

**ADDRESSING EMERGING AUTHENTICITY
ISSUES USING NOVEL DATA-DRIVEN ANALYTICS**

Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

an der Fakultät für Biologie, Chemie und Geowissenschaften der

Universität Bayreuth

vorgelegt von

Stefan Gerd Bindereif

geboren in Heilbronn

Bayreuth 2026

Die vorliegende Arbeit wurde in der Zeit von April 2017 bis Februar 2026 in Bayreuth am BayCEER - Labor für Isotopen-Biogeochemie unter Betreuung von Herrn Prof. Dr. Gerhard Gebauer angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.).

Die Dissertation wurde von April 2017 bis Mai 2020 aus Mitteln des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) aufgrund eines Beschlusses des deutschen Bundestages gefördert. Die Projektträgerschaft erfolgte über die Bundesanstalt für Landwirtschaft und Ernährung (BLE) im Rahmen des Programms zur Innovationsförderung (Förderkennzeichen: 2816502414).

Art der Dissertation: Kumulative Dissertation

Dissertation eingereicht am: 25.02.2026

Zulassung durch die Promotionskommission: 04.03.2026

Wissenschaftliches Kolloquium: 22.06.2026

Amtierender Dekan: Prof. Dr. Janosch Hennig

Prüfungsausschuss:

Prof. Dr. Gerhard Gebauer (Gutachter)

Prof. Dr. Johanna Pausch (Gutachterin)

Prof. Dr. Andreas Römpf (Vorsitz)

Prof. Dr. Susanne Baldermann

"[...] potiusque sero quam nunquam [...]" - Liv. 4, 2, 11

Titus Livius (59 BCE – 17 CE)

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES	I
LIST OF ABBREVIATIONS	II
SUMMARY	1
ZUSAMMENFASSUNG.....	5
1 INTRODUCTION	9
1.1 Understanding Authenticity: Origins and Modern Relevance	9
1.2 Cocoa as a Case Study: Authenticity in a Globalized Food System	10
1.3 From Documentation to Verification: The Rise of Analytical Solutions	12
1.4 Stable Isotope Analysis: Foundations and Applications in Authenticity Testing.....	12
1.4.1 The Role of Stable Isotopes in Verifying Agricultural Practices.....	15
1.4.2 Expanding the Scope: Isotopic Techniques for Origin Tracing and Adulteration Detection	16
1.5 Nuclear Magnetic Resonance: From Theory to Real-World Food Authentication.....	18
1.5.1 Beyond Composition: NMR's Evolving Role in Cereal Quality and Authenticity Control	20
1.6 Optical Approaches in Cereal Testing: Near-Infrared Spectroscopy as a Leading Tool ..	21
1.7 Elemental Profiling Across Sectors: Geochemical Tracing in Food and Gemstones	24
1.8 Ensuring Robustness in Authenticity Science: The Role of Data and Validation.....	25
1.8.1 The Growing Need for Robustness in Cereal Authentication.....	26
1.9 Objectives of This Thesis	29
2 SYNOPSIS.....	31
2.1 Foundation: Establishing the Analytical and Statistical Backbone	31
2.1.1 Project Background and Dataset Generation	31
2.1.2 Software Development and Statistical Methods.....	33
2.1.2.1 Manufacturer-Independent Import of Raw Data Sets From Various Sources and Analytical Techniques	33
2.1.2.2 Flexible Handling of Diverse Analytical Data Formats.....	34
2.1.2.3 Comprehensive Data Pre-Processing Capabilities	34
2.1.2.4 Built-In Algorithms for Exploratory and Predictive Data Analysis	38
2.1.2.5 Storage of Models and Associated Processing Parameters for Full Reproducibility	39
2.1.2.6 Automated Generation of Performance Metrics and Cross-Validation Outputs to Evaluate Model Robustness.....	39
2.1.2.7 Workflow Automation with User-Friendly Navigation	41
2.2 Discriminating Harvest Years of Spring Barley Exposed to Extreme Weather.....	41
2.3 Classifying Demantoid Origin Using Chemometric Modeling of LA-ICP-MS Data.....	44
2.4 Differentiating Processing Sites of Barley Malt Using Water Isotope Signatures	46

TABLE OF CONTENTS

2.5	Outlook	49
3	REFERENCES	53
4	AUTHOR CONTRIBUTIONS TO THE PUBLICATIONS AND MANUSCRIPT'	83
5	PUBLICATION LIST.....	86
6	PUBLICATIONS OF THIS THESIS.....	87
6.1	PUBLICATION 1.....	87
	Impact of Global Climate Change on the European Barley Market Requires Novel Multi- Method Approaches to Preserve Crop Quality and Authenticity	87
6.2	PUBLICATION 2.....	101
	Chemometric Modeling of Trace Element Data for Origin Determination of Demantoid Garnets.....	101
6.3	MANUSCRIPT 3	127
	Stable Hydrogen and Oxygen Isotope Abundance Analysis of Barley Malt and Its Potential to Trace the Malting Location	127
7	ACKNOWLEDGEMENTS.....	145
8	(Eidesstattliche) Versicherungen und Erklärungen	147

LIST OF TABLES AND FIGURES

Table 1 Key attributes defining authenticity in commercial goods	9
Figure 1 Schematic illustration of the project workflow.....	33
Figure 2 Chemometric analysis of combined FT-NIR and IRMS data from two barley varieties.	43
Figure 3 PCA-LDA score plot with all classes (countries) including the log-transformed concentrations of the six relevant elements Mg, Al, Ti, V, Cr and Mn.	45
Figure 4 Line plots showing the natural (A) hydrogen and (B) oxygen isotope abundances of the analysed barley samples and malt samples from two malthouses (Munich and Berlin) and results of the multiple paired t-tests.....	48

LIST OF ABBREVIATIONS

LIST OF ABBREVIATIONS

AACC	American Association of Cereal Chemists
AI	Artificial intelligence
AOAC	Association of Official Agricultural Chemists
ASBC	American Society of Brewing Chemists
C	Combustion
CAM	Crassulacean acid metabolism
cv.	Cultivar
FT	Fourier transform
GC	Gas chromatography
<i>H.</i>	<i>Hordeum</i>
HRMS	High-resolution mass spectrometry
ICC	International Association for Cereal Science and Technology
ICCO	International Cocoa Organization
ICP	Inductively coupled plasma
IRMS	Isotope-ratio mass spectrometry
L.	Linné
LA	Laser ablation
LC	Liquid chromatography
LDA	Linear discriminant analysis
MHz	Megahertz
MS	Mass spectrometry
m/z	mass-to-charge ratio
NIR	Near-infrared
NMR	Nuclear magnetic resonance
PCA	Principal component analysis
VIP	Variable importance in projection

SUMMARY

As globalization continues to grow, trade structures are becoming increasingly complex, which creates numerous challenges for businesses. One of these is ensuring the authenticity of high-priced products. Recent developments have made it clear that established control mechanisms are often inadequate for ensuring protection against fraud or mislabeling, so the need for new analytical solutions is growing rapidly. In its individual studies, this dissertation examines the application of modern data-driven approaches for the authentication of various goods of agricultural and non-agricultural origin. At the heart of all investigations is a software that was developed specifically for this research project but, thanks to its modular design, has great potential for future applications in diverse fields. It enables the automated processing, modeling, and evaluation of spectroscopic and spectrometric data. To date, there has been no solution that enables streamlined analysis of data from, for example, isotope ratio mass spectrometry (IRMS), nuclear magnetic resonance spectroscopy (NMR), near-infrared spectroscopy (NIR), and the associated metadata in a fully reproducible and transparent manner. One area of investigation focused on the authentication of spring barley with respect to the harvest year. Grain samples were taken directly from active trade flows and then analyzed to ensure that they reflected realistic conditions in the supply chain, in contrast to numerous studies that relied on samples from strictly controlled experimental fields and are therefore hardly transferable to real-world conditions. The aim was to develop a novel analytical approach that allows differentiation of the two harvest years 2017 and the extremely dry 2018 for spring barley samples by means of multiple modern analytical techniques and to link the observed metabolic differences to the physiological responses of plants to drought stress. Samples from the 2018 harvest year showed significantly less negative $\delta^{13}\text{C}$ values than those from 2017, likely reflecting reduced stomatal conductance and altered carbon assimilation during the grain-filling stage. These findings were supported by NIR, which showed lower protein, starch, and moisture content and an increase in lipid content for 2018. NMR analysis revealed elevated concentrations of the osmolyte betaine, a

SUMMARY

well-known biochemical marker for abiotic stress adaptation. By merging the isotope and NIR datasets and applying multivariate statistical analysis, a classification model was built that achieved 100% separation accuracy between the two investigated years. Although not directly linked to authenticity, differences in variety-specific stress responses were also observed, demonstrating the broader applicability of the approach for selecting stress-resilient varieties in the future. Subsequently, the established chemometric framework for authenticity testing was adapted and transferred to a fundamentally different sample matrix, namely gemstones. In this context, demantoid garnets were analyzed, a material whose market value is strongly influenced by its geographical origin. A total of 136 samples from six different countries were analyzed using laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS), with the aim of distinguishing origin based on trace element profiles. Logarithmic transformation of the recorded data proved critical for appropriate statistic modeling, enabling clear separation for five of the six countries of origin with classification accuracies being especially high for samples from Namibia, Pakistan, and Madagascar. Analysis of variable importance identified the elements Al and Cr as key discriminants across most models. Importantly, this work was among the first in gemological research to comprehensively address issues of database design, data preprocessing, and model validation, providing a reproducible and transparent workflow that can serve as a blueprint for future studies. In response to a specific industry need in the food sector, the final study of this dissertation was initiated with the intent to trace the processing location of barley malt, an important intermediate product in beer production. Stable isotope analysis of hydrogen and oxygen was used to investigate whether water isotope signatures in malt reflect the location of the malting facility. Therefore, thirty-four barley samples were processed in two different malthouses located approximately 500 km apart. Using IRMS analysis, significant differences in $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values between malts from the two investigated locations were observed with all malt samples showing significantly less negative $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values compared to the raw material barley. It is the first study to document systematic shifts in isotope ratios during the malting process.

SUMMARY

Notably, the direction of these differences matched expected regional trends in precipitation isotope values, supporting the hypothesis that local water sources influence the final isotopic signature of the malt. The work illustrates that authenticity verification, when grounded in state-of-the-art analytics and supported by rigorous data science, can evolve into a powerful, flexible toolset applicable across industries. Such capabilities are becoming ever more critical in a world where supply chain complexity and the demand for verifiable product integrity continue to grow.

ZUSAMMENFASSUNG

Mit dem fortschreitenden Prozess der Globalisierung werden Handelsstrukturen zunehmend komplexer, was für Unternehmen zahlreiche Herausforderungen mit sich bringt. Eine davon ist die Sicherstellung der Authentizität hochpreisiger Produkte. Jüngste Entwicklungen haben gezeigt, dass etablierte Kontrollmechanismen häufig nicht ausreichen, um einen wirksamen Schutz vor Betrug oder Falschdeklaration zu gewährleisten. Der Bedarf an neuen analytischen Lösungen wächst daher rasant. Diese Dissertation untersucht in ihren Einzelarbeiten die Anwendung moderner, datengetriebener Ansätze zur Authentifizierung verschiedener Waren landwirtschaftlichen und nicht-landwirtschaftlichen Ursprungs. Im Zentrum aller Untersuchungen steht eine eigens für dieses Forschungsprojekt entwickelte Software, die dank ihres modularen Aufbaus großes Potenzial für zukünftige Anwendungen in unterschiedlichen Bereichen bietet. Sie ermöglicht die automatisierte Verarbeitung, Modellierung und Auswertung spektroskopischer und spektrometrischer Daten. Bislang existierte keine Lösung, die eine effiziente, reproduzierbare und transparente Analyse von Daten aus beispielsweise der Isotopenverhältnis-Massenspektrometrie (IRMS), der Kernspinresonanzspektroskopie (NMR), der Nahinfrarotspektroskopie (NIR) sowie den zugehörigen Metadaten ermöglicht. Ein Schwerpunkt der Untersuchungen lag auf der Authentifizierung von Sommergerste im Hinblick auf das Erntejahr. Die Getreideproben wurden direkt aus laufenden Handelsströmen entnommen, um realistische Bedingungen innerhalb der Lieferkette abzubilden, im Gegensatz zu zahlreichen Studien, die auf Proben aus streng kontrollierten Versuchsflächen beruhen und daher nur eingeschränkt auf die Praxis übertragbar sind. Ziel war es, einen neuartigen analytischen Ansatz zu entwickeln, der es ermöglicht, Sommergerstenproben der beiden Erntejahre 2017 und dem extrem trockenen Jahr 2018 mithilfe moderner analytischer Techniken zu unterscheiden und die beobachteten metabolischen Unterschiede auf physiologische Reaktionen der Pflanzen auf Trockenstress zurückzuführen. Die Proben aus dem Erntejahr 2018 zeigten deutlich weniger negative $\delta^{13}\text{C}$ -Werte als jene aus 2017, vermutlich durch eine verringerte stomatare Leitfähigkeit

ZUSAMMENFASSUNG

und eine veränderte Kohlenstoffassimilation während der Kornfüllungsphase. Diese Befunde wurden durch die NIR-Analyse gestützt, die für 2018 niedrigere Gehalte an Protein, Stärke und Feuchtigkeit sowie einen erhöhten Lipidgehalt zeigte. Die NMR-Analyse offenbarte zudem erhöhte Konzentrationen des Osmolyten Betain, ein bekannter biochemischer Marker für die Anpassung an abiotischen Stress. Durch die Kombination der Isotopen- und NIR-Daten und die Anwendung multivariater statistischer Verfahren konnte ein Klassifikationsmodell erstellt werden, das eine vollständige Trennung der beiden Erntejahre mit 100 %iger Genauigkeit erreichte. Obwohl dies nicht unbedingt mit Authentizitätsfragen verknüpft ist, wurden auch sortenspezifische Unterschiede in der Stressantwort festgestellt, die auf eine breitere Anwendbarkeit des Ansatzes, etwa zur Auswahl stresstoleranter Sorten in der Zukunft, hindeuten. Im nächsten Schritt wurde das etablierte chemometrische Rahmenwerk zur Authentizitätsprüfung auf eine grundlegend andere Probenmatrix übertragen, nämlich auf Edelsteine. In diesem Zusammenhang wurden Demantoid-Granate analysiert, deren Marktwert stark vom geographischen Ursprung abhängt. Insgesamt wurden 136 Proben aus sechs verschiedenen Ländern mittels Laserablation-induktiv gekoppelter Plasma-Massenspektrometrie (LA-ICP-MS) untersucht, mit dem Ziel, die Herkunft anhand der Spurenelementprofile zu differenzieren. Die logarithmische Transformation der erfassten Daten erwies sich als entscheidend für eine geeignete statistische Modellierung und ermöglichte eine klare Trennung für fünf der sechs Herkunftsländer, wobei die Klassifikationsgenauigkeit insbesondere für Proben aus Namibia, Pakistan und Madagaskar sehr hoch war. Die Analyse der Variablenbedeutung identifizierte die Elemente Al und Cr als zentrale Unterscheidungsmerkmale in den meisten Modellen. Besonders hervorzuheben ist, dass diese Arbeit zu den ersten in der gemmologischen Forschung zählt, die Fragen des Datenbankdesigns, der Datenvorverarbeitung und der Modellvalidierung umfassend adressiert und damit einen reproduzierbaren und transparenten Workflow bereitstellt, der als Blaupause für zukünftige Studien dienen kann. Auf einen konkreten Bedarf aus der Lebensmittelbranche reagierend, wurde die letzte Studie dieser Dissertation mit

ZUSAMMENFASSUNG

dem Ziel initiiert, den Verarbeitungsort von Gerstenmalz - einem wichtigen Zwischenprodukt in der Bierherstellung - zu bestimmen. Durch Stabilisotopenanalysen von Wasserstoff und Sauerstoff sollte ermittelt werden, ob die im Malz nachweisbaren Isotopensignaturen das verwendete Wasser widerspiegeln und dadurch Rückschlüsse auf den Standort der Mälzerei zulassen. Zu diesem Zweck wurden 34 Gerstenproben in zwei Mälzereien verarbeitet, die rund 500 km voneinander entfernt liegen. Die IRMS-Analyse zeigte signifikante Unterschiede in den $\delta^2\text{H}$ - und $\delta^{18}\text{O}$ -Werten zwischen den Malzproben der beiden Standorte. Zudem wiesen alle Malzproben im Vergleich zur Ausgangsgerste deutlich weniger negative Werte auf. Es handelt sich um die erste Studie, die systematische Verschiebungen der Isotopenverhältnisse während des Mälzprozesses dokumentiert. Bemerkenswert ist, dass die Richtung dieser Verschiebungen mit den erwarteten regionalen Trends der Niederschlagsisotopie übereinstimmte, was ein deutlicher Hinweis darauf ist, dass lokale Wasserquellen die finale Isotopensignatur des Malzes beeinflussen.

Die Arbeit zeigt, dass Authentizitätsprüfung, gestützt auf modernste Analytik und fundierte Datenwissenschaft, zu einem leistungsfähigen, flexiblen Werkzeugkasten heranwachsen kann, der branchenübergreifend Anwendung findet. Solche Fähigkeiten werden in einer Welt, in der Lieferketten immer komplexer werden und die Nachfrage nach überprüfbarer Produktintegrität stetig steigt, zunehmend unverzichtbar.

1 INTRODUCTION

1.1 Understanding Authenticity: Origins and Modern Relevance

What actually is authenticity? Mankind has been dealing with this admittedly rather philosophical question for many decades, if not centuries. And yet, no clear and uniform definition can be found. Depending on what is to be described by the term, there are the most diverse interpretations (Carroll, 2015). Nonetheless, there is widespread consensus that something is considered “authentic” (from Ancient Greek: *authentikós*) if it is genuine, meaning it does not deviate from the original. In other words, when people seek authenticity, they look for consistency with established expectations or standards. To better understand what this means in the context of commercial goods, it is helpful to know the main attributes commonly associated with it (Table 1).

Table 1 Key attributes defining authenticity in commercial goods

Authenticity attribute	Description	Literature example(s)
Geographical origin	Association with a specific region (e.g., Champagne, Parmigiano Reggiano, etc.)	(Charters & Spielmann, 2014; Cozzi et al., 2019)
Production method	Unique or traditional production process (e.g., Haymilk).	(Paredes et al., 2018)
Purity	Genuine, unaltered ingredients or materials without substitution or adulteration (e.g., honey, fruit juices).	(García & Enghelmayer, 2025; L. Xu et al., 2022)
Brand reputation	Established businesses or producers with a reputation for quality and consistency.	(Napoli et al., 2014, 2016)

INTRODUCTION

Certifications or labels	Third-party certifications that verify authenticity attributes (e.g., Marine Stewardship Council certification standard program for fisheries products)	(Cundy et al., 2023; de Melo et al., 2024)
Scarcity	Limited availability or exclusivity that enhances the perceived uniqueness and value of the product (e.g., argan oil or tanzanite).	(Mohammed et al., 2021; Vincent et al., 2021)
Sustainability	Use of environmentally responsible practices and fair labor standards that contribute to the product's integrity and consumer trust (e.g., specialty coffee, craft chocolate).	(Barreto Peixoto et al., 2023; Cadby et al., 2021)

1.2 Cocoa as a Case Study: Authenticity in a Globalized Food System

Among the authenticity attributes listed, some are especially prominent in the context of food products. Cocoa beans are not only the most critical raw product for chocolate production, but they also perfectly illustrate how origin, production practices, and ethical considerations influence the consumers' perceptions of authenticity.

Cocoa (*Theobroma cacao* L.) beans are one of the world's most important tropical crops. The vast majority of cocoa trees are exclusively cultivated by small-scale family farmers located in West Africa and South America. While cocoa is mainly traded on the commodity exchanges in London and New York, not all cocoa is valued equally. Some of this is due to the quality of the beans, which is heavily influenced by standard cocoa production methods, including fermentation and

INTRODUCTION

drying techniques (Beckett et al., 2017). Likewise, transparent sourcing and the use of certifications signal integrity along the supply chain, increasing consumer trust and competitive advantage. Increasingly, traceability has become a key marker of authenticity, as buyers and consumers seek clarity about where and how cocoa was produced, by whom, and under what conditions. The International Cocoa Organization (ICCO, <https://www.icco.org/>), which is a global organization with members from both cocoa-producing and cocoa-consuming countries, distinguishes between high-quality "fine or flavor" cocoa beans and the regular "bulk" or "ordinary" cocoa beans. Countries whose cocoa has been evaluated by an official ICCO panel can receive approval to produce "fine or flavor" cocoa and subsequently charge significantly higher prices for their goods (Bindereif et al., 2019). From this point in time, the cocoa's origin drastically changes the wages of farmers, the margins of middlemen and processors, but also the price that consumers ultimately pay when buying chocolate (Cadby & Araki, 2021). Even within the fine-flavor cocoa sector, the exclusivity of particularly rare cultivars can lead to significantly higher market values, positioning them at the very top end of this premium category. The cocoa cultivar Criollo Porcelana, grown in Venezuela, is a perfect example of this phenomenon: its low yield and limited geographical distribution make it one of the most sought-after varieties in the world (Brunetto et al., 2020; Lachenaud & Motamayor, 2017). The ICCO program also tries to promote fair labor standards, supports farmer education and technical training, and encourages independence from middlemen through the creation of direct-trade markets. Moreover, buyers of "fine or flavor" cocoa are usually more environmentally conscious, incentivizing sustainable practices that preserve biodiversity, improve soil health and limit the use of agrochemicals. Having said that, there is still a long way to go, as farmers remain exposed to the influence of intermediaries and international distributors who continue to control market access (Salazar et al., 2023). Consequently, authenticity now carries connotations of social responsibility, and products sourced from regions associated with labor exploitation may face reputational risks (Iyer & Kalita, 1997). In the worst-case scenario, the origin of raw materials can lead to lasting damage to

the reputation of major international businesses. This issue is not exclusive to the cocoa industry and similar patterns are observed in other global sectors. For example, many well-known fashion brands have faced criticism for sourcing textiles from countries with widespread poor working conditions and human rights violations. The media attention is enormous, and voices are quickly raised that wish for more transparency (Fraser & van der Ven, 2022; Islam & Deegan, 2010).

1.3 From Documentation to Verification: The Rise of Analytical Solutions

It follows that the authenticity of products must be ensured at any point in the supply chain in order to protect all parties involved. At the same time, this very task is becoming an increasingly difficult challenge in our globalized world (Aung & Chang, 2014; L. Li, 2013). In many cases, companies rely largely on documents provided by intermediaries or contractors, but such records offer only limited and indirect insight; ultimately, authenticity must be measurable and verifiable. Fortunately, the scientific community, consisting of academia and the official bodies, but also commercial laboratories and equipment manufacturers, has been continuously searching for solid analytical solutions to authenticity issues for decades and today provides a plethora of valuable tools (Böhme et al., 2019; Hong et al., 2017; Melcher et al., 2021). These developments are not limited to theoretical research; many of the resulting techniques have been successfully translated into applied settings across a variety of industries. Accredited governmental laboratories increasingly rely on scientific testing methods once they have undergone formal validation and have been standardized through recognized procedures (O'Mahony et al., 2024). These steps are fairly time-consuming but essential as the techniques should not only be scientifically sound but also compliant with legal requirements.

1.4 Stable Isotope Analysis: Foundations and Applications in Authenticity Testing

One technology has proven to be particularly powerful in the field of authenticity testing (Picó & Barceló, 2023) and is still regularly considered the “gold standard”: the determination of stable isotope ratios, which has a long and impressive history of development. The British scientist and later Nobel Prize winner Frederick Soddy laid the foundation more than 100 years ago when he

INTRODUCTION

introduced the concept of isotopes in 1913 and recognized that most chemical elements in nature do not occur as pure isotopes, but as mixtures of stable nuclides which differ in their atomic mass but have the same chemical properties (Wolfsberg et al., 2009). The difference in mass between isotopes arises because they have the same number of protons (which defines the element) but differ in the number of neutrons in the atomic nucleus. Once the necessary technical prerequisites were established, research into the relative abundance of isotopes and mean stable isotope ratios for elements was initiated (Nier, 1950; Rossmann, 2001). While research in the first years concentrated on unstable isotopes (radioisotopes) such as uranium and thorium for age determination in geosciences, this changed from the 1950s onwards and the light bioelements hydrogen ($^2\text{H}/^1\text{H}$), carbon ($^{13}\text{C}/^{12}\text{C}$), nitrogen ($^{15}\text{N}/^{14}\text{N}$), oxygen ($^{18}\text{O}/^{16}\text{O}$) and sulphur ($^{34}\text{S}/^{32}\text{S}$) came to the fore (Craig, 1953; Emiliani, 1955; Epstein & Mayeda, 1953; I. Friedman, 1953; Hoering, 1955; McCrea, 1950; Thode et al., 1949). Despite numerous developments and optimizations, the instruments used today for isotope analysis of light elements are based on the same principles developed back at that time (Akoka & Remaud, 2020; Jézéquel et al., 2017). In isotope ratio mass spectrometry, the sample material to be analyzed is converted into simple measuring gases (H_2 , CO_2 , N_2 , SO_2), which are then isolated and finally transferred to a specialized mass spectrometer using the carrier gas helium (Bigeleisen et al., 1952; H. Craig, 1953; Kelly et al., 2018; Preston, 1992). This is achieved by combustion (Brand et al., 2014), pyrolysis (Gehre & Strauch, 2003), or chemical oxidation (Krummen et al., 2004), depending on the sample type and the isotopes being analyzed. In the mass spectrometer, the measuring gases are ionized in the ion source and then accelerated into a magnetic field, which separates them based on their mass-to-charge ratio (m/z). Finally, the separated ions are directed to detectors (Faraday cups), where their relative abundances are measured (Brown & Tautfest, 1956; Kelly et al., 2018). The mass spectrometer detects the ion ratios corresponding to different isotopic forms (isotopologues) of the gases. For instance, in carbon isotope ratio analysis, the mass spectrometer simultaneously monitors ions with mass-to-charge ratios (m/z) of 44, 45, and

INTRODUCTION

46, representing ions from CO₂ molecules that contain ¹²C, ¹³C, ¹⁶O, ¹⁷O, and ¹⁸O in different combinations (Muccio & Jackson, 2009). This precise monitoring allows scientists to detect even small differences in isotopic ratios, which can reveal important information about the history and origin of a sample. Variations in natural isotope abundances are typically expressed in per mil (‰) using the so-called delta (δ) notation, which compares the isotope ratio of a sample to an internationally agreed-upon standard reference (zero-point) (Brand et al., 2014; Coplen, 2011). It has several advantages over the use of absolute values, including sensitivity to minimal differences and better interpretability. International standards are also used to ensure consistency and comparability of isotope ratio measurements across different laboratories and instruments (Meier-Augenstein, 2017). The isotopic abundances of the named elements were fixed during the formation of the Earth and have stayed unchanged globally since then, but subtle variations in the natural isotopic composition of various sample matrices are introduced through so-called fractionation processes. These can be of a biological, chemical, or physical nature. Through many years of research into these processes, we are now able to interpret them accurately and understand their significance (Hoefs, 2021). For example, because of the pioneering work on carbon isotope discrimination during photosynthesis, we know how and why plants using different photosynthetic pathways, namely C₃ plants, C₄ plants, and plants utilizing the crassulacean acid metabolism (CAM) differ in their natural carbon isotope compositions (Cerling et al., 1997; Cernusak et al., 2013; Farquhar et al., 1989; G. Schmidt et al., 1993). However, not only does the biomass of the plants themselves carry a photosynthetic pathway-dependent signature, but also the products obtained from these plants (Meier-Augenstein, 2017). As the technology advanced, IRMS found broader applications in testing various food products, such as olive oil (Camin et al., 2010), dairy products (Bontempo et al., 2011; Brescia et al., 2005; Perini et al., 2022), and meats (Boner & Förstel, 2004; Heaton et al., 2008; Horacek & Min, 2010). Scientists around the globe quickly discovered that the recorded isotopic fingerprints of these

products contain valuable information about their geographical origin, production methods, and potential adulteration (Camin et al., 2016; Kelly et al., 2005).

1.4.1 The Role of Stable Isotopes in Verifying Agricultural Practices

Especially in the field of cereal science, global interest in solutions based on stable isotope analysis is increasing (Li et al., 2016; Liu et al., 2023). For example, the nitrogen isotope signatures ($\delta^{15}\text{N}$) of plants are one of the key indicators for organic farming (Giannioti, Ogrinc, et al., 2024; Laursen et al., 2013; Paolini et al., 2015). This is largely related to the main nitrogen source of the plants, i.e. in many cases the fertilizer used. In conventional agriculture, predominantly synthetic fertilizers are applied (Penuelas et al., 2023). Nitrogen in these fertilizers is obtained via the Haber-Bosch process, i.e. by the chemical fixation of atmospheric nitrogen (N_2), followed by a conversion to ammonia (NH_3) (Haber & Le Rossignol, 1913). The process does not lead to any significant fractionation processes and consequently the $\delta^{15}\text{N}$ of the fertilizers is close to the value of atmospheric nitrogen ($\delta^{15}\text{N}_{\text{atm.}} = 0\text{‰}$) (Bateman & Kelly, 2007; Choi et al., 2017). Organic farming differs mainly in that pesticides and synthetic fertilizers are prohibited, so organic fertilizers such as animal manure, compost, or green manure are used. As there is a strong isotopic discrimination of the heavy nitrogen isotope ^{15}N due to N-transformation processes prior to plant uptake (ammonia volatilization, denitrification, nitrification etc.), plants supplied with animal manure and compost usually show higher $\delta^{15}\text{N}$ values (Choi et al., 2017; Laursen et al., 2013). Nonetheless, numerous publications in the field of agricultural science repeatedly show that this supposedly clear connection between the use of different fertilizers and the nitrogen isotope signature of the resulting plant material is anything but reliable and that the situation appears to be much more complex, as an extremely large number of factors have an influence (Inácio et al., 2015; Schmidt et al., 2005). To cope with this complexity, cereal research is increasingly moving towards multi-method approaches and combining nitrogen isotope data with other promising technologies to eventually achieve a reliable approach to differentiate between cereals from organic and conventional farming

INTRODUCTION

(Lozano-Castellón et al., 2024). Recently, a working group from Italy used a multi-method approach to differentiate between wheat samples from conventional and organic cultivation (Giannioti, Suman, et al., 2024). As a first step, they analyzed wheat grain samples for bulk nitrogen, carbon, and sulfur isotope values and the isotopic compositions of several amino acids by gas chromatography isotope ratio mass spectrometry (GC-C-IRMS). The samples were also analyzed for the presence of mycotoxins and pesticides. To cover the whole production chain, the grain samples were subsequently milled into flour and then processed into baked goods and pasta, which were also analyzed using the methods stated above. The results from the nitrogen isotope testing showed the aforementioned trend with more positive $\delta^{15}\text{N}$ values in organic wheat grains compared to the conventional samples. However, even in this study there were numerous samples from organic cultivation that showed values close to 0‰, which resulted in a large overlap with the samples from conventional cultivation. The authors concluded that a differentiation was only made possible by the addition of pesticide data, as crop-protecting agents were exclusively detected in conventional samples. Interestingly, a significant reduction of mycotoxins and pesticides was shown in the processed products, while nearly all isotope values remained unchanged, with very few amino acids showing deviations. This confirms the strength of multi-method approaches, especially when aiming to investigate entire production chains. While pesticide data enabled separation by agricultural cultivation method for cereal grains, this information was lost through the processing steps, and isotope abundances were better suited due to their stability.

1.4.2 Expanding the Scope: Isotopic Techniques for Origin Tracing and Adulteration Detection

Oxygen and hydrogen isotope abundances are also frequently used in combination with other techniques in cereal science, as they are a powerful tool for determining the geographic origin of plant material or products derived from it. $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values in water vary geographically due to environmental factors like temperature, altitude, latitude, and proximity to bodies of water

INTRODUCTION

(Carter & Chesson, 2017). As water undergoes cycles of evaporation, condensation, and precipitation, fractionation occurs and lighter isotopes preferentially evaporate, while heavier isotopes condense more readily (Hoefs, 2021). When plants then use this water, they incorporate the corresponding isotope signature into their biomass. However, plant material does not always directly reflect the local water isotope signature due to additional processes in the soil and within the plant that can alter the isotopic abundances (Barbour, 2007; Gebauer et al., 2016; Ziegler, 1989). Nietner et al. (2014) utilized these relationships in a comprehensive study to determine the origin of distillers dried grains and solubles, a cereal by-product of alcohol production commonly used as a protein-rich feed in livestock farming. Using a multi-element approach ($^2\text{H}/^1\text{H}$, $^{13}\text{C}/^{12}\text{C}$, $^{15}\text{N}/^{14}\text{N}$, $^{18}\text{O}/^{16}\text{O}$, and $^{34}\text{S}/^{32}\text{S}$ isotope ratios), samples from the USA, China, and the EU were successfully differentiated, with hydrogen and oxygen abundances being the decisive variables in the course of the statistical discriminant analysis. Another noteworthy example of the application of stable isotope techniques in food authentication is the testing of honey for adulteration (Ziegler et al., 1977). Unlike many analytical methods in the field, which are still under development or lack formal regulatory acceptance, the use of isotope ratio mass spectrometry (IRMS) in this context stands out as it is officially recognized by the Association of Official Agricultural Chemists (AOAC) method 998.12 and can be used in court proceedings (Rogers et al., 2013; White, 1992; White & Winters, 1989). A European Directive (2001/110/EC) defines honey as “the natural sweet substance produced by *Apis mellifera* bees from the nectar of plants or from secretions of living parts of plants or excretions of plant-sucking insects on the living parts of plants, which the bees collect, transform by combining with specific substances of their own, deposit, dehydrate, store and leave in honeycombs to ripen and mature.” Unfortunately, honey is frequently adulterated with low-cost sugar syrup to illegally increase the margin. Of course, this means that the product no longer meets the official definition of honey and is therefore not considered marketable (García & Schwarzingner, 2021).

INTRODUCTION

Most honeys available on the market are mainly based on C_3 plants, thus their sugars but also pollen have a $\delta^{13}C$ between -22‰ and -33‰, where $\delta^{13}C$ is a measure of the isotopic ratio of the stable carbon isotopes ^{13}C and ^{12}C with respect to an internationally recognized standard. When a honey is adulterated with a sugar syrup made from C_4 plants, which naturally show a $\delta^{13}C$ between -10‰ and -20‰, there is a measurable difference between the isotopic signature of the bulk honey and its pollen. A difference of 1‰ between honey and pollen corresponds to a C_4 sugar content of about 7 %, which simultaneously is the internationally tolerated limit to consider a honey as being genuine (White, 1992). As this detection method for C_4 sugar syrups has been established for several decades, many fraudulent activities have shifted towards the use of C_3 sugar syrups (e.g., from rice or sugar beet), which are difficult to detect using IRMS. Often the suppliers of the syrups even advertise that the products are guaranteed to pass the official IRMS test. There is a frequently used advancement of the AOAC 998.12 method, in which the sugars contained in honey are separated by liquid chromatography before the carbon isotope signature is measured (LC-C-IRMS), but the added value of this approach is questionable (Cabañero et al., 2006; Elflein & Ræzke, 2008). In addition, the detection of adulteration using this technique is oftentimes based on the presence of oligosaccharides instead of differences in carbon isotope values, and various syrups are detected only in quantities that are not relevant for the honey market (Aries et al., 2021; QSI – Quality Services International, 2025). No matter what type of carbon isotope analysis is performed in the end, it is not sufficient in any case to assess the overall authenticity of a honey sample (Biswas et al., 2023). Therefore, complementary tests are necessary.

1.5 Nuclear Magnetic Resonance: From Theory to Real-World Food Authentication

Another advanced analytical method that has gained increasing attention in the field of authenticity testing is nuclear magnetic resonance (NMR) spectroscopy. The theoretical groundwork for this technique was done by a physicist named Isidor Isaac Rabi. He first demonstrated the magnetic resonance of atomic nuclei in a molecular beam and finally

INTRODUCTION

introduced the basic principles of the technique back in 1938 (Rabi et al., 1938). For this work, he was awarded the Nobel Prize in Physics in 1944. Building on these pioneering findings, two scientists independently performed the first NMR experiments in condensed matter in 1946. With that, Felix Bloch (Bloch, 1946), and Edward M. Purcell (Purcell et al., 1946) marked the formal birth of NMR spectroscopy as we know it today. Just like Rabi, they were jointly awarded the Nobel Prize in Physics in 1952 for their important contributions to the development of this influential analytical method (Giunta & Mainz, 2020). In the following years, it evolved from a basic technique for evaluating molecular structures in organic chemistry to a high-throughput method that enables non-targeted screening in complex matrices. For almost all types of available NMR testing, the basic principle lies in the behavior of atomic nuclei with an odd number of protons or neutrons, which therefore possess a property called spin. When placed in an external magnetic field, such nuclei align with or against it (Günther, 2013). Through irradiation with radiofrequency pulses, those nuclei absorb energy and are promoted to a higher energy state, but as soon as the irradiation stops, they relax back to their original state. While doing so, they emit electromagnetic signals that can be detected and transformed into NMR spectra. These provide valuable information about the molecular environment of the nuclei that can be used to identify and quantify a wide range of compounds (Friebolin, 2013). Until today, major advancements in instrumentation, such as Fourier-transform processing, pulse sequence optimization, as well as the introduction of superconducting magnets, have significantly increased spectral resolution, sensitivity, and speed of measurements (Freeman, 1993; Wikus et al., 2022). Initially confined to academic research, NMR quickly found its way into applied routine authenticity testing, including the analysis of juices (Spraul et al., 2009), wines (Solovyev et al., 2021), oils (Maestrello et al., 2022) and also honey (Kuballa et al., 2018; Spiteri et al., 2015).

1.5.1 Beyond Composition: NMR's Evolving Role in Cereal Quality and Authenticity Control

In the field of cereal analytics, the vast majority of publications on NMR-based research is focused on metabolic profiling, which aims to investigate biochemical compositions rather than addressing questions of authenticity (Coulomb et al., 2015; Glawischnig et al., 2002; Shewry et al., 2017). As an example, Poudel et al. (2021) looked for compositional changes in wheat cultivars that have been released between 1870 and 2013. They wanted to determine how wheat cultivation over more than a century has affected grain metabolites. They analyzed extracts of flour from 21 cultivars grown in two harvest years by ^1H NMR spectroscopy. In modern cultivars, the working group detected significant increases in the concentrations of various metabolites such as tryptophan, sum of the measured amino acids, chlorogenic acid, ferulic acid, vanillic acid, and sum of the measured phenolic acids, while protein content declined. In addition, they performed complex multivariate statistical analyses that enabled separation between historical and modern cultivars. The study demonstrated NMR's capability to track long-term biochemical shifts in plant material. At the same time, the application of NMR for the authenticity testing of cereal and cereal products remains very limited. One of the main reasons for the slow gain of traction is that sample preparation for NMR testing of cereal samples is elaborate. Unlike liquid or semi-liquid foods that often can be analyzed directly or with minor preparation, cereal products require robust and detailed extraction protocols to obtain suitable liquid-phase samples and reproducible results. Such standardized and validated extraction procedures for cereal matrices are still lacking, complicating cross-study comparability (Deborde et al., 2019). This could be circumvented by the use of solid-state NMR, but although technically possible, the technique is not yet routinely applied in food authenticity testing due to factors like high initial investment cost and a more complex data evaluation (Freitas et al., 2023; Nowacka-Perrin et al., 2022).

While not solving the issues regarding sample preparation, recent developments in low-field benchtop technology may contribute to overcoming some limitations in NMR cereal testing.

These compact instruments that operate at lower magnetic field strengths, like 60 or 80 MHz, offer several key advantages. Those include significantly reduced initial and running costs, less expert knowledge needed to operate, and low infrastructure requirements (Salimi Khorshidi et al., 2018). While the spectral resolution of high-field NMR systems is superior, the potential of benchtop NMR for rapid screening due to portability and on-site analysis makes them particularly attractive for industrial applications. As protocols and chemometric models tailored to cereal matrices continue to develop, low-field NMR is likely to gain increasing relevance in routine authenticity testing of cereal products (Gautam et al., 2025).

