NWWa-04	2015	NRW	Walbeck	Walbeck	Gi
15-DE-RP-01	2015	RP	-	Kandel	Gi
15-DE-RP-02	2015	RP	-	Neustadt an der Weinstraße	Ba
15-DE-RP-03	2015	RP	-	Weisenheim/Sand	Ba
15-DE-RP-04	2015	RP	-	Weisenheim/Sand	Gi
15-DE-RP-05	2015	RP	-	Weisenheim/Sand	Ba
15-DE-SH-01	2015	SH	-	Buchholz am Ratzeburger See	Gi
15-DE-SH-02	2015	SH	-	Freienwill	Gi
15-DE-SH-03	2015	SH	-	Wiemersdorf	Rmr
15-DE-SH-04	2015	SH	-	Wiemersdorf	Gi
15-DE-SH-05	2015	SH	-	Wiemersdorf	Gi
15-DE-SH-06	2015	SH	-	Wiemersdorf	Gi
15-DE-SH-07	2015	SH	-	Wiemersdorf	Gi
15-DE-SH-08	2015	SH	-	Wiemersdorf	Rmr
15-DE-SH-09	2015	SH	-	Wiemersdorf	Gi
15-DE-SH-10	2015	SH	-	Wiemersdorf	Ba
15-DE-SH-11	2015	SH	-	Wiemersdorf	Gi
16-DE-BB-01	2016	BB	-	Beelitz	Gr
16-DE-BB-02	2016	BB	-	Beelitz	Gr
16-DE-BB-03	2016	BB	-	Beelitz	NA
16-DE-BB-04	2016	BB	-	Beelitz	NA
16-DE-BB-05	2016	BB	-	Beelitz	NA

**Tabelle 20:** Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
16-DE-BB-06	2016	BB	-	Beelitz	Cu
16-DE-BB-07	2016	BB	-	Beelitz	NA
16-DE-BB-08	2016	BB	-	Beelitz	Gi
16-DE-BB-09	2016	BB	-	Beelitz	NA
16-DE-BWN-14	2016	BW	Nord	Rastatt	Ba
16-DE-BWN-01	2016	BW	Nord	Forst	Ba
16-DE-BWN-02	2016	BW	Nord	Rastatt	Rps
16-DE-BWN-03	2016	BW	Nord	Rastatt	Rmd
16-DE-BWN-04	2016	BW	Nord	Rastatt	Gr
16-DE-BWN-05	2016	BW	Nord	Rastatt	Rmr
16-DE-BWN-06	2016	BW	Nord	Rastatt	Rfl
16-DE-BWN-07	2016	BW	Nord	Rastatt	Rvl
16-DE-BWN-08	2016	BW	Nord	Rastatt	Cu
16-DE-BWN-09	2016	BW	Nord	Rastatt	Gi
16-DE-BWN-10	2016	BW	Nord	Rastatt	Av
16-DE-BWN-11	2016	BW	Nord	Wiesental	Ba
16-DE-BWN-12	2016	BW	Nord	Wiesental	Gi
16-DE-BWN-13	2016	BW	Nord	Wiesental	Gr
16-DE-BWS-01	2016	BW	Süd	Freiburg	Ba
16-DE-BWS-02	2016	BW	Süd	Hartheim- Feldkirch	Ba
16-DE-BWS-03	2016	BW	Süd	Hartheim- Feldkirch	Cu
16-DE-BWS-04	2016	BW	Süd	Hartheim- Feldkirch	Gi
16-DE-BWS-05	2016	BW	Süd	Schlatt	Ba
16-DE-BYAb-01	2016	Bayern	Abensberg	Abensberg	Ba
16-DE-BYAb-02	2016	Bayern	Abensberg	Abensberg	Gr
16-DE-BYAb-03	2016	Bayern	Abensberg	Offenstetten	Ba
16-DE-BYAb-04	2016	Bayern	Abensberg	Offenstetten	Gi
16-DE-BYAb-05	2016	Bayern	Abensberg	Train	Av
16-DE-BYAb-06	2016	Bayern	Abensberg	Train	Gi
16-DE-BYFr-01	2016	Bayern	Franken	Hausen	Ba
16-DE-BYFr-02	2016	Bayern	Franken	Nürnberg	Gi
16-DE-BYFr-03	2016	Bayern	Franken	Nürnberg	Ba

**Tabelle 20:** Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
16-DE-BYFr-04	2016	Bayern	Franken	Prichsenstadt	Ba
16-DE-BYFr-05	2016	Bayern	Franken	Prichsenstadt	Cu
16-DE-BYFr-06	2016	Bayern	Franken	Prichsenstadt	Gr
16-DE-BYFr-07	2016	Bayern	Franken	Prichsenstadt	Gi
16-DE-BYSr-01	2016	Bayern	Schrobenhausen	Inchenhofen	Ba
16-DE-BYSr-02	2016	Bayern	Schrobenhausen	Inchenhofen	Gi
16-DE-BYSr-03	2016	Bayern	Schrobenhausen	Inchenhofen	Cu
16-DE-BYSr-04	2016	Bayern	Schrobenhausen	Schrobenhausen	Ba
16-DE-BYSr-05	2016	Bayern	Schrobenhausen	Schrobenhausen	Gi
16-DE-BYSr-06	2016	Bayern	Schrobenhausen	Schrobenhausen	Gi
16-DE-BYSr-07	2016	Bayern	Schrobenhausen	Schrobenhausen	Cu
16-DE-BYSr-08	2016	Bayern	Schrobenhausen	Schrobenhausen	Gr
16-DE-BYSr-09	2016	Bayern	Schrobenhausen	Schrobenhausen	Ba
16-DE-HE-01	2016	Hessen	-	Darmstadt	He
16-DE-HE-02	2016	Hessen	-	Darmstadt	Ba
16-DE-HE-03	2016	Hessen	-	Griesheim	Gi
16-DE-HE-04	2016	Hessen	-	Griesheim	Ba
16-DE-HE-05	2016	Hessen	-	Griesheim	Gr
16-DE-HE-06	2016	Hessen	-	Weiterstadt	Ba
16-DE-HE-07	2016	Hessen	-	Weiterstadt	Gr
16-DE-HE-08	2016	Hessen	-	Weiterstadt	Rfl
16-DE-HE-09	2016	Hessen	-	Weiterstadt	He
16-DE-NI-01	2016	NI	-	Fuhrberg	Gi
16-DE-NI-02	2016	NI	-	Fuhrberg	Ba
16-DE-NI-03	2016	NI	-	Hoyerhagen	Ba
16-DE-NI-04	2016	NI	-	Hoyerhagen	Gi
16-DE-NI-05	2016	NI	-	Kirchdorf	Ba
16-DE-NI-06	2016	NI	-	Kirchdorf	Gi
16-DE-NI-07	2016	NI	-	Kirchdorf	Ba
16-DE-NI-08	2016	NI	-	Kirchdorf	Cu
16-DE-NI-09	2016	NI	-	Kirchdorf	Gr
16-DE-NI-10	2016	NI	-	Kirchdorf	Gi
16-DE-NI-11	2016	NI	-	Kirchdorf	Cu
16-DE-NW-01	2016	NRW	Westfalen	Dorsten	Ba
16-DE-NW-02	2016	NRW	Westfalen	Rahden	Gr

**Tabelle 20:** Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
16-DE-NW-03	2016	NRW	Westfalen	Rahden	Ba
16-DE-NW-04	2016	NRW	Westfalen	Rahden	Cu
16-DE-NW-05	2016	NRW	Westfalen	Rahden	Gi
16-DE-NW-06	2016	NRW	Westfalen	Sassenberg	Gi
16-DE-NW-07	2016	NRW	Westfalen	Sassenberg	Ba
16-DE-NW-08	2016	NRW	Westfalen	Sassenberg	Gr
16-DE-NW-09	2016	NRW	Westfalen	Sassenberg	Cu
16-DE-NWBo-01	2016	NRW	Bornheim	Bornheim	Gr
16-DE-NWBo-02	2016	NRW	Bornheim	Bornheim	Gr
16-DE-NWBo-03	2016	NRW	Bornheim	Bornheim	Ba
16-DE-NWBo-04	2016	NRW	Bornheim	Bornheim	Ba
16-DE-NWBo-05	2016	NRW	Bornheim	Bornheim	Gr
16-DE-NWWa-01	2016	NRW	Walbeck	Goch	He
16-DE-NWWa-02	2016	NRW	Walbeck	Goch	Rmr
16-DE-NWWa-03	2016	NRW	Walbeck	Walbeck	Gr
16-DE-NWWa-04	2016	NRW	Walbeck	Walbeck	Ba
16-DE-NWWa-05	2016	NRW	Walbeck	Walbeck	Gi
16-DE-NWWa-06	2016	NRW	Walbeck	Walbeck	Ba
16-DE-NWWa-07	2016	NRW	Walbeck	Walbeck	Cu
16-DE-NWWa-08	2016	NRW	Walbeck	Walbeck	Gi
16-DE-NWWa-09	2016	NRW	Walbeck	Walbeck	Cu
16-DE-RP-01	2016	RP	-	Kandel	Ba
16-DE-RP-02	2016	RP	-	Kandel	Cu
16-DE-RP-03	2016	RP	-	Kandel	Gr
16-DE-RP-04	2016	RP	-	Kandel	Gi
16-DE-RP-05	2016	RP	-	Neustadt an der Weinstraße	Ba
16-DE-RP-06	2016	RP	-	Weisenheim/Sand	NA
16-DE-SH-01	2016	SH	-	Buchholz am Ratzeburger See	Ba
16-DE-SH-02	2016	SH	-	Buchholz am Ratzeburger See	Gi
16-DE-SH-03	2016	SH	-	Buchholz am Ratzeburger See	Gr
16-DE-SH-04	2016	SH	-	Freienwill	Gi

**Tabelle 20:** Übersicht der akquirierten deutschen Spargelproben - Fortsetzung.

Proben-ID	Jahr	Region	Unterregion	Ort	Sorte
16-DE-SH-05	2016	SH	-	Freienwill	Ba
16-DE-SH-06	2016	SH	-	Wiemersdorf	Ba
16-DE-SH-07	2016	SH	-	Wiemersdorf	Gi

**Tabelle 21:** Übersicht der akquirierten niederländischen Spargelproben.

Proben-ID	Jahr	Ort	Sorte
16-NL-01	2016	Veghel	Cu
16-NL-02	2016	Sint-Michielsgestel	NA
16-NL-03	2016	Prinsenbeek	Cu
16-NL-04	2016	Ossendrecht	He
16-NL-05	2016	Budel	Gi
16-NL-06	2016	Asten	NA
16-NL-07	2016	Sevenum	NA
16-NL-08	2016	Terheijden	Ba
16-NL-09	2016	Berlicum	NA
16-NL-10	2016	Udenhout	Ba
17-NL-01	2017	Udenhout	Ba
17-NL-02	2017	Sint-Michielsgestel	He
17-NL-03	2017	Veghel	Ba
17-NL-04	2017	Breda	NA
17-NL-05	2017	Diessen	NA
17-NL-06	2017	Terheijden	Cu
17-NL-07	2017	Oirschot	NA
17-NL-08	2017	Alphen	Ba
17-NL-09	2017	Prinsenbeek	Cu
17-NL-10	2017	Berlicum	Cu

**Tabelle 22:** Übersicht der akquirierten ausländischen Spargelproben (ohne NL).

<b>Proben-ID</b>	<b>Jahr</b>	<b>Herkunftsland</b>
14-CHN-01	2014	China
14-CHN-02	2014	China
14-PE-01	2014	Peru
15-CHN-01	2015	China
15-CHN-02	2015	China
15-CHN-03	2015	China
15-ES-01	2015	Spanien
15-ES-02	2015	Spanien
15-ES-03	2015	Spanien
15-ES-04	2015	Spanien
15-GR-01	2015	Griechenland
15-GR-02	2015	Griechenland
15-PE-01	2015	Peru
15-PE-02	2015	Peru
15-PE-03	2015	Peru
15-PE-04	2015	Peru
15-PL-01	2015	Polen
16-ES-01	2016	Spanien
16-ES-02	2016	Spanien
16-ES-03	2016	Spanien
16-GR-01	2016	Griechenland
16-GR-02	2016	Griechenland
16-GR-03	2016	Griechenland
16-GR-04	2016	Griechenland
16-GR-05	2016	Griechenland
16-GR-06	2016	Griechenland
16-GR-07	2016	Griechenland
16-GR-08	2016	Griechenland
16-GR-09	2016	Griechenland
16-GR-10	2016	Griechenland
16-GR-11	2016	Griechenland
16-GR-12	2016	Griechenland
16-GR-13	2016	Griechenland
16-GR-14	2016	Griechenland
16-GR-15	2016	Griechenland