1.6 Optical Approaches in Cereal Testing: Near-Infrared Spectroscopy as a Leading Tool

Although analytical techniques such as low-field NMR are improving, optical methods such as Raman spectroscopy (Kniese et al., 2021) and hyperspectral imaging (Femenias et al., 2022) currently still dominate when it comes to cereal authenticity testing (Badaró et al., 2022; Zareef et al., 2021). Their omnipresence in this field is mainly attributed to several key advantages, including rapid results, comparatively low instrument cost, and minimal or no need for sample preparation. One optical technique has become particularly prominent, which is near-infrared (NIR) spectroscopy (Zhang et al., 2022). The emergence of this popular technology can be traced back to the beginning of the 20th century. The physicist William W. Coblentz conducted pioneering studies on the absorption of electromagnetic radiation in the near-infrared region at Cornell University in Ithaca, New York, USA (Coblentz, 1906, 1949). Despite his early discoveries, technical constraints led to the technique remaining almost irrelevant throughout the first half of the century, during which it was almost exclusively used for fundamental research. It was not until the middle of the 20th century that several improvements in electronics and optical components marked a major turning point (Hindle & Ciurczak, 2021). While working for the United States Department of Agriculture in the 1960s and 1970s, Karl Norris played a central role in demonstrating the great potential of NIR spectroscopy in the areas of food science and

INTRODUCTION

agricultural analysis (Ben-Gera & Norris, 1968; Hindle & Ciurczak, 2021). Today, many internationally recognized standard operating procedures for cereal analysis are based on this technique, including the American Society of Brewing Chemists (ASBC) Barley-5 for the measurements of moisture content in whole grains, AOAC Method 997.06-2005 for crude protein in wheat, as well as multiple American Association of Cereal Chemists (AACC) and International Association for Cereal Science and Technology (ICC) guidelines that describe NIR-based routine analyses. These are just a few of the many examples that reflect the widespread acceptance of NIR for both quality control and compliance testing in the cereal industry (Squeo et al., 2024). NIR spectroscopy examines the absorption of electromagnetic radiation in the range of wavelengths between approximately 780 and 2500 nanometers. The energy absorption is related to overtones and combination bands of specific molecular vibrations. This particularly involves bonds such as C–H, O–H, and N–H (Siesler, 2007), which are vibrational modes that are abundant in organic molecules. This clearly makes the technique a perfect candidate for the analysis of major food components such as water, proteins, lipids, and carbohydrates (Cen & He, 2007). A major drawback is that the resulting spectra are typically broad and have overlapping absorption bands due to the very complex nature of overtone transitions. This often leads to individual peaks not being sharply resolved. Nonetheless, the recorded spectral patterns usually provide a reproducible chemical fingerprint of the sample’s chemical makeup (Czarnecki et al., 2015). In recent years, near-infrared spectroscopy has frequently been explored for its potential to become an alternative to traditional analytical methods in cereal testing that are commonly labor-intensive. This is often accomplished by creating statistical regression models that find correlations between NIR spectral data and reference values determined using the traditional analytical techniques (Ezenarro & Schorn-García, 2025; Ravichandran et al., 2022). After rigorous validation, these models enable the prediction of the investigated parameters directly from the spectra. In a best-case scenario, the created models effectively are a quick and non-destructive alternative to the original reference method. It is no coincidence that such approaches are

INTRODUCTION

increasingly explored in industrial and high-throughput settings, where time and cost efficiency are critical. Recently established safety and quality guidelines bind agricultural companies to organize their entire production chain so that they can trace every intermediate step, but are also able to react to inconsistencies at any time. Based on these newly implemented guidelines, companies should have the goal and obligation to prevent food fraud by actively developing authenticity testing methods beyond the scope of quality assurance (Kotsanopoulos & Arvanitoyannis, 2017; Popping et al., 2022). As previously mentioned, this certainly includes replacing many traditional analysis methods that are still in daily use due to a lack of modern alternatives that allow application directly in the field. Without prior time-consuming transport of the samples to an external laboratory, personnel responsible for quality assurance can react more rapidly (Medina et al., 2019). A very prominent example in the field of cereal testing is the so-called falling number (or Hagberg number). It is an important parameter mainly used to assess the baking capability of cereal flours. It involves rather complex sample preparation steps including the production of a standardized piece of dough. Despite this, the test is still carried out globally on a daily basis due to the lack of an adequate alternative (Hu et al., 2022). The parameter indirectly determines the alpha-amylase activity, which is significantly affected by weather-induced sprouting or pre-germination. The higher the alpha-amylase activity in the flour, the less viscous the respective dough, sometimes resulting in completely different baking properties (Hagberg, 1960). A recent study investigated the potential of NIR spectroscopy for predicting the falling number in barley grains. For this product type, the parameter is not critical for the determination of baking properties but for the assessment of malting suitability (Gerisch et al., 2022). Having access to samples from a breeding program, the group of scientists was able to analyze a large and genetically diverse set of barley cultivars. They calculated multiple partial least squares regression models to estimate falling number values based on spectral data. Although some spectral correlation with falling numbers was detected during the study, the models demonstrated limited predictive accuracy and lacked robustness. Ultimately, the authors

determined that, at least in this particular industry case, NIR spectroscopy may not reliably replace the traditional reference method. These findings were later confirmed by two very similar studies (Chen et al., 2023; Orth et al., 2025). However, such supposed negative examples do not alter the fact that the method remains a promising and universally applicable tool that will undoubtedly be explored further. This is particularly true in light of improvements in calibration strategies and spectral preprocessing techniques, which will certainly help in the search for alternatives to traditional grain quality tests.

1.7 Elemental Profiling Across Sectors: Geochemical Tracing in Food and Gemstones

Besides spectroscopic approaches, elemental analysis has become an increasingly valuable tool in authenticity testing. Although certainly not rapid or field-deployable like the just introduced near-infrared spectroscopy, inductively coupled plasma (ICP)-based methods offer exceptional sensitivity, precision, and multielement capability (Mazarakioti et al., 2022). Often coupled with mass spectrometry or optical emission spectrometry, the technology enables accurate determination of major and trace elements after appropriate sample preparation. Moreover, it can quantify several elements in one single run while at the same time exhibiting very low detection limits (Aceto, 2016). When combined with multivariate statistical analysis, these fingerprints have proven effective in discriminating geographic origin in various food matrices, including wine (Bronzi et al., 2020; Geana et al., 2013; Šelih et al., 2014) and honey (Drivelos et al., 2021; Silva et al., 2021). Their elemental profile reflects influences from environmental factors like soil composition, local water pathways, and anthropogenic processes, often allowing for origin differentiation at various levels (Aceto, 2016). Importantly, the same analytical principles can also be extended beyond organic matrices. In recent years, the use of elemental profiling to evaluate the geographical origin of gemstones has gained momentum (Groat et al., 2019). While a gemstone's quality is a primary determinant of its price, confirmation of its origin from a defunct mine can lead to a substantial increase in value. Provenance claims are also relevant for ethical and legal considerations, including trade restrictions and sourcing transparency. For a long time,

the classification according to provenance depended on expert judgement, which gave an estimate via inclusions, color, size, or shape. However, it is not uncommon for experts to come to different conclusions about the provenance of a sample (Schwarzinger, 2019; Thomsen & Hess, 2022). Therefore, robust analytical detection methods that are based on geochemistry are being developed in order to be able to issue reliable certificates. Here, crystallographic data in combination with data on major and trace elements or isotopes are most promising. ICP is thus the perfect candidate; for gemstones it is usually applied in the form of LA-ICP-MS (Abduriyim & Kitawaki, 2006). This technique was developed by Gray (1985) and later became commercially available. Nevertheless, it took quite a long time until the potential of the method was recognized in gemology, partly because of the high initial costs for the instrument. One of the first mentions is from Günther & Kane (1999). In the following 20 years, the number of scientific publications using LA-ICP-MS for the investigation of gemstones has increased rapidly. A large part of them deals with corundum, that is, rubies and sapphires (Palke & Breeding, 2017; Sorokina et al., 2017), but also other specimens such as demantoids (Adamo et al., 2015; Schwarzinger, 2019). During the analysis, the gemstones are damaged, but the surface damage caused by the laser is only a few μm in size and therefore justifiable. Early difficulties with the quality and reproducibility of measurements are no longer a limiting factor (Abduriyim & Kitawaki, 2006), and current work is increasingly focused on database design and robust data evaluation using multivariate statistical approaches (Su & Yang, 2022, 2023; Winderbaum et al., 2012).

1.8 Ensuring Robustness in Authenticity Science: The Role of Data and Validation

The increasing reliance on analytical geochemical data for origin testing of gemstones reflects a general trend in the field of authenticity testing. Across diverse sample types, from food to non-food products, there is a clear movement away from reliance on single parameters or judgements based on expert knowledge toward more trusted data-driven strategies (Wang et al., 2025). As analytical techniques become more powerful and the datasets they generate more complex, multivariate statistical evaluations and multi-method approaches are gaining central importance in

ensuring reliable and reproducible authenticity assessments (Borràs et al., 2015; Brettschneider & Seifert, 2025; Spiteri et al., 2016; Xu et al., 2023; You et al., 2023). Multivariate methods offer the possibility of recording a kind of chemical fingerprint of a sample and are therefore more holistic in nature (Esslinger et al., 2014; Martinello et al., 2022). However, parallel consideration of numerous variables (sometimes several thousand) and samples generates considerable amounts of data and requires substantial computing power and sophisticated statistical evaluation procedures (Brereton, 2009; de Araújo Gomes et al., 2022). Mathematical algorithms either search for correlations and patterns in data in an exploratory manner or perform a targeted analysis with a given question and prior knowledge on metadata such as group membership by geographical and botanical origin (González-Domínguez et al., 2022). At the same time, multivariate methods are only as good as their underlying database, which must not only have a certain number of data points suitable for statistical evaluation, but whose structure and content must also be appropriate to the analytical question (Donarski et al., 2019). For example, a database covering only a narrow geographical range would not be suitable for global origin classification. For this reason, various validation techniques exist to test the basic assumptions for performing the analyses. This is critical because sophisticated algorithms will find patterns in almost any dataset, which can lead to misinterpretation of results if not properly validated (Brereton & Lloyd, 2014). In the case of the previously mentioned NMR screening method for honey analysis, for example, a strictly validated database with currently more than 28,500 reference spectra of authentic honey samples is used to not only quantify multiple substances but also to verify geographical origin, honey type, and to detect the addition of foreign sugars (Bruker BioSpin, 2025).

1.8.1 The Growing Need for Robustness in Cereal Authentication

The huge potential of multivariate analyses is just being discovered in many research areas (Dayananda & Cozzolino, 2023), such as the field of cereal science (Heredia et al., 2022). This growing interest is driven in part by recent developments in global agriculture, which are placing increasing pressure on the reliability of authenticity testing. Climate change and the associated rise

INTRODUCTION

in extreme weather events, such as prolonged droughts, heatwaves, and irregular rainfall, are causing significant fluctuations in crop yields and qualities (Brás et al., 2021; Moravec et al., 2021; Trnka et al., 2014). In particular, recurring dry periods across large parts of Europe have not only affected overall production volumes but also altered the biochemical composition of cereal grains, leading to greater batch variability and challenges in classification and quality grading (Fioravanti et al., 2025; Santini et al., 2022). These environmental stressors increase the risk of fraudulent practices, such as blending or substitution of lower-quality material, making comprehensive and scientifically validated authenticity testing methods more indispensable than ever (Faller et al., 2021; Hoffman et al., 2024; Thorsen et al., 2025).

1.9 Objectives of This Thesis

Analytical authenticity testing remains a highly scientific challenge. In the trade of many types of goods, whether food or non-food, which must address issues of transparency and traceability, it continuously presents new obstacles. This thesis aims to tackle several of these through targeted investigations. A wide spectrum of analytical techniques was utilized, including stable isotope analysis, NIR spectroscopy, LA-ICP-MS, and NMR, applied across diverse sample matrices.

A dedicated chemometrics software was co-developed with project partners to enable the integrated evaluation of data from multiple analytical platforms. The two publications and one manuscript forming the basis of this thesis aimed to explore multi-method analytical strategies for the differentiation of barley samples by harvest year [publication 1]; the chemometric classification of demantoid garnets based on trace element data for geographic origin assignment [publication 2]; and the applicability of isotope ratio analysis to determine the malthouse in which a batch of barley was processed [manuscript 3]. More specifically, studies were conducted to achieve the following objectives:

[1] Development of a novel analytical approach that allows differentiation of the two harvest years 2017 and the extremely dry 2018 for spring barley samples by means of multiple modern analytical techniques and consideration of weather data. An additional objective was to interpret the observed chemical differences in relation to plant physiological responses to environmental stress through an extensive literature review.

[2] Application of modern chemometrics (PCA-LDA) to trace element profiles of demantoid garnets from Russia, Pakistan, Namibia, Iran, Madagascar, and Italy for geographic origin discrimination. The objective included assessing the robustness of the resulting classification models using multiple performance metrics.

INTRODUCTION

[3] Estimation of the potential of stable isotope analysis ($\delta^2\text{H}$ and $\delta^{18}\text{O}$) for tracing the origin of malt production, based on barley samples processed at two different malting facilities in Berlin and Munich.

2 SYNOPSIS

This dissertation presents a novel approach to the development of analytical procedures for authenticity testing, establishing a foundation for addressing future challenges in this and related fields. It is among the first comprehensive studies to implement a multi-method strategy at such a scale, combining advanced analytical techniques with robust chemometric modeling. Publications 1 and 2, along with manuscript 3, collectively describe a progression from an initial broad exploratory investigation of cereal grain authenticity (publication 1), to the development of a multi-method procedure for determining harvest year. This was followed by the application of the chemometric software developed in publication 1 to the sample matrix of gemstones in the context of geographic origin determination (publication 2). Finally, the work culminates in a targeted case study addressing the traceability of barley malt using the method best suited for this purpose, namely stable isotope analysis (manuscript 3).

2.1 Foundation: Establishing the Analytical and Statistical Backbone

2.1.1 Project Background and Dataset Generation

Before the individual publications are summarized in detail and put into context, the background to the initiation of the original research project and its scope should first be briefly outlined. Companies involved in global cereal trade face increasingly complex challenges in preserving the authenticity and safety of their products. Climate change, through its influence on precipitation patterns and the rising frequency of extreme weather events, has a significant impact on crop yield and quality, complicating efforts to maintain consistent product standards (Rezaei et al., 2023; Schmitt et al., 2022). Concurrently, consumer expectations regarding food safety and traceability, risk of food fraud, as well as ongoing legislative and regulatory changes place additional pressure on farmers and traders to implement rigorous monitoring systems. The so-called AgrOr project was thus launched, a collaborative initiative supported by both governmental and industry funding with the aim of developing methods that enable authenticity testing of grain samples. In this project, nearly 3500 grain samples were initially provided by

industry partners. In addition to metadata, traditional grain quality parameters such as falling number, crude protein, moisture, and sedimentation value were provided for a large proportion of the samples. A strict screening for quality defects such as mold, insect infestation, or incomplete documentation of type, variety, harvest year, cultivation system, or geographical origin was applied to all samples. Roughly 500 samples did not pass this screening and were discarded, resulting in approximately 3,000 samples. Most of these samples originated from southern and central Germany (Bavaria, Baden-Württemberg, Thuringia, Saxony), spanned the harvest years 2014 to 2018, and consisted of barley, malting barley, spelt, wheat, rye, and old grain varieties such as einkorn and emmer, as well as a few individual oat samples. The subsequent step involved the development and refinement of a standardized method for preparing flour samples in the laboratory, since most analytical techniques used are incompatible with whole grain material and industry partners did not have the possibility to process small batches. Moreover, standard operating procedures were implemented for analyzing cereal grains with NIR and for flour using IRMS, NMR, and Raman spectroscopy (Figure 1). This important step ensured consistent results throughout the entire project period and, at the same, time reduced the likelihood that artifacts, such as differences in chemical profiles due to deviations in measurement conditions, would interfere with the planned statistical analysis. Especially when performing exploratory data analysis, it is important to always have an idea of why groups of data points are formed and to make sure that the clustering has not been influenced by factors such as processing by two different scientists, temperature fluctuations in the laboratory or use of different software settings. Approximately 2000 samples from the original pool underwent analysis with the earlier mentioned techniques. This resulted in a remarkably extensive and detailed dataset.

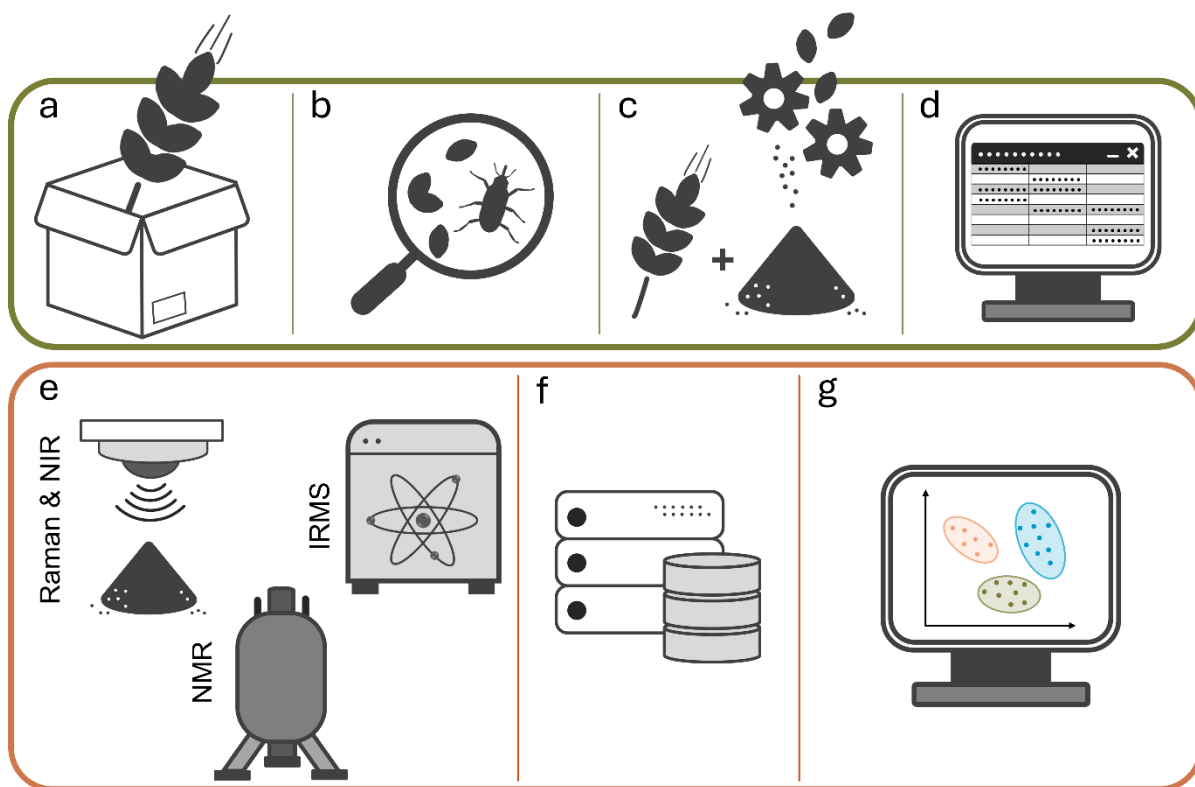


Figure 1 Schematic illustration of the project workflow. (a) Sample delivery by industry partners (b) Inspection for suitability, defects, or infestation (c) Processing of all samples into flour at laboratory scale, with a portion of each sample retained as whole grain (d) Compilation of a comprehensive sample database including all metadata (e) Analysis using multiple analytical technologies (f) Integration of all analytical results into a common database (g) Multivariate statistical evaluation using custom-developed software.

2.1.2 Software Development and Statistical Methods

To enable detailed and scientifically rigorous analysis of the project data and to provide a flexible foundation for future applications, a dedicated software solution was developed. The following requirements were defined based on the specific needs of the AgrOr project:

2.1.2.1 Manufacturer-Independent Import of Raw Data Sets From Various Sources and Analytical Techniques

Nowadays, it is commonplace for manufacturers to offer comprehensive software along with an analytical instrument, which can be used for recording, displaying, processing, and in some cases, evaluating data. Often, however, the focus is clearly on recording and display, and the algorithms supplied for evaluation are kept basic. In addition, the evaluation software is often a “black box” in which the underlying mathematical formulas cannot be accessed. This can be problematic

because even if two software solutions seem to perform the same function because of the same name, a “principal component analysis” conducted in software A, for example, may produce completely different results from software B, making comparison of outputs nearly impossible. Fortunately, most manufacturers offer the option of exporting raw data in an open format. The software developed within the project is capable of directly reading these formats, thereby avoiding transfer-related errors.

2.1.2.2 Flexible Handling of Diverse Analytical Data Formats

With a multi-method approach such as the one employed in the AgrOr project, a diverse array of data types must be handled. The project database, for instance, included categorical nominal data, meaning data used solely for sample identification or classification into groups (e.g., grain species, harvest year). Most of the data from traditional flour analysis was continuous numerical data, i.e. quantitative measurements expressed as numerical values. Examples include protein content (expressed in %), as well as isotope signatures or element concentrations. However, spectral data comprised by far the largest portion of the dataset. Strictly speaking, these too fall under the category of continuous numerical data, but importing such data is not always straightforward due to the large volume of individual data points. For instance, ^1H NMR spectra can easily contain tens of thousands of data points per spectrum. This is where many commercial software packages reach their limitations, especially when several hundred or even thousands of datasets must be processed simultaneously.

2.1.2.3 Comprehensive Data Pre-Processing Capabilities

Data pre-processing in multivariate statistics involves all methods used to prepare and refine data before conducting statistical analysis. Proper data pre-processing is essential for accurate chemometric analysis (Dayananda et al., 2024), and selecting the right pre-processing methods must be done carefully, as improper techniques can impair model performance and cause misinterpretation. Different data pre-processing methods were utilized across this dissertation and its associated studies. Without exceeding the scope of this work, the most relevant

techniques will be briefly introduced below. The paragraph will focus specifically on spectral data, where pre-processing is essential. First steps, such as normalization, scaling, baseline correction, and similar methods serve the main purpose of making samples comparable (Liland, 2011; Skov et al., 2014). Spectral data typically consists of multiple thousands of individual data points. Thus, their statistical evaluation requires substantial computing power. Furthermore, they contain a high degree of redundancy, especially at the flanks of peaks or in regions that only contain baseline information (Westad & Marini, 2022). This inevitably results in high collinearity but also a strong imbalance between the number of variables and the number of observations. This is unfavorable because most statistical algorithms are not designed to perform well under such conditions (Savorani, Rasmussen, Rinnan, et al., 2013). To circumvent this issue, multiple techniques exist that aim to split the data into fewer units while keeping essential information. Well-recognized methods include peak picking, where selected data points are manually or automatically identified and retained, or binning (bucketing), which solely divides the spectra into either uniform or variable-width segments (Sousa et al., 2013). In this thesis, binning was used for both NMR and NIR spectra. Through extensive recalculation of models with parallel adjustment of bin sizes, a balanced "sweet spot" was identified for both the number and size of bins. At this point, it should be noted that while variable reduction techniques are applied to spectral data, signal shifts across samples must be considered, and often, a pre-aligned spectrum becomes necessary. When performing NMR analyses of cereal flour samples, factors like temperature fluctuations or differences in pH adjustment significantly affect the resulting signal positions. Here, we used automated pH adjustment and temperature-controlled instrumentation, but also added an internal reference compound to each sample to ensure minimal signal shifts. During data-splitting, these precautions prevent the misallocation of data points to wrong bins across multiple samples. In cases where signal shifts are unavoidable, for example, due to matrix effects, peak picking can be a viable alternative, as it can compensate for the shifts to a certain extent (Chai et al., 2023). However, this approach also comes with its drawbacks and was ultimately

rejected for two main reasons: First, manual peak picking was infeasible given the large number of spectra. Second, automated peak detection often results in inconsistent variable counts because some applied algorithms may interpret spectral noise as actual peaks or ignore relevant signals below a certain detection threshold (Jarman et al., 2003). These effects can be mitigated to some extent by smoothing the spectra. Still, it will not fully resolve the problem of a fluctuating number of variables, which is not desirable for robust statistical modeling (Karaman, 2017).

In certain cases, a baseline correction may be necessary for spectral data (Wartewig, 2006). For NIR spectra, the baseline is often not flat, but shows deviations due to stray light, instrument noise, or other disturbances. For example, a typical raw NIR spectrum frequently shows pronounced baseline drift. If such a spectrum were divided into bins without pre-treatment, and the resulting sections were then integrated, the data would be rendered unusable for statistical modeling. Multivariate models rely on properly pre-processed input, and the absence of baseline correction can significantly reduce predictive accuracy (Laxalde et al., 2011). When analyzing grain samples using NIR, we used the so-called rubber band method. It corrects for baseline drift and background noise by stretching an imaginary line (“rubber band”) from one end of the spectrum to the other, aligning with the lower edge of the spectral data. The actual baseline is then replaced by that new smooth line, preserving actual peaks while greatly reducing interference effects (Baek et al., 2015; Shen et al., 2018). To ensure that identified patterns and clusters are reflective of true underlying chemical differences rather than artifacts of the data collection process, all spectral data were normalized, scaled, and mean-centered. Normalization and mean-centering are essential pre-processing steps that adjust the data for consistent comparison: normalization corrects for variations in sample concentration and measurement scale, while mean-centering shifts the data around zero to emphasize relative deviations (A. Craig et al., 2006). In our MATLAB-based software, this was achieved by utilizing the built-in MATLAB function “normalize” that by default will mean-center (subtracts the mean of a variable from every observation on that variable in the dataset) (Iacobucci et al., 2016) and normalize to

standard deviation (divides each data point by the standard deviation of the dataset) (Martinez et al., 2004). And last but not least, using NMR spectra and ICP data as examples, this section briefly discusses the role of "scaling", which was used repeatedly throughout this work. The ability to measure precisely over large concentration ranges, a key advantage of NMR technology, can often pose challenges for statistical methods. Important information may be contained in both small signals near the baseline noise level and large, dominant peaks. It is therefore essential to prepare data sets with large differences in scale for statistical evaluation, to avoid overestimating or underestimating the contribution of individual signals. Scaling addresses such differences between variables by dividing each variable by a specific factor.

In this work, the so-called Pareto scaling was applied to all NMR data. It is one of the most commonly used methods in spectral fingerprinting. Pareto scaling is very similar to autoscaling but preserves more of the original data structure: instead of dividing by the standard deviation, it uses the square root of the standard deviation as the scaling factor (Ebrahimi et al., 2017; van den Berg et al., 2006). As a result, after Pareto scaling, variables do not all influence the statistical models equally (as in autoscaling), but their relative magnitudes are partially retained (Savorani, Rasmussen, Mikkelsen, et al., 2013). In one part of this dissertation, the application of an appropriate scaling method proved indispensable for statistical analysis. The trace element concentrations of demantoids determined by LA-ICP-MS spanned several orders of magnitude and were heavily skewed. In addition, complex geochemical interactions and heterogeneous formation conditions introduced non-linear characteristics to the data, which linear multivariate statistical methods struggle to handle (Dalpé et al., 2010; Möller et al., 1981). By applying the natural logarithm as scaling factor, we successfully minimized the impact of outliers, achieved scale invariance, and linearized the relationships in the data (Huelin et al., 2006; Templ & Gonzalez-Rodriguez, 2024; Winderbaum et al., 2012).

2.1.2.4 Built-In Algorithms for Exploratory and Predictive Data Analysis

In large data sets such as that of the AgrOr project, it is almost impossible to gain an overview or recognize patterns simply by examining the raw data. Without appropriate mathematical processing for simplified visualization, important relationships or structures remain undetected. To address this challenge, one of the most widely used techniques, namely principal component analysis (PCA), was implemented in our software as an initial analytical tool.

In spite of the fact that PCA primarily is a method intended for dimensionality reduction, scientists also often utilize it for exploratory data analysis (Ivosev et al., 2008). The idea behind the method is to transform a data set with many variables into a smaller set of features. These so-called principal components are new linear combinations of the original variables that aim to explain the maximum variance in the data while reducing dimensionality at the same time (Greenacre et al., 2022). Depending on the software used, principal components can be visualized in two- or three-dimensional plots, which often offer an intuitive overview of sample distribution. Furthermore, so-called loading plots can highlight the importance of individual variables, so the user can get an idea of which parameter leads to variance in the dataset. PCA belongs to the category of unsupervised methods, meaning it does not consider class labels. Therefore, the structures or patterns found in the data build clusters because of similarity and not because of any prior “knowledge” on group membership (Ballabio, 2015). In the course of this work, PCA was applied following variable selection using the Kruskal–Wallis rank sum test and in combination with Euclidean distance calculations for classification purposes (publication 1), as well as for dimensionality reduction prior to performing linear discriminant analysis (publication 2). Apart from PCA and Euclidean distance analysis, no additional unsupervised methods have been implemented in the software to date. However, a wide range of supervised methods and combinations thereof are available. Supervised methods do utilize *a priori* class information and try to find features that best separate two or more groups (Lindon et al., 2001). Consequently, they enable the confirmation of a predefined group membership and also the identification of key

variables that correlate with these defined groups (Vu et al., 2019). The following classification algorithms have been implemented in the software developed for this work: linear discriminant analysis (Fisher, 1936; Xanthopoulos et al., 2013), orthogonal partial least squares discriminant analysis (Trygg & Wold, 2002), partial least squares discriminant analysis (Barker & Rayens, 2003; Wold et al., 2001), quadratic discriminant analysis (Friedman, 1989; Li & Shao, 2015), k-nearest neighbors analysis (Cover & Hart, 1967). During statistical testing for publication 2, trace element data were analyzed using a combination of PCA and LDA (often referred to as PCA-LDA), aiming to create models capable of provenance determination for gemstones. This specific combination is well established in the field of chemometrics, as it offers an outstanding balance between simplicity, reproducibility, and overall performance. When validated thoroughly with solid cross-validation, PCA-LDA has been shown to produce highly robust results (Ackermann et al., 2019; de Almeida et al., 2021; Gottstein et al., 2024; Kuballa et al., 2018).

2.1.2.5 Storage of Models and Associated Processing Parameters for Full Reproducibility

Regardless of which statistical test is used, it should always deliver the same results for the same data set, even when used at a different time or by a different user. Therefore, a main focus was enabling the software to store not only the created classification models and corresponding plots, but also all setup parameters used during model development. With each model calculated, the software creates a so-called initialization (.ini) file that saves important details such as selected variables, preprocessing steps, and the classification algorithm that was applied. When loaded into the software at a later point in time, that file can be used for an identical analysis of data. Only in this way can methodological consistency be ensured and unintended deviations due to changed settings be prevented.

2.1.2.6 Automated Generation of Performance Metrics and Cross-Validation Outputs to Evaluate Model Robustness

Almost all device manufacturer-specific software now includes at least some basic functions that allow multivariate statistical analysis with just a few clicks when using the default settings. At the

same time, software solutions specializing in statistics that for a long time required a certain amount of programming background are now relatively intuitive and user-friendly (Rodionova et al., 2024). Scientists with limited statistical knowledge can easily produce output that may inspire unjustified confidence because of the polished plots. It is not uncommon that such “push-button” solutions lead to incorrect conclusions being drawn from the analyzed data. Such studies typically lack critical reflection and proper validation procedures (Daszykowski & Walczak, 2006; Kjeldahl & Bro, 2010). It was therefore critical step to implement algorithms that are capable of evaluating the generated models using established quality indicators derived from cross-validation. Users should instantly identify when a model is inappropriate, even if the generated result looks convincing, which is typically achieved through the implementation of model cross-validation techniques. Cross-validation procedures try to give an estimate of the predictive ability of a model by using previously unseen data that was not taken into account during model training. That way, it provides a more realistic picture of a model's robustness on new data (Westad & Marini, 2015). There is a plethora of cross-validation techniques, such as K-fold cross-validation (Arlot & Celisse, 2010; Geisser, 1975) or leave-one-out cross-validation, but for our software, we have opted for Monte Carlo cross-validation (Picard & Cook, 1984; Q.-S. Xu & Liang, 2001). It relies on repeated random sampling, meaning the dataset is randomly split into training and testing subsets multiple times. Our default setting is 100 iterations to obtain a reliable estimate of model performance, as it becomes highly probable that every portion of the dataset is included in the test subset at least once and in various combinations. The variability created reduces the risk of statistical bias from a single, potentially unrepresentative portion of data. Ultimately, this kind of cross-validation provides insight into the model's performance on unseen data and, therefore, is close to the gold standard of model validation, which is testing with an entirely independent external dataset (Haddad et al., 2013; Shao, 1993). Confusion matrices that are the output of cross-validation are often complex and challenging to interpret; it can be helpful to calculate performance metrics that summarize classification quality in a single value.

Accordingly, we implemented functions that calculate performance indicators such as accuracy (the percentage of all correct predictions relative to the total number of predictions), sensitivity (the proportion of true positives among all actual positives), specificity (the proportion of true negatives among all actual negatives), and the area under the receiver operating characteristic curve (a global measure of model performance across all classification thresholds) (Bona et al., 2018). Most of these metrics are developed for two-class scenarios and cannot be easily transferred to multi-class questions, which is why the Matthews correlation coefficient was additionally included. It takes into account all four elements of a confusion matrix (true positives, true negatives, false positives, false negatives), and is more reliable compared to commonly used alternatives such as Cohen’s Kappa. This is particularly true for cases with high class imbalance where the number of observations per class differs a lot (Delgado & Tibau, 2019).

2.1.2.7 Workflow Automation with User-Friendly Navigation

Towards the end of the research project, the focus shifted more towards enhancing the user-friendliness of the developed software. Our goal was to make it accessible to as many scientists as possible so that it can be further developed in the future and gain widespread use in other projects. To date, all core functionalities can be accessed through a guided interface. The user is navigated step-by-step via prompts such as “Which folder contains your NMR spectra?” or “Would you like to inspect and optionally remove outliers from your dataset?”. At the same time, the software executes the corresponding code in the background. Newly generated models are automatically converted into publication-ready figures that can be easily customized through MATLAB’s graphical user interface.

2.2 Discriminating Harvest Years of Spring Barley Exposed to Extreme Weather

Having outlined the conceptual framework and practical implementation of the overall project, attention now turns to the specific scientific questions addressed in the individual studies. The following section aims to summarize the most important findings of the individual manuscripts. Publication 1 directly addressed a central challenge in authenticity testing within the agri-food

sector: the verification of harvest year in cereal grains under conditions of environmental variability. The study was motivated by a real-world scenario involving a record-breaking extreme weather event. In 2018, during the cultivation of a batch of spring barley samples meant for this project, Europe faced an exceptional drought (Blauhut et al., 2022). Compared to 2017, 2018 was characterized by higher temperatures and significantly less rainfall. The German Meteorological Service reported that it was the warmest year since modern temperature records began (Mücke & Litvinovitch, 2020; Wang et al., 2020). This caused severe crop losses and low-quality harvested barley (de Brito et al., 2020; Webber et al., 2020). From an authenticity standpoint, such environmental pressures cause serious problems in the post-harvest supply chain, especially when farmers, under economic pressure, are forced to blend lower-quality harvests with leftover stock from previous years. Depending on regulations and intended use, these practices might require official declaration, which presents a typical authenticity issue. To explore this issue under realistic market conditions, a total of 86 spring barley samples were analyzed, representing two cultivars, namely Quench (*H. vulgare* L. cv. Quench, n=39) and Steffi (*H. vulgare* L. cv. Steffi, n=47), all from the same agricultural region and grown under rain-fed conditions. This approach provided a real-world scenario of commercial grain production while carefully controlling variables like soil type, irrigation and geography. By trying to limit variation, it became possible to mainly link observed chemical differences to the drought conditions in 2018. In order to enable a deeper understanding and better comparability of weather conditions during the periods under investigation, a thorough analysis of meteorological data from 2017 and 2018 was performed for the specific cultivation site. The analysis showed that during the summer months of 2018, rainfall at the sampling location was only 42% of the long-term average of the reference period (1981–2010), so the plants were affected during the critical stage of grain filling.

One objective was to test whether modern analytical techniques could reliably capture these environmental impacts in a way that would allow statistical classification of grain samples by harvest year. Isotopic analysis of the elements carbon, nitrogen, and oxygen revealed that drought

effects were most pronounced in the natural carbon isotope abundance of the spring barley samples. Significantly less negative $\delta^{13}\text{C}$ values were found for samples of the two investigated barley varieties, Quench and Steffi, harvested in 2018, in comparison to the respective samples from 2017, indicative of stomatal closure under drought stress (Farquhar et al., 1982; Farquhar & Richards, 1984; Laverne et al., 2020). NIR spectroscopy provided complementary information on macronutrient composition. Both varieties exhibited notable reductions in moisture, starch, and protein contents and an increase in lipid levels, consistent with metabolic reprogramming under stress conditions (Högy et al., 2013; Wu et al., 2017). By merging the isotopic and NIR data sets and applying multivariate statistical methods, including the Kruskal–Wallis test for dimensionality reduction and principal component analysis to explore clustering, cross-validated classification models were developed. These models enabled a complete separation of the barley samples by harvest year, achieving 100 % correct classification (Figure 2).

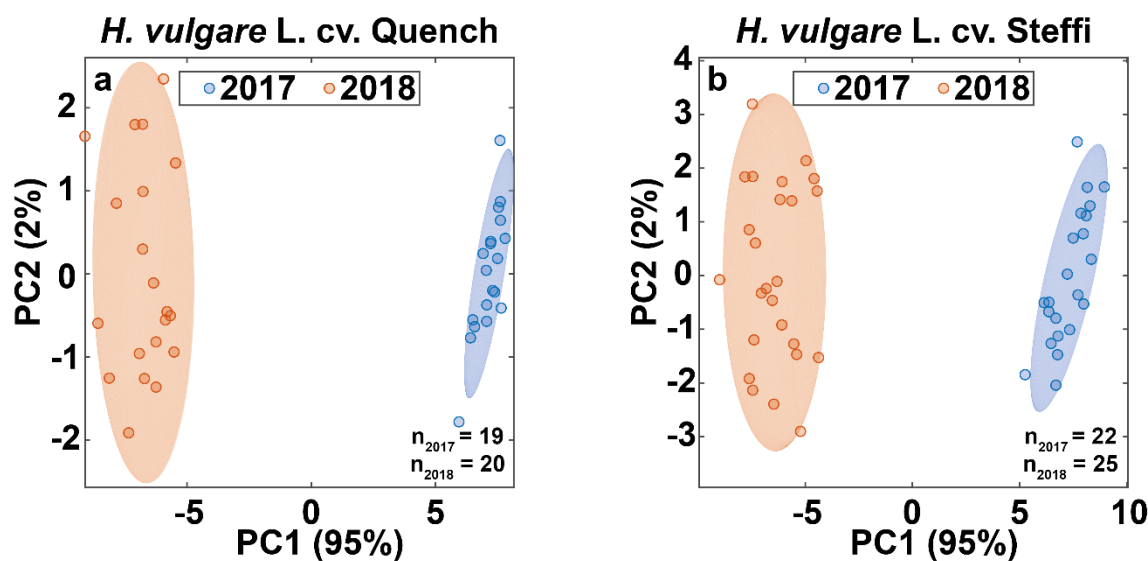


Figure 2 Chemometric analysis of combined FT-NIR and IRMS data from two barley varieties. Score plots including 95% confidence ellipses of the first two principal components from the PCA of fused data (FT-NIR and IRMS). Plots are displayed for the two spring barley varieties (a) Quench and (b) Steffi, and the samples are colored by the crop years 2017 (blue) and 2018 (red). For each principal component, the respective percentage of variance explained is given in the axis label.

By far the most important variables contributing to the separation of the studied groups were lipid and protein contents, along with $\delta^{13}\text{C}$ values. In addition, all samples from the harvest year

2018 showed elevated concentrations of the osmolyte betaine in the NMR spectra, which was a particularly important finding. It provided molecular confirmation that drought-induced metabolic adaptations drove the observed group separation. Interestingly, the weaker stress response observed in the cultivar 'Steffi' highlights its potential as a cultivar that is more resilient under drought conditions. While breeding for drought-resilience is not necessarily related to authenticity, this insight is a noteworthy secondary outcome of the study. In the direct context of authenticity, frequent extreme weather events due to climate change will impact the yield and quality of cereals. As a result, they will naturally create conditions that demand analytical verification of temporal identity. The multi-method approach developed here offers a reproducible and scalable model for tackling such challenges across various agricultural products.

2.3 Classifying Demantoid Origin Using Chemometric Modeling of LA-ICP-MS Data

Building on the first publication, the software framework and statistical modeling approaches developed initially for cereal authenticity testing were subsequently applied to a fundamentally different analytical context: the verification of gemstone provenance. In this field of research, robust and reproducible analytical methods for determining origin are urgently needed, as current assessments mostly rely on expert evaluations and certificates (if available). Information on gemstone provenance not only influences their price but can also provide insights into sourcing practices and regulatory compliance (Schwarzinger, 2019). LA-ICP-MS was used to measure the chemical compositions of 136 demantoid samples from Tubussis (Namibia), Ufaley (Russia), Karkodin (Russia), Poldnevskoye (Russia), Bobrovka (Russia), Val Malenco (Italy), Baluchistan (Pakistan), Kerman (Iran) and Antetazambato (Madagascar). A key finding of this study was that linearization of data by logarithmization is a crucial step during chemometric analysis of trace element profiles. This transformation step tackles the problem of skewed, multiplicative, and heteroscedastic data spanning over multiple orders of magnitude, allowing the application of linear modeling. Separation of samples from five out of six countries of origin was achieved (Figure 3). In particular, demantoids from Namibia, Pakistan, and Madagascar can be classified

very well with accuracies over 90 %. Accuracies for assignments of samples to the origin Russia were slightly lower, with a percentage of around 80 %. This was to be expected, as the respective sample material showed a more inhomogeneous composition. Classification accuracy for the classes Italy and Iran was 70 %, and samples from Italy tended to be falsely assigned to the class Russia. This finding can be explained by the more similar trace element profiles of these classes. At the same time, it highlights the importance of well-designed databases that ensure balanced class sizes and realistic representation of the natural variability within and between origins.

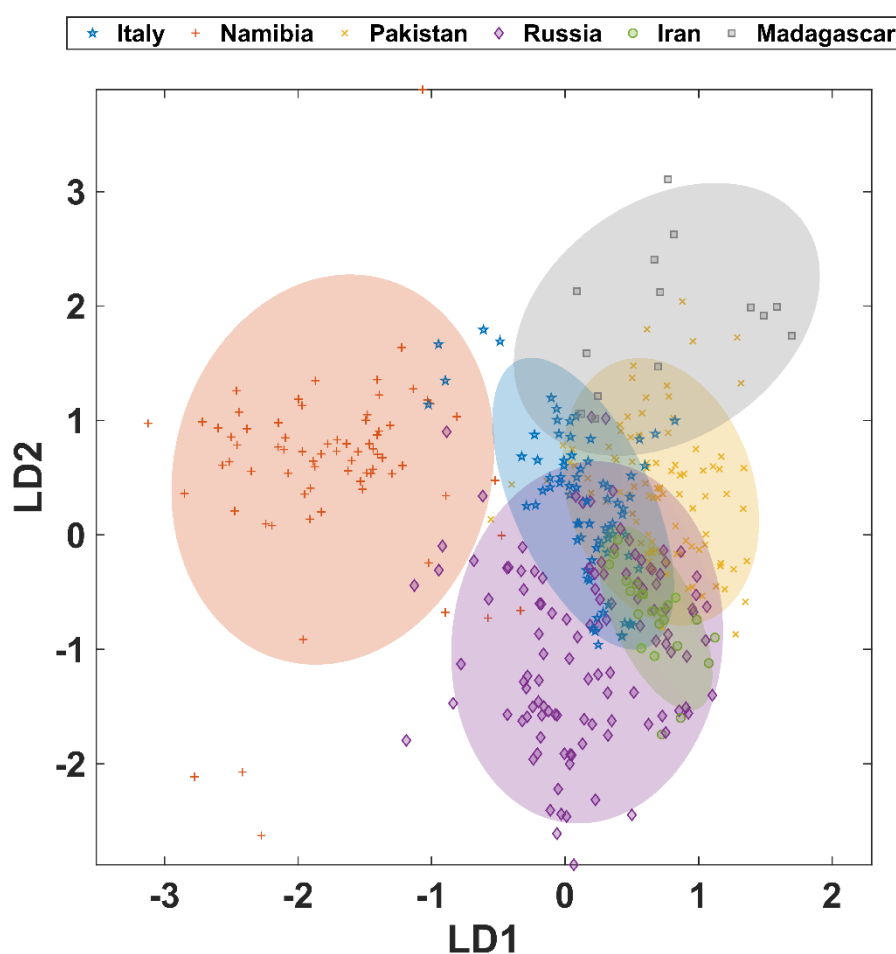


Figure 3 PCA-LDA score plot with all classes (countries) including the log-transformed concentrations of the six relevant elements Mg, Al, Ti, V, Cr and Mn. Each dot represents a single element profile from an individual measurement after Monte-Carlo cross-validation. Axis shown represent the first two linear discriminants. Ellipses confer to the 95% confidence interval according to Hotellings T^2 , which is the multivariate analog to the t-test. Note that graphics represent a projection from 6-dimensional space, i.e., overlap in this projection does not necessarily mean samples cannot be separated.

Separation significantly improved when pairwise origin comparisons were performed, indicating that such models may offer an attractive solution in specific gemological applications. Variable importance in projection (VIP) scores confirmed that Al and Cr play a dominant role in the discriminant models, followed by Mg and Mn, which also contributed significantly. The elements V and Ti were less prominent in the overall VIP rankings. However, they proved essential for distinguishing specific country pairs like Russia–Madagascar or Pakistan–Iran, marking a key principle of multivariate authenticity testing. Even variables with seemingly limited general importance may carry critical discriminatory power in specific scientific questions. This underscores the value of correct usage of preprocessing algorithms that even capture subtle but meaningful variations across complex datasets.

This study not only yielded promising results in terms of classification accuracy but also contributed to methodological transparency. To our knowledge, it is the first work in gemological research to systematically address the role of database design, data preprocessing (including transformation and scaling), and statistical validation in the construction of multivariate models for provenance determination. Finally, it demonstrated that multivariate pipelines developed initially for agricultural traceability can be successfully transferred to other commodities, which highlights the broad applicability of the developed software solution.

2.4 Differentiating Processing Sites of Barley Malt Using Water Isotope Signatures

The final part of this dissertation (manuscript 3) resulted from an acute need within the food industry. In this sector, authenticity has dramatically gained prominence in recent years. This has also been acknowledged by international certification bodies that quickly revised and strengthened their audit frameworks. Pretty much all of them now explicitly address food fraud as a critical component of supply chain integrity (Barrere et al., 2020). To become certified, companies must convincingly demonstrate that they regularly evaluate their vulnerability to fraud and implement appropriate countermeasures. This is primarily realized by monitoring existing systems but also requires proactive development of analytical methods that improve traceability

and transparency within companies' production chains (Huisman & van Ruth, 2022). Against this backdrop, the focus of this study shifted toward the downstream products of barley, which had already been extensively investigated in earlier parts of the project. Barley is the principal raw material in beer production and is transformed into the intermediate product malt (Bahmani et al., 2021). Here, we collaborated with a partner from the brewing industry and aimed to explore the feasibility of using isotope ratio analysis to improve traceability in the part of the production process where barley is turned into malt. The primary goal was to determine whether the water isotope composition of a malt sample could indicate the specific malting facility where it was processed. To address this, 34 barley samples were each processed at two geographically distant malting plants, one located in Munich and the other in Berlin, yielding a total of 68 malt samples. Including the original barley material, a total of 102 samples were analyzed using IRMS, with statistical evaluation of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values performed to assess discriminatory potential.

We found statistically significant differences among all compared groups for both hydrogen and oxygen isotope ratios (Figure 4). In general, malt samples showed enriched isotope values compared to their barley counterparts, with $\delta^2\text{H}$ values being particularly increased. This enrichment likely reflects systematic fractionation processes occurring during malting. Additionally, statistically significant yet more subtle differences were found between malt samples from the two malting facilities. Specifically, samples processed in Berlin showed slightly more positive $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values compared to those from Munich. A likely explanation for this pattern is based on the well-known geographic trends in precipitation isotope ratios found throughout Germany (Förstel et al., 1997). Local tap water usually retains the isotopic signature of local precipitation. Because malting facilities usually utilize local tap water during the steeping stage of the malting process, this might cause distinct regional isotopic variations in the malt's biomass. We observed more positive values in malt samples from Berlin, which is consistent with known isotopic characteristics of precipitation in that region.

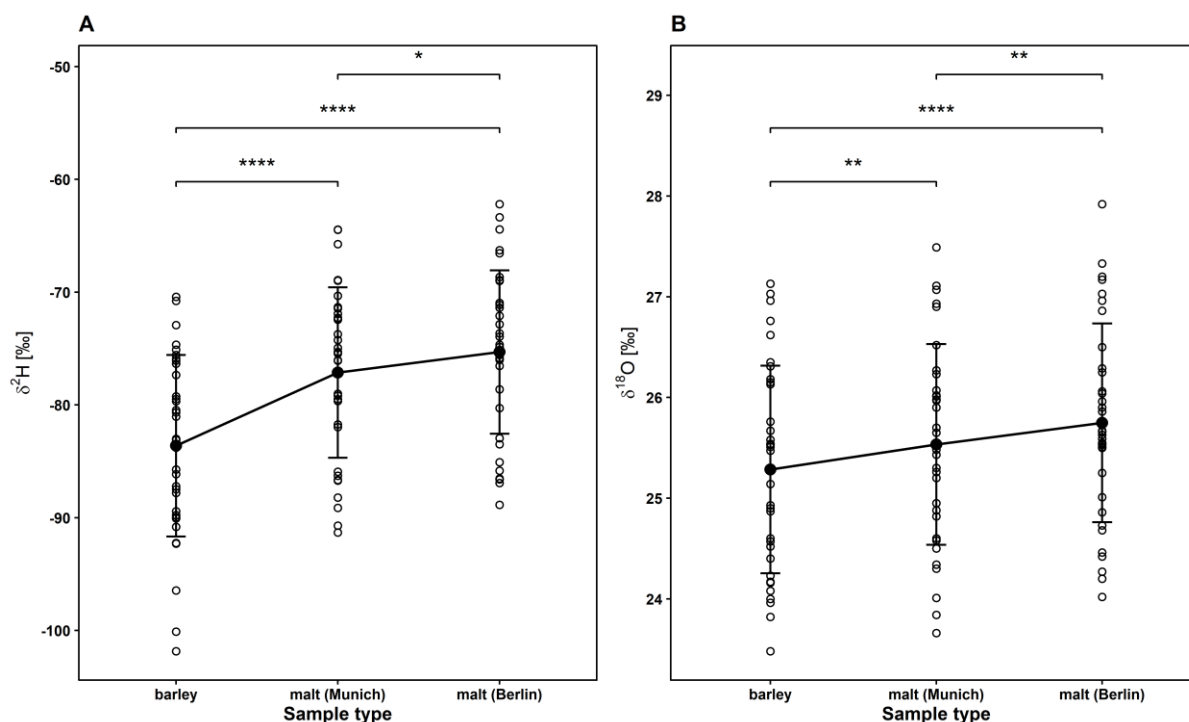


Figure 4 Line plots showing the natural (A) hydrogen and (B) oxygen isotope abundances of the analysed barley samples and malt samples from two malthouses (Munich and Berlin) and results of the multiple paired t-tests. Unfilled circles display the δ values for each sample while filled circles represent means. Error bars show the standard deviation. Significance levels: Highly significant differences ($p \leq 0.001$; ****), very significant differences ($p \leq 0.01$; **), significant differences ($p \leq 0.05$; *).

Although the absolute differences between the two locations were small, their statistical significance suggests they could be useful for verifying processing origin in specific industrial settings. The findings of this work add a layer of traceability to the barley-to-malt process and is a great example of how authenticity testing can be transferred from primary agricultural products to intermediate food processing stages.

In conclusion, the three studies of this dissertation demonstrate the adaptability of data-driven authenticity testing across diverse materials and stages of the value chain. Through a combination of advanced analytical methods and transparent statistical modeling, the work provides a robust framework for addressing authenticity challenges in both food and non-food sectors.

2.5 Outlook

One of the central outcomes of this dissertation is a software that enables transparent, reproducible chemometric modeling across diverse data types. Thanks to its modularity, it can be adapted to many scientific and industrial needs while always remaining user-friendly. Organizations aiming to establish their own data-driven authenticity testing pipelines can access the software from the online repository (AgrOr project, 2019) and use it as a solid starting point. Not only the software itself but also the statistical frameworks developed in this dissertation offer great potential for extension to other raw and/or processed commodities.

For example, the methods used for spring barley samples in publication 1 could easily be adapted for other cereals like wheat, spelt, oats, or even processed products further along the production chain. In the same way, the LA-ICP-MS-based testing utilized for origin discrimination of demantoids in publication 2 could be transferred to other gem types. This is, of course, only possible if the relevant reference materials are available. Robust classification models could eventually replace time-consuming traditional techniques in both areas of research. Over time, these approaches may become valuable alternatives that align with the growing demands of supply chain transparency.

An important step in advancing authenticity testing will be the establishment of large, public multi-method reference databases. Including various data from multiple techniques, such as isotopic, elemental, and spectroscopic data, in shared platforms would reveal the full potential of data science in authenticity testing and uncover hidden variables of importance. New concepts, including open sharing of anonymized data among researchers, regulators, and industry, could reveal new insights into many product characteristics. At the same time, further questions will be examined, as exploratory data analyses will lead to exciting observations. For example, as demonstrated in publication 1, drought stress leaves measurable isotopic and metabolic fingerprints in barley. If systematically recorded across crops, regions, and years, such “chemical fingerprints” could serve as biomarkers in both authenticity testing and agricultural resilience

research. Such opportunities are further amplified by the various recent technical advances in instrumentation. Portable and more economical devices such as benchtop NMR (Colnago et al., 2021; Galvan et al., 2021) or handheld NIR spectrometers (Czaja & Engelsen, 2025) could quickly increase the global volume and diversity of data generated, specifically outside of typical laboratory settings.

While this dissertation focused on transparent statistical modeling using long-established algorithms such as PCA and LDA, there is clear potential to improve chemometric evaluations by incorporating artificial intelligence (AI) approaches, many of which are currently in their early stages. Random forests, support vector machines, and deep learning could be used to model even more complex, nonlinear interactions, especially in very large datasets. It is without question that AI will prove essential in future applications involving even more subtle authenticity distinctions, in which currently available solutions reach their limits (García-Pérez et al., 2024; Gu et al., 2025; Houhou & Bocklitz, 2021). Along with these developments in data evaluations, the general conditions for authenticity testing could be further improved by integrating secure digital systems. Combining analytical testing with blockchain technology could enable samples to be digitally linked to unalterable and traceable records. Such procedures are already well established in the financial sector, for example, and would significantly improve trust and traceability along the entire value chain (Galvez et al., 2018; Hassoun et al., 2024; Puthenveetil & Sappati, 2024).

Each part of this dissertation contributes to the broader aim of enabling cross-domain authenticity testing by advanced chemometric analyses. Publication 1 showed how isotopic and spectroscopic data can identify drought-induced changes in metabolism that can be used to answer questions of authenticity, such as determining the harvest year. Publication 2 used a data-driven approach in gemmology and demonstrated that trace element profiles are suitable for creating statistical models for provenance verification. As a result, the software programmed in this work may serve as a flexible blueprint for authenticity testing across diverse domains. Manuscript 3, though still exploratory, offered insights into how stable isotope analysis could

SYNOPSIS

enhance traceability, not just for raw products but also for intermediate ones. It furthermore revealed a limitation of traditional bulk isotope analysis: isotope signatures of specific substances or tissues can be masked by the averaging effects of proportionally dominant fractions. Future work should address this through compound-specific isotope analysis, examining specific target analytes such as proteins or sugars to enhance resolution (Liu et al., 2023). This is especially promising in malt traceability, where local water signatures may integrate more distinctly into select compound classes.

Taken together, the individual steps of this dissertation illustrate a clear path: from methodological development to data processing to practical application. As authenticity concerns continue to evolve in global supply chains, novel data-driven analytics that can address them will become increasingly critical. In this context, the groundwork laid in this dissertation provides a strong basis for future studies.

3 REFERENCES

Abduriyim, A., & Kitawaki, H. (2006). Applications of laser ablation—Inductively coupled plasma—Mass spectrometry (LA-ICP-MS) to gemology. *Gems & Gemology*, 42(2), 98–118. <https://doi.org/10.5741/GEMS.42.2.98>

Aceto, M. (2016). The use of ICP-MS in food traceability. In M. Espiñeira & F. J. Santaclara (Eds.), *Advances in food traceability techniques and technologies* (pp. 137–164). Woodhead Publishing. <https://doi.org/10.1016/B978-0-08-100310-7.00008-9>

Ackermann, S. M., Lachenmeier, D. W., Kuballa, T., Schütz, B., Spraul, M., & Bunzel, M. (2019). NMR-based differentiation of conventionally from organically produced chicken eggs in Germany. *Magnetic Resonance in Chemistry*, 57(9), 579–588. <https://doi.org/10.1002/mrc.4838>

Adamo, I., Bocchio, R., Diella, V., Caucia, F., & Schmetzer, K. (2015). Demantoid from Balochistan, Pakistan: Gemmological and mineralogical characterization. *The Journal of Gemmology*, 34(5), 428–433. <https://doi.org/10.15506/JoG.2015.34.5.428>

Akoka, S., & Remaud, G. S. (2020). NMR-based isotopic and isotopomic analysis. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 120–121, 1–24. <https://doi.org/10.1016/j.pnmrs.2020.07.001>

AgrOr project. (2019). *AI(OMICS)* [Source code]. GitLab. Retrieved February 20, 2026, from <https://gitlab.com/ai-omics>

Aries, E., De Rudder, O., Kaklamanos, G., Maquet, A., & Ulberth, F. (2021). Results of an interlaboratory comparison of a liquid chromatography–isotope ratio mass spectrometry method for the determination of $^{13}\text{C}/^{12}\text{C}$ ratios of saccharides in honey. *Journal of AOAC INTERNATIONAL*, 104(6), 1698–1702. <https://doi.org/10.1093/jaoacint/qsab091>

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79. <https://doi.org/10.1214/09-SS054>

Aung, M. M., & Chang, Y. S. (2014). Traceability in a food supply chain: Safety and quality perspectives. *Food Control*, 39, 172–184. <https://doi.org/10.1016/j.foodcont.2013.11.007>

REFERENCES

- Badaró, A. T., Hebling e Tavares, J. P., Blasco, J., Aleixos-Borrás, N., & Barbin, D. F. (2022). Near infrared techniques applied to analysis of wheat-based products: Recent advances and future trends. *Food Control*, *140*, 109115. <https://doi.org/10.1016/j.foodcont.2022.109115>
- Back, S.-J., Park, A., Ahn, Y.-J., & Choo, J. (2015). Baseline correction using asymmetrically reweighted penalized least squares smoothing. *Analyst*, *140*(1), 250–257. <https://doi.org/10.1039/C4AN01061B>
- Bahmani, M., O’Lone, C. E., Juhász, A., Nye-Wood, M., Dunn, H., Edwards, I. B., & Colgrave, M. L. (2021). Application of mass spectrometry-based proteomics to barley research. *Journal of Agricultural and Food Chemistry*, *69*(31), 8591–8609. <https://doi.org/10.1021/acs.jafc.1c01871>
- Ballabio, D. (2015). A MATLAB toolbox for principal component analysis and unsupervised exploration of data structure. *Chemometrics and Intelligent Laboratory Systems*, *149*, 1–9. <https://doi.org/10.1016/j.chemolab.2015.10.003>
- Barbour, M. M. (2007). Stable oxygen isotope composition of plant tissue: A review. *Functional Plant Biology*, *34*(2), 83–94. <https://doi.org/10.1071/FP06228>
- Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*(3), 166–173. <https://doi.org/10.1002/cem.785>
- Barrere, V., Everstine, K., Théolier, J., & Godefroy, S. (2020). Food fraud vulnerability assessment: Towards a global consensus on procedures to manage and mitigate food fraud. *Trends in Food Science & Technology*, *100*, 131–137. <https://doi.org/10.1016/j.tifs.2020.04.002>
- Barreto Peixoto, J. A., Silva, J. F., Oliveira, M. B. P. P., & Alves, R. C. (2023). Sustainability issues along the coffee chain: From the field to the cup. *Comprehensive Reviews in Food Science and Food Safety*, *22*(1), 287–332. <https://doi.org/10.1111/1541-4337.13069>
- Bateman, A. S., & Kelly, S. D. (2007). Fertilizer nitrogen isotope signatures. *Isotopes in Environmental and Health Studies*, *43*(3), 237–247. <https://doi.org/10.1080/10256010701550732>
- Beckett, S. T., Fowler, M. S., & Ziegler, G. R. (2017). *Beckett’s industrial Chocolate Manufacture and Use* (5th ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118923597>

REFERENCES

- Ben-Gera, I., & Norris, K. H. (1968). Direct spectrophotometric determination of fat and moisture in meat products. *Journal of Food Science*, 33(1), 64–67. <https://doi.org/10.1111/j.1365-2621.1968.tb00885.x>
- Bigeleisen, J., Perlman, M. L., & Prosser, H. C. (1952). Conversion of hydrogenic materials to hydrogen for isotopic analysis. *Analytical Chemistry*, 24(8), 1356–1357. <https://doi.org/10.1021/ac60068a025>
- Bindereif, S. G., Brauer, F., Schubert, J.-M., Schwarzing, S., & Gebauer, G. (2019). Complementary use of ¹H NMR and multi-element IRMS in association with chemometrics enables effective origin analysis of cocoa beans (*Theobroma cacao* L.). *Food Chemistry*, 299, 125105. <https://doi.org/10.1016/j.foodchem.2019.125105>
- Biswas, A., Naresh, K., Jaygadkar, S. S., & Chaudhari, S. R. (2023). Enabling honey quality and authenticity with NMR and LC-IRMS based platform. *Food Chemistry*, 416, 135825. <https://doi.org/10.1016/j.foodchem.2023.135825>
- Blauhut, V., Stoelzle, M., Ahopelto, L., Brunner, M. I., Teutschbein, C., Wendt, D. E., Akstinas, V., Bakke, S. J., Barker, L. J., Bartošová, L., Briede, A., Cammalleri, C., Kalin, K. C., De Stefano, L., Fendeková, M., Finger, D. C., Huysmans, M., Ivanov, M., Jaagus, J., Jakubínský, J., Krakovska, S., Laaha, G., Lakatos, M., Manevski, K., Neumann Andersen, M., Nikolova, N., Osuch, M., van Oel, P., Radeva, K., Romanowicz, R. J., Toth, E., Trnka, M., Urošev, M., Urquijo Reguera, J., Sauquet, E., Stevkov, A., Tallaksen, L. M., Trofimova, I., Van Loon, A. F., van Vliet, M. T. H., Vidal, J.-P., Wanders, N., Werner, M., Willems, P., & Živković, N. (2022). Lessons from the 2018–2019 European droughts: A collective need for unifying drought risk management. *Natural Hazards and Earth System Sciences*, 22(6), 2201–2217. <https://doi.org/10.5194/nhess-22-2201-2022>
- Bloch, F. (1946). Nuclear induction. *Physical Review*, 70(7–8), 460–474. <https://doi.org/10.1103/PhysRev.70.460>
- Böhme, K., Calo-Mata, P., Barros-Velázquez, J., & Ortea, I. (2019). Review of recent DNA-based methods for main food-authentication topics. *Journal of Agricultural and Food Chemistry*, 67(14), 3854–3864. <https://doi.org/10.1021/acs.jafc.8b07016>
- Bona, E., Março, P. H., & Valderrama, P. (2018). Chapter 4—Chemometrics applied to food control. In A. M. Holban & A. M. Grumezescu (Eds.), *Food Control and Biosecurity* (pp. 105–133). Academic Press. <https://doi.org/10.1016/B978-0-12-811445-2.00004-0>

REFERENCES

- Boner, M., & Förstel, H. (2004). Stable isotope variation as a tool to trace the authenticity of beef. *Analytical and Bioanalytical Chemistry*, 378(2), 301–310. <https://doi.org/10.1007/s00216-003-2347-6>
- Bontempo, L., Larcher, R., Camin, F., Hölzl, S., Rossmann, A., Horn, P., & Nicolini, G. (2011). Elemental and isotopic characterisation of typical Italian alpine cheeses. *International Dairy Journal*, 21(6), 441–446. <https://doi.org/10.1016/j.idairyj.2011.01.009>
- Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., & Busto, O. (2015). Data fusion methodologies for food and beverage authentication and quality assessment – A review. *Analytica Chimica Acta*, 891, 1–14. <https://doi.org/10.1016/j.aca.2015.04.042>
- Brand, W. A., Douthitt, C. B., Fourel, F., Maia, R., Rodrigues, C., Maguas, C., & Prohaska, T. (2014). Chapter 16: Gas source isotope ratio mass spectrometry (IRMS). In T. Prohaska, J. Irrgeher, A. Zitek, & N. Jakubowski (Eds.), *Sector Field Mass Spectrometry for Elemental and Isotopic Analysis* (pp. 500–549). Royal Society of Chemistry. <https://doi.org/10.1039/9781849735407-00500>
- Brás, T. A., Seixas, J., Carvalhais, N., & Jägermeyr, J. (2021). Severity of drought and heatwave crop losses tripled over the last five decades in Europe. *Environmental Research Letters*, 16(6), 065012. <https://doi.org/10.1088/1748-9326/abf004>
- Brereton, R. G. (2009). Preprocessing. In *Chemometrics for pattern recognition* (pp. 107–176). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470746462.ch4>
- Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: Taking the magic away. *Journal of Chemometrics*, 28(4), 213–225. <https://doi.org/10.1002/cem.2609>
- Brescia, M. A., Monfreda, M., Buccolieri, A., & Carrino, C. (2005). Characterisation of the geographical origin of buffalo milk and mozzarella cheese by means of analytical and spectroscopic determinations. *Food Chemistry*, 89(1), 139–147. <https://doi.org/10.1016/j.foodchem.2004.02.016>
- Brettschneider, K. C., & Seifert, S. (2025). Fusion of food profiling data from very different analytical techniques. *Current Opinion in Food Science*, 61, 101256. <https://doi.org/10.1016/j.cofs.2024.101256>

REFERENCES

- Bronzi, B., Brilli, C., Beone, G. M., Fontanella, M. C., Ballabio, D., Todeschini, R., Consonni, V., Grisoni, F., Parri, F., & Buscema, M. (2020). Geographical identification of Chianti red wine based on ICP-MS element composition. *Food Chemistry*, 315, 126248. <https://doi.org/10.1016/j.foodchem.2020.126248>
- Brown, K. L., & Tautfest, G. W. (1956). Faraday-cup monitors for high-energy electron beams. *Review of Scientific Instruments*, 27(9), 696–702. <https://doi.org/10.1063/1.1715674>
- Bruker BioSpin. (2025, August 10). *NMR Honey-Profilig*. <https://www.bruker.com/en/products-and-solutions/mr/nmr-food-solutions/honey-profilig.html>
- Brunetto, M. del R., Galignani, M., Orozco, W., Clavijo, S., Delgado, Y., Ayala, C., & Zambrano, A. (2020). The effect of fermentation and roasting on free amino acids profile in Criollo cocoa (*Theobroma cacao* L.) grown in Venezuela. *Brazilian Journal of Food Technology*, 23, e2019150. <https://doi.org/10.1590/1981-6723.15019>
- Cabañero, A. I., Recio, J. L., & Rupérez, M. (2006). Liquid chromatography coupled to isotope ratio mass spectrometry: A new perspective on honey adulteration detection. *Journal of Agricultural and Food Chemistry*, 54(26), 9719–9727. <https://doi.org/10.1021/jf062067x>
- Cadby, J., & Araki, T. (2021). Towards ethical chocolate: Multicriterial identifiers, pricing structures, and the role of the specialty cacao industry in sustainable development. *SN Business & Economics*, 1(3), 44. <https://doi.org/10.1007/s43546-021-00051-y>
- Cadby, J., Araki, T., & Villacis, A. H. (2021). Breaking the mold: Craft chocolate makers prioritize quality, ethical and direct sourcing, and environmental welfare. *Journal of Agriculture and Food Research*, 4, 100122. <https://doi.org/10.1016/j.jafr.2021.100122>
- Camin, F., Bontempo, L., Perini, M., & Piasentier, E. (2016). Stable isotope ratio analysis for assessing the authenticity of food of animal origin. *Comprehensive Reviews in Food Science and Food Safety*, 15(5), 868–877. <https://doi.org/10.1111/1541-4337.12219>
- Camin, F., Larcher, R., Nicolini, G., Bontempo, L., Bertoldi, D., Perini, M., Schlicht, C., Schellenberg, A., Thomas, F., Heinrich, K., Voerkelius, S., Horacek, M., Ueckermann, H., Froeschl, H., Wimmer, B., Heiss, G., Baxter, M., Rossmann, A., & Hoogewerff, J. (2010). Isotopic and elemental data for tracing the origin of European olive oils. *Journal of Agricultural and Food Chemistry*, 58(1), 570–577. <https://doi.org/10.1021/jf902814s>

REFERENCES

- Carroll, G. R. (2015). Authenticity: Attribution, value, and meaning. In *Emerging Trends in the Social and Behavioral Sciences* (pp. 1–13). John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781118900772.etrds0020>
- Carter, J. F., & Chesson, L. A. (2017). *Food forensics: Stable isotopes as a guide to authenticity and origin*. CRC Press.
- Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science & Technology*, 18(2), 72–83.
<https://doi.org/10.1016/j.tifs.2006.09.003>
- Cerling, T. E., Harris, J. M., MacFadden, B. J., Leakey, M. G., Quade, J., Eisenmann, V., & Ehleringer, J. R. (1997). Global vegetation change through the Miocene/Pliocene boundary. *Nature*, 389(6647), 153–158. <https://doi.org/10.1038/38229>
- Cernusak, L. A., Ubierna, N., Winter, K., Holtum, J. A. M., Marshall, J. D., & Farquhar, G. D. (2013). Environmental and physiological determinants of carbon isotope discrimination in terrestrial plants. *New Phytologist*, 200(4), 950–965. <https://doi.org/10.1111/nph.12423>
- Chai, X., Liu, C., Fan, X., Huang, T., Zhang, X., Jiang, B., & Liu, M. (2023). Combination of peak-picking and binning for NMR-based untargeted metabonomics study. *Journal of Magnetic Resonance*, 351, 107429. <https://doi.org/10.1016/j.jmr.2023.107429>
- Charters, S., & Spielmann, N. (2014). Characteristics of strong territorial brands: The case of champagne. *Journal of Business Research*, 67(7), 1461–1467.
<https://doi.org/10.1016/j.jbusres.2013.07.020>
- Chen, C.-P. J., Hu, Y., Li, X., Morris, C. F., Delwiche, S., Carter, A. H., Steber, C., & Zhang, Z. (2023). An independent validation reveals the potential to predict Hagberg–Perten falling number using spectrometers. *The Plant Phenome Journal*, 6(1), e20070. <https://doi.org/10.1002/ppj.20070>
- Choi, W.-J., Kwak, J.-H., Lim, S.-S., Park, H.-J., Chang, S. X., Lee, S.-M., Arshad, M. A., Yun, S.-I., & Kim, H.-Y. (2017). Synthetic fertilizer and livestock manure differently affect $\delta^{15}\text{N}$ in the agricultural landscape: A review. *Agriculture, Ecosystems & Environment*, 237, 1–15.
<https://doi.org/10.1016/j.agee.2016.12.020>

REFERENCES

- Coblentz, W. W. (1906). *Radiometric investigations of infra-red absorption and reflection spectra* (No. 45). U.S. Department of Commerce and Labor, Bureau of Standards.
- Coblentz, W. W. (1949). Early history of infrared spectroradiometry. *The Scientific Monthly*, 68(2), 102–107.
- Colnago, L. A., Wiesman, Z., Pages, G., Musse, M., Monaretto, T., Windt, C. W., & Rondeau-Mouro, C. (2021). Low field, time domain NMR in the agriculture and agrifood sectors: An overview of applications in plants, foods and biofuels. *Journal of Magnetic Resonance*, 323, 106899. <https://doi.org/10.1016/j.jmr.2020.106899>
- Coplen, T. B. (2011). Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results. *Rapid Communications in Mass Spectrometry*, 25(17), 2538–2560. <https://doi.org/10.1002/rcm.5129>
- Coulomb, M., Gombert, A., & Moazzami, A. (2015). Metabolomics study of cereal grains reveals the discriminative metabolic markers associated with anatomical compartments. *Italian Journal of Food Science*, 27(2), 14–22. <https://doi.org/10.14674/1120-1770/ijfs.v180>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Cozzi, E., Donati, M., Mancini, M. C., Guareschi, M., & Veneziani, M. (2019). PDO Parmigiano Reggiano Cheese in Italy. In F. Arfini & V. Bellassen (Eds.), *Sustainability of European food quality schemes: Multi-performance, structure, and governance of PDO, PGI, and organic agri-food systems* (pp. 427–449). Springer International Publishing. https://doi.org/10.1007/978-3-030-27508-2_22
- Craig, A., Cloarec, O., Holmes, E., Nicholson, J. K., & Lindon, J. C. (2006). Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Analytical Chemistry*, 78(7), 2262–2267. <https://doi.org/10.1021/ac0519312>
- Craig, H. (1953). The geochemistry of the stable carbon isotopes. *Geochimica et Cosmochimica Acta*, 3(2), 53–92. [https://doi.org/10.1016/0016-7037\(53\)90001-5](https://doi.org/10.1016/0016-7037(53)90001-5)

REFERENCES

- Cundy, M. E., Santana-Garcon, J., McLennan, A. G., Ayad, M. E., Bayer, P. E., Cooper, M., Corrigan, S., Harrison, E., & Wilcox, C. (2023). Seafood label quality and mislabelling rates hamper consumer choices for sustainability in Australia. *Scientific Reports*, 13(1), 10146. <https://doi.org/10.1038/s41598-023-37066-4>
- Czaja, T. P., & Engelsens, S. B. (2025). Why nothing beats NIRS technology: The green analytical choice for the future sustainable food production. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 325, 125028. <https://doi.org/10.1016/j.saa.2024.125028>
- Czarnecki, M. A., Morisawa, Y., Futami, Y., & Ozaki, Y. (2015). Advances in molecular structure and interaction studies using near-infrared spectroscopy. *Chemical Reviews*, 115(18), 9707–9744. <https://doi.org/10.1021/cr500013u>
- Dalpé, C., Hudon, P., Ballantyne, D. J., Williams, D., & Marcotte, D. (2010). Trace element analysis of rough diamond by LA-ICP-MS: A case of source discrimination? *Journal of Forensic Sciences*, 55(6), 1443–1456. <https://doi.org/10.1111/j.1556-4029.2010.01509.x>
- Daszykowski, M., & Walczak, B. (2006). Use and abuse of chemometrics in chromatography. *TrAC Trends in Analytical Chemistry*, 25(11), 1081–1096. <https://doi.org/10.1016/j.trac.2006.09.001>
- Dayananda, B., & Cozzolino, D. (2023). Emerging food authentication methodologies using GC/MS. In K. Pastor (Ed.), *Emerging Food Authentication Methodologies Using GC/MS* (pp. 33–43). Springer International Publishing. https://doi.org/10.1007/978-3-031-30288-6_2
- Dayananda, B., Owen, S., Kolobaric, A., Chapman, J., & Cozzolino, D. (2024). Pre-processing applied to instrumental data in analytical chemistry: A brief review of the methods and examples. *Critical Reviews in Analytical Chemistry*, 54(8), 2745–2753. <https://doi.org/10.1080/10408347.2023.2199864>
- de Almeida, V. E., de Sousa Fernandes, D. D., Diniz, P. H. G. D., de Araújo Gomes, A., Vêras, G., Galvão, R. K. H., & Araujo, M. C. U. (2021). Scores selection via Fisher's discriminant power in PCA-LDA to improve the classification of food data. *Food Chemistry*, 363, 130296. <https://doi.org/10.1016/j.foodchem.2021.130296>
- de Araújo Gomes, A., Azcarate, S. M., Diniz, P. H. G. D., de Sousa Fernandes, D. D., & Veras, G. (2022). Variable selection in the chemometric treatment of food data: A tutorial review. *Food Chemistry*, 370, 131072. <https://doi.org/10.1016/j.foodchem.2021.131072>

REFERENCES

- de Brito, M. M., Kuhlicke, C., & Marx, A. (2020). Near-real-time drought impact assessment: A text mining approach on the 2018/19 drought in Germany. *Environmental Research Letters*, 15(10), 1040a9. <https://doi.org/10.1088/1748-9326/aba4ca>
- de Melo, M. T. G., da Silva Barros, J. M., Ribeiro, A. R. B., Lima, T. L. de A., & Sobral, M. F. F. (2024). The role of certifications and eco-labels in fisheries: A systematic literature review of their benefits and challenges. *Environmental Sciences Europe*, 36(1), 200. <https://doi.org/10.1186/s12302-024-01018-0>
- Deborde, C., Fontaine, J.-X., Jacob, D., Botana, A., Nicaise, V., Richard-Forget, F., Lecomte, S., Decourtil, C., Hamade, K., Mesnard, F., Moing, A., & Molinié, R. (2019). Optimizing 1D 1H-NMR profiling of plant samples for high throughput analysis: Extract preparation, standardization, automation and spectra processing. *Metabolomics*, 15(3), 28. <https://doi.org/10.1007/s11306-019-1488-3>
- Delgado, R., & Tibau, X.-A. (2019). Why Cohen's Kappa should be avoided as performance measure in classification. *PLOS ONE*, 14(9), e0222916. <https://doi.org/10.1371/journal.pone.0222916>
- Donarski, J., Camin, F., Fauhl-Hassek, C., Posey, R., & Sudnik, M. (2019). Sampling guidelines for building and curating food authenticity databases. *Trends in Food Science & Technology*, 90, 187–193. <https://doi.org/10.1016/j.tifs.2019.02.019>
- Drivelos, S. A., Danezis, G. P., Halagarda, M., Popek, S., & Georgiou, C. A. (2021). Geographical origin and botanical type honey authentication through elemental metabolomics via chemometrics. *Food Chemistry*, 338, 127936. <https://doi.org/10.1016/j.foodchem.2020.127936>
- Ebrahimi, P., Viereck, N., Bro, R., & Engelsen, S. (2017). Chemometric analysis of NMR spectra. In G. A. Webb (Ed.), *Modern Magnetic Resonance* (pp. 1649–1668). Springer. https://doi.org/10.1007/978-3-319-28275-6_20-1
- Elflein, L., & Ræzke, K.-P. (2008). Improved detection of honey adulteration by measuring differences between ¹³C/¹²C stable carbon isotope ratios of protein and sugar compounds with a combination of elemental analyzer–isotope ratio mass spectrometry and liquid chromatography–isotope ratio mass spectrometry (δ¹³C-EA/LC-IRMS). *Apidologie*, 39(5), 574–587. <https://doi.org/10.1051/apido:2008042>

REFERENCES

- Emiliani, C. (1955). Pleistocene temperatures. *The Journal of Geology*, 63(6), 538–578. <https://doi.org/10.1086/626295>
- Epstein, S., & Mayeda, T. (1953). Variation of O18 content of waters from natural sources. *Geochimica et Cosmochimica Acta*, 4(5), 213–224. [https://doi.org/10.1016/0016-7037\(53\)90051-9](https://doi.org/10.1016/0016-7037(53)90051-9)
- Esslinger, S., Riedl, J., & Fahl-Hassek, C. (2014). Potential and limitations of non-targeted fingerprinting for authentication of food in official control. *Food Research International*, 60, 189–204. <https://doi.org/10.1016/j.foodres.2013.10.015>
- Ezenarro, J., & Schorn-García, D. (2025). How are chemometric models validated? A systematic review of linear regression models for NIRS data in food analysis. *Journal of Chemometrics*, 39(6), e70036. <https://doi.org/10.1002/cem.70036>
- Faller, A. C., Kesanakurti, P., & Arunachalam, T. (2021). Chapter 14—Fraud in grains and cereals. In R. S. Hellberg, K. Everstine, & S. A. Sklare (Eds.), *Food fraud* (pp. 281–308). Academic Press. <https://doi.org/10.1016/B978-0-12-817242-1.00007-5>
- Farquhar, G. D., Ehleringer, J. R., & Hubick, K. T. (1989). Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, 40(1), 503–537. <https://doi.org/10.1146/annurev.pp.40.060189.002443>
- Farquhar, G. D., O’Leary, M. H., & Berry, J. A. (1982). On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. *Functional Plant Biology*, 9(2), 121–137. <https://doi.org/10.1071/pp9820121>
- Farquhar, G. D., & Richards, R. A. (1984). Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. *Functional Plant Biology*, 11(6), 539–552. <https://doi.org/10.1071/pp9840539>
- Femenias, A., Gatiús, F., Ramos, A. J., Teixido-Orries, I., & Marín, S. (2022). Hyperspectral imaging for the classification of individual cereal kernels according to fungal and mycotoxins contamination: A review. *Food Research International*, 155, 111102. <https://doi.org/10.1016/j.foodres.2022.111102>

REFERENCES

- Fioravanti, G., Toreti, A., Cammalleri, C., Muñoz, C. A., Bavera, D., De Jager, A., Hrast Essenfelder, A., Di Ciollo, C., Masante, D., Magni, D., Navarro, J. A., Mazzesch, M., & Maetens, W. (2025). A dataset for monitoring agricultural drought in Europe. *Scientific Data*, 12(1), 308. <https://doi.org/10.1038/s41597-024-04199-8>
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2), 179–188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>
- Förstel, H., Houbé, J., & Hützen, H. (1997). Use of tap water samples for monitoring the geographical variation of stable isotopes used in authenticity studies. *Zeitschrift Für Lebensmitteluntersuchung Und -Forschung A*, 204(2), 103–108. <https://doi.org/10.1007/s002170050045>
- Fraser, E., & van der Ven, H. (2022). Increasing transparency in global supply chains: The case of the fast fashion industry. *Sustainability*, 14(18), 11520. <https://doi.org/10.3390/su141811520>
- Freeman, R. (1993). The Fourier transform revolution NMR spectroscopy. *Analytical Chemistry*, 65(17), 743–753. <https://doi.org/10.1021/ac00065a003>
- Freitas, J. C. C., Ejaz, M., Toci, A. T., Romão, W., & Khimyak, Y. Z. (2023). Solid-state NMR spectroscopy of roasted and ground coffee samples: Evidences for phase heterogeneity and prospects of applications in food screening. *Food Chemistry*, 409, 135317. <https://doi.org/10.1016/j.foodchem.2022.135317>
- Friebolin, H. (2013). *Ein- und zweidimensionale NMR-Spektroskopie: Eine Einführung*. John Wiley & Sons.
- Friedman, I. (1953). Deuterium content of natural waters and other substances. *Geochimica et Cosmochimica Acta*, 4(1), 89–103. [https://doi.org/10.1016/0016-7037\(53\)90066-0](https://doi.org/10.1016/0016-7037(53)90066-0)
- Friedman, J. H. (1989). Regularized discriminant analysis. *Journal of the American Statistical Association*, 84(405), 165–175. <https://doi.org/10.1080/01621459.1989.10478752>
- Galvan, D., Tanamati, A. A. C., Casanova, F., Danieli, E., Bona, E., & Killner, M. H. M. (2021). Compact low-field NMR spectroscopy and chemometrics applied to the analysis of edible oils. *Food Chemistry*, 365, 130476. <https://doi.org/10.1016/j.foodchem.2021.130476>