**Tabelle 22:** Übersicht der akquirierten ausländischen Spargelproben (ohne NL) - Fortsetzung.

Proben-ID	Jahr	Herkunftsland
16-GR-16	2016	Griechenland
16-PE-01	2016	Peru
16-PE-02	2016	Peru
16-PE-03	2016	Peru
16-PE-04	2016	Peru
16-PE-05	2016	Peru
16-PE-06	2016	Peru
16-PE-07	2016	Peru
16-PE-08	2016	Peru
16-PE-09	2016	Peru
16-PE-10	2016	Peru
16-PL-01	2016	Polen
16-PL-02	2016	Polen
16-PL-03	2016	Polen
16-PL-04	2016	Polen
16-PL-05	2016	Polen
16-PL-06	2016	Polen
16-PL-07	2016	Polen
16-PL-08	2016	Polen
16-PL-09	2016	Polen
16-PL-10	2016	Polen
17-CHN-01	2017	China
17-CHN-02	2017	China
17-ES-01	2017	Spanien
17-GR-01	2017	Griechenland
17-GR-02	2017	Griechenland
17-GR-03	2017	Griechenland
17-GR-04	2017	Griechenland
17-GR-05	2017	Griechenland
17-GR-06	2017	Griechenland
17-GR-07	2017	Griechenland
17-PE-01	2017	Peru
17-PE-02	2017	Peru
17-PE-03	2017	Peru

**Tabelle 22:** Übersicht der akquirierten ausländischen Spargelproben (ohne NL) - Fortsetzung.

<b>Proben-ID</b>	<b>Jahr</b>	<b>Herkunftsland</b>
17-PE-04	2017	Peru
17-PE-05	2017	Peru
17-PE-06	2017	Peru
17-PE-07	2017	Peru
17-PE-08	2017	Peru
17-PL-01	2017	Polen
17-PL-02	2017	Polen
17-PL-03	2017	Polen
17-PL-04	2017	Polen
17-PL-05	2017	Polen
17-PL-06	2017	Polen
17-PL-07	2017	Polen
17-PL-08	2017	Polen
17-PL-09	2017	Polen
17-PL-10	2017	Polen

## 8.2 Akquirierte Bodenproben

Tabelle 23: Übersicht der akquirierten Bodenproben.

Proben-ID	Region	Unterregion	Ort	Bodentiefe
14-DE-BYAb-01-030	Bayern	Abensberg	Offenstetten	0 - 30 cm
14-DE-BYAb-01-3060				30 - 60 cm
14-DE-BYAb-01-6090				60 - 90 cm
14-DE-BYAb-02-030	Bayern	Abensberg	Abensberg	0 - 30 cm
14-DE-BYAb-02-3060				30 - 60 cm
14-DE-BYAb-02-6090				60 - 90 cm
14-DE-BYAb-03-030	Bayern	Abensberg	Train	0 - 30 cm
14-DE-BYAb-03-3060				30 - 60 cm
14-DE-BYAb-03-6090				60 - 90 cm
14-DE-BYSr-01-030	Bayern	BYSr	Inchenhofen	0 - 30 cm
14-DE-BYSr-01-3060				30 - 60 cm
14-DE-BYSr-01-6090				60 - 90 cm
14-DE-BYSr-02-030	Bayern	Schrobenhausen	Hohenwart	0 - 30 cm
14-DE-BYSr-02-3060				30 - 60 cm
14-DE-BYSr-02-6090				60 - 90 cm
14-DE-BYSr-03-030	Bayern	Schrobenhausen	Schrobenhausen	0 - 30 cm
14-DE-BYSr-03-3060				30 - 60 cm
14-DE-BYSr-03-6090				60 - 90 cm
14-DE-BYFr-01-030	Bayern	Schrobenhausen	Hausen	0 - 30 cm
14-DE-BYFr-01-3060				30 - 60 cm
14-DE-BYFr-01-6090				60 - 90 cm
14-DE-BYFr-02-030	Bayern	Franken	Prichsenstadt	0 - 30 cm
14-DE-BYFr-02-3060				30 - 60 cm
14-DE-BYFr-02-6090				60 - 90 cm
14-DE-BYFr-03-030	Bayern	Franken	Nürnberg	0 - 30 cm
14-DE-BYFr-03-3060				30 - 60 cm
14-DE-BYFr-03-6090				60 - 90 cm
14-DE-NWWa-01-030	NRW	Walbeck	Walbeck	0 - 30 cm
14-DE-NWWa-01-3060				30 - 60 cm
14-DE-NWWa-01-6090				60 - 90 cm

**Tabelle 23:** Übersicht der akquirierten Bodenproben - Fortsetzung.

Proben-ID	Region	Unterregion	Ort	Bodentiefe
14-DE-NWWa-02-030	NRW	Walbeck	Goch	0 - 30 cm
14-DE-NWWa-02-3060				30 - 60 cm
14-DE-NWWa-02-6090				60 - 90 cm
14-DE-NWWa-03-030	NRW	Walbeck	Walbeck	0 - 30 cm
14-DE-NWWa-03-3060				30 - 60 cm
14-DE-NWWa-03-6090				60 - 90 cm
14-DE-NWBo-01-030	NRW	Bornheim	Bornheim	0 - 30 cm
14-DE-NWBo-01-3060				30 - 60 cm
14-DE-NWBo-01-6090				60 - 90 cm
14-DE-NWBo-02-030	NRW	Bornheim	Zülpich	0 - 30 cm
14-DE-NWBo-02-3060				30 - 60 cm
14-DE-NWBo-02-6090				60 - 90 cm
14-DE-NWBo-03-030	NRW	Bornheim	Dorsten	0 - 30 cm
14-DE-NWBo-03-3060				30 - 60 cm
14-DE-NWBo-03-6090				60 - 90 cm
14-DE-NW-01-030	NRW	Westfalen	Sassenberg	0 - 30 cm
14-DE-NW-01-3060				30 - 60 cm
14-DE-NW-01-6090				60 - 90 cm
14-DE-NW-02-030	NRW	Westfalen	Sassenberg	0 - 30 cm
14-DE-NW-02-3060				30 - 60 cm
14-DE-NW-02-6090				60 - 90 cm
14-DE-NW-03-030	NRW	Westfalen	Rahden	0 - 30 cm
14-DE-NW-03-3060				30 - 60 cm
14-DE-NW-03-6090				60 - 90 cm
14-DE-SH-01-030	SH	-	Freienwill	0 - 30 cm
14-DE-SH-01-3060				30 - 60 cm
14-DE-SH-01-6090				60 - 90 cm
14-DE-SH-02-030	SH	-	Buchholz am Ratzeburger See	0 - 30 cm
14-DE-SH-02-3060				30 - 60 cm
14-DE-SH-02-6090				60 - 90 cm
14-DE-SH-03-030	SH	-	Wiemersdorf	0 - 30 cm
14-DE-SH-03-3060				30 - 60 cm
14-DE-SH-03-6090				60 - 90 cm

**Tabelle 23:** Übersicht der akquirierten Bodenproben - Fortsetzung.

Proben-ID	Region	Unterregion	Ort	Bodentiefe
14-DE-SH-04-030	SH	-	Wiemersdorf	0 - 30 cm
14-DE-SH-04-3060				30 - 60 cm
14-DE-SH-04-6090				60 - 90 cm
14-DE-NI-01-030	NI	-	Kirchdorf	0 - 30 cm
14-DE-NI-01-3060				30 - 60 cm
14-DE-NI-01-6090				60 - 90 cm
14-DE-NI-02-030	NI	-	Fuhrberg	0 - 30 cm
14-DE-NI-02-3060				30 - 60 cm
14-DE-NI-02-6090				60 - 90 cm
14-DE-NI-03-030	NI	-	Hoyerhagen	0 - 30 cm
14-DE-NI-03-3060				30 - 60 cm
14-DE-NI-03-6090				60 - 90 cm
14-DE-BWS-01-030	BW	Süd	Hartheim- Feldkirch	0 - 30 cm
14-DE-BWS-01-3060				30 - 60 cm
14-DE-BWS-01-6090				60 - 90 cm
14-DE-BWS-02-030	BW	Süd	Freiburg	0 - 30 cm
14-DE-BWS-02-3060				30 - 60 cm
14-DE-BWS-02-6090				60 - 90 cm
14-DE-BWS-03-030	BW	Süd	Schlatt	0 - 30 cm
14-DE-BWS-03-3060				30 - 60 cm
14-DE-BWS-03-6090				60 - 90 cm
14-DE-BWN-01-030	BW	Nord	Forst	0 - 30 cm
14-DE-BWN-01-3060				30 - 60 cm
14-DE-BWN-01-6090				60 - 90 cm
14-DE-BWN-02-030	BW	Nord	Wiesental	0 - 30 cm
14-DE-BWN-02-3060				30 - 60 cm
14-DE-BWN-02-6090				60 - 90 cm
14-DE-RP-01-030	RP	-	Neustadt an der Weinstraße	0 - 30 cm
14-DE-RP-01-3060				30 - 60 cm
14-DE-RP-01-6090				60 - 90 cm
14-DE-RP-02-030	RP	-	Weisenheim/Sand	0 - 30 cm
14-DE-RP-02-3060				30 - 60 cm
14-DE-RP-02-6090				60 - 90 cm

**Tabelle 23:** Übersicht der akquirierten Bodenproben - Fortsetzung.

Proben-ID	Region	Unterregion	Ort	Bodentiefe
14-DE-RP-03-030	RP	-	Kandel	0 - 30 cm
14-DE-RP-03-3060				30 - 60 cm
14-DE-RP-03-6090				60 - 90 cm
14-DE-HE-01-030	HE	-	Griesheim	0 - 30 cm
14-DE-HE-01-3060				30 - 60 cm
14-DE-HE-01-6090				60 - 90 cm
14-DE-HE-02-030	HE	-	Weiterstadt	0 - 30 cm
14-DE-HE-02-3060				30 - 60 cm
14-DE-HE-02-6090				60 - 90 cm
14-DE-HE-03-030	HE	-	Darmstadt	0 - 30 cm
14-DE-HE-03-3060				30 - 60 cm
14-DE-HE-03-6090				60 - 90 cm

### 8.3 Messparameter ICP-MS

**Tabelle 24:** Messparameter der Multielementmethode; ISTD: interner Standard, *Integration time per point*: Integrationszeit in s pro Messpunkt der dreifachen Messung, Helium: Helium-Modus.

Element	Isotop	Messmodus	ISTD	Integration time
Li	7	<i>NoGas</i>	Ge	0.1
Be	9	<i>NoGas</i>	Ge	0.1
Al	27	<i>NoGas</i>	Ge	0.1
Sc	45	Helium	Ge	0.1
V	51	Helium	Ge	0.5
Cr	53	Helium	Ge	1
Co	59	Helium	Ge	0.3
Ni	60	Helium	Ge	1
Zn	66	Helium	Ge	0.3
Ga	69	<i>NoGas</i>	Ge	0.1
As	75	Helium	Ge	1
Se	77	Helium	Ge	1
Se	82	Helium	Ge	1
Rb	85	<i>NoGas</i>	In	0.1
Sr	88	<i>NoGas</i>	In	0.1
Mo	95	Helium	Rh	0.3
Ag	107	<i>NoGas</i>	Rh	0.1
Cd	111	<i>NoGas</i>	Rh	1
Te	125	<i>NoGas</i>	Rh	0.3
Ba	137	<i>NoGas</i>	In	0.1
La	139	<i>NoGas</i>	In	0.1
Ce	140	<i>NoGas</i>	In	0.1
Pr	141	<i>NoGas</i>	In	0.1
Nd	146	<i>NoGas</i>	In	0.1
Sm	147	<i>NoGas</i>	In	0.1
Eu	153	Helium	In	0.1
Gd	157	Helium	In	0.3
Tb	159	<i>NoGas</i>	In	0.1
Dy	163	<i>NoGas</i>	In	0.1
Ho	165	<i>NoGas</i>	In	0.1
Er	166	<i>NoGas</i>	In	0.1

**Tabelle 24:** Messparameter der Multielementmethode - Fortsetzung

<b>Element</b>	<b>Isotop</b>	<b>Messmodus</b>	<b>ISTD</b>	<b>Integration time</b>
Tm	169	<i>NoGas</i>	In	0.1
Yb	172	<i>NoGas</i>	In	0.1
Lu	175	<i>NoGas</i>	In	0.1
Tl	205	<i>NoGas</i>	Re	0.3
Pb	208	<i>NoGas</i>	Re	0.3
Bi	209	<i>NoGas</i>	Re	0.3
Th	232	<i>NoGas</i>	Re	0.1
U	238	<i>NoGas</i>	Re	0.3

## 8.4 Kalibrierstandards

**Tabelle 25:** Zusammensetzung der ICP-Mehrelementstandardlösung VI (Merck, Darmstadt, Deutschland); die grau unterlegten Elemente wurden nicht in die Multielementmethode einbezogen.

Element	Konzentration
Ag	10 mg/L
Al	10 mg/L
As	100 mg/L
B	100 mg/L
Ba	10 mg/L
Be	100 mg/L
Bi	10 mg/L
Ca	1000 mg/L
Cd	10 mg/L
Co	10 mg/L
Cr	10 mg/L
Cu	10 mg/L
Fe	100 mg/L
Ga	10 mg/L
K	10 mg/L
Li	10 mg/L
Mg	10 mg/L
Mn	10 mg/L
Mo	10 mg/L
Na	10 mg/L
Ni	10 mg/L
Pb	10 mg/L
Rb	10 mg/L
Se	100 mg/L
Sr	10 mg/L
Te	10 mg/L
Tl	10 mg/L
U	10 mg/L
V	10 mg/L
Zn	100 mg/L

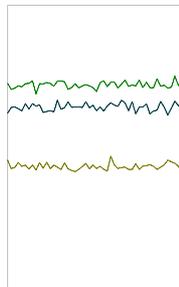
**Tabelle 26:** Zusammensetzung des Multielement-Kalibrierstandards 2 (PerkinElmer, Shelton, USA); die grau unterlegten Elemente wurden nicht in die Multielementmethode einbezogen.

<b>Element</b>	<b>Konzentration</b>
Ce	10 mg/L
Dy	10 mg/L
Er	10 mg/L
Eu	10 mg/L
Gd	10 mg/L
Ho	10 mg/L
La	10 mg/L
Lu	10 mg/L
Nd	10 mg/L
Pr	10 mg/L
Sc	10 mg/L
Sm	10 mg/L
Tb	10 mg/L
Th	10 mg/L
Tm	10 mg/L
Y	10 mg/L
Yb	10 mg/L

## 8.5 Geräteoptimierung ICP-MS

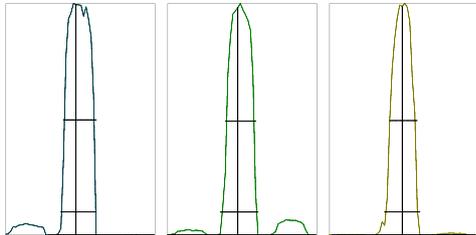
### Tune Report

Tune File : nogas.u



m/z	Range	Count	Mean	RSD%	Background
7	10,000	6311.0	6375.6	2.10	1.10
89	50,000	35747.0	35828.8	1.64	11.10
205	50,000	20405.0	21410.3	2.70	26.00

Integration Time: 0.1000 sec  
 Sampling Period: 0.3100 sec  
 n: 50  
 Oxide: 156/140 1.38%  
 Doubly Charged: 70/140 1.44%



m/z	7	89	205
Height:	6,294	35,575	21,113
Axis:	6.95	88.95	205.00
W-50%:	0.65	0.60	0.55
W-10%:	0.700	0.7500	0.700

Integration Time: 0.1000 sec  
 Acquisition Time: 22.7600 sec  
 Y axis : Linear

#### ===Plasma Condition===

RF Power : 1550 W  
 RF Matching : 1.9 V  
 Smpl Depth : 8 mm  
 Torch-H : -0.2 mm  
 Torch-V : 0.2 mm  
 Carrier Gas : 0.75 L/min  
 Dilution Mode : OFF  
 Makeup Gas : 0.35 L/min  
 Optional Gas : --- %  
 Nebulizer Pump : 0.15 rps  
 Sample Pump : --- rps  
 S/C Temp : 2 degC

#### ===Ion Lenses===

Extract 1 : 0 V  
 Extract 2 : -145 V  
 Omega Bias : -70 V  
 Omega Lens : 7.2 V  
 Cell Entrance : -30 V  
 Cell Exit : -50 V  
 Deflect : 13.2 V  
 Plate Bias : -40 V  
 Octopole Parameters===  
 OctP RF : 180 V  
 OctP Bias : -8 V

#### ===Q-Pole Parameters===

AMU Gain : 129  
 AMU Offset : 128  
 Axis Gain : 1.0001  
 Axis Offset : 0.04  
 QP Bias : -3 V

#### ===Detector Parameters===

Discriminator : 4.7 mV  
 Analog HV : 1837 V  
 Pulse HV : 1046 V

#### ===Reaction Cell===

Reaction Mode : OFF  
 H2 Gas : --- mL/min      He Gas : 0 mL/min      Optional Gas : --- %

Tune File : he.u

He Gas:	5.5 mL/min	m/z	Count (Mean)	RSD%	Integration Time:	0.1000sec
Optional Gas:	--- %	59	2300.0	3.29		
Cell Exit:	-60 V	89	1157.3	4.23		
OctP Bias:	-18 V	205	8971.4	2.27		
QP Bias:	-15 V					

Abbildung 57: Beispielhafter *Tuning*-Report der ICP-MS.

## 8.6 Methodenvalidierung ICP-MS

**Tabelle 27:** Ergebnisse der Analyse des Referenzmaterials BCR 670 (Wasserpflanze); Angaben in mg/kg, <sup>a</sup> Konzentration in µg/kg, <sup>b</sup> Mittelwert zweier Messwerte.

Element	zertifizierter Gehalt	analysierter Gehalt	Wiederfindung
Ce	0.99 ± 0.04	0.95 ± 0.01	0.96
Dy	0.079 ± 0.007	0.077 ± 0.001	0.98
Er	44.0 ± 2.8 <sup>a</sup>	42.0 ± 0.5 <sup>a</sup>	0.96
Eu	23.2 ± 1.5 <sup>a</sup>	24.7 ± 0.3 <sup>a</sup>	1.06
Gd	0.098 ± 0.008	0.096 ± 0.001	0.98
Ho	15.8 ± 1.8 <sup>a</sup>	14.7 ± 0.3 <sup>a</sup>	0.93
La	0.487 ± 0.020	0.478 ± 0.002	0.98
Lu	6.3 ± 0.5 <sup>a</sup>	5.3 ± 0.1 <sup>a</sup>	0.85
Nd	0.473 ± 0.015	0.458 ± 0.003	0.97
Pr	0.121 ± 0.006	0.113 ± 0.001	0.93
Sc	0.191 ± 0.011	0.229 ± 0.002	1.20
Sm	0.094 ± 0.007	0.098 ± 0.001	1.04
Tb	14.0 ± 1.1 <sup>a</sup>	13.9 ± 0.2 <sup>a</sup>	0.99
Tm	5.7 ± 0.7 <sup>a</sup>	5.4 ± 0.1 <sup>a</sup>	0.95
Yb	0.040 ± 0.004	0.035 ± 0.001	0.88
Th	0.159 ± 0.018	0.132 ± 0.002	0.83
U	0.082 ± 0.008	0.075 ± 0.001	0.91
Element	Richtwert	analysierter Gehalt	Wiederfindung
As	1.98 ± 0.19	1.98 ± 0.01	1.00
Cd	75.5 ± 2.5 <sup>a</sup>	75.6 ± 1.6 <sup>a</sup>	1.00
Cr	2.05 ± 0.1	1.89 ± 0.01	0.92
Mo	0.56 ± 0.07	0.63 ± 0.01	1.13
Pb	2.06 ± 0.12	1.96 ± 0.02	0.95
Zn	24.0 ± 2.1	22.0 ± 0.1	0.92
Element	Hinweiswert	analysierter Gehalt	Wiederfindung
Se	0.211 ± 0.062 <sup>b</sup>	0.431 ± 0.016	2.05

**Tabelle 28:** Ergebnisse der Analyse des Referenzmaterials DLA 41/2015 (Gerstengras); Angaben in mg/kg, <sup>a</sup> Konzentration in µg/kg.

Element	zertifizierter Gehalt	analysierter Gehalt	Wiederfindung
Al	161.0 ± 25.9	163.2 ± 10.9	1.01
Ba	3.29 ± 0.268	3.03 ± 0.15	0.92
Cd	64.0 ± 12.2 <sup>a</sup>	61.0 ± 1.0 <sup>a</sup>	0.90
Cr	0.990 ± 0.148	0.996 ± 0.058	1.01
Mn	22.60 ± 2.12	22.40 ± 0.20	1.03
Mo	0.735 ± 0.066	0.761 ± 0.008	1.03
Ni	0.568 ± 0.219	0.359 ± 0.012	0.63
Pb	0.891 ± 0.053	0.853 ± 0.126	0.96
Rb	17.50 ± 1.75	17.27 ± 0.22	0.99
Zn	29.10 ± 2.39	26.80 ± 0.19	0.92
Element	Richtwert	analysierter Gehalt	Wiederfindung
Ce	0.278 ± 0.073	0.309 ± 0.026	1.11
La	0.155 ± 0.012	0.160 ± 0.018	1.03
Nd	0.143 ± 0.036	0.142 ± 0.007	0.99

**Tabelle 29:** Ergebnisse der Analyse des Referenzmaterials BCR 667 (Sediment); Angaben in mg/kg, <sup>b</sup> Mittelwert zweier Messwerte.

Element	zertifizierter Gehalt	analysierter Gehalt	Wiederfindung
Ce	56.7 ± 2.5	43.4 ± 3.6	0.77
Dy	4.01 ± 0.14	3.07 ± 0.24	0.77
Er	2.35 ± 0.15	1.50 ± 0.11	0.64
Eu	1.00 ± 0.05	0.90 ± 0.07	0.90
Gd	4.41 ± 0.12	3.99 ± 0.31	0.90
Ho	0.80 ± 0.06	0.56 ± 0.05	0.70
La	27.8 ± 1.0	21.4 ± 1.7	0.77
Lu	0.325 ± 0.020	0.159 ± 0.013	0.49
Nd	25.0 ± 1.4	20.1 ± 1.5	0.80
Pr	6.1 ± 0.5	5.1 ± 0.4	0.84
Sc	13.7 ± 0.7	10.9 ± 0.9	0.80
Sm	4.66 ± 0.20	4.21 ± 0.31	0.90
Tb	0.682 ± 0.017	0.576 ± 0.048	0.84
Tm	0.326 ± 0.025	0.189 ± 0.016	0.58
Yb	2.20 ± 0.09	1.15 ± 0.09	0.52

**Tabelle 29:** Ergebnisse der Analyse des Referenzmaterials BCR 667 - Fortsetzung.

<b>Element</b>	<b>zertifizierter Gehalt</b>	<b>analysierter Gehalt</b>	<b>Wiederfindung</b>
Th	10.0 ± 0.5	7.2 ± 0.7	0.72
U	2.26 ± 0.15	1.12 ± 0.10	0.50
<b>Element</b>	<b>Richtwert</b>	<b>analysierter Gehalt</b>	<b>Wiederfindung</b>
Cd	0.67 ± 0.11	0.61 ± 0.04	0.91
Co	23.0 ± 1.3	21.5 ± 1.6	0.93
Cr	178 ± 16	163 ± 10	0.92
Ni	128 ± 9	124 ± 6	0.97
Pb	31.9 ± 1.1	30.0 ± 2.2	0.94
Zn	175 ± 13	163 ± 11	0.93
<b>Element</b>	<b>Hinweiswert</b>	<b>analysierter Gehalt</b>	<b>Wiederfindung</b>
As	17.1 ± 2.8	18.2 ± 1.7	1.06
Sr	225 ± 19	194 ± 10	0.86

**Tabelle 30:** Validierungsparameter der finalen Multielementmethode – Teil 1: Nachweis- und Bestimmungsgrenzen (in ng/L), linearer Bereich (in µg/L) und Bestimmtheitsmaß der Kalibriergeraden.