REFERENCES

- Galvez, J. F., Mejuto, J. C., & Simal-Gandara, J. (2018). Future challenges on the use of blockchain for food traceability analysis. *TrAC Trends in Analytical Chemistry*, 107, 222–232. <https://doi.org/10.1016/j.trac.2018.08.011>
- García, N., & Enghelmayer, M. A. (2025). The complexity of honey fraud detection. *Bee World*, 102(1), 8–14. <https://doi.org/10.1080/0005772X.2025.2467474>
- García, N., & Schwarzsinger, S. (2021). Chapter 15—Honey fraud. In R. S. Hellberg, K. Everstine, & S. A. Sklare (Eds.), *Food Fraud* (pp. 309–334). Academic Press. <https://doi.org/10.1016/B978-0-12-817242-1.00019-1>
- García-Pérez, P., Becchi, P. P., Zhang, L., Rocchetti, G., & Lucini, L. (2024). Metabolomics and chemometrics: The next-generation analytical toolkit for the evaluation of food quality and authenticity. *Trends in Food Science & Technology*, 147, 104481. <https://doi.org/10.1016/j.tifs.2024.104481>
- Gautam, V., Garg, V., Meena, N., Kumari, S., Patel, S., Mukesh, Singh, H., Singh, S., & Singh, R. K. (2025). Harnessing NMR technology for enhancing field crop improvement: Applications, challenges, and future perspectives. *Metabolomics*, 21(2), 27. <https://doi.org/10.1007/s11306-025-02229-z>
- Geana, I., Iordache, A., Ionete, R., Marinescu, A., Ranca, A., & Culea, M. (2013). Geographical origin identification of Romanian wines by ICP-MS elemental analysis. *Food Chemistry*, 138(2), 1125–1134. <https://doi.org/10.1016/j.foodchem.2012.11.104>
- Gebauer, G., Preiss, K., & Gebauer, A. C. (2016). Partial mycoheterotrophy is more widespread among orchids than previously assumed. *New Phytologist*, 211(1), 11–15. <https://doi.org/10.1111/nph.13865>
- Gehre, M., & Strauch, G. (2003). High-temperature elemental analysis and pyrolysis techniques for stable isotope analysis. *Rapid Communications in Mass Spectrometry*, 17(13), 1497–1503. <https://doi.org/10.1002/rcm.1076>
- Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70(350), 320–328. <https://doi.org/10.1080/01621459.1975.10479865>

REFERENCES

- Gerisch, N., Guerreiro, R., Wespel, F., & Stich, B. (2022). Development of a near-infrared spectroscopy calibration for Hagberg falling number assessment of barley (*Hordeum vulgare*): A comparison of methods. *Plant Breeding*, 141(3), 355–365. <https://doi.org/10.1111/pbr.13024>
- Giannioti, Z., Ogrinc, N., Suman, M., Camin, F., & Bontempo, L. (2024). Isotope ratio mass spectrometry (IRMS) methods for distinguishing organic from conventional food products: A review. *TrAC Trends in Analytical Chemistry*, 170, 117476. <https://doi.org/10.1016/j.trac.2023.117476>
- Giannioti, Z., Suman, M., Roncone, A., Rollo, E., Tonidandel, L., Barbero, A., Catellani, D., Larcher, R., & Bontempo, L. (2024). Isotopic, mycotoxin, and pesticide analysis for organic authentication along the production chain of wheat-derived products. *Food Chemistry*, 452, 139519. <https://doi.org/10.1016/j.foodchem.2024.139519>
- Giunta, C. J., & Mainz, V. V. (2020). Discovery of Nuclear Magnetic Resonance: Rabi, Purcell, and Bloch. In *Pioneers of magnetic resonance* (Vol. 1349, pp. 3–20). American Chemical Society. <https://doi.org/10.1021/bk-2020-1349.ch001>
- Glawischnig, E., Gierl, A., Tomas, A., Bacher, A., & Eisenreich, W. (2002). Starch biosynthesis and intermediary metabolism in maize kernels: Quantitative analysis of metabolite flux by nuclear magnetic resonance. *Plant Physiology*, 130(4), 1717–1727. <https://doi.org/10.1104/pp.006726>
- González-Domínguez, R., Sayago, A., & Fernández-Recamales, Á. (2022). An overview on the application of chemometrics tools in food authenticity and traceability. *Foods*, 11(23), 3940. <https://doi.org/10.3390/foods11233940>
- Gottstein, V., Lachenmeier, D. W., Kuballa, T., & Bunzel, M. (2024). ¹H NMR-based approach to determine the geographical origin and cultivation method of roasted coffee. *Food Chemistry*, 433, 137278. <https://doi.org/10.1016/j.foodchem.2023.137278>
- Gray, A. L. (1985). Solid sample introduction by laser ablation for inductively coupled plasma source mass spectrometry. *Analyst*, 110(5), 551–556. <https://doi.org/10.1039/AN9851000551>
- Greenacre, M., Groenen, P. J. F., Hastie, T., D’Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, 2(1), 1–21. <https://doi.org/10.1038/s43586-022-00184-w>

REFERENCES

- Groat, L. A., Giuliani, G., Stone-Sundberg, J., Sun, Z., Renfro, N. D., & Palke, A. C. (2019). A review of analytical methods used in geographic origin determination of gemstones. *Gems & Gemology*, 55(4), 512–535. <https://doi.org/10.5741/GEMS.55.4.512>
- Gu, C., Wang, G., Zhuang, W., Hu, J., He, X., Zhang, L., Du, Z., Xu, X., Yin, M., Yao, Y., Sun, X., & Hu, W. (Walter). (2025). Artificial intelligence-enabled analysis methods and their applications in food chemistry. *Critical Reviews in Food Science and Nutrition*, 0(0), 1–22. <https://doi.org/10.1080/10408398.2025.2521648>
- Günther, D., & Kane, R. (1999). Laser ablation-inductively coupled plasma-mass spectrometry: A new way of analyzing gemstones. *Gems & Gemology*, 35(3), 160–161.
- Günther, H. (2013). *NMR spectroscopy: Basic principles, concepts and applications in chemistry*. John Wiley & Sons.
- Haber, F., & Le Rossignol, R. (1913). Über die technische Darstellung von Ammoniak aus den Elementen. *Zeitschrift Für Elektrochemie Und Angewandte Physikalische Chemie*, 19(2), 53–72. <https://doi.org/10.1002/bbpc.19130190201>
- Haddad, K., Rahman, A., Zaman, M. A., & Shrestha, S. (2013). Applicability of Monte Carlo cross validation technique for model development and validation using generalised least squares regression. *Journal of Hydrology*, 482, 119–128. <https://doi.org/10.1016/j.jhydrol.2012.12.041>
- Hagberg, S. (1960). A rapid method for determining alpha-amylase activity. *Cereal Chemistry*, 37(2), 218–222.
- Hassoun, A., Alhaj Abdullah, N., Aït-Kaddour, A., Ghellam, M., Beşir, A., Zannou, O., Önal, B., Aadil, R. M., Lorenzo, J. M., Mousavi Khaneghah, A., & Regenstein, J. M. (2024). Food traceability 4.0 as part of the fourth industrial revolution: Key enabling technologies. *Critical Reviews in Food Science and Nutrition*, 64(3), 873–889. <https://doi.org/10.1080/10408398.2022.2110033>
- Heaton, K., Kelly, S. D., Hoogewerff, J., & Woolfe, M. (2008). Verifying the geographical origin of beef: The application of multi-element isotope and trace element analysis. *Food Chemistry*, 107(1), 506–515. <https://doi.org/10.1016/j.foodchem.2007.08.010>

REFERENCES

- Heredia, J. Z., Montemerlo, A. E., Wagner, M., Azcarate, S. M., & Camiña, J. M. (2022). Authenticity and Chemometrics of Cereals and Grains and Their Derived Products. In *Chemometrics and authenticity of foods of plant origin* (pp. 131–151). CRC Press. <https://doi.org/10.1201/9781003266167>
- Hindle, P. H., & Ciurczak, E. W. (2021). History of NIRS Development. In *Handbook of near-infrared analysis* (4th ed., pp. 1–8). CRC Press.
- Hoefs, J. (2021). Isotope fractionation processes of selected elements. In J. Hoefs (Ed.), *Stable Isotope Geochemistry* (pp. 49–265). Springer International Publishing. https://doi.org/10.1007/978-3-030-77692-3_2
- Hoering, T. (1955). Variations of nitrogen-15 abundance in naturally occurring substances. *Science*, 122(3182), 1233–1234. <https://doi.org/10.1126/science.122.3182.1233>
- Hoffman, L. C., Schreuder, J., & Cozzolino, D. (2024). Food authenticity and the interactions with human health and climate change. *Critical Reviews in Food Science and Nutrition*, 1–14. <https://doi.org/10.1080/10408398.2024.2387329>
- Högy, P., Poll, C., Marhan, S., Kandeler, E., & Fangmeier, A. (2013). Impacts of temperature increase and change in precipitation pattern on crop yield and yield quality of barley. *Food Chemistry*, 136(3), 1470–1477. <https://doi.org/10.1016/j.foodchem.2012.09.056>
- Hong, E., Lee, S. Y., Jeong, J. Y., Park, J. M., Kim, B. H., Kwon, K., & Chun, H. S. (2017). Modern analytical methods for the detection of food fraud and adulteration by food category. *Journal of the Science of Food and Agriculture*, 97(12), 3877–3896. <https://doi.org/10.1002/jsfa.8364>
- Horacek, M., & Min, J.-S. (2010). Discrimination of Korean beef from beef of other origin by stable isotope measurements. *Food Chemistry*, 121(2), 517–520. <https://doi.org/10.1016/j.foodchem.2009.12.018>
- Houhou, R., & Bocklitz, T. (2021). Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. *Analytical Science Advances*, 2(3–4), 128–141. <https://doi.org/10.1002/ansa.202000162>

REFERENCES

- Hu, Y., Sjöberg, S. M., Chen, C. (James), Hauvermale, A. L., Morris, C. F., Delwiche, S. R., Cannon, A. E., Steber, C. M., & Zhang, Z. (2022). As the number falls, alternatives to the Hagberg–Pertén falling number method: A review. *Comprehensive Reviews in Food Science and Food Safety*, 21(3), 2105–2117. <https://doi.org/10.1111/1541-4337.12959>
- Huelin, S. R., Longerich, H. P., Wilton, D. H. C., & Fryer, B. J. (2006). The determination of trace elements in Fe–Mn oxide coatings on pebbles using LA-ICP-MS. *Journal of Geochemical Exploration*, 91(1), 110–124. <https://doi.org/10.1016/j.gexplo.2006.03.001>
- Huisman, W., & van Ruth, S. (2022). Risky business: Food fraud vulnerability assessments. *Crime, Law and Social Change*, 78(5), 559–576. <https://doi.org/10.1007/s10611-021-10008-9>
- Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2016). Mean centering helps alleviate “micro” but not “macro” multicollinearity. *Behavior Research Methods*, 48(4), 1308–1317. <https://doi.org/10.3758/s13428-015-0624-x>
- Inácio, C. T., Chalk, P. M., & Magalhães, A. M. T. (2015). Principles and limitations of stable isotopes in differentiating organic and conventional foodstuffs: 1. Plant products. *Critical Reviews in Food Science and Nutrition*, 55(9), 1206–1218. <https://doi.org/10.1080/10408398.2012.689380>
- Islam, M. A., & Deegan, C. (2010). Media pressures and corporate disclosure of social responsibility performance information: A study of two global clothing and sports retail companies. *Accounting and Business Research*, 40(2), 131–148. <https://doi.org/10.1080/00014788.2010.9663388>
- Ivosev, G., Burton, L., & Bonner, R. (2008). Dimensionality reduction and visualization in principal component analysis. *Analytical Chemistry*, 80(13), 4933–4944. <https://doi.org/10.1021/ac800110w>
- Iyer, G. R., & Kalita, J. K. (1997). The impact of country-of-origin and country-of-manufacture cues on consumer perceptions of quality and value. *Journal of Global Marketing*, 11(1), 7–28. https://doi.org/10.1300/J042v11n01_02
- Jarman, K. H., Daly, D. S., Anderson, K. K., & Wahl, K. L. (2003). A new approach to automated peak detection. *Chemometrics and Intelligent Laboratory Systems*, 69(1), 61–76. [https://doi.org/10.1016/S0169-7439\(03\)00113-8](https://doi.org/10.1016/S0169-7439(03)00113-8)

REFERENCES

- Jézéquel, T., Joubert, V., Giraudeau, P., Remaud, G. S., & Akoka, S. (2017). The new face of isotopic NMR at natural abundance. *Magnetic Resonance in Chemistry*, 55(2), 77–90. <https://doi.org/10.1002/mrc.4548>
- Karaman, I. (2017). Preprocessing and pretreatment of metabolomics data for statistical analysis. In A. Sussulini (Ed.), *Metabolomics: From fundamentals to clinical applications* (pp. 145–161). Springer International Publishing. https://doi.org/10.1007/978-3-319-47656-8_6
- Kelly, S., Brodie, C., & Hilkert, A. (2018). Isotopic-spectroscopic technique: Stable isotope-ratio mass spectrometry (IRMS). In D.-W. Sun (Ed.), *Modern techniques for food authentication* (pp. 349–413). Academic Press. <https://doi.org/10.1016/B978-0-12-814264-6.00011-6>
- Kelly, S., Heaton, K., & Hoogewerff, J. (2005). Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends in Food Science & Technology*, 16(12), 555–567. <https://doi.org/10.1016/j.tifs.2005.08.008>
- Kjeldahl, K., & Bro, R. (2010). Some common misunderstandings in chemometrics. *Journal of Chemometrics*, 24(7–8), 558–564. <https://doi.org/10.1002/cem.1346>
- Kniese, J., Race, A. M., & Schmidt, H. (2021). Classification of cereal flour species using Raman spectroscopy in combination with spectra quality control and multivariate statistical analysis. *Journal of Cereal Science*, 101, 103299. <https://doi.org/10.1016/j.jcs.2021.103299>
- Kotsanopoulos, K. V., & Arvanitoyannis, I. S. (2017). The role of auditing, food safety, and food quality standards in the food industry: A review. *Comprehensive Reviews in Food Science and Food Safety*, 16(5), 760–775. <https://doi.org/10.1111/1541-4337.12293>
- Krummen, M., Hilkert, A. W., Juchelka, D., Duhr, A., Schlüter, H.-J., & Pesch, R. (2004). A new concept for isotope ratio monitoring liquid chromatography/mass spectrometry. *Rapid Communications in Mass Spectrometry*, 18(19), 2260–2266. <https://doi.org/10.1002/rcm.1620>
- Kuballa, T., Brunner, T. S., Thongpanchang, T., Walch, S. G., & Lachenmeier, D. W. (2018). Application of NMR for authentication of honey, beer and spices. *Current Opinion in Food Science*, 19, 57–62. <https://doi.org/10.1016/j.cofs.2018.01.007>

REFERENCES

- Lachenaud, P., & Motamayor, J. C. (2017). The Criollo cacao tree (*Theobroma cacao* L.): A review. *Genetic Resources and Crop Evolution*, 64(8), 1807–1820. <https://doi.org/10.1007/s10722-017-0563-8>
- Laursen, K. H., Mihailova, A., Kelly, S. D., Epov, V. N., Bérail, S., Schjoerring, J. K., Donard, O. F. X., Larsen, E. H., Pedentchouk, N., Marca-Bell, A. D., Halekoh, U., Olesen, J. E., & Husted, S. (2013). Is it really organic? – Multi-isotopic analysis as a tool to discriminate between organic and conventional plants. *Food Chemistry*, 141(3), 2812–2820. <https://doi.org/10.1016/j.foodchem.2013.05.068>
- Lavergne, A., Sandoval, D., Hare, V. J., Graven, H., & Prentice, I. C. (2020). Impacts of soil water stress on the acclimated stomatal limitation of photosynthesis: Insights from stable carbon isotope data. *Global Change Biology*, 26(12), 7158–7172. <https://doi.org/10.1111/gcb.15364>
- Laxalde, J., Ruckebusch, C., Devos, O., Caillol, N., Wahl, F., & Duponchel, L. (2011). Characterisation of heavy oils using near-infrared spectroscopy: Optimisation of pre-processing methods and variable selection. *Analytica Chimica Acta*, 705(1), 227–234. <https://doi.org/10.1016/j.aca.2011.05.048>
- Li, C., Dong, H., Luo, D., Xian, Y., & Fu, X. (2016). Recent developments in application of stable isotope and multi-element analysis on geographical origin traceability of cereal grains. *Food Analytical Methods*, 9(6), 1512–1519. <https://doi.org/10.1007/s12161-015-0328-y>
- Li, L. (2013). Technology designed to combat fakes in the global supply chain. *Business Horizons*, 56(2), 167–177. <https://doi.org/10.1016/j.bushor.2012.11.010>
- Li, Q., & Shao, J. (2015). Sparse quadratic discriminant analysis for high dimensional data. *Statistica Sinica*, 25(2), 457–473.
- Liland, K. H. (2011). Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis. *TrAC Trends in Analytical Chemistry*, 30(6), 827–841. <https://doi.org/10.1016/j.trac.2011.02.007>
- Lindon, J. C., Holmes, E., & Nicholson, J. K. (2001). Pattern recognition methods and applications in biomedical magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 39(1), 1–40. [https://doi.org/10.1016/S0079-6565\(00\)00036-4](https://doi.org/10.1016/S0079-6565(00)00036-4)

REFERENCES

- Liu, H., Nie, J., Liu, Y., Wadood, S. A., Rogers, K. M., Yuan, Y., & Gan, R.-Y. (2023). A review of recent compound-specific isotope analysis studies applied to food authentication. *Food Chemistry*, 415, 135791. <https://doi.org/10.1016/j.foodchem.2023.135791>
- Liu, H.-Y., Wadood, S. A., Xia, Y., Liu, Y., Guo, H., Guo, B.-L., & Gan, R.-Y. (2023). Wheat authentication: An overview on different techniques and chemometric methods. *Critical Reviews in Food Science and Nutrition*, 63(1), 33–56. <https://doi.org/10.1080/10408398.2021.1942783>
- Lozano-Castellón, J., Laveriano-Santos, E. P., Abuhabib, M. M., Pozzoli, C., Pérez, M., Vallverdú-Queralt, A., & Lamuela-Raventós, R. M. (2024). Proven traceability strategies using chemometrics for organic food authenticity. *Trends in Food Science & Technology*, 147, 104430. <https://doi.org/10.1016/j.tifs.2024.104430>
- Maestrello, V., Solovyev, P., Bontempo, L., Mannina, L., & Camin, F. (2022). Nuclear magnetic resonance spectroscopy in extra virgin olive oil authentication. *Comprehensive Reviews in Food Science and Food Safety*, 21(5), 4056–4075. <https://doi.org/10.1111/1541-4337.13005>
- Martinello, M., Stella, R., Baggio, A., Biancotto, G., & Mutinelli, F. (2022). LC-HRMS-based non-targeted metabolomics for the assessment of honey adulteration with sugar syrups: A preliminary study. *Metabolites*, 12(10), 985. <https://doi.org/10.3390/metabo12100985>
- Martinez, W. L., Martinez, A. R., Martinez, W. L., Solka, J., & Martinez, A. (2004). *Exploratory data analysis with MATLAB*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780203483374>
- Mazarakioti, E. C., Zotos, A., Thomatou, A.-A., Kontogeorgos, A., Patakas, A., & Ladavos, A. (2022). Inductively coupled plasma-mass spectrometry (ICP-MS), a useful tool in authenticity of agricultural products' and foods' origin. *Foods*, 11(22), 3705. <https://doi.org/10.3390/foods11223705>
- McCrea, J. M. (1950). On the isotopic chemistry of carbonates and a paleotemperature scale. *The Journal of Chemical Physics*, 18(6), 849–857. <https://doi.org/10.1063/1.1747785>
- Medina, S., Perestrelo, R., Silva, P., Pereira, J. A. M., & Câmara, J. S. (2019). Current trends and recent advances on food authenticity technologies and chemometric approaches. *Trends in Food Science & Technology*, 85, 163–176. <https://doi.org/10.1016/j.tifs.2019.01.017>

REFERENCES

- Meier-Augenstein, W. (2017). *Stable isotope forensics: Methods and forensic applications of stable isotope analysis* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781119080190>
- Melcher, F., Dietrich, V., & Gäbler, H.-E. (2021). Analytical proof of origin for raw materials. *Minerals*, 11(5), 461. <https://doi.org/10.3390/min11050461>
- Mohammed, F., Guillaume, D., Warland, J., & Abdulwali, N. (2021). Analytical methods to detect adulteration of argan oil: A critical review. *Microchemical Journal*, 168, 106501. <https://doi.org/10.1016/j.microc.2021.106501>
- Möller, P., Dulski, P., Schley, F., Luck, J., & Szacki, W. (1981). A new way of interpreting trace element concentrations with respect to modes of mineral formation. *Journal of Geochemical Exploration*, 15(1), 271–284. [https://doi.org/10.1016/0375-6742\(81\)90068-6](https://doi.org/10.1016/0375-6742(81)90068-6)
- Moravec, V., Markonis, Y., Rakovec, O., Svoboda, M., Trnka, M., Kumar, R., & Hanel, M. (2021). Europe under multi-year droughts: How severe was the 2014–2018 drought period? *Environmental Research Letters*, 16(3), 034062. <https://doi.org/10.1088/1748-9326/abe828>
- Muccio, Z., & Jackson, G. P. (2009). Isotope ratio mass spectrometry. *Analyst*, 134(2), 213–222. <https://doi.org/10.1039/B808232D>
- Mücke, H.-G., & Litvinovitch, J. M. (2020). Heat Extremes, Public Health Impacts, and Adaptation Policy in Germany. *International Journal of Environmental Research and Public Health*, 17(21), 7862. <https://doi.org/10.3390/ijerph17217862>
- Napoli, J., Dickinson, S. J., Beverland, M. B., & Farrelly, F. (2014). Measuring consumer-based brand authenticity. *Journal of Business Research*, 67(6), 1090–1098. <https://doi.org/10.1016/j.jbusres.2013.06.001>
- Napoli, J., Dickinson-Delaporte, S., & Beverland, M. B. (2016). The brand authenticity continuum: Strategic approaches for building value. *Journal of Marketing Management*, 32(13–14), 1201–1229.
- Nier, A. O. (1950). A redetermination of the relative abundances of the isotopes of carbon, nitrogen, oxygen, argon, and potassium. *Physical Review*, 77(6), 789–793. <https://doi.org/10.1103/PhysRev.77.789>

REFERENCES

- Nietner, T., Haughey, S. A., Ogle, N., Fauhl-Hassek, C., & Elliott, C. T. (2014). Determination of geographical origin of distillers dried grains and solubles using isotope ratio mass spectrometry. *Food Research International*, 60, 146–153. <https://doi.org/10.1016/j.foodres.2013.11.002>
- Nowacka-Perrin, A., Steglich, T., Topgaard, D., & Bernin, D. (2022). In situ ^{13}C solid-state polarization transfer NMR to follow starch transformations in food. *Magnetic Resonance in Chemistry*, 60(7), 671–677. <https://doi.org/10.1002/mrc.5253>
- O'Mahony, P., Glennon, C., McElhinney, C., Danaher, M., Graham, C., Gutierrez, M., Kelly, A., Kilcawley, K., Lee, D., Malone, E., Molloy, A., O'Callaghan, T., & Kinahan, I. (2024). Report of the Scientific Committee of the Food Safety Authority of Ireland on the suitability of analytical methods for assessing food authenticity. *Food Risk Assess Europe*, 2(4), 0050E. <https://doi.org/10.2903/fr.efsa.2024.FR-0051>
- Orth, S. H., Fox, G. P., Manley, M., & Hayward, S. (2025). Current detection methods and future trends for barley pre-harvest sprouting: Is NIR the preferred method of the future? *Journal of the American Society of Brewing Chemists*, 1–12. <https://doi.org/10.1080/03610470.2025.2489331>
- Palke, A. C., & Breeding, C. M. (2017). The origin of needle-like rutile inclusions in natural gem corundum: A combined EPMA, LA-ICP-MS, and nanoSIMS investigation. *American Mineralogist*, 102(7), 1451–1461. <https://doi.org/10.2138/am-2017-5965>
- Paolini, M., Ziller, L., Laursen, K. H., Husted, S., & Camin, F. (2015). Compound-specific $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ analyses of amino acids for potential discrimination between organically and conventionally grown wheat. *Journal of Agricultural and Food Chemistry*, 63(25), 5841–5850. <https://doi.org/10.1021/acs.jafc.5b00662>
- Paredes, C. L. L., Werteker, M., Rossmann, B., Keplinger, J., Olschewski, I. L., & Schreiner, M. (2018). Discrimination of haymilk and conventional milk via fatty acid profiles. *Journal of Food Measurement and Characterization*, 12(2), 1391–1398. <https://doi.org/10.1007/s11694-018-9753-0>
- Penuelas, J., Coello, F., & Sardans, J. (2023). A better use of fertilizers is needed for global food security and environmental sustainability. *Agriculture & Food Security*, 12(1), 5. <https://doi.org/10.1186/s40066-023-00409-5>
- Perini, M., Thomas, F., Cabañero Ortiz, A. I., Simoni, M., & Camin, F. (2022). Stable isotope ratio analysis of lactose as a possible potential geographical tracer of milk. *Food Control*, 139, 109051. <https://doi.org/10.1016/j.foodcont.2022.109051>

REFERENCES

- Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, 79(387), 575–583. <https://doi.org/10.1080/01621459.1984.10478083>
- Picó, Y., & Barceló, D. (2023). Isotopic mass spectrometry in food and environmental chemistry. In Y. Picó & J. Campo (Eds.), *Mass spectrometry in food and environmental chemistry* (pp. 99–125). Springer International Publishing. https://doi.org/10.1007/698_2022_907
- Popping, B., Buck, N., Bánáti, D., Brereton, P., Gendel, S., Hristozova, N., Chaves, S. M., Saner, S., Spink, J., Willis, C., & Wunderlin, D. (2022). Food inauthenticity: Authority activities, guidance for food operators, and mitigation tools. *Comprehensive Reviews in Food Science and Food Safety*, 21(6), 4776–4811. <https://doi.org/10.1111/1541-4337.13053>
- Poudel, R., Bhinderwala, F., Morton, M., Powers, R., & Rose, D. J. (2021). Metabolic profiling of historical and modern wheat cultivars using proton nuclear magnetic resonance spectroscopy. *Scientific Reports*, 11(1), 3080. <https://doi.org/10.1038/s41598-021-82616-3>
- Preston, T. (1992). The measurement of stable isotope natural abundance variations. *Plant, Cell & Environment*, 15(9), 1091–1097. <https://doi.org/10.1111/j.1365-3040.1992.tb01659.x>
- Purcell, E. M., Torrey, H. C., & Pound, R. V. (1946). Resonance absorption by nuclear magnetic moments in a solid. *Physical Review*, 69(1–2), 37–38. <https://doi.org/10.1103/PhysRev.69.37>
- Puthenveetil, N. R., & Sappati, P. K. (2024). A review of smart contract adoption in agriculture and food industry. *Computers and Electronics in Agriculture*, 223, 109061. <https://doi.org/10.1016/j.compag.2024.109061>
- QSI – Quality Services International. (2025, June 7). *JRC report on using LC-IRMS to determine authenticity of honey*. <https://www.qsi-q3.com/jrc-report-on-using-lc-irms-to-determine-authenticity-of-honey/>
- Rabi, I. I., Zacharias, J. R., Millman, S., & Kusch, P. (1938). A new method of measuring nuclear magnetic moment. *Physical Review*, 53(4), 318–318. <https://doi.org/10.1103/PhysRev.53.318>
- Ravichandran, P., Viswanathan, S., Ravichandran, S., Pan, Y.-J., & Chang, Y. K. (2022). Estimation of grain quality parameters in rice for high-throughput screening with near-infrared spectroscopy and deep learning. *Cereal Chemistry*, 99(4), 907–919. <https://doi.org/10.1002/cche.10546>

REFERENCES

- Rezaei, E. E., Webber, H., Asseng, S., Boote, K., Durand, J. L., Ewert, F., Martre, P., & MacCarthy, D. S. (2023). Climate change impacts on crop yields. *Nature Reviews Earth & Environment*, 4(12), 831–846. <https://doi.org/10.1038/s43017-023-00491-0>
- Rodionova, O. Ye., Oliveri, P., Malegori, C., & Pomerantsev, A. L. (2024). Chemometrics as an efficient tool for food authentication: Golden pillars for building reliable models. *Trends in Food Science & Technology*, 147, 104429. <https://doi.org/10.1016/j.tifs.2024.104429>
- Rogers, K. M., Cook, J. M., Krueger, D., & Beckmann, K. (2013). Modification of AOAC official method SM 998.12 to add filtration and/or centrifugation: Interlaboratory comparison exercise. *Journal of AOAC INTERNATIONAL*, 96(3), 607–614. <https://doi.org/10.5740/jaoacint.12-386>
- Rossmann, A. (2001). Determination of stable isotope ratios in food analysis. *Food Reviews International*, 17(3), 347–381. <https://doi.org/10.1081/FRI-100104704>
- Salazar, O. V., Latorre, S., Godoy, M. Z., & Quelal-Vásquez, M. A. (2023). The challenges of a sustainable cocoa value chain: A study of traditional and “fine or flavour” cocoa produced by the Kichwas in the Ecuadorian Amazon region. *Journal of Rural Studies*, 98, 92–100. <https://doi.org/10.1016/j.jrurstud.2023.01.015>
- Salimi Khorshidi, A., Storsley, J., Malunga, L. N., Thandapilly, S. J., & Ames, N. (2018). Advancing the science of wheat quality evaluation using nuclear magnetic resonance (NMR) and ultrasound-based techniques. *Cereal Chemistry*, 95(3), 347–364. <https://doi.org/10.1002/cche.10040>
- Santini, M., Noce, S., Antonelli, M., & Caporaso, L. (2022). Complex drought patterns robustly explain global yield loss for major crops. *Scientific Reports*, 12(1), 5792. <https://doi.org/10.1038/s41598-022-09611-0>
- Savorani, F., Rasmussen, M. A., Mikkelsen, M. S., & Engelsen, S. B. (2013). A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. *Food Research International*, 54(1), 1131–1145. <https://doi.org/10.1016/j.foodres.2012.12.025>
- Savorani, F., Rasmussen, M. A., Rinnan, Å., & Engelsen, S. B. (2013). Interval-based chemometric methods in NMR foodomics. In F. Marini (Ed.), *Data handling in science and technology* (Vol. 28, pp. 449–486). Elsevier. <https://doi.org/10.1016/B978-0-444-59528-7.00012-0>

REFERENCES

- Schmidt, G., Gebauer, G., Widmann, K., & Ziegler, H. (1993). Influence of nitrogen supply and temperature on stable carbon isotope ratios in plants of different photosynthetic pathways (C3, C4, CAM). *Isotopenpraxis Isotopes in Environmental and Health Studies*, 29(1–2), 9–13. <https://doi.org/10.1080/10256019308046129>
- Schmidt, H., Roßmann, A., Voerkelius, S., Schnitzler, W. H., Georgi, M., Graßmann, J., Zimmermann, G., & Winkler, R. (2005). Isotope characteristics of vegetables and wheat from conventional and organic production. *Isotopes in Environmental and Health Studies*, 41(3), 223–228. <https://doi.org/10.1080/10256010500230072>
- Schmitt, J., Offermann, F., Söder, M., Frühauf, C., & Finger, R. (2022). Extreme weather events cause significant crop yield losses at the farm level in German agriculture. *Food Policy*, 112, 102359. <https://doi.org/10.1016/j.foodpol.2022.102359>
- Schwarzinger, C. (2019). Determination of demantoid garnet origin by chemical fingerprinting. *Monatshefte Für Chemie - Chemical Monthly*, 150(5), 907–912. <https://doi.org/10.1007/s00706-019-02409-3>
- Šelih, V. S., Šala, M., & Drgan, V. (2014). Multi-element analysis of wines by ICP-MS and ICP-OES and their classification according to geographical origin in Slovenia. *Food Chemistry*, 153, 414–423. <https://doi.org/10.1016/j.foodchem.2013.12.081>
- Shao, J. (1993). Linear model selection by cross-validation. *Journal of the American Statistical Association*, 88(422), 486–494. <https://doi.org/10.1080/01621459.1993.10476299>
- Shen, X., Ye, S., Xu, L., Hu, R., Jin, L., Xu, H., Liu, J., & Liu, W. (2018). Study on baseline correction methods for the Fourier transform infrared spectra with different signal-to-noise ratios. *Applied Optics*, 57(20), 5794–5799. <https://doi.org/10.1364/AO.57.005794>
- Shewry, P. R., Corol, D. I., Jones, H. D., Beale, M. H., & Ward, J. L. (2017). Defining genetic and chemical diversity in wheat grain by 1H-NMR spectroscopy of polar metabolites. *Molecular Nutrition & Food Research*, 61(7), 1600807. <https://doi.org/10.1002/mnfr.201600807>
- Siesler, H. W. (2007). Basic principles of near-infrared spectroscopy. In *Handbook of near-infrared analysis* (3rd ed.). CRC Press.

REFERENCES

- Silva, B., Gonzaga, L. V., Maltez, H. F., Samochvalov, K. B., Fett, R., & Costa, A. C. O. (2021). Elemental profiling by ICP-MS as a tool for geographical discrimination: The case of bracatinga honeydew honey. *Journal of Food Composition and Analysis*, 96, 103727. <https://doi.org/10.1016/j.jfca.2020.103727>
- Skov, T., Honoré, A. H., Jensen, H. M., Næs, T., & Engelsen, S. B. (2014). Chemometrics in foodomics: Handling data structures from multiple analytical platforms. *TrAC Trends in Analytical Chemistry*, 60, 71–79. <https://doi.org/10.1016/j.trac.2014.05.004>
- Solovyev, P. A., Fauhl-Hassek, C., Riedl, J., Esslinger, S., Bontempo, L., & Camin, F. (2021). NMR spectroscopy in wine authentication: An official control perspective. *Comprehensive Reviews in Food Science and Food Safety*, 20(2), 2040–2062. <https://doi.org/10.1111/1541-4337.12700>
- Sorokina, E. S., Rösel, D., Häger, T., Mertz-Kraus, R., & Saul, J. M. (2017). LA-ICP-MS U–Pb dating of rutile inclusions within corundum (ruby and sapphire): New constraints on the formation of corundum deposits along the Mozambique belt. *Mineralium Deposita*, 52(5), 641–649. <https://doi.org/10.1007/s00126-017-0732-x>
- Sousa, S. A. A., Magalhães, A., & Ferreira, M. M. C. (2013). Optimized bucketing for NMR spectra: Three case studies. *Chemometrics and Intelligent Laboratory Systems*, 122, 93–102. <https://doi.org/10.1016/j.chemolab.2013.01.006>
- Spiteri, M., Dubin, E., Cotton, J., Poirel, M., Corman, B., Jamin, E., Lees, M., & Rutledge, D. (2016). Data fusion between high resolution 1H-NMR and mass spectrometry: A synergetic approach to honey botanical origin characterization. *Analytical and Bioanalytical Chemistry*, 408(16), 4389–4401. <https://doi.org/10.1007/s00216-016-9538-4>
- Spiteri, M., Jamin, E., Thomas, F., Rebours, A., Lees, M., Rogers, K. M., & Rutledge, D. N. (2015). Fast and global authenticity screening of honey using 1H-NMR profiling. *Food Chemistry*, 189, 60–66. <https://doi.org/10.1016/j.foodchem.2014.11.099>
- Spraul, M., Schütz, B., Humpfer, E., Mörtter, M., Schäfer, H., Koswig, S., & Rinke, P. (2009). Mixture analysis by NMR as applied to fruit juice quality control. *Magnetic Resonance in Chemistry*, 47(S1), S130–S137. <https://doi.org/10.1002/mrc.2528>

REFERENCES

- Squeo, G., Cruz, J., De Angelis, D., Caponio, F., & Amigo, J. M. (2024). Considerations about the gap between research in near-infrared spectroscopy and official methods and recommendations of analysis in foods. *Current Opinion in Food Science*, 59, 101203. <https://doi.org/10.1016/j.cofs.2024.101203>
- Su, Y., & Yang, M. (2022). Combining rare earth element analysis and chemometric method to determine the geographical origin of nephrite. *Minerals*, 12(11), 1399. <https://doi.org/10.3390/min12111399>
- Su, Y., & Yang, M. (2023). A multi-strategy linear discriminant analysis (LDA) method coupled with laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) for discriminating the geographical origin of nephrite. *Journal of Analytical Atomic Spectrometry*, 38(3), 667–680. <https://doi.org/10.1039/D2JA00386D>
- Templ, M., & Gonzalez-Rodriguez, J. (2024). Advancing forensic research: An examination of compositional data analysis with an application on petrol fraud detection. *Science & Justice*, 64(1), 9–18. <https://doi.org/10.1016/j.scijus.2023.11.003>
- Thode, H. G., Macnamara, J., & Collins, C. B. (1949). Natural variations in the isotopic content of sulphur and their significance. *Canadian Journal of Research*, 27b(4), 361–373. <https://doi.org/10.1139/cjr49b-038>
- Thomsen, L., & Hess, M. (2022). Dialectics of association and dissociation: Spaces of valuation, trade, and retail in the gemstone and jewelry sector. *Economic Geography*, 98(1), 49–67. <https://doi.org/10.1080/00130095.2021.1989302>
- Thorsen, M., Hill, J., Farber, J., Yiannas, F., Rietjens, I. M. C. M., Venter, P., Lues, R., & Bremer, P. (2025). Megatrends and emerging issues: Impacts on food safety. *Comprehensive Reviews in Food Science and Food Safety*, 24(3), e70170. <https://doi.org/10.1111/1541-4337.70170>
- Trnka, M., Rötter, R. P., Ruiz-Ramos, M., Kersebaum, K. C., Olesen, J. E., Žalud, Z., & Semenov, M. A. (2014). Adverse weather conditions for European wheat production will become more frequent with climate change. *Nature Climate Change*, 4(7), 637–643. <https://doi.org/10.1038/nclimate2242>
- Trygg, J., & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, 16(3), 119–128. <https://doi.org/10.1002/cem.695>

REFERENCES

- van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. (2006). Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics*, 7, 142. <https://doi.org/10.1186/1471-2164-7-142>
- Vincent, F., Coste-Manière, I., & Basseporte, M. (2021). Tanzanites: The Maasäi sustainable dilemma for the rarest gemstones. In I. Coste-Manière & M. Á. Gardetti (Eds.), *Sustainable luxury and jewelry* (pp. 117–144). Springer. https://doi.org/10.1007/978-981-16-2454-4_7
- Vu, T., Siemek, P., Bhinderwala, F., Xu, Y., & Powers, R. (2019). Evaluation of multivariate classification models for analyzing NMR metabolomics data. *Journal of Proteome Research*, 18(9), 3282–3294. <https://doi.org/10.1021/acs.jproteome.9b00227>
- Wang, S., Zhang, Y., Ju, W., Porcar-Castell, A., Ye, S., Zhang, Z., Brümmer, C., Urbaniak, M., Mammarella, I., Juszczak, R., & Folkert Boersma, K. (2020). Warmer spring alleviated the impacts of 2018 European summer heatwave and drought on vegetation photosynthesis. *Agricultural and Forest Meteorology*, 295, 108195. <https://doi.org/10.1016/j.agrformet.2020.108195>
- Wang, Z., He, T., Fang, Y., Lan, Z., Liu, B., Kong, K. W., Sun, J., & He, X. (2025). Impact of processing methods in shaping taste, flavor, antioxidants, and metabolites in teas (*Camellia sinensis*): A multi-method analysis. *Food Research International*, 208, 116060. <https://doi.org/10.1016/j.foodres.2025.116060>
- Wartewig, S. (2006). *IR and Raman spectroscopy: Fundamental processing*. John Wiley & Sons.
- Webber, H., Lischeid, G., Sommer, M., Finger, R., Nendel, C., Gaiser, T., & Ewert, F. (2020). No perfect storm for crop yield failure in Germany. *Environmental Research Letters*, 15(10), 104012. <https://doi.org/10.1088/1748-9326/aba2a4>
- Westad, F., & Marini, F. (2015). Validation of chemometric models – A tutorial. *Analytica Chimica Acta*, 893, 14–24. <https://doi.org/10.1016/j.aca.2015.06.056>
- Westad, F., & Marini, F. (2022). Variable selection and redundancy in multivariate regression models. *Frontiers in Analytical Science*, 2, 897605. <https://doi.org/10.3389/frans.2022.897605>

REFERENCES

- White, J. W. (1992). Internal standard stable carbon isotope ratio method for determination of C-4 plant sugars in honey: Collaborative study, and evaluation of improved protein preparation procedure. *Journal of AOAC INTERNATIONAL*, 75(3), 543–548. <https://doi.org/10.1093/jaoac/75.3.543>
- White, J. W., & Winters, K. (1989). Honey protein as internal standard for stable carbon isotope ratio detection of adulteration of honey. *Journal of Association of Official Analytical Chemists*, 72(6), 907–911. <https://doi.org/10.1093/jaoac/72.6.907>
- Wikus, P., Frantz, W., Kümmerle, R., & Vonlanthen, P. (2022). Commercial gigahertz-class NMR magnets. *Superconductor Science and Technology*, 35(3), 033001. <https://doi.org/10.1088/1361-6668/ac4951>
- Winderbaum, L., Ciobanu, C. L., Cook, N. J., Paul, M., Metcalfe, A., & Gilbert, S. (2012). Multivariate analysis of an LA-ICP-MS trace element dataset for pyrite. *Mathematical Geosciences*, 44(7), 823–842. <https://doi.org/10.1007/s11004-012-9418-1>
- Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)
- Wolfsberg, M., Van Hook, W. A., & Paneth, P. (2009). A short history of early work on isotopes. In M. Wolfsberg, W. A. Hook, P. Paneth, & L. P. N. Rebelo (Eds.), *Isotope effects: In the chemical, geological, and bio sciences* (pp. 1–36). Springer Netherlands. https://doi.org/10.1007/978-90-481-2265-3_1
- Wu, X., Cai, K., Zhang, G., & Zeng, F. (2017). Metabolite profiling of barley grains subjected to water stress: To explain the genotypic difference in drought-induced impacts on malting quality. *Frontiers in Plant Science*, 8, 1547. <https://doi.org/10.3389/fpls.2017.01547>
- Xanthopoulos, P., Pardalos, P. M., & Trafalis, T. B. (2013). Linear discriminant analysis. In P. Xanthopoulos, P. M. Pardalos, & T. B. Trafalis (Eds.), *Robust data mining* (pp. 27–33). Springer. https://doi.org/10.1007/978-1-4419-9878-1_4
- Xu, L., Xu, Z., & Liao, X. (2022). A review of fruit juice authenticity assessments: Targeted and untargeted analyses. *Critical Reviews in Food Science and Nutrition*, 62(22), 6081–6102. <https://doi.org/10.1080/10408398.2021.1895713>

REFERENCES

Xu, Q.-S., & Liang, Y.-Z. (2001). Monte Carlo cross validation. *Chemometrics and Intelligent Laboratory Systems*, 56(1), 1–11. [https://doi.org/10.1016/S0169-7439\(00\)00122-2](https://doi.org/10.1016/S0169-7439(00)00122-2)

Xu, Y., Zhang, J., & Wang, Y. (2023). Recent trends of multi-source and non-destructive information for quality authentication of herbs and spices. *Food Chemistry*, 398, 133939. <https://doi.org/10.1016/j.foodchem.2022.133939>

You, Q., Wang, Z., Tian, X., & Xu, X. (2023). A multi-block data approach to assessing beef quality: ComDim analysis of hyperspectral imaging, ¹H NMR, electronic nose and quality parameters data. *Food Chemistry*, 425, 136469. <https://doi.org/10.1016/j.foodchem.2023.136469>

Zareef, M., Arslan, M., Hassan, M. M., Ahmad, W., Ali, S., Li, H., Ouyang, Q., Wu, X., Hashim, M. M., & Chen, Q. (2021). Recent advances in assessing qualitative and quantitative aspects of cereals using nondestructive techniques: A review. *Trends in Food Science & Technology*, 116, 815–828. <https://doi.org/10.1016/j.tifs.2021.08.012>

Zhang, S., Liu, S., Shen, L., Chen, S., He, L., & Liu, A. (2022). Application of near-infrared spectroscopy for the nondestructive analysis of wheat flour: A review. *Current Research in Food Science*, 5, 1305–1312. <https://doi.org/10.1016/j.crfs.2022.08.006>

Ziegler, H. (1989). Hydrogen isotope fractionation in plant tissues. In P. W. Rundel, J. R. Ehleringer, & K. A. Nagy (Eds.), *Stable Isotopes in Ecological Research* (pp. 105–123). Springer. https://doi.org/10.1007/978-1-4612-3498-2_8

Ziegler, H., Stichler, W., Maurizio, A., & Vorwohl, G. (1977). Die Verwendung stabiler Isotope zur Charakterisierung von Honigen, ihrer Herkunft und ihrer Verfälschung. *Apidologie*, 8(4), 337–347. <https://doi.org/10.1051/apido:19770404>

4 AUTHOR CONTRIBUTIONS TO THE PUBLICATIONS AND MANUSCRIPT

PUBLICATION 1

Authors	Stefan Gerd Bindereif , Felix Rüll, Peter Kolb, Lucas Köberle, Holger Willms, Simon Steidele, Stephan Schwarzingger, Gerhard Gebauer
Title	Impact of Global Climate Change on the European Barley Market Requires Novel Multi-Method Approaches to Preserve Crop Quality and Authenticity
Journal	Foods , 10(7), 1592 (2021), doi: 10.3390/foods10071592
Contributions	Conceptualization: Stefan Gerd Bindereif, Gerhard Gebauer, and Stephan Schwarzingger; methodology: Stefan Gerd Bindereif; software: Stefan Gerd Bindereif, Felix Rüll, Stephan Schwarzingger, and Peter Kolb; validation: Stefan Gerd Bindereif, Simon Steidele, and Peter Kolb; formal analysis: Stefan Gerd Bindereif; investigation: Stefan Gerd Bindereif, Peter Kolb, Stephan Schwarzingger, and Lucas Köberle; resources: Holger Willms, Stephan Schwarzingger, and Gerhard Gebauer; data curation: Stefan Gerd Bindereif, Holger Willms, Felix Rüll, Lucas Köberle, and Peter Kolb; writing—original draft preparation: Stefan Gerd Bindereif; writing—review and editing: Stefan Gerd Bindereif, Stephan Schwarzingger, and Gerhard Gebauer; visualization: Stefan Gerd Bindereif; supervision: Stephan Schwarzingger, Gerhard Gebauer, Lucas Köberle, and Holger Willms; project administration: Stephan Schwarzingger; funding acquisition: Stephan Schwarzingger.