<b>Element</b>	<b>LOD</b>	<b>LOQ</b>	<b>linearer Bereich</b>	<b>R<sup>2</sup></b>
Li	25.5	76.6	0.01 - 1000	0.9997
Be	30.1	90.4	0.025 - 1000	0.9995
Al	322	966	1 - 1000	1.0000
V	1.2	3.6	0.025 - 1000	0.9999
Cr	14.3	43	0.25 - 1000	0.9997
Co	8	24	0.025 - 1000	1.0000
Ni	12.4	37.3	0.05 - 1000	1.0000
Zn	2.2	6.5	0.25 - 1000	0.9999
Ga	2.2	6.5	0.01 - 1000	1.0000
As	18	53.8	0.025 - 500	1.0000
Rb	2.3	6.8	0.05 - 1000	1.0000
Sr	2.4	7.3	0.25 - 1000	1.0000
Mo	5.9	17.6	0.025 - 1000	1.0000
Ag	3.5	10.5	0.01 - 500	0.9999
Cd	6	18.1	0.025 - 500	1.0000
Te	6	17.9	0.025 - 1000	0.9998
Ba	12.7	38.2	0.05 - 1000	1.0000

**Tabelle 30:** Validierungsparameter der finalen Multielementmethode – Teil 1 - Fortsetzung.

Element	LOD	LOQ	linearer Bereich	R <sup>2</sup>
La	0.2	0.5	0.0025 - 5	0.9998
Ce	0.2	0.7	0.0025 - 5	0.9998
Pr	0.1	0.3	0.001 - 5	0.9999
Nd	0.4	1.3	0.0025 - 5	1.0000
Sm	0.2	0.6	0.0025 - 5	0.9999
Eu	0.1	0.3	0.0025 - 5	1.0000
Gd	0.2	0.7	0.0025 - 5	0.9999
Tb	0.1	0.3	0.0025 - 5	1.0000
Dy	0.3	0.8	0.001 - 5	1.0000
Ho	0.1	0.3	0.0025 - 5	1.0000
Er	0.2	0.6	0.001 - 5	1.0000
Tm	0.1	0.2	0.001 - 5	0.9999
Yb	0.2	0.6	0.001 - 5	0.9999
Lu	0.2	0.5	0.001 - 5	0.9999
Tl	42.5	127.4	0.0025 - 1000	0.9999
Pb	73.2	219.5	0.1 - 1000	0.9998
Bi	4.5	13.5	0.01 - 1000	0.9993
Th	0.3	0.9	0.0025 - 5	0.9995
U	1.8	5.4	0.005 - 1000	0.9999

**Tabelle 31:** Validierungsparameter der finalen Multielementmethode – Teil 2: Wiederholpräzision und Wiederholbarkeit (in %).

Element	Wiederhol- präzision	Wiederholbarkeit Spargel	Wiederholbarkeit Boden
Li	5.3	13.9	11.2
Be	4.8	98.1	7.3
Al	5.9	9.3	6.7
V	2.3	7.5	6.5
Cr	4.1	12.2	5.9
Co	2.3	10.7	7.5
Ni	4.9	5.3	5.0
Zn	3.3	4.4	6.9
Ga	2.4	9.1	9.7

**Tabelle 31:** Validierungsparameter der finalen Multielementmethode – Teil 2 - Fortsetzung.

Element	Wiederhol- präzision	Wiederholbarkeit Spargel	Wiederholbarkeit Boden
As	1.9	28.1	9.5
Rb	3.7	3.7	6.3
Sr	2.9	4.2	5.2
Mo	2.8	3.7	6.2
Ag	2.4	37.7	13.6
Cd	2.9	5.2	6.0
Te	2.4	31.8	17.8
Ba	2.7	5.1	6.1
La	2.0	12.8	7.8
Ce	2.1	13.0	8.2
Pr	2.1	13.7	7.3
Nd	2.1	3.6	7.3
Sm	2.3	13.8	7.4
Eu	2.3	10.5	8.3
Gd	2.3	10.1	7.8
Tb	2.4	14.0	8.4
Dy	2.3	10.7	8.0
Ho	2.8	13.3	8.2
Er	2.7	11.3	7.6
Tm	3.0	13.8	8.3
Yb	3.0	14.5	7.7
Lu	3.1	14.5	8.0
Tl	2.6	12.6	8.9
Pb	2.6	11.5	7.3
Bi	2.4	54.8	8.2
Th	2.8	16.4	9.5
U	3.1	9.8	9.1

**Tabelle 32:** Ergebnisse der Laborvergleichsuntersuchung DLA 32/2017;  
Gehalte in mg/kg.

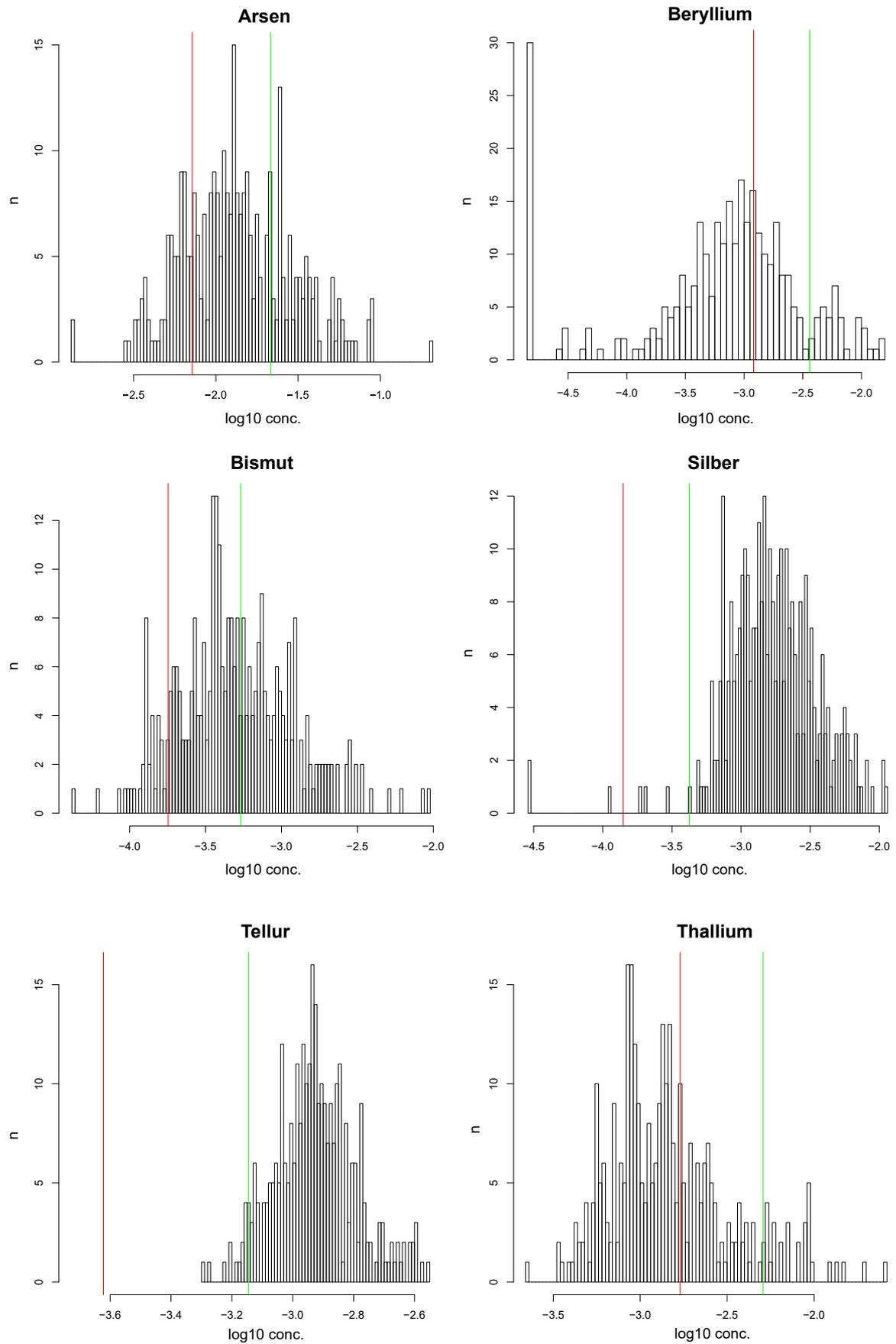
Element	robuster Mittelwert	robuste Standard-abweichung	analysierter Wert	z-Score
As	0.378	0.034	0.410	0.46
Cd	0.464	0.066	0.460	-0.05
Pb	0.446	0.052	0.398	-0.6

**Tabelle 33:** Ergebnisse der Laborvergleichsuntersuchung DLA 46/2017;  
Gehalte in mg/kg.

Element	robuster Mittelwert	robuste Standard-abweichung	analysierter Wert	z-Score
Al	0.527	0.283	0.570	0.2
Ba	0.220	0.0559 <sup>a</sup>	0.202	-0.4
Cd	0.0399	0.0029	0.0390	-0.1
Co	0.0110	0.0022	0.0094	-0.5
Cr	0.0423	0.0104	0.0350	<sup>b</sup>
Cu	1.98	0.12	2.05	0.3
Mo	0.197	0.016	0.203	0.1
Ni	0.0398	0.0065	0.0377	-0.2
Rb	2.76	0.06	2.71	<sup>b</sup>
Sr	0.720	0.034	0.710	-0.1
Zn	7.83	0.73	7.55	-0.3

<sup>a</sup> Vergleichsstandardabweichung  $s_R$

<sup>b</sup> aufgrund zu geringer Teilnehmerzahl (4) erfolgte keine statistische Auswertung



**Abbildung 58:** Verteilung der Spargelproben (n) über den Konzentrationsbereich im Vergleich zu LOD (rot) und LOQ (grün) für die Elemente Arsen, Beryllium, Bismut, Silber, Tellur und Thallium.

## 8.7 Konzentrationen der erfassten Elemente

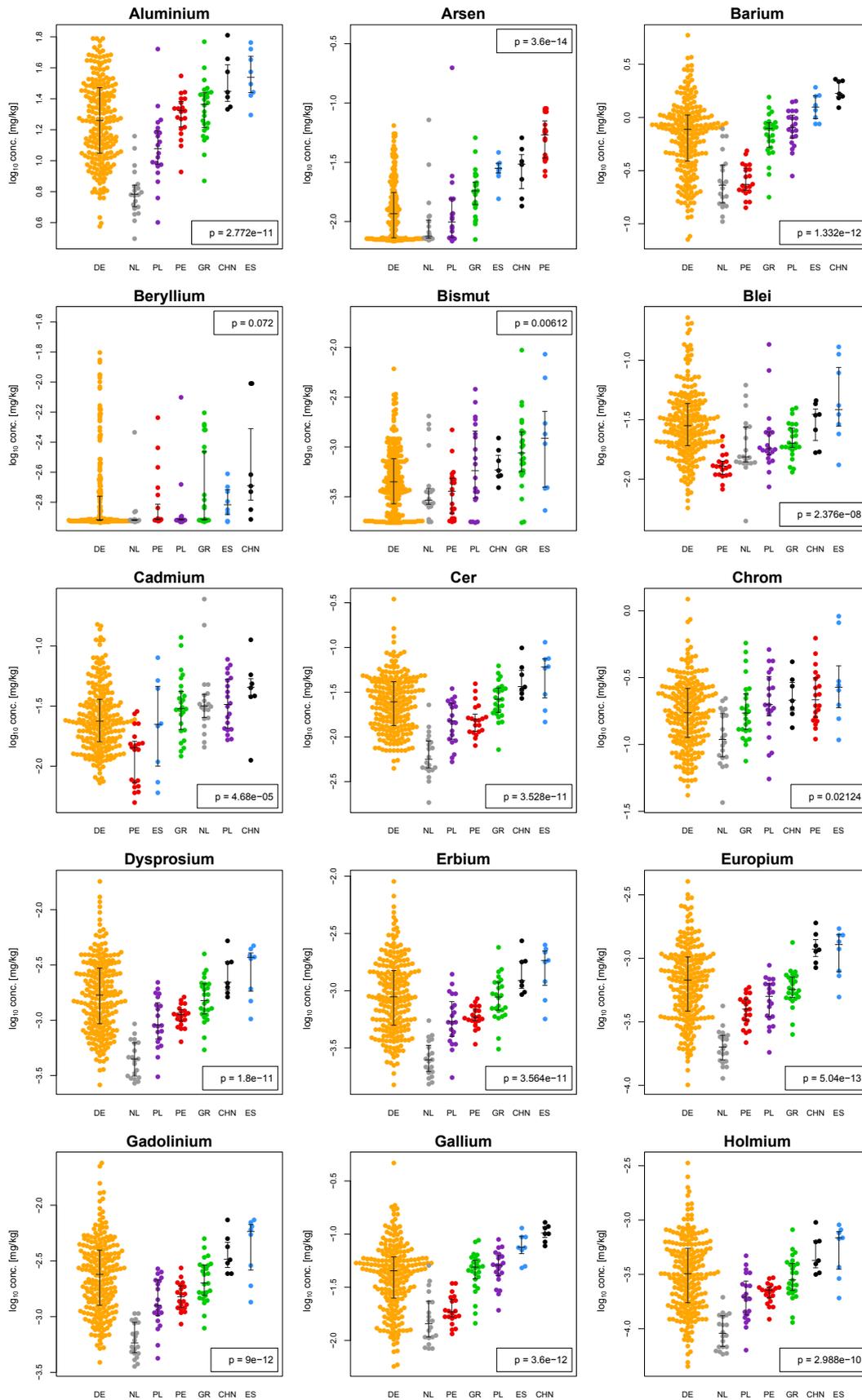


Abbildung 59: Konzentrationsprofile der untersuchten Elemente – Teil 1; p-Werte aus dem Kruskal-Wallis-Test.

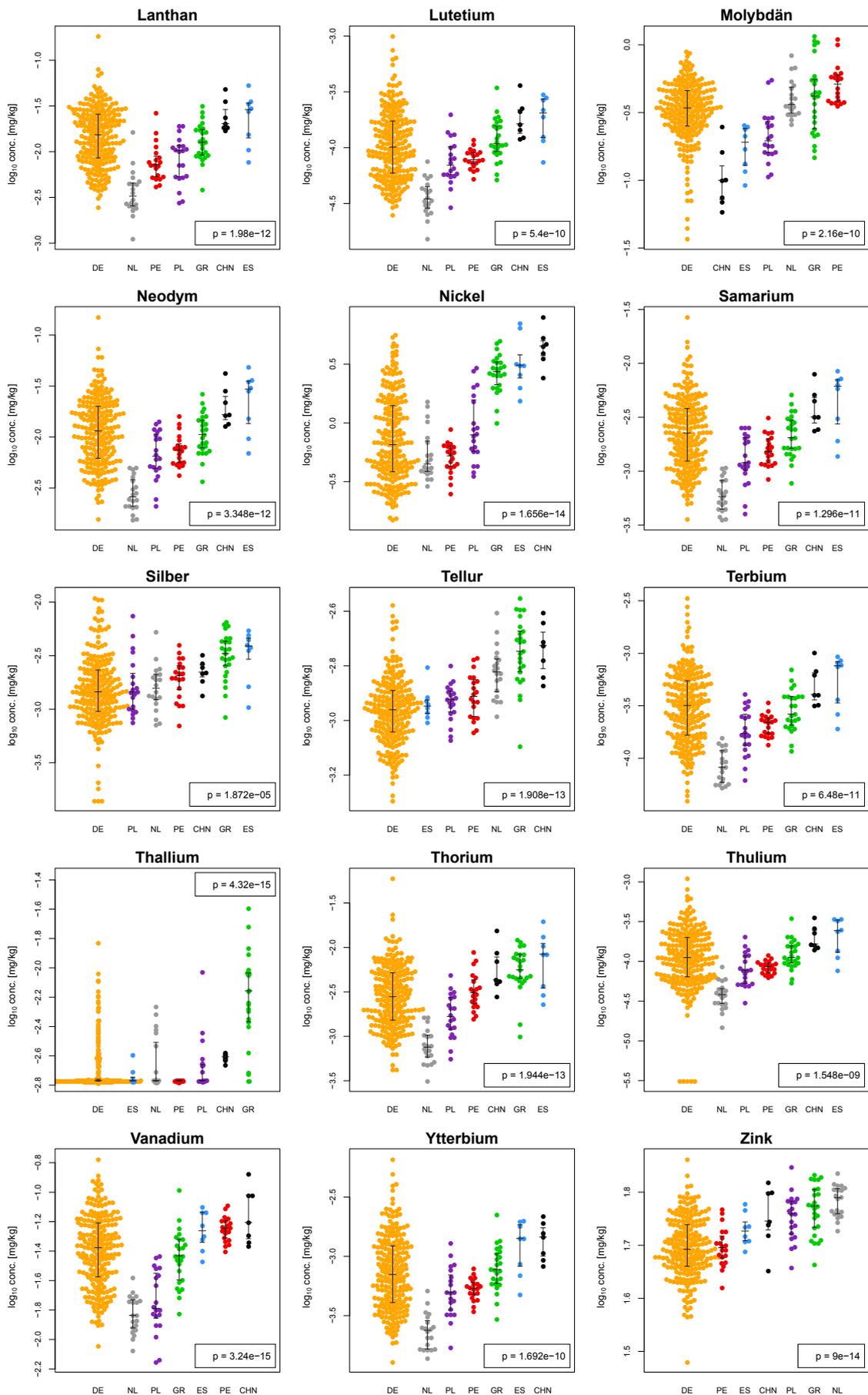


Abbildung 60: Konzentrationsprofile der untersuchten Elemente – Teil 2; p-Werte aus dem Kruskal-Wallis-Test.

**Tabelle 34:** Mittlere Elementkonzentrationen und Standardabweichungen der Spargelproben in  $\mu\text{g}/\text{kg}$ ; <sup>a</sup> in  $\text{mg}/\text{kg}$ .

Element	DE	PL	NL	ES	GR	CHN	PE
Li	$34.0 \pm$	$31.2 \pm$	$31.3 \pm$	$113.1 \pm$	$76.4 \pm$	$246.1 \pm$	$509.8 \pm$
	33.3	14.3	20.4	23.6	43.2	38.7	160.9
Be	$1.7 \pm 2.6$	$1.4 \pm 2.7$	$0.7 \pm 1.0$	$1.4 \pm 0.7$	$1.9 \pm 1.9$	$4.1 \pm 3.7$	$1.3 \pm 1.3$
Al <sup>a</sup>	$21.3 \pm$	$14.3 \pm$	$6.8 \pm 2.7$	$37.6 \pm$	$23.4 \pm$	$35.2 \pm$	$21.0 \pm$
	13.2	10.0		12.7	10.4	14.5	6.1
V	$45.9 \pm$	$20.2 \pm$	$15.8 \pm$	$57.2 \pm$	$39.5 \pm$	$75.2 \pm$	$57.8 \pm$
	25.9	9.1	4.4	15.7	18.0	31.3	10.7
Cr <sup>a</sup>	$0.203 \pm$	$0.239 \pm$	$0.124 \pm$	$0.381 \pm$	$0.210 \pm$	$0.244 \pm$	$0.252 \pm$
	0.139	0.123	0.054	0.291	0.125	0.090	0.126
Co	$21.9 \pm$	$21.1 \pm$	$16.5 \pm$	$112.4 \pm$	$45.4 \pm$	$93.2 \pm$	$55.7 \pm$
	11.9	11.0	16.9	36.9	11.1	11.4	16.1
Ni <sup>a</sup>	$1.05 \pm$	$1.11 \pm$	$0.62 \pm$	$3.63 \pm$	$2.85 \pm$	$4.64 \pm$	$0.52 \pm$
	1.01	0.91	0.36	1.89	0.96	1.61	0.14
Zn <sup>a</sup>	$50.4 \pm$	$56.8 \pm$	$61.0 \pm$	$53.8 \pm$	$58.8 \pm$	$57.1 \pm$	$49.8 \pm$
	6.8	5.6	4.0	3.6	6.1	6.8	4.1
Ga	$51.2 \pm$	$51.3 \pm$	$18.5 \pm$	$78.4 \pm$	$47.6 \pm$	$104.3 \pm$	$20.8 \pm$
	44.3	17.7	11.5	21.6	17.8	17.2	6.7
As	$14.6 \pm$	$19.9 \pm$	$12.0 \pm$	$27.9 \pm$	$19.3 \pm$	$26.1 \pm$	$56.9 \pm$
	11.0	40.7	15.4	6.1	9.5	12.6	23.6
Rb <sup>a</sup>	$7.06 \pm$	$5.67 \pm$	$15.8 \pm$	$8.13 \pm$	$26.4 \pm$	$19.1 \pm$	$7.64 \pm$
	3.98	2.23	3.6	2.53	15.6	2.8	1.50
Sr <sup>a</sup>	$1.79 \pm$	$1.75 \pm$	$1.985 \pm$	$4.89 \pm$	$1.99 \pm$	$13.5 \pm$	$2.83 \pm$
	1.65	0.52	0.936	1.41	0.63	6.7	1.15
Mo <sup>a</sup>	$0.358 \pm$	$0.234 \pm$	$0.430 \pm$	$0.184 \pm$	$0.490 \pm$	$0.116 \pm$	$0.535 \pm$
	0.158	0.117	0.152	0.061	0.304	0.063	0.197
Ag	$2.0 \pm 1.8$	$2.0 \pm 1.6$	$1.8 \pm 1.0$	$3.6 \pm 1.4$	$3.5 \pm 1.5$	$2.3 \pm 0.6$	$2.1 \pm 0.8$
Cd	$31.6 \pm$	$38.1 \pm$	$48.6 \pm$	$30.9 \pm$	$37.8 \pm$	$50.7 \pm$	$13.56 \pm$
	24.7	19.0	54.9	24.7	25.7	29.0	7.06
Te	$1.1 \pm 0.3$	$1.2 \pm 0.3$	$1.5 \pm 0.3$	$1.2 \pm 0.2$	$1.8 \pm 0.5$	$1.9 \pm 0.4$	$1.3 \pm 0.2$
Ba <sup>a</sup>	$0.867 \pm$	$0.856 \pm$	$0.302 \pm$	$1.30 \pm$	$0.749 \pm$	$1.83 \pm$	$0.277 \pm$
	0.732	0.307	0.204	0.36	0.305	0.38	0.094
La	$19.2 \pm$	$9.52 \pm$	$4.35 \pm$	$26.8 \pm$	$14.2 \pm$	$25.9 \pm$	$8.62 \pm$
	16.9	4.91	3.22	14.1	6.5	10.8	5.02
Ce	$31.3 \pm$	$16.9 \pm$	$7.51 \pm$	$56.7 \pm$	$28.5 \pm$	$48.6 \pm$	$17.17 \pm$
	31.1	8.4	4.75	31.9	13.0	23.5	7.77
Pr	$3.81 \pm$	$1.92 \pm$	$0.776 \pm$	$6.65 \pm$	$3.09 \pm$	$5.65 \pm$	$1.98 \pm$
	3.66	0.92	0.293	3.51	1.43	2.65	0.77

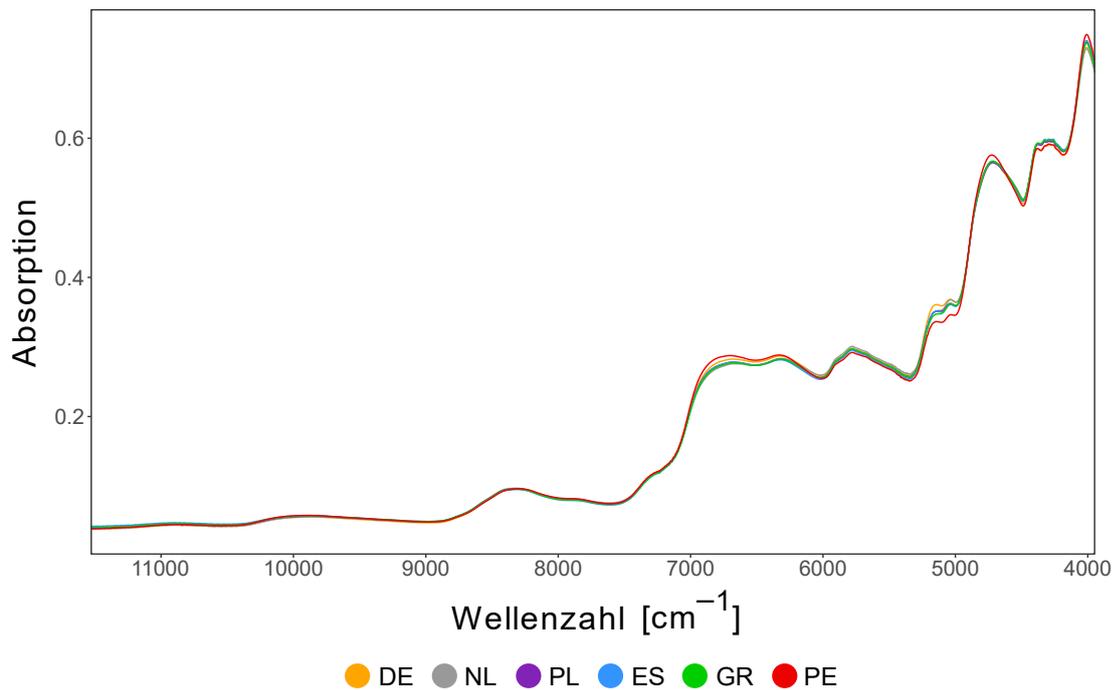
**Tabelle 34:** Mittlere Elementkonzentrationen und Standardabweichungen der Spargelproben in µg/kg; <sup>a</sup> in mg/kg - Fortsetzung.