PUBLICATION 2

Authors	Stefan Gerd Bindereif , Felix Rüll, Stephan Schwarzingner, Clemens Schwarzingner
Title	Chemometric Modeling of Trace Element Data for Origin Determination of Demantoid Garnets
Journal	Minerals , 10(12), 1046 (2020), doi: 10.3390/min10121046
Contributions	Conceptualization: Clemens Schwarzingner and Stephan Schwarzingner; methodology: Clemens Schwarzingner and Stephan Schwarzingner; software: Felix Rüll and Stephan Schwarzingner; validation: Stefan Gerd Bindereif, Felix Rüll, Stephan Schwarzingner and Clemens Schwarzingner; formal analysis: Stefan Gerd Bindereif, Felix Rüll, Stephan Schwarzingner and Clemens Schwarzingner; investigation: Stefan Gerd Bindereif, Felix Rüll, Stephan Schwarzingner and Clemens Schwarzingner; resources: Clemens Schwarzingner and Stephan Schwarzingner; writing—original draft preparation: Clemens Schwarzingner and Stephan Schwarzingner and Stefan Gerd Bindereif; writing—review and editing: Clemens Schwarzingner and Stephan Schwarzingner and Stefan Gerd Bindereif and Felix Rüll; visualization: Stefan Gerd Bindereif; project administration: Clemens Schwarzingner and Stephan Schwarzingner; funding acquisition: Clemens Schwarzingner and Stephan Schwarzingner.

MANUSCRIPT 3

Authors	Stefan Gerd Bindereif , Christopher Porkert, Gerhard Gebauer
Title	Stable Hydrogen and Oxygen Isotope Abundance Analysis of Barley Malt and Its Potential to Trace the Malting Location
Journal	Isotopes in Environmental and Health Studies , 1–8 (2026), doi: 10.1080/10256016.2026.2721629
Status	Published online (not yet assigned to a volume/issue)
Contributions	Conceptualization: Stefan Gerd Bindereif, Gerhard Gebauer; methodology: Stefan Gerd Bindereif software: Stefan Gerd Bindereif; validation: Stefan Gerd Bindereif, Gerhard Gebauer; formal analysis: Stefan Gerd Bindereif, Christopher Porkert; investigation: Stefan Gerd Bindereif, Christopher Porkert, Gerhard Gebauer; resources: Stefan Gerd Bindereif, Gerhard Gebauer; writing—original draft preparation: Stefan Gerd Bindereif; writing—review and editing: Stefan Gerd Bindereif, Gerhard Gebauer; visualization: Stefan Gerd Bindereif; project administration: Stefan Gerd Bindereif, Gerhard Gebauer; funding acquisition: Stefan Gerd Bindereif, Gerhard Gebauer.

5 PUBLICATION LIST

Bindereif, S. G., Rüll, F., Kolb, P., Köberle, L., Willms, H., Steidele, S., Schwarzing, S., & Gebauer, G. (2021). Impact of Global Climate Change on the European Barley Market Requires Novel Multi-Method Approaches to Preserve Crop Quality and Authenticity. *Foods*, 10(7), 1592. <https://doi.org/10.3390/foods10071592>

Bindereif, S. G., Rüll, F., Schwarzing, S., & Schwarzing, C. (2020). Chemometric Modeling of Trace Element Data for Origin Determination of Demantoid Garnets. *Minerals*, 10(12), 1046. <https://doi.org/10.3390/min10121046>

Bindereif, S. G., Porkert, C., & Gebauer, G. (2026). Stable Hydrogen and Oxygen Isotope Abundance Analysis of Barley Malt and Its Potential to Trace the Malting Location. *Isotopes in Environmental and Health Studies*, 1–8. <https://doi.org/10.1080/10256016.2026.2721629>
Published online (not yet assigned to a volume/issue)

The following publication is not part of this dissertation:

Bindereif, S. G., Brauer, F., Schubert, J. M., Schwarzing, S., & Gebauer, G. (2019). Complementary use of ¹H NMR and multi-element IRMS in association with chemometrics enables effective origin analysis of cocoa beans (*Theobroma cacao* L.). *Food Chemistry*, 299, 125105. <https://doi.org/10.1016/j.foodchem.2019.125105>

6 PUBLICATIONS OF THIS THESIS

6.1 PUBLICATION 1

Impact of Global Climate Change on the European Barley Market Requires Novel Multi-Method Approaches to Preserve Crop Quality and Authenticity

Stefan G. Bindereif¹, Felix Rüll², Peter Kolb², Lucas Köberle³, Holger Willms⁴,
Simon Steidele², Stephan Schwarzingler², and Gerhard Gebauer¹

Foods 2021, 10(7), 1592; <https://doi.org/10.3390/foods10071592>

¹BayCEER—Laboratory of Isotope Biogeochemistry, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

²NBNC—North Bavarian NMR Centre, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

³ALNuMed GmbH, Gottlieb-Keim Straße 60, 95448 Bayreuth, Germany

⁴IREKS GmbH, Lichtenfelser Straße 20, 95326 Kulmbach, Germany

Received: 27 May 2021 / Revised: 1 July 2021 / Accepted: 6 July 2021 / Published: 8 July 2021

Keywords: barley; stable isotopes; proton nuclear magnetic resonance; Fourier transform near-infrared spectroscopy; drought; climate change; Europe; Germany; multi-method approach; beer

This is an open access article published by MDPI distributed under the terms of the Creative Commons CC BY license. No special permission is required to reuse all or part of article published by MDPI, including figures and tables. Any part of the article may be reused without permission provided that the original article is clearly cited. Reuse of an article does not imply endorsement by the authors or MDPI.

Article

Impact of Global Climate Change on the European Barley Market Requires Novel Multi-Method Approaches to Preserve Crop Quality and Authenticity

Stefan G. Bindereif ¹, Felix Rüll ², Peter Kolb ², Lucas Köberle ³, Holger Willms ⁴, Simon Steidele ², Stephan Schwarzingen ^{2,*} and Gerhard Gebauer ^{1,*} 

¹ BayCEER—Laboratory of Isotope Biogeochemistry, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany; stefan.bindereif@uni-bayreuth.de

² NBNC—North Bavarian NMR Centre, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany; Felix.Ruell@uni-bayreuth.de (F.R.); peter.kolb.mail@googlemail.com (P.K.); simon-steidele@outlook.de (S.S.)

³ ALNuMed GmbH, Gottlieb-Keim Straße 60, 95448 Bayreuth, Germany; lucasausberg@gmx.de

⁴ IREKS GmbH, Lichtenfelser Straße 20, 95326 Kulmbach, Germany; holger.willms@ireks.com

* Correspondence: s.schwarzingen@uni-bayreuth.de (S.S.); gerhard.gebauer@uni-bayreuth.de (G.G.); Tel.: +49-(0)-9-2155-2046 (S.S.); +49-(0)-9-2155-2060 (G.G.)

Abstract: Most recently in 2018 and 2019, large parts of Europe were affected by periods of massive drought. Resulting losses in cereal yield pose a major risk to the global supply of barley, as more than 60% of global production is based in Europe. Despite the arising price fluctuations on the cereal market, authenticity of the crop must be ensured, which includes correct declaration of harvest years. Here, we show a novel approach that allows such differentiation for spring barley samples, which takes advantage of the chemical changes caused by the extreme drought. Samples from 2018 were successfully differentiated from those of 2017 by analysis of changes in near-infrared spectra, enrichment in the isotope ¹³C, and strong accumulation of the plant-physiological marker betaine. We demonstrate that through consideration of multiple modern analysis techniques, not only can fraudulent labelling be prevented, but indispensable knowledge on the drought tolerance of crops can be obtained.

Keywords: barley; stable isotopes; proton nuclear magnetic resonance; Fourier transform near-infrared spectroscopy; drought; climate change; Europe; Germany; multi-method approach; beer



Citation: Bindereif, S.G.; Rüll, F.; Kolb, P.; Köberle, L.; Willms, H.; Steidele, S.; Schwarzingen, S.; Gebauer, G. Impact of Global Climate Change on the European Barley Market Requires Novel Multi-Method Approaches to Preserve Crop Quality and Authenticity. *Foods* **2021**, *10*, 1592. <https://doi.org/10.3390/foods10071592>

Received: 27 May 2021

Accepted: 6 July 2021

Published: 8 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recent climate changes in Europe are characterized by frequent and severe weather extremes, which pose a major threat to agricultural production and are predicted to increase in the near future [1,2]. Particularly during the last two decades, most parts of Europe were impacted by record-breaking droughts leading to losses in both crop yield and quality. Global warming poses a serious challenge to the global barley (*Hordeum vulgare* L.) supply, as more than 60% of the world's barley production is provided by Europe, most of it grown under rainfed conditions [3]. This crop is mainly used as malting barley for the production of alcoholic beverages and for feeding animals. European beverage producers typically use spring barley for malting due to its low protein content and thus more fermentable portion of sugar. However, spring barley has its most sensitive periods during flowering and grain filling around the summer months, hence there is a high risk of yield loss and crop failure in extremely dry years [4]. In 2018, Germany, the largest producer of spring barley in Europe, experienced its warmest year ever recorded in the 138 years of modern temperature records, ultimately resulting in the lowest cereal yield since 1994 and creating existence-threatening conditions for many farmers across the country [5]. Subsequently, farmers were forced to mix low-grade spring barley batches and marketable grains from previous harvest years

to meet the quality requirements of the customers. However, maltings often refuse to accept deliveries of blends from two harvest years, consequently the product can only be offered as significantly less profitable feed barley [6]. Such stages of emergency increase the risk of economically motivated food fraud. The likelihood of these scenarios gains in importance when considering a recently published model by Xie et al., who found that the effects of concurrent drought extremes will cause a large decrease in global barley supply under forecasted climate conditions [7]. To preserve transparency in such difficult periods, the development of a method that is capable of classifying spring barley by harvest year is indispensable. Here, we assess the potential of distinguishing spring barley grain samples according to the very different harvest years 2017 and 2018. We looked at two varieties, namely the modern and high-yielding cultivar ‘Quench’ (*H. vulgare* L. cv. Quench, year of approval: 2006) and the comparatively old cultivar ‘Steffi’ (*H. vulgare* L. cv. Steffi, year of approval: 1989). As results produced by a single analytical method often do not provide enough information for an adequate evaluation, we have developed a novel methodology using fused data from stable isotope analysis (IRMS) and chemical fingerprints provided by Fourier-transform infrared spectroscopy (FT-NIR). Furthermore, metabolite profiles produced by proton nuclear magnetic resonance (^1H -NMR) were applied to substantiate the findings through identification of stress-related plant physiological parameters. For the present work, we hypothesize that the severe drought events of 2018 created systematic variability in the chemical makeup of barley grains that, when evaluated by our multi-methodological approach, should allow effective separation from samples of 2017. Based on previous works found in literature, we further hypothesize that the older barley cultivar ‘Steffi’, though less yielding, should be more drought-tolerant than the modern cultivar ‘Quench’.

2. Materials and Methods

2.1. Origin and Storage of the Spring Barley Samples

Eighty-six spring barley samples of two different varieties (*H. vulgare* L. cv. Quench; *H. vulgare* L. cv. Steffi) and from two consecutive harvest years (Quench, $n_{2017} = 19$, $n_{2018} = 20$ and Steffi, $n_{2017} = 22$, $n_{2018} = 25$) have been provided by IREKS GmbH (Kulmbach, Germany). As the focus of this study was put on describing the effects of years with extreme drought, we were provided with barley grains grown under real conventional farming conditions rather than controlled field experiments. All plants were cultivated in the same region by an agricultural cooperative of the district Soemmerda (Thuringia, Germany) on loamy soil at 140 m elevation to eliminate geographic effects. Samples were received as whole grains and stored in vacuum-sealed bags at room temperature until further processing.

2.2. Weather at the Sampling Site

Weather data were retrieved from the Climate Data Center of the German Meteorological Service [8]. The De Martonne aridity index (I_{DM}) was used as an indicator of drought, which is calculated according to the following equation:

$$I_{DM} = RR / (TM + 10) \quad (1)$$

where RR is the mean annual precipitation in mm and TM is the mean air temperature in $^{\circ}\text{C}$ [9]. I_{DM} values below 30 correspond to a humid to Mediterranean habitat, while regions with I_{DM} values below 20 are characterized by semi-dry to dry conditions.

2.3. Isotope-Ratio Mass Spectrometry (IRMS)

For measurement of natural isotope abundances of the elements C, N, and O, around 4 g of grains were ground to wholemeal flour in a ball mill (MM400, Retsch GmbH, Haan, Germany) at 30 Hz for 2 min. The sample material was dried in a drying oven at 105°C to constant weight and stored in a desiccator until analysis. For carbon and nitrogen measurements, 2.8–3.5 mg of the powder was weighed into tin capsules, while for oxygen

measurements 0.5–1 mg was weighed into silver capsules using a microbalance (CPA2P, Sartorius, Göttingen, Germany).

Carbon and nitrogen stable isotope ratios were measured using an elemental analyzer (NA 1108, Carlo Erba Instruments, Milan, Italy) coupled to a continuous flow isotope-ratio mass spectrometer (delta S, Finnigan MAT, Bremen, Germany) via a ConFlo III open-split interface (Thermo Fisher Scientific, Bremen, Germany) as described by Bidartondo, Burghardt, Gebauer, Bruns, & Read [10]. Oxygen stable isotope ratios were analyzed with thermal conversion through pyrolysis (HTO, HEKAtech, Wegberg, Germany) coupled to a continuous flow isotope-ratio mass spectrometer (delta V advantage, Thermo Fisher Scientific) via a ConFlo IV open-split interface (Thermo Fisher Scientific) as described by Gebauer, Preiss, & Gebauer [11]. Isotope abundances are given in δ values, calculated according to the following equation:

$$\delta^{13}\text{C}, \delta^{15}\text{N}, \text{ or } \delta^{18}\text{O} = \left(R_{\text{sample}} / R_{\text{standard}} - 1 \right) \times 1000 \text{ (‰)} \quad (2)$$

where R_{sample} and R_{standard} are the ratios of heavy to light isotope of the samples and the respective standard. This so called “delta notation” is used to express the difference between the isotope-number ratio of an unknown sample and an internationally agreed zero-point with known isotopic composition. Nowadays, it is commonly used in many fields of isotope research, as it enables a more convenient and accurate expression of small differences in isotope-number ratios that are usually in the order of 10^{-3} or smaller. The standard gases (Riessner-Gase, Lichtenfels, Germany) were calibrated with respect to international standards (CO_2 vs. V-PDB, N_2 vs. N_2 in air, CO vs. V-SMOW) with the reference substances IAEA-CH-6, IAEA-CO-8 and NBS 18 for the C isotopes, N1 and N2 for the N isotopes and IAEA601 and IAEA602 for the O isotopes, all provided by the IAEA (International Atomic Energy Agency, Vienna, Austria). Measurements of the laboratory standard acetanilide were routinely performed to control the reproducibility of C and N isotope abundance analysis [12]. Acetanilide was monitored with variable sample weight at least six times per batch of 50 samples. For oxygen isotope abundance measurements, benzoic acid was analyzed with variable sample weight at least six times within each run of 40 samples. The uncertainty of measurements was always below 0.2‰ for $\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and 0.6‰ for $\delta^{18}\text{O}$. Data evaluation was performed in the software ISODAT (Version 2.0, Thermo Fisher Scientific).

2.4. Fourier-Transform Near-Infrared Spectroscopy (FT-NIR)

FT-NIR measurements were performed on a Bruker MPA spectrometer (Bruker Optik GmbH, Ettlingen, Germany) equipped with an interferometer, an integrating sphere for diffused reflection and a lead sulfide (PbS) detector. The sample cup was filled to the brim with whole grains and placed on a sample rotator. Spectra were recorded in triplicates using 64 scans in the range of 12,500 to 3600 cm^{-1} at a resolution of 16 cm^{-1} . Data were collected and preprocessed applying baseline correction using the concave rubber-band method with 10 iterations on 64 baseline points in OPUS (Bruker Optik GmbH). Analysis of FT-NIR spectra was limited to the range between 11,000 to 4000 cm^{-1} , and therefore excluded the flanking regions that display a strongly reduced signal-to-noise ratio.

2.5. Proton Nuclear Magnetic Resonance (^1H -NMR)

For ^1H -NMR measurements, grains were ground to wholemeal flour in a ball mill (MM400, Retsch GmbH) at 30 Hz for 2 min. 50 mg of wholemeal flour were weighed in a 2 mL reaction tube (Carl Roth GmbH, Karlsruhe, Germany) and 2 mL of HPLC-grade water was added. The sample was extracted for 2 h in an overhead shaker (Tube Revolver D-6050, neoLab Migge GmbH, Heidelberg, Germany) at 40 rpm before centrifugation for 45 min at 13,300 rpm (MIKRO 185, Andreas Hettich GmbH, Tuttlingen, Germany). 900 μL of the supernatant was transferred in a 2 mL reaction tube (Carl Roth GmbH) and 100 μL of buffer was added [13]. The pH was adjusted to 3.10 (± 0.02) using an automated titration

system (BTpH Unit, Bruker BioSpin, Rheinstetten, Germany). The sample was centrifuged again for 45 min at 13,300 rpm (Andreas Hettich GmbH) and 600 μ L of the supernatant was transferred in a 5 mm NMR tube. A 400 MHz FoodScreenerTM spectrometer equipped with a BBI probe head with z gradients was used to acquire the spectra (Bruker Biospin) with a spectral width of 20.458 ppm. All samples were handled with the IconNMR and TopSpin 3.5 software. A 1D-NOESY-experiment (noesyggpr1d) was performed with an acquisition time of 4 s and 32 scans. The excitation pulse was determined before each measurement and the temperature was set to 300 K. The receiver gain was set to a constant value of 32. An exponential window function with an LB of 0.3 was applied during automatic processing. Spectra were referenced to the signal of the internal standard 3-(trimethylsilyl)-propanoic acid- d_4 sodium salt (TSP, $\delta = 0$ ppm). To ensure the qualitative comparability of the barley flour extracts, the recovery rate of at least 95% was determined using the signal of the compound betaine, which was found to be a promising marker for drought response of barley in the course of this study, after double extraction and measurement of a representative sample.

2.6. Data Preprocessing and Statistical Analysis

Plots for temperature and precipitation data were done in R version 3.5.3 using additional packages (ggplot2, ggridges, RColorBrewer). Maps displaying the aridity index grids with a spatial resolution of 1 km $\times$ 1 km were reproduced graphically in the open-source software QGIS version 3.1 (Quantum Geographic Information System). Univariate statistical analysis of isotope data was also conducted in the computing environment R. As the assumptions for the use of a parametric test were not met, differences between groups were evaluated by a non-parametric two-sided Mann–Whitney–Wilcoxon test ($\alpha = 0.05$). Chemometric analysis (Kruskal–Wallis test, PCA) of FT-NIR spectra and combined data were performed in MATLAB (Version R2019a, TheMathWorks, Natick, MA, USA). Dimensionality reduction was achieved by splitting the FT-NIR-spectra into bins. Binning (also called bucketing) is a very common pre-processing step for spectral data, that often contain thousands of data points/variables. To ease statistical analysis, the number of variables is reduced by dividing each spectrum into for example equidistant (uniform) so-called bins/buckets and calculating the integral (total area within each bucket) of each bin. In our particular case, each FT-NIR spectrum was divided into 300 equidistant regions in the range between 11,000 and 4000 cm^{-1} . Afterwards, the IRMS variables ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$) were attached to the data set. Feature selection was performed for combined data, extracting the 50 most relevant variables based on ranked class-related variance using a Kruskal–Wallis test. Specifically, the test statistic (H statistic) of the Kruskal–Wallis test was calculated and used as a filter score, where higher values of the test statistic indicate that the values of the corresponding variable differ more between the classes. Subsequently, the test statistic values were sorted from the highest value to the lowest value to reveal the most relevant variables. Even if statistical approaches are sophisticated, they should not be blindly relied upon. For this reason, we looked at the extracted variables manually to verify that they indeed belong to meaningful spectral regions rather than, e.g., noise. Here, all of the 50 extracted variables were found to be relevant, hence they were used for further statistical analysis. Selected features were then autoscaled (mean-centering followed by division of each column/variable by the standard deviation of that column) and processed using principal component analysis (PCA), to obtain an impression of the differences between spectra among harvest years. As PCA is an exploratory technique rather than a classification technique, we defined a criterium based on which a sample is assigned to a class. Here, we have classified the observations according to their position in the PCA score subspace by defining a threshold based on Euclidean distance to the calculated center of each group at a 95% confidence level ($\alpha = 5\%$). Using this classification technique, model validation using Monte Carlo cross-validation was performed. Specifically, 100 random models were generated with 90% of the data set as training data and the remaining 10% as test set.

3. Results and Discussion

3.1. Weather at the Sampling Site Differed between 2017 and 2018

Total precipitation at the sampling site in 2017 was 573.9 L/m², which corresponds to 106% of the mean of the 30-year reference period (1981–2010). In comparison, total precipitation in 2018 was 70% of the mean of the reference period with a total of 377.1 L/m². During the three summer months (June–August), the sum of rainfall was 285.1 L/m² (155%) in 2017, while in 2018, only 77 L/m² were recorded, meeting 42% of the reference period mean (Figure 1a). In addition, the 2018 summer months had 48 days with maximum temperatures above 25 °C (2017: 28 days), which very likely led to co-occurring heat and drought stress (Figure 1b).

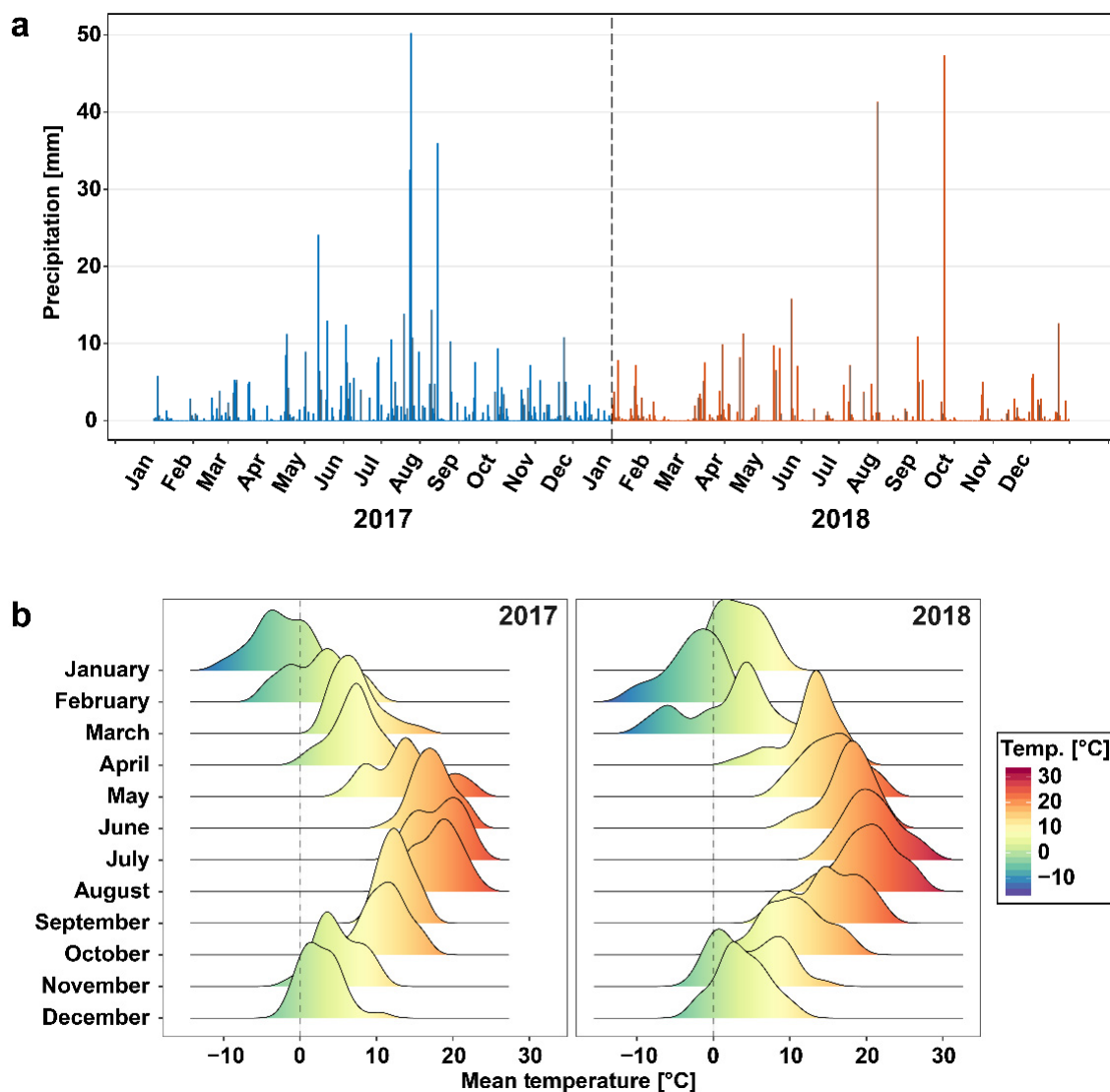


Figure 1. Precipitation and temperature at the sampling site. (a) The graph displays the daily precipitation totals (in mm) at the sampling site for the crop years 2017 (blue) and 2018 (red). (b) Density plots show the distribution of daily average air temperatures (in °C) for the crop years 2017 (left) and 2018 (right). Raw data were retrieved from the weather station (Erfurt-Weimar, German Meteorological Service) with the closest proximity to the sample origin.

Recent findings showed that most barley varieties are particularly sensitive to a combination of these stresses in comparison to heat or drought individually [14,15]. Both precipitation and temperature are included in the calculation of the De Martonne's aridity index (I_{DM}), which was plotted on a map of Germany. Based on the classes of this frequently

used index, the barley experienced humid to Mediterranean ($I_{DM} \leq 30$) conditions in 2017. In contrast, the year 2018 in the region the samples were sourced was characterized by semi-dry to dry conditions ($I_{DM} \leq 20$) [9] (Figure 2).

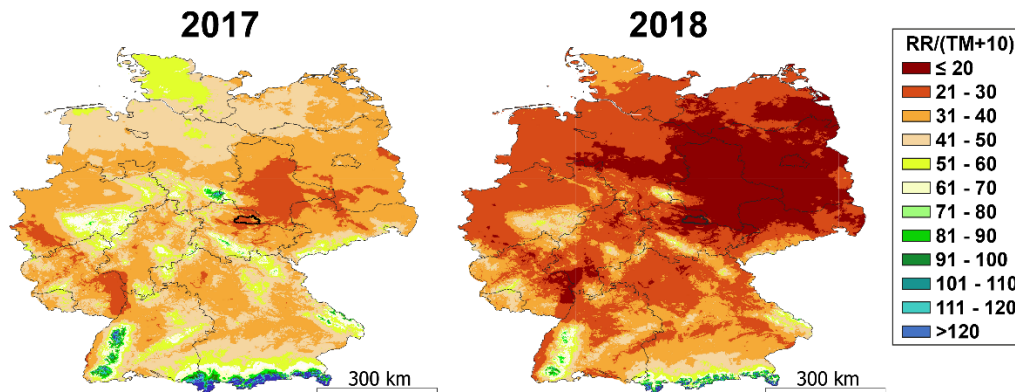


Figure 2. Comparison of the aridity in Germany. Maps show the spatial variation in annual drought index (De Martonne aridity index, I_{DM}) for Germany in the years 2017 (left) and 2018 (right) and the map scale. All values are calculated using the formula: $RR/(TM + 10)$, where RR is the annual sum of precipitation (in mm) and TM is the annual mean air temperature (in °C). Lower index values refer to more pronounced aridity at respective location. The sample origin is highlighted with a thick line and shows index values below 30 in 2017 (humid to Mediterranean) and values below 20 in 2018 (semi-dry to dry).

3.2. Stable Isotope Abundances Indicate Severe Drought Stress

Carbon isotope abundances of all analyzed barley samples ranged from -24.5‰ to -27.6‰ with a mean of -26.3‰ , which perfectly fits the typical average ratio for the C_3 crop barley of $-25.9 \pm 0.7\text{‰}$ [16]. Samples from the harvest year 2018 showed significantly less negative $\delta^{13}\text{C}$ values for both Quench (Mdn, $N_{2017} = -27.1\text{‰}$, $n = 19$; Mdn, $N_{2018} = -25.2\text{‰}$, $n = 20$; $U = 0$, $p < 0.0001$) and Steffi (Mdn, $N_{2017} = -26.7\text{‰}$, $n = 22$; Mdn, $N_{2018} = -26.0\text{‰}$, $n = 25$; $U = 52.5$, $p < 0.0001$) compared to samples from 2017 with a more pronounced difference for the Quench variety. The results match the values of studies from Ferrio et al., who found less negative $\delta^{13}\text{C}$ values for bread wheat kernels grown under increased drought stress [17]. Between the two trials that varied the most in water availability, they observed a difference in carbon isotope composition of 2.2‰ . The less negative values in kernels grown under heavy drought stress match our expectations as the relationship between water status and the carbon isotopic composition of plants and the underlying mechanism are well researched [18]. During assimilation of CO_2 , plants utilizing the C_3 photosynthetic pathway (e.g., barley) discriminate against the heavy isotope of carbon (^{13}C). Under normal conditions, the CO_2 -fixing enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) favors the light $^{12}\text{CO}_2$. Stomatal closure in response to drought stress reduces water loss caused by transpiration but also leads to CO_2 being ‘trapped’ inside the leaf’s internal air spaces. The closed stomata force the enzyme to process more of the heavy $^{13}\text{CO}_2$, as the trapped CO_2 is recycled [19,20]. The resulting change in isotopic composition of the biomass can be detected in various parts of crops but appears to be most pronounced in mature grains, as they might serve as a sink for different photosynthetic plant organs [21]. Compared to Quench, which is a high-yielding barley cultivar, Steffi is a rather old and less cultivated variety for malting. Therefore, differences in genotype and reaction to drought might have attributed to the more pronounced difference in ^{13}C abundance for Quench (Figure 3a,d) [22].

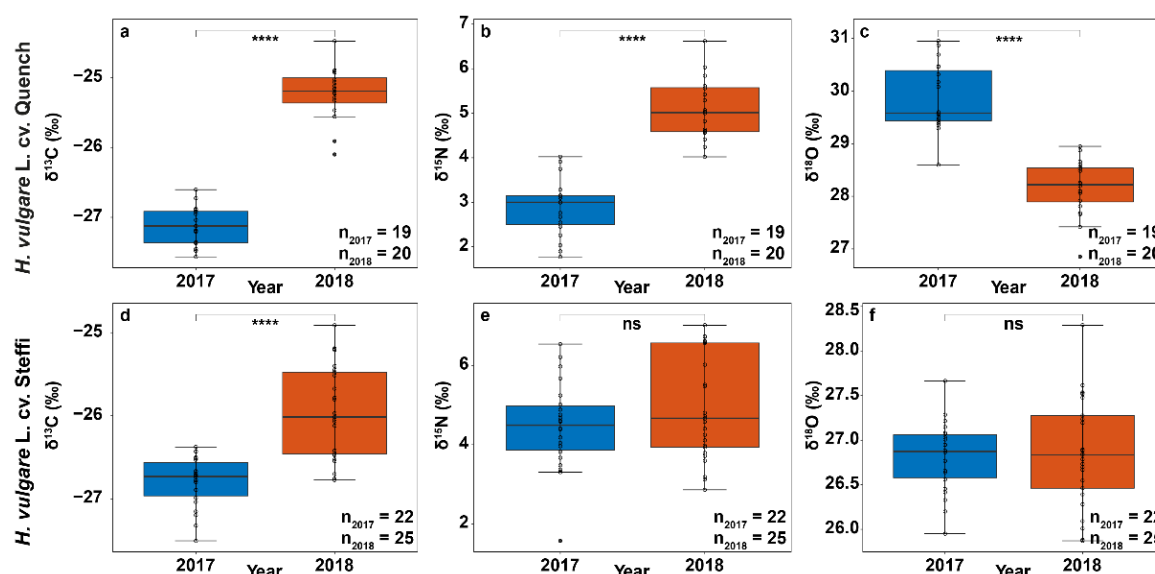


Figure 3. Natural stable isotope abundances of the analyzed spring barley samples. Box and whisker plot graphs with interquartile ranges, medians, error bars (non-outlier minimum and maximum) and results of the Mann–Whitney–Wilcoxon test on the isotopic values ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$) of spring barley samples from the two varieties (a–c) Quench and (d–f) Steffi. Samples are grouped based on the crop years 2017 (blue) and 2018 (red). Highly significant differences ($p < 0.0001$) are highlighted with four stars (****), while non-significant differences are marked with 'ns'. Outliers are denoted as filled circles.

In fact, there is growing evidence that some barley genotypes, which carry more desirable wild-type traits, have a higher abiotic stress tolerance to environmental factors such as drought and heat, thus raising interest in targeted breeding of cultivated barley, especially in view of the challenges posed by climate change [23,24]. The reported N isotopic compositions showed significant interannual differences for the Quench variety (Mdn, $N_{2017} = 3.0\text{‰}$, $n = 19$; Mdn, $N_{2018} = 5.0\text{‰}$, $n = 20$; $U = 1$, $p < 0.0001$) with a higher $\delta^{15}\text{N}$ in the dry year 2018, while no difference was found for Steffi (Mdn, $N_{2017} = 4.5\text{‰}$, $n = 22$; Mdn, $N_{2018} = 4.7\text{‰}$, $n = 25$; $U = 229.5$, $p > 0.05$) (Figure 3). The connection between plant $\delta^{15}\text{N}$ values and aridity are not straightforward, as multiple factors such as microorganism activity, N source, and plant metabolism influence the signature. The observed increase in $\delta^{15}\text{N}$ for Quench might be attributed to differences in agricultural practice, response of nitrate reductase activity to drought or due to limited uptake and transfer of N to the upper plant parts [25]. Apart from the information that the barley comes from conventional cultivation, details on fertilization are missing, which makes interpretation of the $\delta^{15}\text{N}$ values rather challenging. At the same time, such obstacles reflect the intention of the study, namely to create real conditions that can be transferred to the market. In our approach, these uncertainties can, to some extent, be compensated by using several measurement techniques for a more holistic point of view. As for nitrogen isotope composition, $\delta^{18}\text{O}$ was significantly different between the two harvest years for the variety Quench (Mdn, $N_{2017} = 29.6\text{‰}$, $n = 19$; Mdn, $N_{2018} = 28.2\text{‰}$, $n = 20$; $U = 376$, $p < 0.0001$) with lower $\delta^{18}\text{O}$ values in 2018, while no significant difference was found for the Steffi variety (Mdn, $N_{2017} = 26.9\text{‰}$, $n = 22$; Mdn, $N_{2018} = 26.8\text{‰}$, $n = 25$; $U = 270$, $p > 0.05$) (Figure 3). A missing interannual difference in $\delta^{18}\text{O}$ values is rather unexpected as similar to $\delta^{13}\text{C}$ values, oxygen isotopic signatures are expected to increase with higher evaporative demand and lower stomatal conductance [17]. However, several studies suggest that there is a lack of a clear relationship in grains compared to, e.g., leaf material or whole plant material. For spring wheat, Barbour, Fischer, Sayre, & Farquhar reported a significant positive correlation of $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ in flag leaves, but the relationship was less clear in grains [26]. We conclude, while grain $\delta^{13}\text{C}$ seems to be a good marker for water status during grain filling, other factors seem to specifically affect $\delta^{18}\text{O}$. The signal could be dominated by variations in

$\delta^{18}\text{O}$ signature of the source water, or it could be influenced by biochemical fractionation during synthesis of grain organic matter from remobilized assimilates [17]. Considering the heavy drought in 2018, a higher contribution of pre-anthesis reserves to grain filling is very likely, which would lead to lower $\delta^{18}\text{O}$ values [27,28].