Element	DE	PL	NL	ES	GR	CHN	PE
Nd	15.1 ±	7.59 ±	3.02 ±	26.4 ±	11.9 ±	21.7 ±	7.92 ±
	14.5	3.60	1.14	13.7	5.4	9.9	2.94
Sm	3.03 ±	1.47 ±	0.648 ±	5.19 ±	2.35 ±	4.13 ±	1.65 ±
	2.97	0.67	0.229	2.51	1.04	1.86	0.54
Eu	0.811 ±	0.498 ±	0.217 ±	1.18 ±	0.613 ±	1.27 ±	0.407 ±
	0.637	0.180	0.069	0.42	0.205	0.34	0.107
Gd	3.19 ±	1.50 ±	0.672 ±	4.87 ±	2.32 ±	4.01 ±	1.59 ±
	3.30	0.68	0.228	2.27	1.00	1.65	0.45
Tb	0.431 ±	0.204 ±	0.092 ±	0.619 ±	0.314 ±	0.537 ±	0.214 ±
	0.465	0.096	0.032	0.276	0.138	0.233	0.048
Dy	2.34 ±	1.10 ±	0.488 ±	3.14 ±	1.70 ±	2.81 ±	1.14 ±
	2.45	0.53	0.183	1.33	0.76	1.20	0.24
Ho	0.439 ±	0.220 ±	0.102 ±	0.592 ±	0.329 ±	0.533 ±	0.220 ±
	0.453	0.107	0.036	0.250	0.152	0.217	0.043
Er	1.23 ±	0.634 ±	0.280 ±	1.67 ±	0.946 ±	1.52 ±	0.616 ±
	1.22	0.311	0.101	0.68	0.436	0.60	0.122
Tm	0.155 ±	0.088 ±	0.041 ±	0.228 ±	0.133 ±	0.210 ±	0.087 ±
	0.152	0.043	0.016	0.100	0.060	0.074	0.016
Yb	0.974 ±	0.560 ±	0.254 ±	1.33 ±	0.865 ±	1.43 ±	0.548 ±
	0.917	0.272	0.091	0.55	0.395	0.48	0.110
Lu	0.141 ±	0.083 ±	0.039 ±	0.197 ±	0.130 ±	0.194 ±	0.082 ±
	0.137	0.040	0.014	0.079	0.061	0.080	0.016
Tl	1.7 ± 1.7	2.0 ± 1.8	2.3 ± 1.4	1.6 ± 0.4	7.7 ± 5.5	2.4 ± 0.1	0.8 ± 0.2
Pb	37.9 ±	28.0 ±	21.9 ±	58.7 ±	22.7 ±	31.5 ±	13.4 ±
	34.2	29.1	14.5	41.4	7.7	10.9	3.3
Bi	0.6 ± 0.7	1.3 ± 1.8	0.5 ± 0.6	2.4 ± 2.8	1.4 ± 1.8	0.7 ± 0.3	0.4 ± 0.3
Th	3.94 ±	1.97 ±	0.863 ±	8.62 ±	6.32 ±	6.65 ±	3.64 ±
	4.88	1.07	0.385	5.49	2.87	4.09	1.86
U	2.52 ±	1.59 ±	0.456 ±	4.26 ±	6.08 ±	119.9 ±	5.29 ±
	4.65	1.09	0.116	1.93	3.16	50.2	3.20

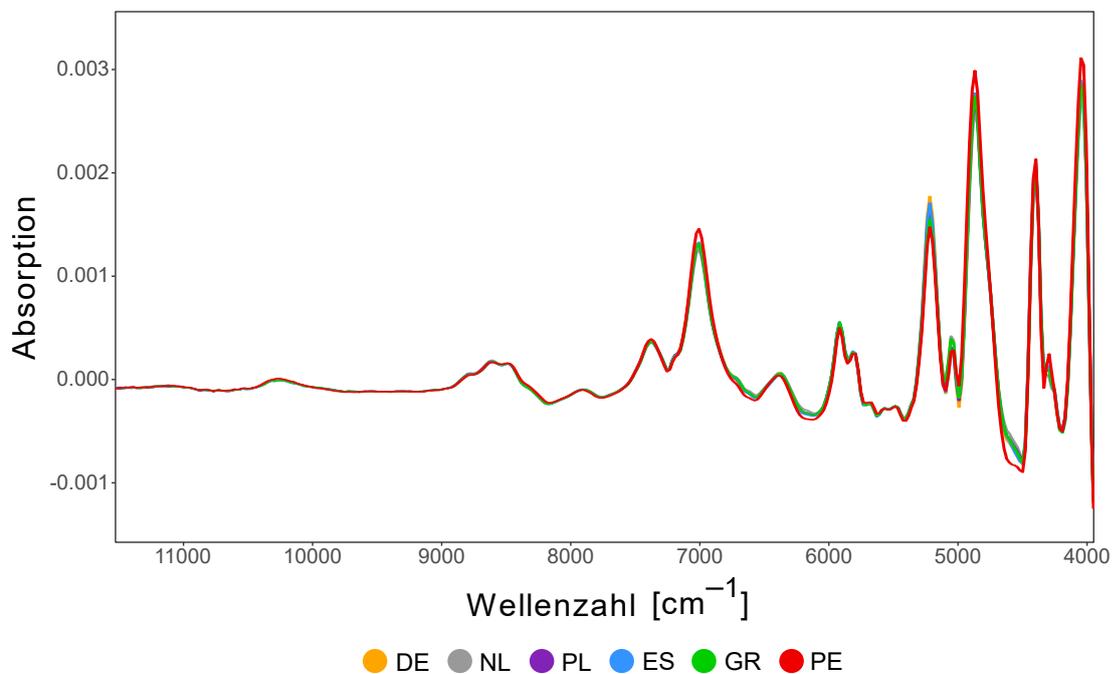




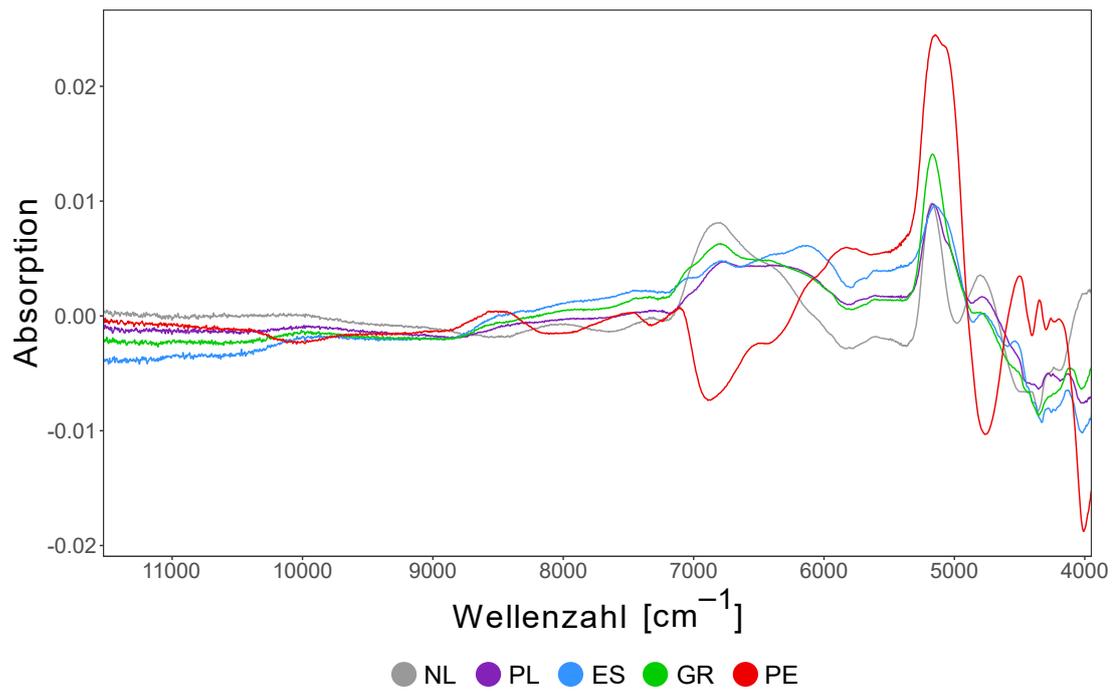
## 8.9 FT-NIR-Spektren



**Abbildung 63:** Mittlere Spektren der Spargelproben nach MSC-Korrektur basierend auf NIR-Daten (gesamtes Spektrum).



**Abbildung 64:** Mittlere Spektren der Spargelproben nach MSC-Korrektur und 1. Ableitung basierend auf NIR-Daten (gesamtes Spektrum).



**Abbildung 65:** Differenzspektren nach Abzug des deutschen vom mittleren NIR-Spektrum des jeweiligen Landes (gesamtes Spektrum).

## 8.10 Klassifikation anhand von FT-NIR-Daten

**Tabelle 35:** Konfusionsmatrices und Gesamtgenauigkeit zur 6-Klassen-Klassifikation des Herkunftslandes mittels (a) SVM mit RBF-Kern und (b) RF basierend auf NIR-Daten.

(a) SVM (RBF): 86.5 %

	DE	NL	PL	ES	GR	PE
DE	<b>175</b>	2	1	0	1	0
NL	7	<b>13</b>	0	0	0	0
PL	9	0	<b>10</b>	0	1	0
ES	2	0	1	<b>0</b>	5	0
GR	2	0	2	0	<b>21</b>	0
PE	2	0	0	0	0	<b>19</b>

(b) RF: 76.4 %

	DE	NL	PL	ES	GR	PE
DE	<b>174</b>	3	1	0	1	1
NL	10	<b>10</b>	0	0	0	0
PL	18	0	<b>3</b>	0	0	0
ES	6	0	0	<b>0</b>	2	0
GR	17	0	0	0	<b>8</b>	0
PE	6	0	0	0	0	<b>15</b>

## 8.11 Formelübersicht

Korrelationskoeffizient nach Pearson: 
$$r_P = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$x, y$  zu vergleichende Messreihen

$\bar{x}, \bar{y}$  Mittelwerte

$n$  Anzahl der Wertepaare

Prüfgröße t für Test auf signifikante Korrelation: 
$$t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}}$$

## 8.12 Chemikalienübersicht

**Tabelle 36:** Übersicht der verwendeten Chemikalien; GHS-Symbole laut Tabelle 37.

Substanz	CAS-Nr.	Hersteller	GHS-Symbol	H-Sätze	P-Sätze
Salpetersäure, Supra-Qualität, ROTIPURAN® Supra 69 %	7697-37-2	Carl Roth, Karlsruhe, Deutschland	1, 3, 4	272, 290, 331	280, 301+330+331, 305+351+338, 308+310
Salpetersäure, ROTIPURAN® 65 %, p.a.	7697-37-2	Carl Roth, Karlsruhe, Deutschland	1, 3, 4	272, 290, 331	280, 301+330+331, 305+351+338, 308+310
Wasserstoff- peroxid 30 %, Suprapur®	7722-84-1	Merck, Darmstadt, Deutschland	3, 5	302, 318, 413	273, 280, 305+351+338, 313
Salzsäure 30 %, Suprapur®	7647-01-0	Merck, Darmstadt, Deutschland	3, 5	290, 314, 335	260, 280, 303+361+353, 304+340+310, 305+351+338
Trockeneis	124-38-9	Linde, Pullach, Deutschland	-	281	-
Stickstoff, flüssig	7727-37-9	Linde, Pullach, Deutschland	2	280	403
Argon 5.0, Reinheit 99.999 Vol. %	7440-37-1	Westfalen/ Praxair	2	280	403
ICP- Mehrelement- standardlösung VI	Stoff- gemisch <sup>a</sup>	Merck, Darmstadt, Deutschland	3, 6	350i, 290, 314, EUH071	201, 280, 301+330+331, 305+351+338, 308+310
Multielement- Kalibrier- standard 2	Stoff- gemisch <sup>b</sup>	PerkinElmer, Shelton, USA	3	314, 318	260, 280, 303+361+353, 305+351+338, 310, 304+340, 405, 501

<sup>a</sup> Stoffgemisch enthält unter anderem: Salpetersäure (7697-37-2), Berylliumacetat (19049-40-2), Silbernitrat (7761-88-8). Zudem enthalten sind alle in Tabelle 25 aufgeführten Elemente.

<sup>b</sup> Stoffgemisch enthält: Salpetersäure (7697-37-2), Erbium (7440-52-0), Europium (7440-53-1), Gadolinium (7440-54-2), Holmium (7440-60-0), Lanthanoxid (1312-81-8), Lutetium (7439-94-3), Neodym (7440-00-8), Dysprosium (7427-91-6), Praseodym (7440-10-0), Samarium (7440-19-9), Scandiumoxid (12060-08-1), Terbium(III,IV)-oxid (12037-01-3), Thorium (7440-29-1), Thulium (7440-30-4), Ytterbium (7440-64-4), Yttriumoxid (1314-36-9), Cer (7440-45-1).

**Tabelle 37:** Definition der GHS-Symbole in Tabelle 36.

Symbol-Nummer	GHS-Symbol	Bedeutung
1		brandfördernd
2		Gase unter Druck, verdichtete/verflüssigte/ tiefgekühlt verflüssigte Gase
3		ätzend
4		giftig
5		reizend
6		gesundheitsschädlich

## 8.13 Geräteübersicht

**Tabelle 38:** Übersicht der verwendeten Geräte.

Gerätebezeichnung	Hersteller
Analysenwaage XA105DU	Mettler-Toledo GmbH, Wien, Österreich
Entladungspistole Zerostat 3	Milty/armourhome, Bishop's Stortford, Vereinigtes Königreich
Gefriertrocknungsanlage Christ Gamma 1-20 mit Trockenrechen	Martin Christ Gefriertrocknungsanlagen, Osterode am Harz, DE
Gefriertrocknungsanlage Christ Beta 1-16 mit Trockenrechen	Martin Christ Gefriertrocknungsanlagen, Osterode am Harz, DE
ICP-MS 7700x	Agilent Technologies, Inc., Santa Clara, CA, USA
Kolbenhubpipette, 10 - 100 µL	Eppendorf, Hamburg, DE
Kolbenhubpipette, 100 - 1000 µL	Eppendorf, Hamburg, DE
Kolbenhubpipette, 1 - 10 mL	Eppendorf, Hamburg, DE
Kühl- und Gefrierschränke	verschiedene
Messer mit Keramikschneidefläche	Böker Manufaktur, Solingen, DE
Messermühle GM300	Retsch, Haan, Deutschland
Mikrowellenaufschluss ETHOS.lab	MLS, Leutkirch, Deutschland
Reinstwasseranlage Direct-Q 3 UV-R	Millipore, Bedford, MA, USA
Sieb 0.5 mm Maschenweite	Haver & Boecker, Oelde, DE
Sieb 2 mm Maschenweite	Haver & Boecker, Oelde, DE
TANGO FT-NIR-Spektrometer	Bruker Optik GmbH, Ettlingen, DE
Ultraschallgerät USC300TH	VWR International, Leuven, Belgien
gekühlte Universal-Zentrifuge 3-16PK	Sigma, Osterode, DE

## 8.14 Verbrauchsmaterialien

**Tabelle 39:** Übersicht der verwendeten Materialien.

<b>Materialbezeichnung</b>	<b>Hersteller</b>
Filterpapier 640 m, Durchmesser 90 mm	Macherey-Nagel, Düren, DE
Pipettenspitzen, 10 - 100 $\mu$ L	Eppendorf, Hamburg, DE
Pipettenspitzen 100 - 1000 $\mu$ L	Eppendorf, Hamburg, DE
Pipettenspitzen 1 - 10 mL	Eppendorf, Hamburg, DE
Schraubdeckelröhren 15 mL, PP, steril, Spitzboden	Sarstedt, Nümbrecht, DE
Schraubdeckelröhren 30 mL, PP, Spitzboden mit Stehrand	Sarstedt, Nümbrecht, DE
Schraubdeckelröhren 50 mL, PP, steril, Spitzboden	Sarstedt, Nümbrecht, DE

# Abbildungsverzeichnis

1	(a) Violetter, weißer und grüner Spargel [25]; (b) Wurzelstock einer Spargelstaude mit Sprossen, adaptiert nach [26]. . . . .	8
2	Blüten (a), Früchte (b) und Samen (c) einer Spargelpflanze [32, 33]. . .	9
3	Zugelassene geschützte geographische Angaben (ggA) für Spargel in Deutschland [37, 39]. . . . .	10
4	(a) Spargelfeld mit aufgehäuften Dämmen [60], (b) schematische Darstellung des Querschnitts durch einen Spargeldamm [61]. . . . .	12
5	Entwicklung der Erntemengen der vier wichtigsten Anbauländer für Spargel von 1961 bis 2017 [36]. . . . .	15
6	Schematischer Aufbau eines Quadrupol-basierten ICP-MS-Gerätes, modifiziert nach [164]. . . . .	22
7	Schematischer Aufbau eines konzentrischen Zerstäubers, modifiziert nach [162]. . . . .	23
8	Schematischer Aufbau einer <i>Scott-like double pass</i> -Sprühkammer, modifiziert nach [171]. . . . .	24
9	(a) Schematischer Aufbau der Plasmafackel, modifiziert nach [185]; (b) Temperaturzonen des Plasmas, modifiziert nach [186]. . . . .	25
10	Schematische Darstellung des Übergangs der Analyten vom Plasma durch die <i>cones</i> hin zum Linsensystem, modifiziert nach [195]. . . . .	27
11	Schematische Darstellung der Funktionsweise eines Sekundärelektronenvervielfachers (SEV), modifiziert nach [211]. . . . .	31
12	Prinzipieller Aufbau eines FT-IR-Spektrometers, modifiziert nach [227].	35
13	Vergleich des Strahlengangs an flachen Spiegeln (oben) und <i>CubeCorner</i> -Spiegeln (unten) bei Neigung des Spiegels [228]. . . . .	36
14	(a) Diffuse Reflexion an Partikeloberflächen, modifiziert nach [233]; (b) Schematischer Schnitt durch eine Integrationskugel, modifiziert nach [234].	36
15	Nicht-lineare Klassentrennung mittels <i>Support Vector Machine</i> . . . . .	43
16	Beispielhafte Darstellung von Hyperebene und Stützvektoren bei der <i>Support Vector Machine</i> . . . . .	44
17	Ablauf einer <i>nested cross-validation</i> ; modifiziert nach [289]. . . . .	47

18	Verlauf der internen Standards (ISTD) bei der Messung der Kalibrierreihe in <i>NoGas</i> -Modus [1] und Helium-Modus [2] in Anwesenheit aller Standardelemente. . . . .	55
19	Verlauf der internen Standards (ISTD) bei der Messung der Kalibrierreihe im <i>NoGas</i> -Modus [1] und Helium-Modus [2] in Abwesenheit der Lanthanoide in den Standards $> 5 \mu\text{g/L}$ . . . . .	56
20	Anzahl und Herkunft der Proben, die für die Herkunftsbestimmung anhand des Elementprofils herangezogen wurden; schwarze Punkte: Markierung der Ernteorte der deutschen Proben [290]. . . . .	64
21	Hauptkomponentenanalyse mit (a) der ersten und zweiten PC und (b) der ersten und dritten PC zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten. . . . .	65
22	Unterscheidung des Herkunftslandes mittels t-SNE basierend auf Multielementdaten. . . . .	66
23	Biplots der PCA zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten für (a) die erste und zweite sowie (b) die erste und dritte Hauptkomponente. . . . .	67
24	<i>Heatmap</i> kombiniert mit hierarchischem Clustering zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten. . . . .	69
25	Korrelationen nach Pearson aller untersuchten Elemente. . . . .	70
26	<i>Heatmap</i> kombiniert mit hierarchischem Clustering ohne korrelierende Elemente zur Unterscheidung des Herkunftslandes basierend auf Multielementdaten. . . . .	71
27	Mittels Welch-Test ermittelte p-Werte nach Bonferroni-Korrektur; grau/schwarz: $p\text{-Wert} \geq 0.05$ , dunkelrot bis gelb: $p\text{-Wert} < 0.05$ in abnehmender Reihenfolge. . . . .	75
28	<i>Beeswarm</i> -Plots der Konzentrationsverteilungen relevanter Elemente für die Herkunftsunterscheidung; p-Werte aus dem Kruskal-Wallis-Test nach Bonferroni-Korrektur. . . . .	76
29	Klasseneinteilung der deutschen Spargelproben in die 5 Regionen Nord (N, grün), Ost (O, orange), West (W, gelb), Südost (SO, blau) und Südwest (SW, rot) sowie Standorte der bei der Probenakquise beteiligten Spargelhöfe (schwarze Rauten). . . . .	86
30	Hierarchisches Clustering mit Markierung des Erntejahres basierend auf Multielementdaten. . . . .	90
31	PCA der deutschen Proben aus den Erntejahren 2014-2016 basierend auf Multielementdaten. . . . .	90
32	Hierarchisches Clustering zur Unterscheidung der Spargelsorte basierend auf Multielementdaten. . . . .	91
33	Hierarchisches Clustering mit Markierung des Messdatums basierend auf Multielementdaten; unten: Darstellung der gemessenen Probenzahlen als Histogramm und Farbverlauf über die gesamte Messzeit. . . . .	92

---

34	Elementkorrelationen der deutschen Spargelproben des Erntejahres 2014 (Korrelationskoeffizient nach Pearson). . . . .	94
35	Elementkorrelationen der Bodenproben des Erntejahres 2014 (Korrelationskoeffizient nach Pearson). . . . .	95
36	Mittlere logarithmierte Konzentrationen aller Elemente in Spargel und Boden sowie die natürliche Häufigkeit in der Erdkruste. . . . .	97
37	Korrelationen nach Pearson für die seltenen Erden in (a) Spargel- und (b) Bodenproben des Erntejahres 2014. . . . .	99
38	Logarithmierte mittlere Konzentrationen der seltenen Erden bezogen auf die Lanthan-Konzentration für die deutschen Spargel- und Bodenproben des Erntejahres 2014 im Vergleich zur natürlichen Häufigkeit. . . . .	100
39	NIR-Rohdaten der Spargelproben ohne Datenpräprozessierung. . . . .	105
40	MSC-korrigierte NIR-Spektren der Spargelproben. . . . .	106
41	NIR-Spektren der Spargelproben nach Anwendung von MSC-Korrektur und 1. Ableitung. . . . .	106
42	Mittlere MSC-korrigierte FT-NIR-Spektren der Spargelproben aller Herkunftsländer (a) vor und (b) nach Anwendung der 1. Ableitung. . . . .	108
43	Differenzspektrum der MSC-korrigierten Spektren nach Abzug des Spektrums der deutschen Proben vom mittleren NIR-Spektrum des jeweiligen Landes. . . . .	109
44	Differenzspektrum nach MSC-Korrektur und 1. Ableitung als Mittelwertspektrum eines Landes nach Abzug des deutschen Spektrums; relevante Wellenzahlen ( <i>selected Features</i> mittels Boruta) sind in grau hervorgehoben. . . . .	110
45	<i>Scores Plot</i> der PCA für die Herkunftsunterscheidung basierend auf NIR-Daten. . . . .	111
46	PLS für die Herkunftsunterscheidung basierend auf NIR-Daten. . . . .	112
47	PLS für die Unterscheidung einer deutschen oder ausländischen Herkunft basierend auf NIR-Daten. . . . .	117
48	Vorgehensweise bei der Abschätzung des Vertrauensbereiches: Verteilung der Klassifikationsentscheidungen anhand der Vorhersagewahrscheinlichkeiten der Klasse der deutschen Proben dargestellt als Histogramm für richtige (blau) und falsche (rot) Klassifikationen; die Mehrzahl der Fehlklassifikationen wird eliminiert (graue Punkte), indem Proben mit geringer Konfidenz ausgeschlossen werden (Intervall innerhalb der gestrichelten Linien). . . . .	119
49	Hierarchisches Clustering der Spargelproben mit Markierung der Erntejahre basierend auf NIR-Daten. . . . .	125
50	(a) PCA und (b) PLS der deutschen Spargelproben zur Untersuchung des Erntejahreinflusses basierend auf NIR-Daten. . . . .	125