3.3. FT-NIR Spectra Show Additional Drought-Related Differences

For visual presentation of the differences in the FT-NIR spectra between the samples from the two harvest years, average spectra were calculated, and signal assignments of fundamental regions were included (Figure 4).

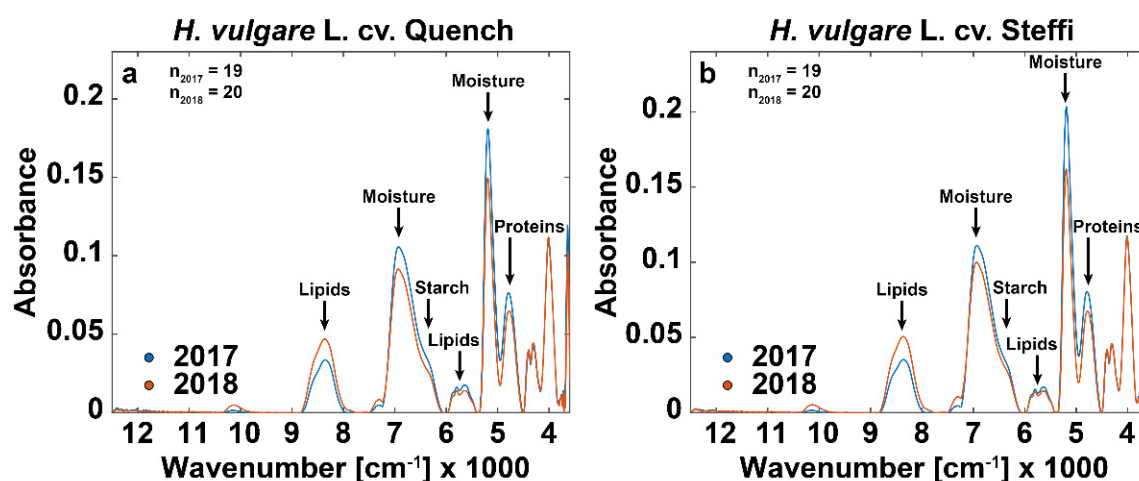


Figure 4. Comparison of spring barley kernel FT-NIR spectra. Figures display baseline-corrected average FT-NIR spectra (4000 to 12,000 cm^{-1}) for whole kernels of the spring barley varieties (a) Quench and (b) Steffi. Samples are grouped based on the crop years 2017 (blue) and 2018 (red). Assignments of relevant spectral bands are indicated at the respective absorbance peak.

It is noteworthy that the spectra of both cultivars show the same changes for the year 2018, indicating similar metabolite responses to drought stress. The absorbance peak at around 8264 cm^{-1} (1200 nm) was related to the second overtone of C-H stretching, whereas the absorbance peak at 5665 cm^{-1} (1765 nm) was related to the first harmonic of C-H stretching. Both peaks are attributed with lipids and showed noteworthy differences between the harvest years, which is consistent with results from a study using GC/MS-based metabolite profiling [29–31]. Further, we observed two absorbance bands with outstanding separation between harvest years that were linked to moisture at around 6800 cm^{-1} (1470 nm , first overtone of O-H stretching) [29,32,33] and 5170 cm^{-1} (1934 nm , O-H stretching) [29,30,32,34,35]. Based on a nationwide report by the Federal Ministry of Food and Agriculture, in 2018 all cereals, including barley, showed significant reduction in average moisture content in comparison to the previous year, which fits our observations [36]. The absorbance band around 6300 cm^{-1} (1587 nm , first overtone of O-H stretching) is a bit lower for the grain samples from 2018 and was associated with starch, whose synthesis is known to be hindered by drought stress [29,30,34,37]. Finally, we could observe differences between harvest years at the absorbance peak around 4854 cm^{-1} (2060 nm). The band is associated with a combination of C-H stretching, N-H stretching, and O-H stretching and is related to variation in protein characteristics potentially attributed to osmotic adjustments [29,31,32,34].

3.4. Data Fusion Enables Differentiation of Barley Harvest Years

In a next step, we statistically evaluated the FT-NIR spectra and data produced by IRMS, using Kruskal–Wallis tests and principal component analysis (PCA). The created models achieved 100% correct classification of harvest years for both barley varieties (Figure 5).

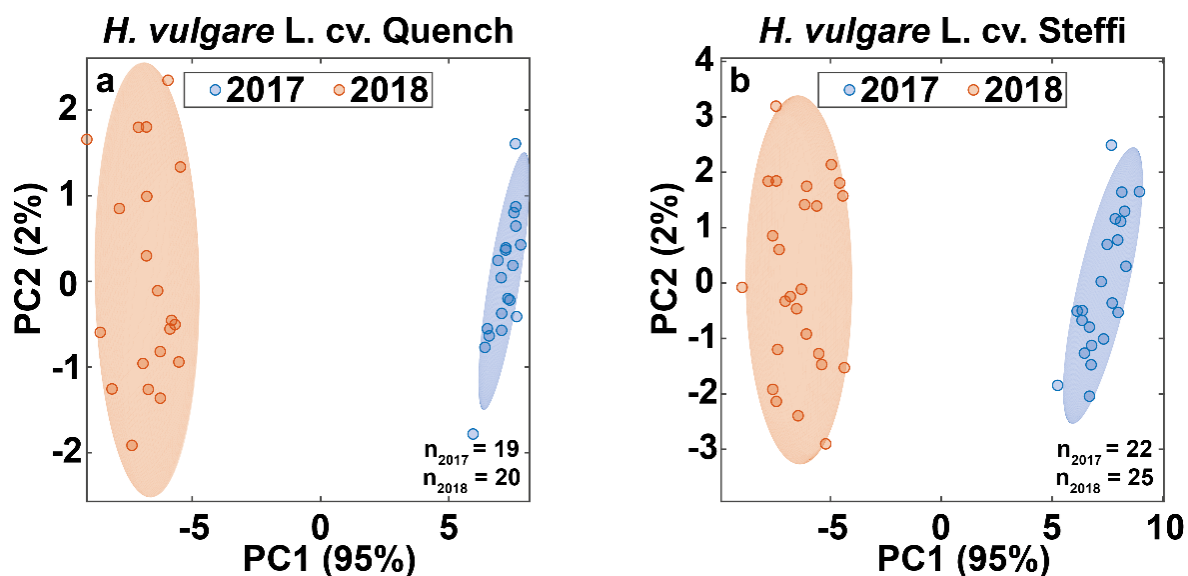


Figure 5. Chemometric analysis of combined FT-NIR and IRMS data. Score plots including 95% confidence ellipses of the first two principal components from the PCA of fused data (FT-NIR and IRMS). Plots are displayed for the two spring barley varieties (a) Quench and (b) Steffi, and the samples are colored by the crop years 2017 (blue) and 2018 (red). For each principal component, the respective percentage of variance explained is given in the axis label.

To gain insight into which of the included variables contributed most to the separation, we ranked them by importance using a Kruskal–Wallis test. We found that the abundance peak at around 8264 cm^{-1} (1200 nm) was the most important, followed by the absorbance bands around 4854 cm^{-1} (2060 nm), 5170 cm^{-1} (1934 nm), 5665 cm^{-1} (1765 nm), and 6800 cm^{-1} (1470 nm). This is consistent with findings from the literature, which state that the composition and content of lipids and proteins in barley change significantly under stress conditions such as heat and drought [24,38]. Interestingly, in the statistical analysis of the Quench variety, the carbon isotopic composition appeared among the most important variables for the separation of crop years, but not for the Steffi variety. Given that the difference in the natural ^{13}C abundance is a marker for water-use efficiency, there is growing evidence that the old Steffi variety is better adapted to drought [20].

3.5. ^1H -NMR Confirms Observations by Means of the Marker Betaine

Finally, we used ^1H -NMR spectra of sample extracts that serve as quantitative metabolite fingerprints. These can not only be analyzed in a non-targeted approach, but also in high-resolution identification of molecular markers. In our case, we focused on plant physiological markers such as osmolytes, which might indicate a plant's response to stress conditions. A direct comparison of the ^1H -NMR spectra from the two investigated harvest years revealed a significantly increased signal intensity at 3.255 ppm chemical shift in all samples from the year 2018. Importantly, this is regardless of which of the two barley varieties was considered (Figure 6).

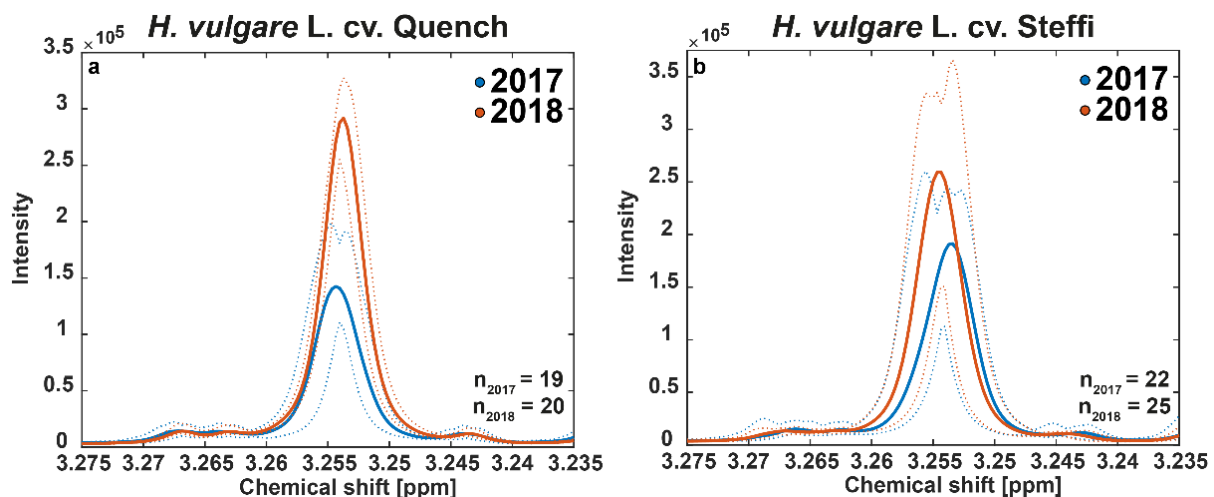


Figure 6. Comparison of the content of the plant-physiological marker betaine. Plots show the ^1H -NMR spectra region (3.275–3.235 ppm), including the signal of betaine at 3.255 ppm chemical shift. The two spring barley varieties (a) Quench and (b) Steffi were analyzed as wholemeal flour extracts. Average spectra are displayed as solid lines and the 95% confidence intervals are represented as dotted lines. Samples are colored by the harvest years 2017 (blue) and 2018 (red). Note that there are minor differences in chemical shift due to slight variations in the sample conditions, which lead to apparent multiplets when the included spectra are averaged. This also creates multiple maxima for the confidence intervals (dotted lines). However, the betaine signal at 3.255 ppm is a singlet.

A comparison with common databases and subsequent spiking experiments for confirmation with the pure substance revealed that the signal refers to betaine (trimethylglycine). Accumulation of this methyl donor is a well-known and critical response of barley to environmental stress such as drought [39]. We would like to mention here that for the purposes of this paper, we consider only this section of the spectrum, as it provides the marker for drought stress and thus the biological background. However, it can be assumed that an elaborate statistical analysis of the complete NMR spectra could reveal other exciting relationships. On a molecular level, the connection between drought stress and betaine accumulation can be explained by the function of the plant hormone abscisic acid (ABA), which not only induces stomatal closure to prevent water loss but was also found to increase expression of betaine aldehyde dehydrogenase (BADH), catalyzing the last step in the synthesis of betaine [40]. The relationship is underpinned by the fractionation we observed for the carbon isotopic composition, which is caused by stomata regulation. Although this important plant physiological process has often been demonstrated by experiments with whole plants or leaves, elevated concentrations of betaine resulting from high temperatures and drought have also been observed in studies with grains of wheat and barley [31,41]. In respective results, there were clear differences between individual genotypes, on the one hand in the basic content of betaine, but also in the degree of accumulation in response to stress. This might explain the difference observed between the two investigated varieties, with the old cultivar once more showing a milder response, possibly due to a higher tolerance to environmental stress.

4. Conclusions

Global warming creates circumstances for the agricultural sector that require rapid action and unprecedented scientific approaches. A central component is the transfer of knowledge from controlled field experiments to real market samples. Our results show that spring barley undergoes noticeable changes in its chemical composition in an extremely dry and hot year like 2018, enabling a solid discriminability from samples of 2017. The significantly less negative $\delta^{13}\text{C}$ values for both analyzed varieties proved to be a suitable marker for water status during grain filling. Furthermore, FT-NIR and ^1H -NMR spectra showed significant differences between the two harvest years, especially in the absorbance

bands related to lipids and proteins and the stress marker betaine. All observed chemical changes could be explained and linked by means of plant physiology, which is an integral part of authenticity testing of plant-based foods. Furthermore, the weaker stress response of the old cultivar “Steffi” has shown that the data obtained can also be transferred to targeted breeding of varieties that tolerate future climate conditions. We are convinced that additional research on our initial findings will show the broad applicability to unknown systems and will be the go-to practice in agricultural science.

Author Contributions: Conceptualization, S.G.B., G.G., and S.S. (Stephan Schwarzingler); methodology, S.G.B.; software, S.G.B., F.R., S.S. (Stephan Schwarzingler), and P.K.; validation, S.G.B., S.S. (Simon Steidele), and P.K.; formal analysis, S.G.B.; investigation, S.G.B., P.K., S.S., and L.K.; resources, H.W., S.S. (Stephan Schwarzingler), and G.G.; data curation, S.G.B., H.W., F.R., L.K., and P.K.; writing—original draft preparation, S.G.B.; writing—review and editing, S.G.B., S.S. (Stephan Schwarzingler), and G.G.; visualization, S.G.B.; supervision, S.S. (Stephan Schwarzingler), G.G., L.K., and H.W.; project administration, S.S. (Stephan Schwarzingler); funding acquisition, S.S. (Stephan Schwarzingler). All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program (Reference number: 2816502414). This publication was funded by the German Research Foundation (DFG) and the University of Bayreuth in the Open Access Publishing funding program.

Data Availability Statement: All weather-related data can be accessed from the Climate Data Center, provided by the German Meteorological Service (<https://cdc.dwd.de/portal>), https://opendata.dwd.de/climate_environment/CDC/). The MATLAB code used for the chemometric analysis of data is available at the gitlab repository of the co-author Felix Rüll (<https://gitlab.com/FelR/ai-omics>). Detailed raw data for IRMS, FT-NIR and ¹H-NMR measurements are available upon request from the corresponding author.

Acknowledgments: First: we would like to thank Paul Rösch (University of Bayreuth) for his great support in the project. We also thank Carina Bauer, Petra Eckert and Christine Tiroch (University of Bayreuth) for performing the stable isotope abundance measurements. Further, we thank Felix Brauer (ALNuMed GmbH, Bayreuth, Germany) for technical assistance with ¹H-NMR analyses and Stephen Strassburg (University of Bayreuth) for comments on language. Special thanks to all AgrOr-project partners who made this work possible.

Conflicts of Interest: S.S. (Stephan Schwarzingler) is co-founder and former CEO (approved secondary employment) of the ALNuMed GmbH (Bayreuth, Germany). P.K. and L.K. are former employees of the ALNuMed GmbH (Bayreuth, Germany). H.W. is employed by the IREKS GmbH (Kulmbach, Germany). All authors declare no conflicts of interest.

References

1. Sun, Q.; Gilgen, A.K.; Signarbieux, C.; Klaus, V.H.; Buchmann, N. Cropping Systems Alter Hydraulic Traits of Barley but not Pea Grown in Mixture. *Plant Cell Environ.* **2021**, 1–13. [CrossRef]
2. Trnka, M.; Rötter, R.P.; Ruiz-Ramos, M.; Kersebaum, K.C.; Olesen, J.E.; Žalud, Z.; Semenov, M.A. Adverse weather conditions for European wheat production will become more frequent with climate change. *Nat. Clim. Chang.* **2014**, 4, 637–643. [CrossRef]
3. Giraldo, P.; Benavente, E.; Manzano-Agugliaro, F.; Gimenez, E. Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy* **2019**, 9, 352. [CrossRef]
4. Paynter, B.H.; Young, K.J. Grain and malting quality in two-row spring barley are influenced by grain filling moisture. *Aust. J. Agric. Res.* **2004**, 55, 539–550. [CrossRef]
5. Federal Ministry of Food and Agriculture (BMEL)—Erntebericht 2018. Available online: https://www.bmel.de/SharedDocs/Downloads/DE/_Landwirtschaft/Pflanzenbau/Ernte-Bericht/ernte-2018.pdf?sessionId=E982D6918788080F847DB05B67B27E07.live921?__blob=publicationFile&v=2 (accessed on 23 May 2021).
6. Heinrich Durst Malzfabriken GmbH & Co. KG. Requirements and Conditions of Barley Deliveries to Durst Malz—Crop 2020. Available online: http://www.durst-malz.de/uploads/media/Requirements_Crop_20_EN.pdf (accessed on 23 May 2021).
7. Xie, W.; Xiong, W.; Pan, J.; Ali, T.; Cui, Q.; Guan, D.; Meng, J.; Mueller, N.D.; Lin, E.; Davis, S.J. Decreases in global beer supply due to extreme drought and heat. *Nat. Plants* **2018**, 4, 964–973. [CrossRef] [PubMed]

8. Kaspar, F.; Müller-Westermeier, G.; Penda, E.; Mächel, H.; Zimmermann, K.; Kaiser-Weiss, A.; Deutschländer, T. Monitoring of climate change in Germany—data, products and services of Germany’s National Climate Data Centre. *Adv. Sci. Res.* **2013**, *10*, 99–106. [\[CrossRef\]](#)
9. de Martonne, E. Une nouvelle fonction climatologique: L’indice d’aridité. *Meteorologie* **1926**, *2*, 449–458.
10. Bidartondo, M.I.; Burghardt, B.; Gebauer, G.; Bruns, T.D.; Read, D.J. Changing partners in the dark: Isotopic and molecular evidence of ectomycorrhizal liaisons between forest orchids and trees. *Proc. R. Soc. B Biol. Sci.* **2004**, *271*, 1799–1806. [\[CrossRef\]](#)
11. Gebauer, G.; Preiss, K.; Gebauer, A.C. Partial mycoheterotrophy is more widespread among orchids than previously assumed. *New Phytol.* **2016**, *211*, 11–15. [\[CrossRef\]](#)
12. Gebauer, G.; Schulze, E.D. Carbon and nitrogen isotope ratios in different compartments of a healthy and a declining *Picea abies* forest in the Fichtelgebirge, NE Bavaria. *Oecologia* **1991**, *87*, 198–207. [\[CrossRef\]](#)
13. Godelmann, R.; Fang, F.; Humpfer, E.; Schütz, B.; Bansbach, M.; Schäfer, H.; Spraul, M. Targeted and nontargeted wine analysis by ¹H NMR spectroscopy combined with multivariate statistical analysis. Differentiation of important parameters: Grape variety, geographical origin, year of vintage. *J. Agric. Food Chem.* **2013**, *61*, 5610–5619. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Mahalingam, R. Phenotypic, physiological and malt quality analyses of US barley varieties subjected to short periods of heat and drought stress. *J. Cereal Sci.* **2017**, *76*, 199–205. [\[CrossRef\]](#)
15. Mahalingam, R.; Bregitzer, P. Impact on physiology and malting quality of barley exposed to heat, drought and their combination during different growth stages under controlled environment. *Physiol. Plant.* **2019**, *165*, 277–289. [\[CrossRef\]](#)
16. Brooks, J.R.; Buchmann, N.; Phillips, S.; Ehleringer, B.; Evans, R.D.; Lott, M.; Martinelli, L.A.; Pockman, W.T.; Sandquist, D.; Sparks, J.P.; et al. Heavy and light beer: A carbon isotope approach to detect C₄ carbon in beers of different origins, styles, and prices. *J. Agric. Food Chem.* **2002**, *50*, 6413–6418. [\[CrossRef\]](#)
17. Ferrio, J.P.; Mateo, M.A.; Bort, J.; Abdalla, O.; Voltas, J.; Araus, J.L. Relationships of grain $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ with wheat phenology and yield under water-limited conditions. *Ann. Appl. Biol.* **2007**, *150*, 207–215. [\[CrossRef\]](#)
18. Hubick, K.; Farquhar, G. Carbon isotope discrimination and the ratio of carbon gained to water lost in barley cultivars. *Plant Cell Environ.* **1989**, *12*, 795–804. [\[CrossRef\]](#)
19. Farquhar, G.D.; O’Leary, M.H.; Berry, J.A. On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. *Aust. J. Plant Physiol.* **1982**, *9*, 121–137. [\[CrossRef\]](#)
20. Farquhar, G.D.; Richards, R.A. Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. *Aust. J. Plant Physiol.* **1984**, *11*, 539–552. [\[CrossRef\]](#)
21. Sánchez-Díaz, M.; García, J.L.; Antolín, M.C.; Araus, J.L. Effects of soil drought and atmospheric humidity on yield, gas exchange, and stable carbon isotope composition of barley. *Photosynthetica* **2002**, *40*, 415–421. [\[CrossRef\]](#)
22. Liu, H.; Guo, B.; Wei, Y.; Wei, S.; Ma, Y.; Zhang, W. Effects of region, genotype, harvest year and their interactions on $\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and δD in wheat kernels. *Food Chem.* **2015**, *171*, 56–61. [\[CrossRef\]](#)
23. Ivandic, V.; Hackett, C.A.; Zhang, Z.J.; Staub, J.E.; Nevo, E.; Thomas WT, B.; Forster, B.P. Phenotypic responses of wild barley to experimentally imposed water stress. *J. Exp. Bot.* **2000**, *51*, 2021–2029. [\[CrossRef\]](#)
24. Wu, X.; Cai, K.; Zhang, G.; Zeng, F. Metabolite profiling of barley grains subjected to water stress: To explain the genotypic difference in drought-induced impacts on malting quality. *Front. Plant Sci.* **2017**, *8*, 1547. [\[CrossRef\]](#)
25. Lopes, M.S.; Araus, J.L. Nitrogen source and water regime effects on durum wheat photosynthesis and stable carbon and nitrogen isotope composition. *Physiol. Plant.* **2006**, *126*, 435–445. [\[CrossRef\]](#)
26. Barbour, M.M.; Fischer, R.A.; Sayre, K.D.; Farquhar, G.D. Oxygen isotope ratio of leaf and grain material correlates with stomatal conductance and grain yield in irrigated wheat. *Aust. J. Plant Physiol.* **2000**, *27*, 625–637. [\[CrossRef\]](#)
27. Flood, R.G.; Martin, P.J.; Gardner, W.K. Dry matter accumulation and partitioning and its relationship to grain yield in wheat. *Aust. J. Exp. Agric.* **1995**, *35*, 495–502. [\[CrossRef\]](#)
28. Slafer, G.A.; Satorre, E.H.; Andrade, F.H. Increases in grain yield in bread wheat from breeding and associated physiological changes. In *Genetic Improvement of Field Crops*, 1st ed.; Slafer, G.A., Ed.; Marcel Dekker Inc.: New York, NY, USA, 1994; pp. 1–54.
29. De Girolamo, A.; Cortese, M.; Cervellieri, S.; Lippolis, V.; Pascale, M.; Logrieco, A.F.; Suman, M. Tracing the geographical origin of durum wheat by FT-NIR spectroscopy. *Foods* **2019**, *8*, 450. [\[CrossRef\]](#)
30. Shenk, J.S.; Workman, J.J.; Westerhaus, M.O. Application of NIR spectroscopy to agricultural products. In *Handbook of Near-Infrared Analysis*, 3rd ed.; Burns, D.A., Ciurczak, E.W., Eds.; Marcel Dekker Inc.: New York, NY, USA, 2001; pp. 419–474.
31. Wenzel, A.; Frank, T.; Reichenberger, G.; Herz, M.; Engel, K.H. Impact of induced drought stress on the metabolite profiles of barley grain. *Metabolomics* **2015**, *11*, 454–467. [\[CrossRef\]](#)
32. Cozzolino, D.; Roumeliotis, S.; Eglinton, J. Exploring the use of near infrared (NIR) reflectance spectroscopy to predict starch pasting properties in whole grain barley. *Food Biophys.* **2013**, *8*, 256–261. [\[CrossRef\]](#)
33. Zhao, H.; Guo, B.; Wei, Y.; Zhang, B. Effects of grown origin, genotype, harvest year, and their interactions of wheat kernels on near infrared spectral fingerprints for geographical traceability. *Food Chem.* **2014**, *152*, 316–322. [\[CrossRef\]](#) [\[PubMed\]](#)
34. McGoverin, C.M.; Engelbrecht, P.; Geladi, P.; Manley, M. Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics. *Anal. Bioanal. Chem.* **2011**, *401*, 2283. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Sohn, M.; Himmelsbach, D.S.; Barton, F.E.; Griffey, C.A.; Brooks, W.; Hicks, K.B. Near-infrared analysis of whole kernel barley: Comparison of three spectrometers. *Appl. Spectrosc.* **2008**, *62*, 427–432. [\[CrossRef\]](#) [\[PubMed\]](#)

-
36. Federal Ministry of Food and Agriculture (BMEL)—Besondere Ernte- und Qualitätsermittlung 2018. Available online: <https://www.bmel-statistik.de/fileadmin/daten/EQB-1002000-2018.pdf> (accessed on 23 May 2021).
 37. Chaves, M.M. Effects of water deficits on carbon assimilation. *J. Exp. Bot.* **1991**, *42*, 1–16. [[CrossRef](#)]
 38. Högy, P.; Poll, C.; Marhan, S.; Kandeler, E.; Fangmeier, A. Impacts of temperature increase and change in precipitation pattern on crop yield and yield quality of barley. *Food Chem.* **2013**, *136*, 1470–1477. [[CrossRef](#)] [[PubMed](#)]
 39. Kebede, A.; Kang, M.S.; Bekele, E. Advances in mechanisms of drought tolerance in crops, with emphasis on barley. In *Advances in Agronomy*; Sparks, D.L., Ed.; Academic Press: London, UK, 2019; Volume 156, pp. 265–314. [[CrossRef](#)]
 40. Ishitani, M.; Nakamura, T.; Han, S.Y.; Takabe, T. Expression of the betaine aldehyde dehydrogenase gene in barley in response to osmotic stress and abscisic acid. *Plant Mol. Biol.* **1995**, *27*, 307–315. [[CrossRef](#)] [[PubMed](#)]
 41. Corol, D.I.; Ravel, C.; Raksegi, M.; Bedo, Z.; Charmet, G.; Beale, M.H.; Shewry, P.R.; Ward, J.L. Effects of genotype and environment on the contents of betaine, choline, and trigonelline in cereal grains. *J. Agric. Food Chem.* **2012**, *60*, 5471–5481. [[CrossRef](#)]

6.2 PUBLICATION 2

Chemometric Modeling of Trace Element Data for Origin Determination of Demantoid Garnets

Stefan G. Bindereif¹, Felix Rüll¹, Stephan Schwarzingner¹ and Clemens Schwarzingner²

Minerals 2020, 10(12), 1046; <https://doi.org/10.3390/min10121046>

¹North Bavarian NMR-Centre, University of Bayreuth, 95447 Bayreuth, Germany

²Institute for Chemical Technology of Organic Materials, Johannes Kepler University Linz, 4040 Linz, Austria

Received: 21 October 2020 / Revised: 20 November 2020 / Accepted: 22 November 2020 /
Published: 24 November 2020

Keywords: multivariate statistics; chemometrics; origin determination; garnets; demantoid; LA-ICP-MS

This is an open access article published by MDPI distributed under the terms of the Creative Commons CC BY license. No special permission is required to reuse all or part of article published by MDPI, including figures and tables. Any part of the article may be reused without permission provided that the original article is clearly cited. Reuse of an article does not imply endorsement by the authors or MDPI.

Article

Chemometric Modeling of Trace Element Data for Origin Determination of Demantoid Garnets

Stefan G. Bindereif ¹, Felix Rüll ¹ , Stephan Schwarzingner ¹ and Clemens Schwarzingner ^{2,*} 

¹ North Bavarian NMR-Centre, University of Bayreuth, 95447 Bayreuth, Germany; stefan.bindereif@uni-bayreuth.de (S.G.B.); Felix.Ruell@uni-bayreuth.de (F.R.); s.schwarzingner@uni-bayreuth.de (S.S.)

² Institute for Chemical Technology of Organic Materials, Johannes Kepler University Linz, 4040 Linz, Austria

* Correspondence: Clemens.schwarzingner@jku.at; Tel.: +43-732-2468-9011

Received: 21 October 2020; Accepted: 22 November 2020; Published: 24 November 2020



Abstract: The determination of country of origin poses a common problem in the appraisal of gemstones and is in many cases still based on the observation of inclusions and growth features of a gem, whereas chemical analysis is only done by major labs. We have used Laser Ablation Inductively Coupled Plasma Mass Spectrometry to analyze the trace element profiles of demantoid garnets from six different countries and identified a set of six elements, which are magnesium, aluminum, titanium, vanadium, chromium, and manganese, that are necessary to assign the mining regions with a good certainty. By using the logarithms of the trace element concentrations and subjecting them to chemometric modeling, we were able to separate the demantoids originating from Russia, Pakistan, Namibia, Iran, and Madagascar very well, leaving only Italy with some uncertainty. Results are presented for an “all origins” model as well as pair-wise comparison of two locations at a time, which lead to even better results.

Keywords: multivariate statistics; chemometrics; origin determination; garnets; demantoid; LA-ICP-MS

1. Introduction

The correct determination of origin of gemstones has gained more and more importance recently. The desire to know where a gem comes from is not only scientifically driven but to a very large extent by the trade. For a scientist it might be hard to understand that for example two blue sapphires, which have the same size, weight, color, clarity and cut are appraised differently only because they hail from different sources. The trade, however, introduces factors such as rarity and history into the appraisal process and thus the blue sapphire originating from e.g., Madagascar (currently producing) will be less valuable than the one that has been found in the famed Kashmir region (found in 1881 after a landslide and now exhausted) [1,2]. Additionally, ethical issues associated with gemstone production are becoming increasingly important. Hence, it will be necessary in certain cases to prove that gems do not come from conflict areas as those might be banned from trade in some countries, such as jade and rubies originating from Burma which were banned in the United States from 2008 to 2016.

In classical gemology, mainly optical characteristics have been used to determine origins, and above all, inclusions. Tiny crystals which grew in the host gem material are dependent on the geological setting and thus provide a good tool to eliminate certain sources. For instance, in sapphires it is possible to assign certain origins by identifying inclusions [3]. However, the fact that inclusions are not always present, especially in high-priced gemstones, renders this approach problematic. In the last decades, to optimize the inclusion-based process, major labs have started to exploit more advanced technologies, among them the determination of chemical composition. By determining major and minor elements

conclusions about e.g., sapphire genesis can be drawn, or the origin of copper bearing (Paraiba-type) tourmalines be determined [4]. The analytical arsenal used today includes techniques such as X-ray fluorescence (XRF), laser induced breakdown spectroscopy (LIBS), or laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS). Each of these technologies has its merits but also weaknesses: whereas XRF gives robust values for major elements, the other two techniques focus on trace elements. Having the lowest detection limits, typically in the ppb range, LA-ICP-MS further exhibits an excellent linearity of up to eight orders of magnitude. By contrast, higher concentrations very often present a problem for this technique. However, having access to both, major components as well as trace elements, is desirable for obtaining the overall chemical composition in a greater context.

In our study we have focused on the determination of geographic origin of demantoids, which is the well accepted trade name of the green variety of andradite ($\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$) garnets. Demantoids have first been found in the Russian Urals in the river Bobrovka near the village Elizavetinskoye, close to Niznij Tagil (or often just called Tagil) in the middle of the 19th century [5,6]. Further deposits in the Urals were later found in Poldnevskoye (Poldnevaya), Korkodin, and Ufalety—all within a range of 100 km of Ekaterinburg. Only in the 1990s a second major deposit was found in the Erongo region of Namibia [7], which together with Russia represents most of the cut gems available in today's market (Figure 1) [8]. Smaller deposits of gem quality material can also be found in Val Malenco in Italy [9], Antetazambato in Madagascar [10,11], Khuzdar, Baluchistan in Pakistan [12,13], as well as in Belqeys and Kerman, both located in Iran [14,15]. Certainly, there are many more deposits of demantoid garnets on the globe which are not mentioned here as to this date they do not play a major role in the gem market.



Figure 1. Demantoid gemstones (0.75–3.23 ct) from Russia (top row) and Namibia (bottom row) represent the majority of today's market.

Geologically, these deposits can be separated into two distinct groups, namely in (1) serpentinite hosted demantoids, such as the ones from the Urals, Val Malenco and Baluchistan and (2) those originating from skarns, such as the sites in Namibia and Madagascar. First successful attempts in determining the origin of demantoids based on chemical fingerprinting have shown promising results, although only simple characteristics such as the Mn/Ti ratio and Cr and/or V content have been used [8,16]. In this study we have carried out systematic quantitative trace-element analysis of demantoid garnets from six different origins using LA-ICP-MS and subsequent chemometric modeling for targeted classification without using inclusion-based assumptions.

2. Materials and Methods

Representative rough demantoid samples have been collected from Tubussis (Namibia), Ufaley (Russia), Karkodin (Russia), Poldnevskoye (Russia), Bobrovka (Russia), Val Malenco (Italy), Baluchistan (Pakistan), Kerman (Iran) and Antetazambato (Madagascar). Cut stones have either been cut by one of the authors (CS) or were taken from the collection of the Austrian Gemmological Society or the collection of Gerhard Brandstetter. Italian demantoids were in part supplied by Thomas Engeli, other samples were taken from the collections of the authors (CS and StS (Stephan Schwarzingger)). In particular for many samples originating from Russian deposits no further site of collection was provided. A total of 136 samples (403 individual points) have been analyzed, 39 from Russia (117 points), 25 from Namibia (73 points), 33 from Pakistan (98 points), 5 from Madagascar (14 points), 8 from Iran (23 points) and 26 from Italy (78 points).

Concentrations of ^{24}Mg , ^{27}Al , ^{47}Ti , ^{51}V , ^{52}Cr , and ^{55}Mn were measured by Laser ablation—inductively coupled plasma mass spectrometry (LA-ICP-MS) using a Cetac LSX-213 G2+ laser ablation system (Teledyne Cetac Technologies, Omaha, NE, USA) connected to a Thermo Fisher Scientific X-Series II ICP MS system (Thermo Fisher Scientific, Waltham, MA, USA). ^{43}Ca , ^{28}Si , and ^{56}Fe were also measured and utilized to monitor the smoothness of the Laser ablation process. Other elements (^7Li , ^{59}Co , ^{60}Ni , ^{63}Cu , ^{64}Zn , ^{71}Ga , ^{74}Ge , and ^{88}Sr) were measured as well for representative samples but did not show any significant concentrations. Tuning and calibration were carried out using NIST 610, 612, and 614 glass standards; reference values were taken from Jochum et al. [17]. The laser ablation was operated with a 100 μm spot size, a laser fluence of 9.2 J cm^{-2} and a repetition rate of 10 Hz. 130–230 shots were made on each spot and the ablated products transferred to the ICP-MS with a flow rate of 500 $\text{mL}\cdot\text{min}^{-1}$ He gas. Each sample was measured at three different spots except for two pieces from Madagascar that have been sampled 5 times each.

Concentration profiles were further analyzed with the publicly available software AI(OMICS)ⁿ (publicly available in-house written software), which provides a MatLab-based workflow for robust chemometric modelling [18]. Specifically, a table with sample IDs, meta data on the geographic origin of the sample as class information, and the respective element concentration profile for each sample were loaded in AI(OMICS)ⁿ, which was operated in MatLab (version R2018b, The Mathworks, Natick, MA, USA). Different strategies were applied, either considering all classes or considering just pair-wise comparison of classes. As a first step, native element concentrations were subjected to the data expansion module of AI(OMICS)ⁿ, which automatically generates, e.g., combinations of pair-wise ratios of parameters, logarithms of single parameters as well as parameter combinations for improved screening of non-linear events contributing variance between groups. To calculate logarithms in this context, concentrations below the limit of detection (LOD) were set to a fixed value of 0.1 ppm since negative and zero values cannot be logarithmized. The resulting variables were then subjected to a Kruskal-Wallis *H*-test in order to detect class-related variance and ranked according to their relevance [19]. A maximum of 50 of the most relevant variables got selected for further dimensionality reduction by principal component analysis (PCA). Subsequently, linear discriminant analysis (LDA) was applied to the extracted principal components (PCs). The number of PCs to be taken into account for LDA was determined by a cut-off at a threshold of 95% explained variance. This is a precaution in order to avoid overfitting and thus the consideration of random error (noise) rather than relationships between variables. Models were validated using the Monte-Carlo cross-validation (MCCV) method [20]. Specifically, 100 random models were generated with 90% of the data set as training data and the remaining 10% as test set. The overall quality of the models, i.e., the correctness of data predicted during MCCV, was judged using the multivariate correlation coefficient R^2 , the multivariate correlation coefficient after cross validation Q^2 , the accuracy (ACC), the Matthews correlation coefficient (MCC), the prevalence threshold as well as the confusion matrix. In particular the MCC has been found to be a very reliable metric for judging the quality of chemometric models. In contrast to the ACC, which constitutes the ratio of the sum of true positive and negative classifications and the total population, MCC also takes into account the false positive and false negative classifications. It can also be applied

well in cases where group sizes differ significantly. The prevalence threshold is applied to test whether a particular class is represented with enough samples under consideration of the accuracy and precision of the respective model. To assess model robustness averaged class accuracies and their standard deviations were calculated based on the 100-fold MCCV. The importance of each element for the respective model was judged by calculating the variable importance in projection (VIP). Based on this information box plots were generated for these elements in R version 3.5.4 using the 'ggplot2' package [21–27].

3. Results and Discussion

Garnets can be divided into 6 groups, which are pyrope, almandine, spessartine (pyralspite-series) and uvarovite, grossular, and andradite (ugrandite-series). Very often garnets are not pure endmembers of a series but represent mixtures of the members of the above-mentioned series and therefore do not exhibit well defined, narrow ranges of element-concentrations. We have therefore decided not to use any internal standard for the ICP-MS experiments, rather, we have used three different NIST glasses which cover concentrations of about 1–400 ppm for most elements in question. A comparison of our data with the ones published by Adamo et al. [9] for Italian demantoids using an internal standard approach shows very good agreement given the fact that the chemistry within a single crystal can vary dramatically as has been shown for an andradite slice from Ufalay [8]. The actual values are Cr: 2–5468 vs. 0–1134 ppm, Ti: 7–649 vs. 3–545 ppm, and V: 12–86 vs. 1–102 ppm for 31 points analyzed by Adamo and 78 points in this study.

As a first result this study reveals that the samples analyzed here all represent rather pure end members of the andradite series except for the samples originating from Madagascar. This class of samples exhibits from low to comparatively high aluminum concentrations of 40 ppm up to 8.5% (most being above 0.2%) indicating presence of a certain grossular content for this site. Hence, within the group of locations analyzed here aluminum serves as a marker element for Madagascar. Another element of interest is chromium, where highest concentrations can be found in samples from Russia and Iran, as low as 100 ppm but mostly in the range of 0.1 to 4%. Namibia has generally the lowest amounts of chromium; in Pakistani, Madagasy and Italian samples concentrations from below detection limit up to 1000 ppm were found. Manganese is high in Namibian and Russian samples (280–500 ppm) whereas magnesium is found to have highest concentrations in Pakistani and in part Russian material (2000 ppm—1.1%) and lowest concentrations in Namibian samples (generally below 500 ppm). A table listing all concentrations can be found as Table S1 in the Supplementary Materials.

A first approach was using the separation criteria which have been successfully applied to differentiate Namibian from Russian demantoids in our previous work [8]. There, the ratio of manganese/titanium was plotted against the sum of chromium and vanadium, however, this two-dimensional approach soon starts to produce significant overlapping of several origins and was thus discarded.

In this work, a total of 17 elements was analyzed to discover marker elements or reveal internal dependencies of element concentrations with respect to the origin of the samples. For successful multivariate modeling it is important to identify variables significantly contributing the inter-group variance and to eliminate other variables to reduce the risk of overfitting the data, which here is tested for by applying extensive Monte-Carlo cross-validation. In addition, it has to be considered that linear multivariate model approaches may have difficulties with non-linear data structures. Therefore, linearization of data, e.g., by logarithmization, may significantly improve performance of linear modeling algorithms. It should, however, be noted that the degree to which this is relevant will always depend on the particular data structure of the data set investigated. In this study the optimal number of variables as well as the optimal pre-processing of these variables was determined in an iterative manner. Specifically, elements that were not detected at concentrations above LOD were eliminated first. This way, the elements Li, Co, Ni, Cu, Zn, Ga, Ge, and Sr (eight out of 17) were not present in significant concentrations in a manner systematically correlating with the geographic origin and

were subsequently not considered further. Additionally, Ca, Si and Fe—which were primarily used for monitoring of the laser ablation process—were excluded from further analysis as no standards are available in the relevant concentration regime preventing that the system could not be calibrated accordingly and the amounts detected were found to be too high. Following this step the data expansion module of AI(OMICS)ⁿ, which aims at automation of certain steps of data pre-processing, was applied for the remaining six elements Al, Cr, Mg, Mn, Ti, and V. Specifically, inverse concentrations, products and ratios of all element concentrations measured as well as the calculated dependencies are generated to allow a systematic screening. Especially ratios, e.g., plotted in 2D graphs, have been shown in the past to contain important information as compared to using pure concentrations [8]. Similarly, as concentrations of the various elements span several orders of magnitude and are heavily skewed, transformation into their logarithm—as pointed out above—may improve chemometric modeling based on linear approaches, including PCA, PCA-LDA, partial least squares-discriminant analysis (PLS-DA), orthogonal PLS-DA (OPLS-DA) and others. Here, the natural logarithm (base-*e* log) has been chosen for model building with the PCA-LDA approach [28,29]. For this specific operation it was necessary to convert measurements numbers smaller or equal to zero to a small value but not zero. Considering the concentrations detected for the individual elements a fixed value of 0.1 was chosen empirically, which—being in the order of the value of the smallest concentrations of all elements measured—prevents unnaturally small logarithms. Several rounds of data elimination were performed considering the ranks of variables especially in the Kruskal-Wallis *H*-test as well as in the VIP-scores in PCA-LDA models. In addition, model quality was judged by the R^2 and Q^2 values as a function of the number of principal components to exclude overfitting. The Matthews correlation coefficient (MCC), which in contrast to R^2 and Q^2 also takes into consideration the effect of true and false positive/negative predictions as well, is also considered as a relative quality criterion between different models. The prevalence threshold, which takes into account the quality of discriminating power under consideration of the size of the groups, was judged, too. All models were generated as PCA-LDA for optimal comparison. Finally, it was found that logarithmizing of the pure element concentration as single data transformation was sufficient to achieve over 95% explained variance with no more than 3 to 5 principal components for pair-wise comparisons of locations. As logarithmizing of variables is a common way of converting non-linear problems into linear ones, it is not unexpected that LDA performs better with the logarithms of the element concentrations. Variables, such as ratios of concentrations or their logarithms were excluded to reduce multi-collinearity, which may significantly impact modeling by overfitting. Of note, PCA-LDA itself was chosen because it is known to be less prone to overfitting due to multi-collinear variables. Additionally, in contrast to, e.g., a PLS-DA (where dimensionality reduction is supervised while it is un-supervised in the PCA approach), it allows better assessment of samples that do not belong to groups used to generate models through judgment of the Mahalanobis distance [30] of a sample relative to the group center in multidimensional space [31]. Note, too, that LDA in principle requires normality of the data, which is often not given in real-life data sets, in particular when only small sample numbers are available. Using logarithms of element concentrations is an important step towards normality, which, however, is not met for every element/country combination. Nevertheless, the LDA classification has been shown in the past to perform very robust, despite the assumption of normality is violated [32,33]. If the data were perfectly normally distributed LDA would determine the optimal solution for each hyperplane. In absence of normality several solutions for a hyperplane may exist, each describing only one local optimum. This in turn would lead to an increased number of mismatches in the cross-validation, which was not observed here. We wish to note that there are many other classification algorithms which in principle may perform similarly provided proper verification. In this study, however, the focus is not to find the best performing statistical model but to gain insights into whether trace element concentration allows classification of the origin at all and which elements play a role in this approach.

Table 1 shows the confusion matrix for a representative model “All origins”. Shown are the model specific benchmark parameters accuracy (ACC), Matthews correlation coefficient (MCC), false

classifications, correct classifications and the prevalence threshold (minimum calculated fraction/true fraction). Numbers in parenthesis are percent fractions for countries and correct and false classifications, respectively. $n(\text{total})$ denotes the number of total element profiles, $n(\text{class})$ the number of element profiles, and $n(\text{latent variables})$ the number of principal components used to achieve 95% explained variance. In addition, cumulative R^2 and Q^2 as well as the averaged group accuracy (mean $\pm$ standard deviation) from the 100-fold MCCV are shown to allow judgement of the model robustness. Diagonal elements in the confusion matrix represent the number (fraction) of correctly assigned cases. Vertical off-diagonal elements correspond to false negative predictions (i.e., wrong origin assigned; e.g., from the 73 true Namibian Demantoids two are mis-assigned to Russia and three to Italy (first row)), while horizontal off-diagonal elements are false negative results (i.e., wrong origin is assumed; e.g., five samples predicted to be from Namibia are indeed from Russia (first line)). The asterisk indicates that the prevalence threshold expected for the accuracy and precision of the model is slightly missed for the origin Iran.

Table 1. Confusion matrix showing prediction accuracy of the method (for details refer to the text).

Model	ACC = 0.84	MCC = 0.78	$R^2_{\text{cum}} = 1$	$Q^2_{\text{cum}} = 0.85$			
All Origins							
$n(\text{total}) = 403$	True Namibia	True Pakistan	True Madagascar	True Russia	True Iran	True Italy	prevalence threshold (calc/observed)
Predicted Namibia	68 (93.2%)	0	0	5 (4.3%)	0	0	0.12/0.18
Predicted Pakistan	0	91 (92.9%)	1 (7.1%)	7 (6.0%)	0	7 (9.0%)	0.12/0.18
Predicted Madagascar	0	0	13 (92.9%)	0	0	0	0.03/0.03
Predicted Russia	2 (2.7%)	1 (1.0%)	0	96 (82.1%)	3 (13.0%)	18 (23.1%)	0.24/0.29
Predicted Iran	0	1 (1.0%)	0	3 (2.6%)	17 (73.9%)	0	0.10/0.06 *
Predicted Italy	3 (4.1%)	5 (5.1%)	0	6 (5.1%)	3 (13.0%)	53 (67.9%)	0.22/0.19
$n(\text{class})$	73	98	14	117	23	78	
false classifications ($n(\%)$)	65 (16.1%)						
correct classifications ($n(\%)$)	338 (83.9 %)						
$n(\text{latent variables})$	6						
Averaged Group Accuracy from MC cross validation	0.92 $\pm$ 0.10	0.91 $\pm$ 0.10	0.94 $\pm$ 0.15	0.81 $\pm$ 0.12	0.72 $\pm$ 0.31	0.68 $\pm$ 0.18	

*: Prevalence threshold missed for location Iran—see text for details. Model performance (correctly assigned) indicated by background color: >90%: green; >80 %: blue; >70 % yellow; rest: red.

The model with all six origins shows that at least three origins can be predicted with a considerable high accuracy and precision (Figure 2, Table 1). In particular, demantoids from Namibia and Pakistan can be classified very well. Specimen from Madagascar can also be classified well, but the low prevalence leads to a higher degree of variability as evidenced by the larger standard deviation of the group accuracy resulting from MCCV. Although samples from Russian deposits can be predicted with accuracies around 80% some sample are mis-assigned to all other origins, except for Madagascar. This indicates a larger variability in the composition of these demantoids. Samples from Iran and Italy could also be assigned only with accuracies around 70%. However, in contrast to the wrongly classified samples from Russia the samples from Iran and Italy are mis-assigned to only two other origins. In particular demantoids from Italy exhibit a tendency to be predicted in a false-positive manner as Russian samples. This leads to the conclusion that a significant fraction of Italian demantoid samples shares a similar trace-element profile with the majority of the Russian samples. It should be mentioned that—although the prevalence index for five of the origins is met or even exceeded and only closely missed for the origin Iran—the smaller number of samples available for Iran and Madagascar adds to uncertainty. The question here is whether natural variability of these very difficult to come by samples is already adequately represented. Especially the narrow distribution of the profiles for Madagascar contributes to excellent classification properties for this particular deposit, which appears to source from an increased grossular content in the samples studied here. The small total number of data points also contributes to comparable large standard deviation of the class accuracy achieved during the 100-fold Monte Carlo cross-validation.

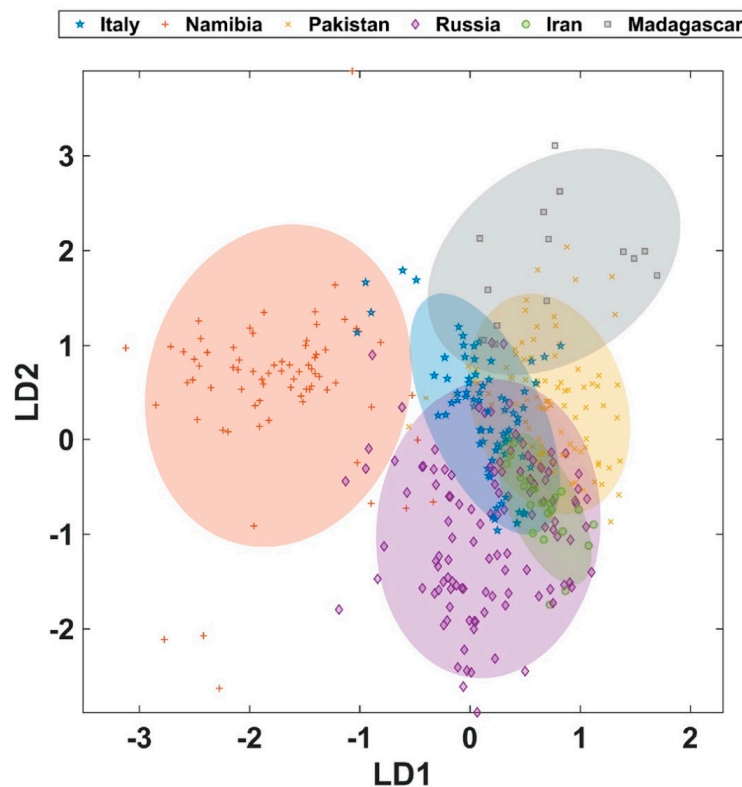


Figure 2. PCA-LDA score plot with all classes including the log-transformed concentrations of the six relevant elements Mg, Al, Ti, V, Cr and Mn. Each dot represents a single element profile from an individual measurement after MCCV. Axis shown represent the first two linear discriminants. Ellipses confer to the 95% confidence interval according to Hotellings T^2 , which is the multivariate analog to the t -test [34]. Note that graphics represent a projection from 6-dimensional space, i.e., overlap in this projection does not necessarily mean samples cannot be separated (cf. confusion matrix in Table 1).

As already indicated above the number of samples has a critical influence on the model quality and should always be tested for using the prevalence threshold. In the above model the prevalence threshold for Iran was slightly missed as indicted by an asterisk in Table 2. This indicates that more samples will be needed to improve future models. However, at the same time this case represents an interesting test for what happens when locations with only small numbers of samples are included in modeling. While ideally all demantoid deposits in the world with all available specimens should be in a model this is not possible in reality, simply for sample availability, but also financial reasons associated with sample acquisition and analysis. As demantoids from Iran are present in the market we decided to include the available specimens into this model rather than to omit it and investigate the effects. While the model with six deposits does not allow satisfactory assignment of all samples in a single test the results from cross validation allow deepened insights into sample confusion. In fact, testing all classes in one model is only one of several possible testing strategies. A different strategy is to test one deposit in a pair-wise manner against all other deposits (i.e., testing location X vs. (combined others)). However, in this case the group consisting of all but one location displays a very high variability preventing improved results for the particular classification problem studied here (data not shown). It should be noted that it is difficult to delineate a generally applicable approach, as this will be determined by the properties of the population investigated. For the present case another possibility is inspired by the directed acyclic graph (DAG) trees [35]. Here, models are constructed for only two deposits at a time, but for all possible combination of deposits. By doing so the model can take optimal advantage of the differences between a particular pair of groups, which may lead to improved classification performance.

Table 2. Confusion matrix for a representative model for pair-wise testing of the country of origin (15 individual models).

Model	Italy vs Iran			True Italy	Italy vs Russia			True Italy	Italy vs Pakistan			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy	Italy vs Nambia			True Italy
-------	---------------	--	--	------------	-----------------	--	--	------------	-------------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------	-----------------	--	--	------------

Hence, in order to better understand these pair-wise relationships we further established 15 models comparing only two selected origins (Figure 3). This is particularly interesting as now for each case a different set of variables can be used to describe the inter-group variance in an optimized way. The approach also has the potential that separations between two particular deposits can be much clearer than in the multi-origin model. However, caution is advised as an unknown sample subjected to classification with only two classes will be assigned to one of the two available choices—regardless if it is indeed a member of one of the two. For instance, if a garnet from Madagascar is investigated for class membership in a pair-wise model trained with samples from Italy and Pakistan, it will be assigned to one of these classes.

From a purely mathematical point of view it is non-sense to use something in a model for which it is not trained for. However, in real-life this is a realistic scenario when testing an unknown sample, especially if there is the suspicion of fraud. The unknown sample will be assigned to the class with the highest similarity. In case the sample does not belong to either of the groups tested for the Mahalanobis distance, which is a very useful metric for detecting multivariate outliers, for such a sample will likely be very large in comparison to samples that indeed belong to classes used for the training of the data set [36]. This is possible, for instance, through the QQ-plots (quantile-quantile) provided by the AI(OMICS)ⁿ package, where Mahalanobis distance of all samples used for model building and the sample under investigation is plotted versus their respective Euclidean distance at a 95% confidence level ($\alpha = 5\%$). In general, if an unknown sample is initially run in the multi-origin data set, one can obtain a first estimate of the samples' country of origin. In a subsequent step this can be verified with the finer pair-wise models. In any case it is necessary to closely investigate the Mahalanobis distance for the sample in all models and to compare it with the typical confidence intervals obtained for true in-class samples.

Figure 3 clearly shows that even with only two depicted linear discriminators most origins can be clearly distinguished. In particular, Italy-Namibia, Italy-Iran, Italy-Madagascar, Namibia-Pakistan, Namibia-Iran, Namibia-Madagascar, Pakistan-Iran, and Iran-Madagascar can be separated essentially in full. For the remaining classification problems, it must not be forgotten that the graphs only represent 2D-projections of a multi-dimensional space and that the confusion matrix as well as additional metrics should be considered. Table 2 reveals that models for Madagascar typically exhibit a reasonable high accuracy which, however, varies considerably as a function of cross-validation. This can be explained to be a more frequent mis-assignment of only a few samples of the group, which due to the small group size accumulates to a large relative variation. Furthermore, Table 2 shows that the distinction of Italy vs. Russia and Russia vs. Iran works least as expressed by their low MCC of only 0.68 and 0.67, respectively. In both cases one group (Italy and Iran, respectively) shows a large number of mis-assigned samples. The first case is particularly interesting, as both groups are also present with comparable and larger sample numbers (Russia: 117, Italy: 78). While in this model Russian samples are mis-classified at a rate of only ~7% (8 out of 117), Italian samples are classified as being Russian demantoids at a rate of ~26% (20 out of 78). This means that the accuracy for both origins is different in this model, which has an impact on the interpretation of the results. Here, it is important to consider the testing hypothesis. This means that if the null hypothesis, H_0 , is Russia (we assume the sample is from Russia) there is a 93% chance the result is correct, and the gem is indeed from Russia. More than 9 out of 10 tests will deliver a correct result. However, in the case H_0 is Italy (assuming the gem is from Italy) chances for the specimen being indeed an Italian demantoid are as low as 75% meaning one out of four tests will be mis-assigned. In such a case it is particularly important to critically evaluate all other models by attempting to rule out other solutions. Ultimately, there will always be a certain risk that a new sample is not well represented in the model. In such instances it is indicated to utilize additional approaches, such as analysis of inclusions (if present) to gain more confidence in the results obtained.

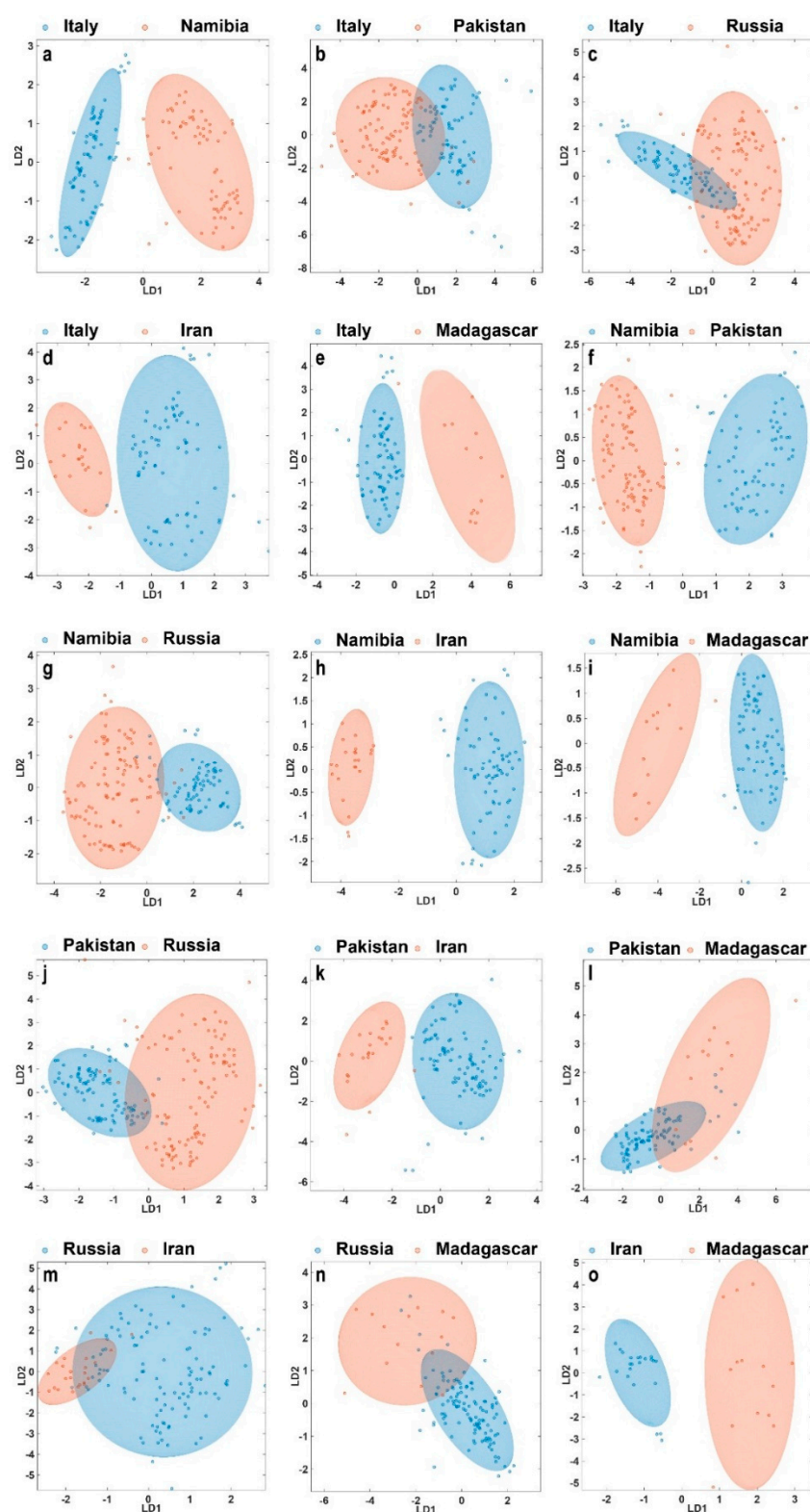


Figure 3. Representative PCA-LDAs for the pair-wise models (a) Italy-Namibia; (b) Italy-Pakistan; (c) Italy-Russia; (d) Italy-Iran; (e) Italy-Madagascar; (f) Namibia-Pakistan; (g) Namibia-Russia, (h) Namibia-Iran; (i) Namibia-Madagascar, (j) Pakistan-Russia, (k) Pakistan-Iran, (l) Pakistan-Madagascar, (m) Russia-Iran, (n) Russia-Madagascar, (o) Iran-Madagascar. Each dot represents a single element profile from an individual measurement after MC cross-validation. Axis shown represent the first two linear discriminants. Ellipses confer to the 95% confidence interval according to Hotellings T^2 .

The confusion matrix for a representative model for pair-wise testing of the country of origin is shown in Table 2. Displayed are the model specific benchmark parameters accuracy (ACC), Matthews correlation coefficient (MCC), false classifications, correct classifications and the prevalence threshold (minimum calculated fraction/true fraction). Numbers in parenthesis are percent fractions for countries and correct and false classifications, respectively. $n(\text{total})$ denotes the number of total element profiles, $n(\text{class})$ the number of element profiles, and $n(\text{latent variables})$ the number of principal components used to achieve 95% explained variance. In addition, cumulative R^2 and Q^2 as well as the averaged group accuracy (mean $\pm$ standard deviation) from the 100-fold MCCV are shown to allow judgement of the model robustness. Diagonal elements in the confusion matrix represent the number (fraction) of correctly assigned cases. Models highlighted in red and blue denote perfectly (exclusively correct classifications) respectively almost perfectly performing models.

A closer inspection of the VIP scores, i.e., the natural logarithms of elements, of the individual elements provides insights into their importance for a particular model (Tables 3 and 4). Table 3 provides the detailed results for all models generated in this study, with VIP1 having the highest impact on a model and VIP6 having the lowest. Overall, it can be seen that Al and Cr appear to play a major role in the discriminant models, followed by Mg and Mn. However, it would be a wrong conclusion to assume that V and Ti could be omitted from the analysis. Despite these two elements playing a less important role in many of the generated models, they are essential components for the discrimination of Russia-Madagascar, Pakistan-Iran, Italy-Iran, and Iran-Madagascar. Moreover, valuable information about the importance of the elements for the within-class structure of the data is gained as the VIP method takes into account all dimensions of the generated model. Hence, for comprehensive analysis of the origin of a gem it is essential to record all six elements. In fact, it is the large range of concentrations found in natural gems which defines the origin of a specimen, when enough elements are combined. This can be visualized best by boxplots of the natural logarithms of the individual elements as a function of the country of origin. In the best cases, an origin can be selected (or excluded) by simply plotting the concentration profile into the boxplot chart (Figure 4).

Table 3. Listing of the most important variables for each model ranked by their VIP-score.

Model	Type of Deposit	VIP1	VIP2	VIP3	VIP4	VIP5	VIP6
All origins		ln(Mg)	ln(Cr)	ln(Ti)	ln(Al)	ln(Mn)	ln(V)
Italy-Namibia	asbestos vs. skarn	ln(Al)	ln(Cr)	ln(Mn)	ln(Ti)	ln(Mg)	ln(V)
Italy-Pakistan	asbestos vs. asbestos	ln(Mn)	ln(Mg)	ln(Al)	ln(Cr)	ln(Ti)	ln(V)
Italy-Russia	asbestos vs. asbestos	ln(Mg)	ln(Al)	ln(V)	ln(Ti)	ln(Cr)	ln(Mn)
Italy-Iran	asbestos vs. asbestos	ln(Al)	ln(Ti)	ln(Cr)	ln(Mn)	ln(Mg)	ln(V)
Italy-Madagascar	asbestos vs. skarn	ln(Cr)	ln(Al)	ln(Mg)	ln(V)	ln(Ti)	ln(Mn)
Namibia-Pakistan	asbestos vs. skarn	ln(Al)	ln(Cr)	ln(Mg)	ln(Mn)	ln(V)	ln(Ti)
Namibia-Russia	asbestos vs. skarn	ln(Al)	ln(Mn)	ln(V)	ln(Ti)	ln(Cr)	ln(Mg)
Namibia-Iran	asbestos vs. skarn	ln(Mn)	ln(Mg)	ln(Cr)	ln(Ti)	ln(Al)	ln(V)
Namibia-Madagascar	skarn vs. skarn	ln(Cr)	ln(Mg)	ln(Mn)	ln(Al)	ln(Ti)	ln(V)
Pakistan-Russia	asbestos vs. asbestos	ln(Mg)	ln(Al)	ln(Cr)	ln(Ti)	ln(V)	ln(Mn)
Pakistan-Iran	asbestos vs. asbestos	ln(V)	ln(Cr)	ln(Al)	ln(Ti)	ln(Mg)	ln(Mn)
Pakistan-Madagascar	asbestos vs. skarn	ln(Cr)	ln(Mn)	ln(Mg)	ln(Al)	ln(Ti)	ln(V)
Russia-Iran	asbestos vs. asbestos	ln(Cr)	ln(Al)	ln(Mg)	ln(Mn)	ln(Ti)	ln(V)
Russia-Madagascar	asbestos vs. skarn	ln(V)	ln(Cr)	ln(Al)	ln(Mg)	ln(Ti)	ln(Mn)
Iran-Madagascar	asbestos vs. skarn	ln(Cr)	ln(V)	ln(Al)	ln(Ti)	ln(Mg)	ln(Mn)

Table 4. Scores of the individual elements as a function of their VIP-rank. Although it appears that Ti and V are overall of less importance Table 3 clearly indicates their important role for particular models.

Rank VIP	Al	Cr	Mg	Mn	Ti	V
1	4	5	3	2	0	2
2	4	5	3	2	1	1
3	4	3	4	2	1	2
4	3	1	1	3	7	1
5	1	2	4	1	6	2
6	0	0	1	6	1	8

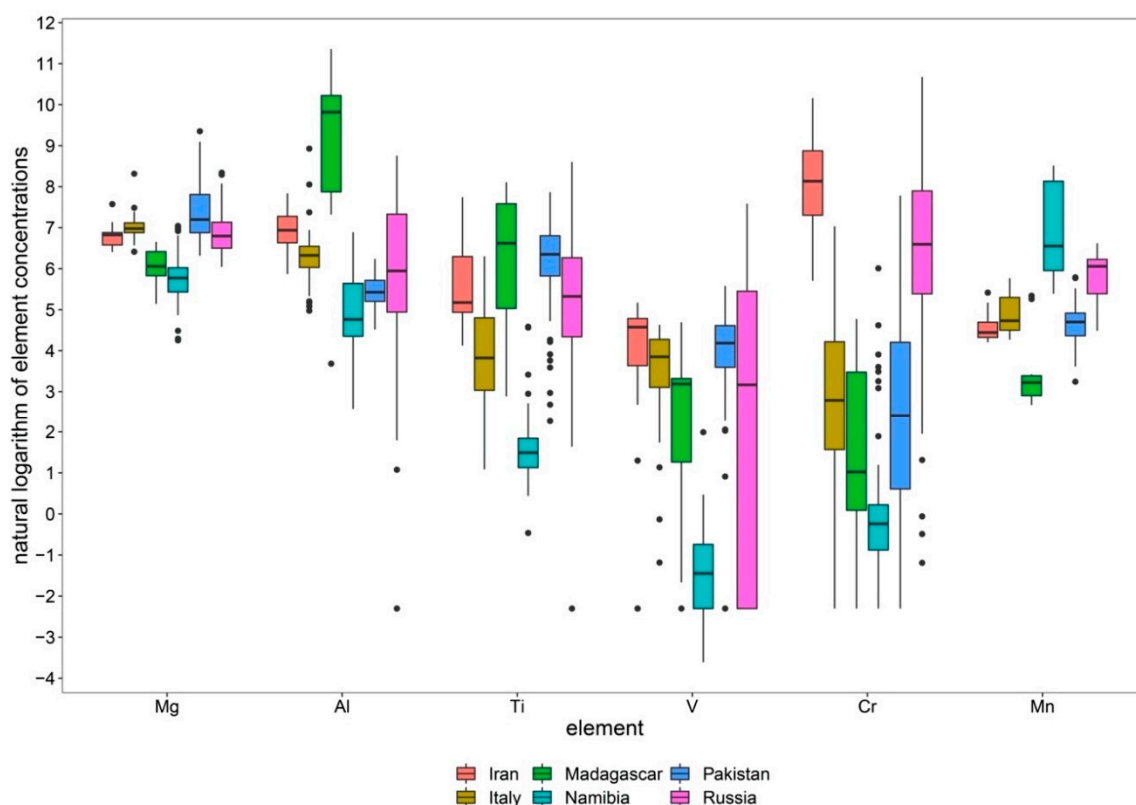


Figure 4. Boxplot showing the natural logarithm of all measured element concentrations. Samples are grouped and colored by origin. The graph displays the median, the first quartile, the third quartile as well as the maximum and minimum of each group. Outliers are represented as black dots.

4. Conclusions

We have successfully shown that the determination of country of origin is possible using a more quantitative approach than identification of inclusions, which is largely based on individual knowledge. By measuring minor and trace elements and applying modern chemometrics an excellent separation of five out of six countries of origin was achieved. The good separation with the complete dataset can even be improved when only a pair-wise comparison of two possible origins is used. This approach can be used when additional information, such as inclusions, can be used to rule out certain origins. Of course, the number of samples is essential for such studies and new samples will be added to the model as they become available. For the future it is planned to include even more elements into multivariate modelling. While it appears not rewarding to conduct large studies on elements that were shown to be not present in significant amounts in this and/or other studies, it seems promising to utilize major elements, such as Ca, Si, and Fe. As discussed for LA-ICP-MS proper calibration is difficult due to lacking standards and instrument limitations. This often prevents comparability

of results obtained on different instruments. However, other technologies might be utilized in the future to obtain reproducible concentrations for these elements. Then, the data resulting from different technologies can be fused and evaluated together. This approach has already been successfully applied to technologies as different as NMR-spectroscopy and multi-isotope ratio mass-spectrometry [37].

Still, any such approach is only as good as the samples that are available. In our study the number of samples from Madagascar and Iran was rather limited with 5 and 8, respectively, so in order to improve the models more samples have to be obtained from those localities. New localities may be added as samples become available in significant numbers. Care has to be taken that selected specimens properly reflect the natural variability at these locations. Nevertheless, when applying such methods, the user has to be aware that likely not any sample population available will (a) ever represent all deposits in the world and (b) in significant numbers. Hence, interpretation shall always be carried out having this in mind. When testing samples for fraud, analytical results often provide hints that something is not as advertised. It is advisable to back up results with independent methods to get more evidence. Here, we have added one more tool to the analytical arsenal.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2075-163X/10/12/1046/s1>, Table S1: Concentrations of Mg, Al, Ti, V, Cr, and Mn of all samples.

Author Contributions: Conceptualization, C.S. and S.S.; methodology, C.S. and S.S.; software, F.R. and S.S.; validation, S.G.B., F.R., S.S. and C.S.; formal analysis, S.G.B., F.R., S.S. and C.S.; investigation, S.G.B., F.R., S.S. and C.S.; resources, C.S. and S.S.; writing—original draft preparation, C.S., S.S. and S.G.B.; writing—review and editing, C.S., S.S., S.G.B. and F.R.; visualization, S.G.B.; project administration, C.S. and S.S.; funding acquisition, C.S. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to Leopold Rößler (Austrian Gemmological Society), Gerhard Brandstetter, Ildar Latypov and Thomas Engeli for providing samples. Open Access Funding by the University of Linz.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Sapphire Market Update 2020 Lotus Gemology. Available online: <https://lotusgemology.com/index.php/library/articles/459-world-sapphire-market-update-2020-lotus-gemology> (accessed on 28 September 2020).
2. Atkinson, D.; Kothavala, R.Z. Kashmir Sapphire. *Gems. Gemol.* **1983**, *19*, 64–76. [CrossRef]
3. Palke, A.C.; Saeseaw, S.; Renfro, N.D.; Sun, Z.; McClure, S.F. Geographic Origin Determination of Blue Sapphire. *Gems. Gemol.* **2019**, *55*, 536–579. [CrossRef]
4. Peretti, A.; Bieri, W.P.; Reusser, E.; Hametner, K.; Günther, D. Chemical Variations in Multicolored “Paraiba”-type Tourmalines from Brazil and Mozambique: Implications for Origin and Authenticity Determination. *Contrib. Gemol.* **2009**, *9*, 1–77.
5. Kolesar, P.; Tvrdý, J. *Zarenschätze*; Rainer Bode: Haltern, Germany, 2006; p. 204.
6. Burlakov, A.; Burlakov, E. Demantoid from Russia. *InColor* **2017**, *36*, 40–46.
7. Reif, S. Green Dragon Mine Demantoid from Namibia. *InColor* **2017**, *36*, 48–52.
8. Schwarzwinger, C. Determination of demantoid garnet origin by chemical fingerprinting. *Monats. Chem.* **2019**, *50*, 907–912. [CrossRef]
9. Adamo, I.; Bocchio, R.; Diella, V.; Pavese, A.; Vignola, P.; Prosperi, L.; Palanza, V. Demantoid from Val Malenco Italy: Review and Update. *Gems. Gemol.* **2009**, *45*, 280–287. [CrossRef]
10. Pezzotta, F. Andradite from antetazambato, North Madagascar. *Miner. Rec.* **2010**, *41*, 209–229.
11. Pezzotta, F.; Adamo, I.; Diella, V. Demantoid and Topazolite from Antetazambato, Northern Madagascar: Review and New Data. *Gems. Gemol.* **2011**, *47*, 2–14. [CrossRef]

12. Palke, A.C.; Pardieu, V. Demantoid from Baluchistan Province in Pakistan. *Gems. Gemol.* **2014**, *50*, 302–303.
13. Adamo, I.; Bocchio, R.; Diella, V.; Caucia, F.; Schmetzer, K. Demantoid from Balochistan, Pakistan: Gemmological and Mineralogical Characterization. *J. Gemmol.* **2015**, *34*, 428–433. [[CrossRef](#)]
14. Hairapetian, V.; Pelckmans, H. Emerging Mineral Finds from Northern Iran. *Rocks Miner.* **2017**, *6*, 540–549. [[CrossRef](#)]
15. Du Toit, G.; Mayerson, W.; Van Der Bogert, C.; Douman, M.; Befi, R.; Koivula, J.I.; Kiefert, L. Demantoid from Iran. *Gems. Gemol.* **2006**, *42*, 131.
16. Schwarzsinger, C.; Engeli, T. LA-ICP-MS spectrometry and its use in gemology. *Ital. Gemol. Rev.* **2019**, *8*, 50–52.
17. Jochum, K.P.; Weis, U.; Stoll, B.; Kuzmin, D.; Yang, Q.; Raczek, I.; Jacob, D.E.; Stracke, A.; Birbaum, K.; Frick, D.A.; et al. Determination of reference values for NIST SRM 610–617 glasses following ISO guidelines. *Geostand. Geoanal. Res.* **2011**, *35*, 397–429. [[CrossRef](#)]
18. Rüll, F.; Schwarzsinger, S. AI(OMICS)n: A MatLab Based Expert System for Chemometrics and Data Fusion. 2020. Available online: <https://gitlab.com/ai-omics/ai-omics/> (accessed on 20 October 2020).
19. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [[CrossRef](#)]
20. Xu, Q.; Liang, Y.Z.; Du, Y.-P. Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *J. Chemom.* **2004**, *18*, 112–120. [[CrossRef](#)]
21. Henningsson, M.; Sundbom, E.; Armelius, B.Å.; Erdberg, P. PLS model building: A multivariate approach to personality test data. *Scand. J. Psychol.* **2001**, *42*, 399–409. [[CrossRef](#)]
22. Johansson, E.; Eriksson, L.; Sandberg, M.; Wold, S. QSAR Model Validation. In *Molecular Modeling and Prediction of Bioactivity*; Springer: Boston, MA, USA, 2000. [[CrossRef](#)]
23. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
24. Matthews, B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* **1975**, *405*, 442–451. [[CrossRef](#)]
25. Baldi, P.; Brunak, S.; Chauvin, Y.; Andersen, C.A.; Nielsen, H. Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics* **2000**, *16*, 412–424. [[CrossRef](#)] [[PubMed](#)]
26. Balayla, J. Prevalence threshold (ϕ_e) and the geometry of screening curves. *PLoS ONE* **2020**, *15*, e0240215. [[CrossRef](#)] [[PubMed](#)]
27. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020; Available online: <https://www.R-project.org/> (accessed on 20 October 2020).
28. Huelin, S.R.; Longerich, H.P.; Wilton, D.H.; Fryer, B.J. The determination of trace elements in Fe–Mn oxide coatings on pebbles using LA-ICP-MS. *J. Geochem. Expl.* **2006**, *91*, 110–124. [[CrossRef](#)]
29. van den Berg, R.A.; Hoefsloot, H.C.; Westerhuis, J.A.; Smilde, A.K.; van der Werf, M.J. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genom.* **2006**, *7*, 142. [[CrossRef](#)] [[PubMed](#)]
30. Kolb, P.; Bindereif, S.G.; Rüll, F.; Schwarzsinger, S. unpublished data, to be published elsewhere.
31. Mahalanobis, P.C. On the Generalized Distance in Statistics. *Proc. Natl. Inst. Sci. India* **1936**, *2*, 49–55.
32. Moawed, S.A.; Osman, M.M. The robustness of binary logistic regression and linear discriminant analysis for the classification and differentiation between dairy cows and buffaloes. *Int. J. Stat. Appl.* **2017**, *7*, 304. [[CrossRef](#)]
33. Barragués, J.I.; Morais, A.; Guisasola, J. (Eds.) *Probability and Statistics: A didactic Introduction*; CRC Press: Boca Raton, FL, USA, 2014.
34. Hotelling, H. The generalization of Student's ratio. *Ann. Math. Statist.* **1931**, *2*, 360–378. [[CrossRef](#)]
35. Platt, J.C.; Cristianini, N.; Shawe-Taylor, J. Large Margin DAGs for Multiclass Classification. *Adv. Neural Inf. Process. Syst.* **2000**, *12*, 547.
36. Ghorbani, H. Mahalanobis distance and its application for detecting multivariate outliers. *Facta Univ. (NIS) Ser. Math. Inform.* **2019**, *34*, 583. [[CrossRef](#)]

37. Bindereif, S.G.; Brauer, F.; Schubert, J.-M.; Schwarzing, S.; Gebauer, G. Complementary use of ^1H NMR and multi-element IRMS in association with chemometrics enables effective origin analysis of cocoa beans (*Theobroma cacao* L.). *Food Chem.* **2019**, *299*, 125105. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Chemometric Modeling of Trace Element Data for Origin Determination of Demantoid Garnets

Stefan G. Bindereif, Felix Rüll, Stephan Schwarzingger and Clemens Schwarzingger

Minerals 2020, 10(12), 1046; <https://doi.org/10.3390/min10121046>

Table S1. Concentrations of Mg, Al, Ti, V, Cr, and Mn of all samples.