51	Boxplot für die Verteilung des Absorptionslevels der Proben aus unterschiedlichen Herkunftsländern und Erntejahren bei der Wellenzahl $5976.8 \text{ cm}^{-1}$ . . . . .	126
52	Hierarchisches Clustering aller Spargelproben mit Markierung der Spargelsorte basierend auf NIR-Daten. . . . .	127
53	PCA der deutschen Spargelproben der Sorten Backlim, Gijnlim, Grolim und Cumulus basierend auf NIR-Daten. . . . .	128
54	Hierarchisches Clustering aller Spargelproben mit Markierung des Messdatums anhand eines Farbverlaufs basierend auf NIR-Daten; unten: Histogramm mit Anzahl gemessener Proben im gesamten Messzeitraum. .	129
55	Schematische Darstellung eines Handlungsstrangs zur Übertragung der in dieser Arbeit gezeigten ICP-MS-basierten Methodik auf ein anderes Lebensmittel. . . . .	135
56	Schematische Darstellung eines Handlungsstrangs zur Übertragung der in dieser Arbeit gezeigten FT-NIR-basierten Methodik auf ein anderes Lebensmittel. . . . .	136
57	Beispielhafter <i>Tuning</i> -Report der ICP-MS. . . . .	194
58	Verteilung der Spargelproben (n) über den Konzentrationsbereich im Vergleich zu LOD (rot) und LOQ (grün) für die Elemente Arsen, Beryllium, Bismut, Silber, Tellur und Thallium. . . . .	201
59	Konzentrationsprofile der untersuchten Elemente – Teil 1; p-Werte aus dem Kruskal-Wallis-Test. . . . .	202
60	Konzentrationsprofile der untersuchten Elemente – Teil 2; p-Werte aus dem Kruskal-Wallis-Test. . . . .	203
61	Elementkorrelationen der Spargelproben: t-Werte zum Test auf Signifikanz; Werte oberhalb des kritischen t-Wertes von 2.704 sind in grüner Farbe hervorgehoben (40 Freiheitsgrade, Signifikanzniveau 0.01 bei zweiseitiger Verteilung). . . . .	206
62	Elementkorrelationen der Bodenproben: t-Werte zum Test auf Signifikanz; Werte oberhalb des kritischen t-Wertes von 2.728 sind in grüner Farbe hervorgehoben (34 Freiheitsgrade, Signifikanzniveau 0.01 bei zweiseitiger Verteilung). . . . .	207
63	Mittlere Spektren der Spargelproben nach MSC-Korrektur basierend auf NIR-Daten (gesamtes Spektrum). . . . .	208
64	Mittlere Spektren der Spargelproben nach MSC-Korrektur und 1. Ableitung basierend auf NIR-Daten (gesamtes Spektrum). . . . .	208
65	Differenzspektren nach Abzug des deutschen vom mittleren NIR-Spektrum des jeweiligen Landes (gesamtes Spektrum). . . . .	209

# Tabellenverzeichnis

1	Erntemenge, Anbaufläche und Ertrag der zehn wichtigsten Anbauländer für Spargel (2017) [36]. . . . .	14
2	Massenzahl und erste Ionisierungsenergien der als interne Standards verwendeten Elemente. . . . .	54
3	Übersicht der ermittelten p-Werte im Kruskal-Wallis-Test nach Bonferroni-Korrektur. . . . .	74
4	Konfusionsmatrix zur Klassifizierung des Herkunftslandes basierend auf Multielementdaten (Gesamtgenauigkeit: 91.2 %). . . . .	80
5	Klassifikationsergebnisse für die Betrachtung verschiedener Konfidenzbereiche begrenzt durch unterschiedliche Wahrscheinlichkeitswerte p der Klasse der deutschen Proben basierend auf Multielementdaten. . . . .	82
6	Konfusionsmatrix zur Klassifizierung des Herkunftslandes anhand von Proben mit hoher Konfidenz ( $p < 0.2$ oder $p > 0.8$ ) basierend auf Multielementdaten (Gesamtgenauigkeit: 95.6 %). . . . .	83
7	Konfusionsmatrix zur Klassifizierung einer deutschen oder ausländischen Herkunft basierend auf Multielementdaten (Gesamtgenauigkeit: 91.8 %). . . . .	84
8	Konfusionsmatrix zur Klassifizierung der deutschen Herkunftsregionen (Gesamtgenauigkeit: 83.6 %). . . . .	86
9	Konfusionsmatrix zur Klassifizierung des Produzenten (P) mittels SVM basierend auf Multielementdaten (Gesamtgenauigkeit: 82.6 %). . . . .	88
10	Elementkorrelationen in Spargel und Boden: Korrelationskoeffizienten nach Pearson ( $r_P$ ) sowie t-Werte (krit. t-Wert = 2.728, 34 Freiheitsgrade, Signifikanzniveau = 0.01, zweiseitige Verteilung [326]). . . . .	102
11	Konfusionsmatrix zur Klassifizierung des Herkunftslandes mittels linearer SVM basierend auf NIR-Daten (Gesamtgenauigkeit: 86.9 %). . . . .	113
12	Konfusionsmatrices und Gesamtgenauigkeit zur Klassifizierung des Herkunftslandes im 5-Ländermodell mittels (a) SVM mit linearem Kern, (b) SVM mit RBF-Kern und (c) RF basierend auf NIR-Daten. . . . .	115
13	Konfusionsmatrices und Gesamtgenauigkeit zur Klassifizierung einer deutschen oder ausländischen Herkunft mittels (a) SVM mit linearem Kern und (b) SVM mit RBF-Kern basierend auf NIR-Daten. . . . .	117

14	Klassifikationsergebnisse der Ländervorhersage für die Betrachtung verschiedener Konfidenzbereiche begrenzt durch unterschiedliche Wahrscheinlichkeitswerte $p$ der Klasse der deutschen Proben im SVM-Modell mit linearem und RBF-Kern basierend auf NIR-Daten. . . . .	120
15	Klassifikationsergebnisse der Vorhersage einer deutschen oder ausländischen Herkunft für die Betrachtung verschiedener Konfidenzbereiche begrenzt durch Wahrscheinlichkeitswerte $p$ der Klasse der deutschen Proben im SVM-Modell mit linearem und RBF-Kern basierend auf NIR-Daten. . . . .	121
16	Konfusionsmatrices und Gesamtgenauigkeit zur Klassifikation des Herkunftslandes im 5-Ländermodell der Proben mit hoher Konfidenz mittels (a) SVM mit linearem Kern und (b) SVM mit RBF-Kern basierend auf NIR-Daten. . . . .	123
17	Geräteparameter der finalen Multielementmethode. . . . .	143
18	Konzentrationen der Kalibrierpunkte und abweichende Konzentrationen für die Elemente Arsen, Beryllium, Zink sowie die seltenen Erden; Angaben in $\mu\text{g/L}$ . . . . .	145
19	Geräte- und Messparameter des FT-NIR-Spektrometers TANGO. . . .	147
20	Übersicht der akquirierten deutschen Spargelproben. . . . .	175
21	Übersicht der akquirierten niederländischen Spargelproben. . . . .	182
22	Übersicht der akquirierten ausländischen Spargelproben (ohne NL). . .	183
23	Übersicht der akquirierten Bodenproben. . . . .	186
24	Messparameter der Multielementmethode; ISTD: interner Standard, <i>Integration time per point</i> : Integrationszeit in s pro Messpunkt der dreifachen Messung, Helium: Helium-Modus. . . . .	190
25	Zusammensetzung der ICP-Mehrelementstandardlösung VI (Merck, Darmstadt, Deutschland); die grau unterlegten Elemente wurden nicht in die Multielementmethode einbezogen. . . . .	192
26	Zusammensetzung des Multielement-Kalibrierstandards 2 (PerkinElmer, Shelton, USA); die grau unterlegten Elemente wurden nicht in die Multielementmethode einbezogen. . . . .	193
27	Ergebnisse der Analyse des Referenzmaterials BCR 670 (Wasserlilie); Angaben in $\text{mg/kg}$ , <sup>a</sup> Konzentration in $\mu\text{g/kg}$ , <sup>b</sup> Mittelwert zweier Messwerte.	195
28	Ergebnisse der Analyse des Referenzmaterials DLA 41/2015 (Gerstengras); Angaben in $\text{mg/kg}$ , <sup>a</sup> Konzentration in $\mu\text{g/kg}$ . . . . .	196
29	Ergebnisse der Analyse des Referenzmaterials BCR 667 (Sediment); Angaben in $\text{mg/kg}$ , <sup>b</sup> Mittelwert zweier Messwerte. . . . .	196
30	Validierungsparameter der finalen Multielementmethode – Teil 1: Nachweis- und Bestimmungsgrenzen (in $\text{ng/L}$ ), linearer Bereich (in $\mu\text{g/L}$ ) und Bestimmtheitsmaß der Kalibriergeraden. . . . .	197

---

31	Validierungsparameter der finalen Multielementmethode – Teil 2: Wiederholpräzision und Wiederholbarkeit (in %). . . . .	198
32	Ergebnisse der Laborvergleichsuntersuchung DLA 32/2017; Gehalte in mg/kg. . . . .	200
33	Ergebnisse der Laborvergleichsuntersuchung DLA 46/2017; Gehalte in mg/kg. . . . .	200
34	Mittlere Elementkonzentrationen und Standardabweichungen der Spargelproben in µg/kg; <sup>a</sup> in mg/kg. . . . .	204
35	Konfusionsmatrices und Gesamtgenauigkeit zur 6-Klassen-Klassifikation des Herkunftslandes mittels (a) SVM mit RBF-Kern und (b) RF basierend auf NIR-Daten. . . . .	210
36	Übersicht der verwendeten Chemikalien; GHS-Symbole laut Tabelle 37.	211
37	Definition der GHS-Symbole in Tabelle 36. . . . .	212
38	Übersicht der verwendeten Geräte. . . . .	213
39	Übersicht der verwendeten Materialien. . . . .	214



## Danksagung

Bedanken möchte ich mich an erster Stelle bei meinem Doktorvater Herrn Prof. Dr. Markus Fischer für die spannende Fragestellung, die wissenschaftliche Unterstützung und die Möglichkeit, die Elementanalytik im Arbeitskreis zu stärken.

Meinen beiden Diplomanden, Stephanie Gurk und Deniz Wagner, danke ich für ihre Unterstützung und die intensive Mitarbeit an Teilen dieser Arbeit.

Marc Rurik danke ich für die konstruktive und interdisziplinäre Zusammenarbeit sowie den wertvollen Austausch auf den Gebieten der Chemie und Bioinformatik.

Zudem möchte ich einen Dank richten an alle Mitarbeiter im AK Fischer, die in unzähligen Stunden mühevoller Arbeit bei der Akquise und Bearbeitung der Spargelproben mitgewirkt haben.

Mein Dank gilt des Weiteren Herrn Dr. Philipp Brüning und seinen Kollegen, die mir besonders zu Beginn meiner Arbeit mit der ICP-MS mit Rat und Tat sowie praktischen Tipps zur Seite standen.

Ein besonderer Dank gilt meinen Kollegen im Institut für Lebensmittelchemie, für unterhaltsame Stunden während und nach der Arbeitszeit, darunter besonders Juliane Klare für konstruktive und dennoch nie langweilig werdende Spargelarbeiten sowie meinen Büronachbarn Annika Honold, Alexandra Scharf, Stefanie Schelm und Nicolas Cain für eine heitere Büroatmosphäre, endlose (wissenschaftliche) Diskussionen und ein offenes Ohr auch bei eher fachfremden Fragestellungen.

Bedanken möchte ich mich zudem bei meinen beiden Nachfolgern in der Elementanalytik, Torben Segelke und Kristian von Wuthenau, für intensive, kritische und unterhaltsame Diskussionen rund um die Elemente.

Meiner Familie danke ich für ihre Unterstützung während meiner gesamten Studien- und Promotionszeit und für die Ermutigung und das Bestärken, meinen eigenen Weg zu gehen.

Ein ganz besonderer Dank gebührt meinem Mann für viel Geduld, liebevollen Zuspruch und Rückendeckung, besonders in schwierigen Phasen. Dass du immer an mich geglaubt hast, hat diese Arbeit erst möglich gemacht.



## **Eidesstattliche Versicherung**

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.

Datum, Unterschrift