Sample No	Origin	Mg / ppm	Al / ppm	Ti / ppm	V / ppm	Cr / ppm	Mn / ppm
1	Italy	1780	3143	454.3	91	90.2	319.9
2	Italy	1496	1596	389.4	87.35	59.4	282.2
3	Italy	1557	7538	345.5	82.48	37.12	270.9
4	Italy	945.9	411.4	296.1	71.97	22.63	226.4
5	Italy	926.5	485.7	344.4	66.33	10.24	73.06
6	Italy	1066	583.1	379.8	61.11	5.976	76.99
7	Italy	1173	655.4	412.5	65.72	6.515	83.72
8	Italy	939.5	509.9	126.1	84.43	11.12	76.56
9	Italy	822	434.9	110.9	80.48	32.41	82.14
10	Italy	892.3	513.2	110.7	82.62	21.34	71.51
11	Italy	977.3	360.3	475	81.68	19.75	201.9
12	Italy	977.4	356.6	502.9	88.25	17.63	212.6
13	Italy	1060	494.2	544.8	96.61	13.07	194.2
14	Italy	1081	635.6	49.92	46.81	813.1	133.8
15	Italy	996.3	491.8	369.9	71.92	19.19	223.5
16	Italy	1142	678.8	47.07	30.64	1106	133.4
17	Italy	1119	660.6	39.82	37.35	1134	130.5
18	Italy	886.8	460.9	15.69	26.12	11.76	82.98
19	Italy	871.7	430.7	34.12	58.74	42.53	111.5
20	Italy	893.9	483.3	17	30.5	15.01	97.46
21	Italy	712.6	236	16.57	70.2	2.739	166
22	Italy	609.6	176.9	30.61	95.18	8.252	152.9
23	Italy	611.3	161	26	86.8	9.129	155.1
24	Italy	840.8	594.3	3.401	6.5	0.292	86.15
25	Italy	969.6	181.9	35.66	67.65	5.285	166.2
26	Italy	906	289.5	222.6	65.97	29.33	228.9
27	Italy	1576	144.8	34.56	47.19	3.144	207.9
28	Italy	1230	772.8	14.59	19.57	4.105	85.05
29	Italy	1406	746.2	11.52	21.72	2.537	91.11
30	Italy	1199	788.6	12.84	22.68	8.711	91.46
31	Italy	980.6	502.2	62.86	25.11	17.64	94.13
32	Italy	1046	568.7	61.99	21.86	23.07	89.99
33	Italy	1060	563	65.06	22.18	13.76	94.75
34	Italy	1041	974.7	3.182	0.306	0.43	90.24
35	Italy	1105	1038	3.401	0.876	0.514	85.3
36	Italy	1091	746.7	2.962	3.106	bdl*	81.75
37	Italy	985.6	646.7	67.91	33.28	85.55	106.4
38	Italy	1102	553.1	21.83	18.97	83.95	84.42
39	Italy	1183	815.6	21.61	26.89	1.188	90.76
40	Italy	1268	929.4	15.03	18.92	1.863	85.87
41	Italy	1001	599.2	14.37	11.68	7.847	90.63
42	Italy	990.7	569.5	13.71	12.05	6.701	90.01
43	Italy	928.5	539.3	12.62	11.34	5.892	83.25
44	Italy	1134	705.2	39.6	49.05	116.6	128.7
45	Italy	1247	732.1	46.41	56.19	68.85	128.2
46	Italy	1144	666.6	45.09	49.11	97.41	123.5

PUBLICATIONS OF THIS THESIS

47	Italy	1147	244.4	84.59	70.88	106.3	272.5
48	Italy	1007	604	62.86	38.01	64.85	106.4
49	Italy	908.9	161.6	99.07	53.43	104.4	234.1
50	Italy	1245	206.7	121.9	49.22	196.5	259.2
51	Italy	990.6	687.2	8.886	13.55	2.183	85.18
52	Italy	1022	744.3	18.98	17.58	1.323	87.72
53	Italy	1096	704.5	24.03	25.35	23.71	91.22
54	Italy	1097	696.6	12.4	21.25	9.265	89.33
55	Italy	1096	694.6	34.56	28.81	4.645	88.79
56	Italy	1012	641.5	18.76	21.34	4.004	93.99
57	Italy	1158	479.4	39.39	28.67	3.414	105.1
58	Italy	1169	440.8	29.73	28.34	8.336	102.5
59	Italy	1037	632.6	67.25	42.41	110.9	115.3
60	Italy	1169	513.7	35.87	26.34	4.021	98.39
61	Italy	4088	876.6	24.46	31.93	3.211	167.7
62	Italy	1011	328.5	26.22	25.15	0.413	138.9
63	Italy	1618	516.8	32.14	26.72	1.542	191.7
64	Italy	1580	576.9	281.7	73.12	459.2	206.3
65	Italy	1243	529	262.3	67.34	610.8	205.1
66	Italy	1232	513	256.4	75.21	623.5	185.7
67	Italy	1316	224.6	57.16	63.05	51.08	233.7
68	Italy	1344	276.9	57.82	95.89	14.57	220.5
69	Italy	1308	317.8	69.23	83.14	278.5	211.5
70	Italy	935.5	733.1	30.39	5.688	50.8	105.1
71	Italy	882.2	304.7	263	69.08	619.8	238.1
72	Italy	1035	400.4	347.3	67.13	707.8	260.7
73	Italy	1424	352.5	267.4	83.88	473.9	252.6
74	Italy	1445	595.8	73.84	93.23	9.095	175.4
75	Italy	1376	498.5	116	102.5	23.97	177.6
76	Italy	1362	375.1	121.2	94.83	26.26	180.4
77	Italy	870.9	782.1	20.52	13.41	19.05	98.62
78	Italy	830.9	760.7	5.156	11.94	bdl	78.31
79	Namibia	88.74	855.4	3.083	0.292	1.199	1921
80	Namibia	73.08	910.4	6.734	0.567	0.956	2117
81	Namibia	70.34	679.6	4.433	0.697	0.403	2149
82	Namibia	283.6	96.32	6.673	0.459	1.079	258.8
83	Namibia	309	77.61	6.489	0.337	0.944	254.3
84	Namibia	290.1	70.11	4.74	bdl	0.743	248.6
85	Namibia	413	519.7	5.507	1.024	1.682	776.2
86	Namibia	408.8	61.38	3.053	0.547	0.786	793
87	Namibia	408.1	62.73	3.82	0.177	0.315	794.1
88	Namibia	327.5	161.4	5.2	0.433	2.68	329.9
89	Namibia	311	108.6	4.004	0.476	2.463	316.3
90	Namibia	326.6	153.1	4.494	bdl	1.249	311
91	Namibia	129.8	363.3	5.2	0.211	1.236	1442
92	Namibia	136	494	8.636	bdl	0.951	1659
93	Namibia	131.4	374.2	4.648	bdl	0.415	1533
94	Namibia	236.4	89.03	6.305	0.243	1.749	3671
95	Namibia	229	128.3	6.213	0.07	0.883	3429

PUBLICATIONS OF THIS THESIS

96	Namibia	231.1	65.59	4.372	bdl	0.781	3582
97	Namibia	145.2	21.47	3.513	bdl	1.682	450.4
98	Namibia	146.6	13.2	2.5	bdl	1.139	451.7
99	Namibia	156.2	18.89	2.04	bdl	0.773	498
100	Namibia	148.6	294.3	2.838	0.206	1.014	1381
101	Namibia	160.2	281.9	2.776	0.158	0.673	1404
102	Namibia	162.8	650.2	0.629	0.21	0.255	1470
103	Namibia	283.6	76.55	4.096	bdl	1.502	467.8
104	Namibia	301.7	80.95	3.237	0.086	0.999	510.4
105	Namibia	298.4	75.98	2.378	0.1	0.763	486.6
106	Namibia	406.4	44.64	2.194	0.152	0.496	670.1
107	Namibia	380.6	54.45	3.39	0.289	0.466	700
108	Namibia	404.8	80.57	4.249	0.41	0.456	598
109	Namibia	435.9	273.1	4.341	0.207	1.061	229.8
110	Namibia	430.2	294.5	3.237	bdl	0.2	252.2
111	Namibia	369.2	218.5	4.494	0.07	bdl	218.9
112	Namibia	271.7	155.3	1.549	bdl	0.723	446.4
113	Namibia	225.5	142.1	2.623	0.392	0.621	351.1
114	Namibia	322.9	164.2	4.065	bdl	0.485	335.2
115	Namibia	357.7	123.5	2.655	1.291	0.873	4633
116	Namibia	357.6	92.33	12.05	0.746	0.388	4548
117	Namibia	376.5	90.17	2.495	0.415	bdl	4981
118	Namibia	672.5	96.63	3.028	bdl	0.396	343.4
119	Namibia	405.4	82.86	4.175	0.385	0.146	388.9
120	Namibia	626.5	245	4.576	0.438	bdl	348.5
121	Namibia	374.2	387.5	15.11	0.713	0.277	4080
122	Namibia	314.9	208.9	7.377	0.302	bdl	3812
123	Namibia	335	75.66	4.416	0.243	bdl	4246
124	Namibia	287.6	201.2	6.07	0.476	0.854	3826
125	Namibia	260.9	73.56	2.921	bdl	0.684	3972
126	Namibia	219.7	112.5	2.228	1.113	0.656	3984
127	Namibia	220.6	29.78	4.656	bdl	407.1	3812
128	Namibia	193.3	24.84	3.962	bdl	101.5	3853
129	Namibia	206.9	17.08	2.708	bdl	49.72	3967
130	Namibia	321	72.73	6.336	0.802	6.632	4195
131	Namibia	341.8	79.55	5.483	0.416	1.136	4352
132	Namibia	301.6	105.3	6.977	0.679	bdl	4190
133	Namibia	210.7	91.87	2.468	bdl	0.777	3213
134	Namibia	297.5	253.6	8.044	0.127	0.313	3271
135	Namibia	232.5	400.7	4.602	0.048	bdl	3408
136	Namibia	909.4	277.3	7.137	bdl	36.67	415.6
137	Namibia	808.2	301.5	8.257	1.146	1.757	314.7
138	Namibia	862.6	166.6	3.749	0.924	0.765	385.4
139	Namibia	1061	475.6	6.55	0.795	0.55	287.8
140	Namibia	1011	111.8	4.389	0.198	bdl	360.1
141	Namibia	1063	116	4.656	0.441	bdl	358.5
142	Namibia	885.7	44.27	30.46	bdl	32.76	1355
143	Namibia	1145	117.1	95.98	1.509	25.94	1429
144	Namibia	1118	96.27	97.88	0.431	1.761	1407

PUBLICATIONS OF THIS THESIS

145	Namibia	479.7	361.3	4.464	0.235	0.903	458.7
146	Namibia	261.3	78.72	4.004	0.037	0.868	525.1
147	Namibia	364.2	132.9	5.016	0.027	0.766	434.9
148	Namibia	480.4	208.2	5.24	7.306	21.86	429
149	Namibia	519.9	468.4	7.249	0.914	1.358	427.1
150	Namibia	563.4	441.1	2.043	0.417	1.037	481.3
151	Namibia	750.5	981.2	19.09	1.596	3.306	304.1
152	Pakistan	1073	308	547.9	37.04	0.924	63.97
153	Pakistan	1049	366	550.4	71.17	222.2	137.6
154	Pakistan	1684	188.9	1048	61.55	1.291	218.3
155	Pakistan	2244	212.8	647.1	54.51	0.774	220.8
156	Pakistan	1967	209.4	1073	54.03	1.832	214.5
157	Pakistan	995.8	376.2	597.3	bdl	129	115.5
158	Pakistan	1064	365	469.7	68.11	67.69	128.5
159	Pakistan	1012	249.1	539.4	90.34	51.78	110.1
160	Pakistan	1090	267	500.8	63.91	13.39	129.6
161	Pakistan	6035	191	355	88.09	114.3	93.11
162	Pakistan	900.2	210.8	42.94	14.22	1.922	109.1
163	Pakistan	914.3	207.6	67.7	14.37	2.306	110.3
164	Pakistan	890.7	194.3	50.01	15	4.056	109.4
165	Pakistan	1459	411.1	634.2	83.89	10.41	95.46
166	Pakistan	6059	287.7	462.1	39.29	1.676	74.92
167	Pakistan	916.9	388.6	764.8	89.21	20.78	86.02
168	Pakistan	870.9	416.6	750.3	91.71	33.08	86.57
169	Pakistan	921.5	300.2	523.9	81.49	2399	88.03
170	Pakistan	831	304	559.9	95.46	359	79.52
171	Pakistan	872.8	301.5	523.7	92.63	97.98	84.81
172	Pakistan	1235	241.4	1311	153.7	64.76	41.3
173	Pakistan	884.1	232.8	1465	157.3	40.23	49.94
174	Pakistan	945.4	163.3	1060	149.2	2.981	25.64
175	Pakistan	849.6	227.7	1134	138.1	298.7	73.63
176	Pakistan	2566	234.2	1193	131.2	225.2	67.92
177	Pakistan	1799	270.7	477	36.11	1.262	72.74
178	Pakistan	759.5	215.2	997.1	118.5	145.1	72.3
179	Pakistan	2357	174.2	243.6	31.61	2.613	109.3
180	Pakistan	1196	157.2	219.6	30.6	2.244	99.14
181	Pakistan	989.4	154.5	282	32.56	2.536	107.4
182	Pakistan	6910	117.8	1023	75.58	23.41	126.5
183	Pakistan	1099	127	1288	78.59	8.633	113.1
184	Pakistan	1240	113.9	1006	81.45	8.955	106.1
185	Pakistan	1197	159.1	218.5	32.46	1.553	115.2
186	Pakistan	1557	155	222.4	36.19	1.277	114.3
187	Pakistan	990.3	161.5	217.8	34.72	2.122	113.4
188	Pakistan	4842	393.9	649.5	71.33	200.3	77.06
189	Pakistan	1821	192.9	275.8	33.74	6.483	112.7
190	Pakistan	915.7	193.5	311.2	34.38	1.753	106
191	Pakistan	876	186.7	333.8	35.35	7.972	106
192	Pakistan	1211	371.5	577.7	103.5	25.21	118.2
193	Pakistan	1003	322.8	487.4	99.07	13.56	120.9

PUBLICATIONS OF THIS THESIS

194	Pakistan	1093	300.6	612.2	101.2	17.85	119.3
195	Pakistan	3813	477.3	19.48	10.63	0.862	327
196	Pakistan	2462	513.1	9.618	9.712	1.338	321
197	Pakistan	2175	495.6	14.64	7.802	0.678	321.5
198	Pakistan	828.6	226.6	501	103.6	482.5	60.77
199	Pakistan	556.8	339.6	736.3	80.6	75.72	47.74
200	Pakistan	1359	226.3	504	104.6	664.3	73.99
201	Pakistan	758.2	217.2	469.3	96.93	520.9	70.01
202	Pakistan	983.7	143.1	408.6	39.93	8.633	126.6
203	Pakistan	1242	141.7	394.3	41.28	12.33	131.6
204	Pakistan	925.5	130.7	349.4	38.47	11.95	117.9
205	Pakistan	4078	196.4	36.24	2.484	2.936	142.2
206	Pakistan	5606	291.2	71.35	7.583	1.368	239.9
207	Pakistan	11520	230.5	121.7	7.616	1.093	200.8
208	Pakistan	1197	248.4	1612	173.3	52.44	58.24
209	Pakistan	858.8	238.7	1543	176.1	53.63	56.22
210	Pakistan	669.8	386.8	500.1	55.05	36.06	61.07
211	Pakistan	923	218.6	1619	171.6	60.96	52.4
212	Pakistan	2634	206.9	351.6	33.22	6.437	134.2
213	Pakistan	3730	195.1	334.3	31.77	3.243	120.2
214	Pakistan	3103	223	295	36.86	6.744	114.1
215	Pakistan	4666	271.1	637.5	105.3	1281	82.32
216	Pakistan	2898	232.3	573.3	119.2	469.5	77.92
217	Pakistan	4781	239	572.1	170.9	625.8	61.23
218	Pakistan	8904	223.4	210.7	67.54	89.24	117.3
219	Pakistan	3932	155.8	286.8	40.41	9.202	138.4
220	Pakistan	7955	123	539	39.25	26.52	128.3
221	Pakistan	3281	252.3	144.8	40.87	1.031	78.65
222	Pakistan	3014	161.9	933.9	150.9	30.2	41.97
223	Pakistan	989.9	220.8	1354	168.8	40.1	37.08
224	Pakistan	854.6	210.7	1337	162.8	61.78	40.12
225	Pakistan	1711	142	894.7	36.27	0.897	84.35
226	Pakistan	978.6	98.8	682.6	44.66	bdl	122.9
227	Pakistan	968	134.2	504.4	42.12	0.16	109.8
228	Pakistan	2456	358.3	1221	70.32	bdl	193.7
229	Pakistan	1874	275.4	1152	69.01	1.783	195.8
230	Pakistan	1358	354.2	901.3	92.44	63.74	186.8
231	Pakistan	1275	272.4	112.4	29	1.446	92.87
232	Pakistan	1345	359.5	885.9	82.48	68.96	183.9
233	Pakistan	1337	340.8	850.4	87.59	51.44	183.7
234	Pakistan	1860	305	655.2	143.1	829.6	78.72
235	Pakistan	1339	311.2	592.7	106.9	72.03	139.7
236	Pakistan	1385	285.8	795.1	137.2	597.8	83.89
237	Pakistan	1954	249.4	362.8	24.13	1.881	163
238	Pakistan	944.3	91.52	593	60.55	1.291	158.6
239	Pakistan	1094	111.6	775.9	46.03	0.701	160.9
240	Pakistan	4571	129.1	1292	42.12	6.479	64.15
241	Pakistan	5808	180	729.7	63.87	15.12	162.4
242	Pakistan	553.2	268.9	142.5	37.37	1.569	79.39

PUBLICATIONS OF THIS THESIS

243	Pakistan	2406	137.5	799.4	54.83	3.209	177.3
244	Pakistan	4218	448.2	2606	264.9	134.8	105.6
245	Pakistan	1420	325.3	1866	223.1	51.84	123.4
246	Pakistan	1461	392	2508	255.3	77.66	95.16
247	Pakistan	1852	213.6	704	45.05	4.635	249.7
248	Pakistan	1963	218.3	591.8	55.32	0.455	235
249	Pakistan	4112	167.5	1040	81.6	1.856	196.3
250	Russia	1133	183	858.8	89.83	3.714	384.7
251	Russia	859.6	149.5	613.1	79.99	0.942	376.8
252	Russia	996.2	157.1	735.9	88.52	0.613	367.4
253	Russia	1097	131.2	207.5	296.5	13910	513.6
254	Russia	759.2	63.44	316.8	413.6	13710	491.3
255	Russia	837.4	63.89	413.1	3.126	12570	498.3
256	Russia	1038	158.3	729	90.5	118	367.1
257	Russia	999.7	169.4	743.5	92.33	105.2	361.9
258	Russia	1448	215.8	988.7	87.33	97.59	358.9
259	Russia	1273	1039	1229	952.1	3094	161.8
260	Russia	1485	1222	1465	928	2990	167.4
261	Russia	1075	1072	1139	926.1	2650	159.1
262	Russia	1397	902.3	328.5	1702	43370	92.39
263	Russia	1323	882.4	242	1972	43010	95.6
264	Russia	1989	924.4	397.9	1535	39940	88.66
265	Russia	495.9	381.8	279.3	471	564.9	97.41
266	Russia	467.1	340.3	236	473.9	197.1	96.81
267	Russia	665.9	329.2	249	464.5	173.3	97.73
268	Russia	1204	495	1387	85.52	10030	366.2
269	Russia	1380	547	1912	116.1	5534	352.5
270	Russia	1102	470.2	1721	94.56	5094	345.8
271	Russia	919.5	564.5	351.8	34.83	1352	302.5
272	Russia	607.5	301.3	176.2	29.72	1031	286.5
273	Russia	766.6	450.6	297.5	44.01	611	299.3
274	Russia	1298	1459	1390	999.5	540.9	219.3
275	Russia	3225	1910	4859	81.66	531.2	345
276	Russia	878.4	1131	917.4	1040	445.6	211.5
277	Russia	1091	1221	1089	1031	344.5	218.8
278	Russia	2669	1523	3509	82.23	310.1	340.5
279	Russia	3207	1814	5447	93.3	250.9	330.1
280	Russia	1077	1519	305.1	bdl	21770	407.8
281	Russia	4196	292.3	254.9	bdl	21610	378.5
282	Russia	748	1521	200	8.637	21360	384.3
283	Russia	918.2	580.9	109.9	bdl	1520	529.2
284	Russia	950.4	415.1	94.86	bdl	858.5	520.8
285	Russia	1117	512.4	122	4.741	701.1	539.7
286	Russia	854.1	142.3	106.3	bdl	2637	493.7
287	Russia	582.2	49.4	49.35	bdl	2448	457.1
288	Russia	682.3	45.22	77.15	bdl	2090	486
289	Russia	446.5	265	6.87	bdl	30.65	597.6
290	Russia	508	1127	28.35	1.576	7.036	539.7
291	Russia	484.7	653.5	25.55	1.081	0.304	579.2

PUBLICATIONS OF THIS THESIS

292	Russia	642.1	23.73	42.35	bdl	1207	458.1
293	Russia	626.4	26.89	38.94	bdl	817.5	453.1
294	Russia	729	58.89	45	bdl	787	469.8
295	Russia	488.8	208	43.29	16.94	41.83	473.2
296	Russia	828.9	39.98	69.68	18.23	11.68	489
297	Russia	1116	48.96	93.02	16.12	9.97	511.1
298	Russia	500.1	53.59	34.83	10.34	30.75	416.7
299	Russia	676.2	93.24	64.79	9.3	29.34	427.4
300	Russia	848.8	103.3	52.57	9.02	12.21	450
301	Russia	837.9	55.79	36.38	bdl	1680	506.4
302	Russia	663.7	226.1	41.21	bdl	985.8	490.5
303	Russia	422.3	30.78	17.2	bdl	438.7	475.9
304	Russia	755.4	6.47	213	bdl	1832	475.1
305	Russia	964.8	161.4	111.4	4.75	590.7	474.7
306	Russia	618.4	108	0.1	bdl	480.2	451.3
307	Russia	627.9	140	5.135	bdl	455	441.4
308	Russia	669	717.9	13.99	1.499	327.3	467.5
309	Russia	571.1	892.2	79.92	27.85	252.1	444.7
310	Russia	1052	17.09	298.5	15.97	218.6	498.5
311	Russia	568.6	376.6	76.73	23.74	208	439.6
312	Russia	875	246.3	158.8	16.96	164	469.3
313	Russia	1035	2.931	338.5	19.37	113.1	504.3
314	Russia	737.7	0.1	205.5	16.15	74.17	443.7
315	Russia	515.1	130.8	37.21	9.083	35.19	459.9
316	Russia	1047	16.13	76.85	15.77	32.59	478.5
317	Russia	495.7	127.9	9.544	6.694	9.624	441.6
318	Russia	643.3	160.9	17.18	6.114	8.462	456.4
319	Russia	1488	1735	726.4	106.7	3003	346.9
320	Russia	1988	2137	931.6	127.4	1163	351.1
321	Russia	1397	1787	768.6	118.2	815.5	316
322	Russia	909.1	246.9	160.9	bdl	2692	697
323	Russia	1479	425.4	234.9	bdl	2639	704.2
324	Russia	1346	361.6	212.9	8.635	2289	749.8
325	Russia	1538	467.5	138.8	bdl	5972	646.3
326	Russia	1629	185.6	132	bdl	5302	667
327	Russia	1453	176.4	129.9	bdl	4403	615.4
328	Russia	860.8	49.43	64.98	bdl	3568	568.9
329	Russia	1812	154.3	156.4	bdl	3011	680.4
330	Russia	1728	144	135.8	bdl	2452	667
331	Russia	2494	3075	1926	54.77	6682	331.6
332	Russia	3990	1738	850.5	52.08	5935	311.9
333	Russia	1252	1430	705.7	105.6	2834	291.3
334	Russia	1726	232.8	211.3	bdl	2134	689.3
335	Russia	985.2	78.49	56.87	10.68	659.7	610
336	Russia	814.9	63.98	62.54	12.23	448.1	614.7
337	Russia	1148	1531	133.8	30.16	6950	643.4
338	Russia	828	1082	90.33	bdl	6637	619.4
339	Russia	891	1155	83.47	13.64	5278	618.5
340	Russia	606.5	6.001	531.1	bdl	6426	557.2

PUBLICATIONS OF THIS THESIS

341	Russia	854.4	56.67	485.4	bdl	4446	621
342	Russia	848.7	27.38	526.8	bdl	6786	619.2
343	Russia	1246	376.3	68.51	bdl	1521	592.6
344	Russia	599.5	179.9	57.07	bdl	1379	543.5
345	Russia	529.9	204.7	90.78	bdl	1139	534
346	Russia	1465	3021	1079	759.6	1398	137.3
347	Russia	1040	2578	603.4	742.8	454.7	128.3
348	Russia	1428	3385	726.5	651.2	309.7	132.1
349	Russia	1568	3675	805	1038	517.2	127.9
350	Russia	894.4	2232	458.1	1055	481.6	116.1
351	Russia	764.4	2116	336.4	968.8	419.9	117
352	Russia	632.3	2042	349.5	1074	387.4	116.1
353	Russia	711.7	1923	278.1	938.9	360.7	115.1
354	Russia	560.1	1724	226.7	913.9	199.3	109.3
355	Russia	806.3	3494	296.6	383.6	764.8	144.7
356	Russia	971	4478	391.7	365.7	748.6	161.8
357	Russia	1184	6351	453.4	389.4	747.3	160.9
358	Russia	503	1689	66.92	224.4	529.4	135.9
359	Russia	613.7	2216	106.5	232.5	501.9	140.9
360	Russia	583.6	1940	74.42	225.3	452.6	142.8
361	Russia	613.3	2344	87.23	260	748.3	142.6
362	Russia	673.9	2627	106	266.6	731.9	144.6
363	Russia	705.3	2686	67.71	235.1	587.8	150.7
364	Russia	1067	3596	100.4	166.4	13.98	307.7
365	Russia	1327	4202	104.5	162.9	11.77	336.9
366	Russia	934	3254	100.8	165.7	9.827	276.2
367	Iran	999.5	1838	449.3	96.94	2468	176.4
368	Iran	1205	2533	600.4	124.9	2255	224.9
369	Iran	1051	857.2	149.1	42.78	3923	72.23
370	Iran	968.5	848.2	166.2	14.59	5421	75.8
371	Iran	1257	911.8	142.8	96.22	3405	78.75
372	Iran	922.1	2014	332.9	47.82	9504	97.14
373	Iran	665	1547	342	3.661	9286	87.79
374	Iran	976.5	1787	410.2	176.3	10260	101.1
375	Iran	739.9	578.3	528.3	164.1	3578	67.05
376	Iran	919.2	632.1	648.9	161.1	5065	68.88
377	Iran	1949	679.4	555.1	137	5501	69
378	Iran	700.9	1061	61.86	108.1	588	85.04
379	Iran	657.2	1030	72.47	97.41	454.6	84.94
380	Iran	689.7	1254	77.02	96.31	300.5	79.97
381	Iran	750.4	1012	127.2	129.1	1450	81.99
382	Iran	605.4	981.5	123.5	102.8	1417	74.23
383	Iran	611.1	1095	140.8	114.6	1712	75.38
384	Iran	828.3	354.1	2127	bdl	24370	141.5
385	Iran	872.5	360	2306	bdl	22680	142.2
386	Iran	936.7	374.5	2309	bdl	25840	145.3
387	Iran	919.7	1285	137.8	33.6	3078	106.6
388	Iran	924.5	1472	164.1	103.8	1526	112.1
389	Iran	858	1420	176.8	92.66	1450	96.98

PUBLICATIONS OF THIS THESIS

390	Madagascar	775.6	24030	3325	27.29	118.4	26.54
391	Madagascar	619.3	23560	2515	26.76	104.1	24.68
392	Madagascar	541.2	25940	2069	bdl	30.58	30.68
393	Madagascar	592	28070	2341	27.79	38	26.69
394	Madagascar	644.6	1508	63.55	1.081	12.52	17.98
395	Madagascar	728.6	39.76	17.93	0.189	1.612	19.03
396	Madagascar	480.4	2104	129.7	3.924	0.646	14.5
397	Madagascar	380.1	13010	484.7	23.59	bdl	25.56
398	Madagascar	333.9	14310	487.3	24.87	1.029	24.41
399	Madagascar	355.6	1999	74.32	3.405	bdl	14.68
400	Madagascar	372.6	5142	250.8	11.35	1.299	17.58
401	Madagascar	241.1	64090	1157	81.53	33.01	195.5
402	Madagascar	171	85470	1689	109.2	3.118	192.4
403	Madagascar	232.5	76410	1553	86.19	2.473	208.7

* bdl ... below detection limit

6.3 MANUSCRIPT 3

**Stable Hydrogen and Oxygen Isotope Abundance
Analysis of Barley Malt and Its Potential to Trace the
Malting Location**

Stefan G. Bindereif¹, Christopher Porkert¹, and Gerhard Gebauer¹

¹BayCEER—Laboratory of Isotope Biogeochemistry, University of Bayreuth, Universitätsstraße
30, 95447 Bayreuth, Germany

*This is an original manuscript of an article published by Taylor & Francis in Isotopes in
Environmental and Health Studies on 25 August 2026, available at:
<https://doi.org/10.1080/10256016.2026.2721629>*

Stable hydrogen and oxygen isotope abundance analysis of barley malt and its potential to trace the malting location

Stefan Gerd Bindereif ^{a*}, Christopher Porkert ^a and Gerhard Gebauer ^a

*^aBayCEER-Laboratory of Isotope Biogeochemistry, University of Bayreuth,
Universitätsstraße 30, 95447 Bayreuth, Germany*

Corresponding author:

Stefan G. Bindereif

BayCEER-Laboratory of Isotope Biogeochemistry

University of Bayreuth

Universitätsstraße 30

95447 Bayreuth, Germany

Email: bindereif.uni@gmail.com

Stable hydrogen and oxygen isotope abundance analysis of barley malt and its potential to trace the malting location

ABSTRACT

The complexity of global food supply chains has driven demand for analytical tools that support traceability assessments beyond traditional documentation and routine quality control, as increasingly required by food safety standards. This also applies to the brewing industry, where important processing steps such as malting are often outsourced to third parties. To ensure full process transparency, breweries currently lack analytical solutions to verify that the barley supplied for malting was processed at the specified malthouse. In this study, stable hydrogen ($\delta^2\text{H}$) and oxygen ($\delta^{18}\text{O}$) isotope abundance analysis was applied to assess whether malting introduces location-related isotope signatures that could support origin verification. Thirty-four barley samples were analysed as raw barley and after parallel malting at two geographically separated German malthouses in Berlin and Munich using high-temperature conversion isotope-ratio mass spectrometry, resulting in three samples per barley lot and a total of 102 samples. Malting resulted in significant enrichments of both $\delta^2\text{H}$ and $\delta^{18}\text{O}$ values relative to the corresponding unmalted barley, indicating systematic isotopic fractionation during the malting process. In addition, malt samples produced at the two malthouses exhibited small but significant differences in hydrogen and oxygen isotope composition. The direction of these differences is consistent with regional patterns in precipitation isotope ratios and corresponding local brewing water sources. This study provides an analytical basis indicating that hydrogen and oxygen isotope ratios in malt reflect malting-location-related process information. When evaluated relative to supplied barley reference material, this approach can support plausibility-based malt traceability and could complement documentation-based verification.

Keywords: stable isotope analysis; barley malt; hydrogen; oxygen; traceability; brewing industry; food authenticity

1. Introduction

Over the past two decades, there have been numerous high-profile food fraud cases in Europe that have permanently damaged consumer confidence in the food industry [1–3]. In response to these incidents and the growing complexity of food supply chains, private actors have created food safety standards, such as the ‘International Featured Standard (IFS)’ or the ‘Safe

Quality Food (SQF) Standard' to complement EU regulatory frameworks and Global Food Safety Initiative (GFSI) requirements [4,5]. Among other things, these standards require hazard analyses and vulnerability assessments of products and suppliers along the supply chain, as well as the establishment of a food fraud mitigation plan. In this context, analytical detection strategies play a decisive role, since documentation and auditing alone are insufficient to reliably identify fraud at many points in complex supply chains [6,7].

Consequently, the development of analytical monitoring methods has become a major focus in the food sector, driven by both food businesses and public authorities, as these methods need to be in place before irregularities occur [8–10]. Within this context, beer production represents a particularly relevant and analytically challenging case, as raw materials are frequently traded internationally and processed by third parties before reaching breweries, creating supply-chain risks. Beer is produced from cereal grains, most commonly from barley (*Hordeum vulgare* L.). Prior to brewing, barley is processed into malt in malthouses, where it undergoes a controlled incomplete germination process that strongly contributes to the quality and characteristics of the final product [11]. In short, malting aims to promote the formation and activation of enzymes that are later responsible for the breakdown of cell walls, starch, lipids, and protein during brewing, thereby converting the grain's stored reserves into fermentable sugars and other compounds required for yeast metabolism [12]. According to Yousif & Evans [13], the final malt quality is essentially determined by three primary factors: the barley variety selected for malting, the quality of the barley itself, and the malting conditions applied by the maltster [14,15]. Even when malting is outsourced, breweries can largely verify compliance with these key quality determinants upon malt delivery through routine quality controls. For example, parameters including moisture content and protein concentration can be assessed using rapid screening methods like near-infrared spectroscopy, while varietal purity can be tested using DNA-based techniques [16–18]. Beyond quality assurance, breweries also seek to verify that the delivered malt originates from the barley lots

they supplied and was processed at the contracted malthouse. However, to the best of our knowledge, there is currently no established analytical approach that allows an independent verification of the declared malting location of purchased malt. In this context, the present study explores the potential of stable isotope analysis to support malt traceability. Natural abundance analysis of stable isotopes by isotope-ratio mass spectrometry (IRMS) is a well-established approach with officially acknowledged methods in the food control sector, routinely used for authenticity testing of honey, wine, and fruit juice, just to name a few [19]. It measures small changes in isotopic composition that can be used to draw conclusions about the ‘history’ of samples [20,21]. Here, we focus specifically on the isotope ratios of the two water elements hydrogen ($^2\text{H}/^1\text{H}$) and oxygen ($^{18}\text{O}/^{16}\text{O}$). Briefly, there is a well-researched relationship between the isotope composition of the precipitation at a given location and that of local ground water. At the same time, systematic differences in the isotopic composition of precipitation occur between regions due to known fractionation processes, resulting in global and regional patterns that can serve as geographic ‘fingerprints’ [22]. Although the isotopic composition of tap water can be influenced by various factors, it generally represents a suitable proxy for local water sources [23–25]. During the steeping step of the malting process, local tap water is added to the dry barley grains (moisture content around 12 %), which are soaked under controlled conditions to increase their moisture content (40-45%) and to initiate germination [26]. The steeping water is then metabolised for enzyme synthesis and hydrolytic reactions and thereby hydrogen and oxygen from the local water is incorporated into newly formed biomolecules and ultimately into the finished malt. The subsequent heat-drying step (the so-called ‘kilning’) removes most free water but retains these newly synthesized compounds. In this study, we investigated whether stable hydrogen and oxygen isotope ratios can capture this process-related incorporation in a way that supports the traceability and verification of malting location, based on comparisons between malt and the supplied barley raw material.

2. Materials and methods

We analysed 34 barley samples from the 2018 crop year. Each barley sample was analysed in its original (unmalted) form and was divided into two representative subsamples, which were malted in malthouses in Munich and Berlin, respectively. This resulted in three samples per barley lot (raw barley, Munich malt, and Berlin malt) and a total of 102 samples. The two malthouses were deliberately selected because they are located approximately 500 km apart, allowing the investigation of the potential distinguishability of malthouses from different geographical locations. All samples were provided by a collaboration partner from the brewing industry. For the determination of isotope abundances of the elements hydrogen and oxygen, approximately 4 g of grain was ground to a fine powder in a ball mill (MM200, Retsch GmbH, Haan, Germany), dried at 105 °C to constant weight in a drying oven, and subsequently stored in a desiccator until analysis. 0.5-1 mg of the powder were weighed into silver capsules using a microbalance (CPA2P, Sartorius, Göttingen, Germany). Isotope analysis was performed with thermal conversion through pyrolysis (HTO, HEKAtech, Wegberg, Germany) coupled to a continuous flow isotope-ratio mass spectrometer (delta V advantage, Thermo Fisher Scientific, Bremen, Germany) via a ConFlo IV open-split interface (Thermo Fisher Scientific), following the method described by Gebauer et al. [27]. All samples for hydrogen isotope abundance determination were measured four times consecutively and the results of the first three samples were neglected to combat the memory effect of the spectrometer. Isotope abundances are denoted as δ values, calculated according to the equation:

$$\delta^2\text{H or } \delta^{18}\text{O} = \left(\text{R}_{\text{sample}} / \text{R}_{\text{standard}} - 1 \right) \times 1000 \text{ (‰)} \quad (1)$$

where R_{sample} and $\text{R}_{\text{standard}}$ are the ratios of heavy to light isotope of the samples and the respective standard. Standard gases (Riessner-Gase, Lichtenfels, Germany) were calibrated

with respect to international standards (H_2 and CO vs. V-SMOW) with the reference substances CH7, V-SMOW and SLAP for hydrogen isotopes and IAEA601 and IAEA602 for O isotopes, all provided by the International Atomic Energy Agency (IAEA), Vienna, Austria. Reproducibility and accuracy of the O and H isotope abundance measurements were routinely controlled by measuring the laboratory standard benzoic acid with variable sample weight at least six times within each batch of 40 samples. The uncertainty of measurement was always below 4‰ for $\delta^2\text{H}$ and 0.6‰ for $\delta^{18}\text{O}$, respectively. As non-carbon bound hydrogen atoms exchange with ambient moisture, presented abundances cannot be considered ‘true’ $\delta^2\text{H}$ [28]. However, all samples were treated identically and measured in batches composed of barley and malt samples from the two malthouses at the same laboratory and device following strict quality guidelines so that comparability of samples within each batch was ensured [27]. Therefore, all isotope data are directly comparable within the scope of this study. For raw isotope data evaluation, the software ISODAT (Version 3.0, Thermo Fisher Scientific) was used. Statistical analysis and figure preparation were carried out in the computing environment R version 4.5.2 [29] in combination with RStudio version 2026.01.0+392 [30]. To test the assumptions for the use of parametric methods, we applied the Shapiro–Wilk test [31] for normality and Mauchly's test for sphericity. Sphericity can be seen as the repeated-measures equivalent of homogeneity of variances and represents the condition that the variances of the pairwise differences are equal [32]. As neither assumption was violated, we subsequently performed a repeated-measures ANOVA. For pairwise comparisons, we applied multiple paired two-tailed t-tests, and p-values were adjusted using the Holm–Bonferroni method to correct for multiple testing [33].

3. Results and discussion

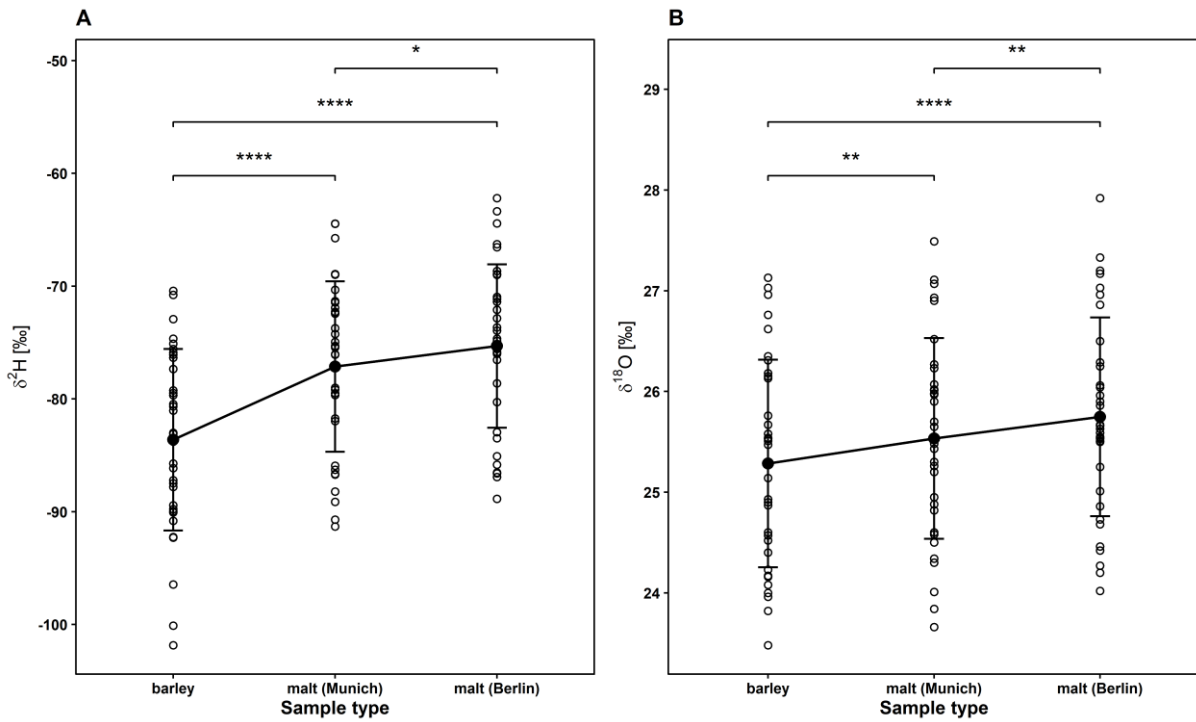


Figure 1: Line plots showing the natural (A) hydrogen and (B) oxygen isotope abundances of the analysed barley samples and malt samples from two malthouses (Munich and Berlin) and results of the multiple paired t-tests. Unfilled circles display the δ values for each sample while filled circles represent means. Error bars show the standard deviation. Significance levels: Highly significant differences ($p \leq 0.001$; ****), very significant differences ($p \leq 0.01$; **), significant differences ($p \leq 0.05$; *).

Statistically significant differences were found between groups for both, hydrogen isotope data ($F_{(2,66)} = 44.1$, $\eta^2G = 0.18$, $p < 0.001$) and oxygen isotope data ($F_{(2,66)} = 14.8$, $\eta^2G = 0.04$, $p < 0.001$) in the repeated-measures ANOVA. Group mean values were $-83.61 \pm 8.05\text{‰}$ for $\delta^2\text{H}$ and $25.28 \pm 1.03\text{‰}$ for $\delta^{18}\text{O}$ in barley, $-77.13 \pm 7.56\text{‰}$ and $25.53 \pm 1.00\text{‰}$ in Munich malt, and $-75.31 \pm 7.24\text{‰}$ and $25.75 \pm 0.99\text{‰}$ in Berlin malt. Furthermore, post-hoc analyses revealed that all pairwise differences were statistically significant (Figure 1). Paired comparisons showed that malt samples from Munich were isotopically enriched (more positive values) in $\delta^2\text{H}$ ($p < 0.001$) and $\delta^{18}\text{O}$ ($p < 0.01$) compared to their corresponding barley samples with a mean enrichment in $\delta^2\text{H}$ of $6.48 \pm 6.22\text{‰}$ and $0.25 \pm 0.52\text{‰}$ in $\delta^{18}\text{O}$. Likewise, malt samples from Berlin were enriched in $\delta^2\text{H}$ ($p < 0.001$) and $\delta^{18}\text{O}$ ($p < 0.001$)

compared to respective barley samples with a mean enrichment in $\delta^2\text{H}$ of $8.31 \pm 5.41\text{‰}$ and $0.46 \pm 0.57\text{‰}$ in $\delta^{18}\text{O}$. Overall, enrichment was observed in the majority of samples for both isotopes, occurring in 82% (Munich) and 91% (Berlin) for $\delta^2\text{H}$ and 71% and 79% for $\delta^{18}\text{O}$, respectively, indicating a robust process-related isotope shift despite considerable inter-sample variability. Several potential mechanisms may contribute to the observed enrichment. In rare cases, barley has an initial moisture content of 15% or higher and must be dried before it can be safely stored, as otherwise there is the risk of premature germination or mould growth. Depending on the drying method used, temperatures of up to 50° C can be reached which could lead to kinetic isotope fractionation, meaning that isotopically lighter water molecules evaporate more readily than heavier isotopologues [34]. In addition, temperatures of 70° C or higher are also reached later in the kilning step of the malting process, which may further enhance this fractionation and could explain the observed isotopic shifts. Interestingly, malt samples from Berlin were slightly but significantly enriched in $\delta^2\text{H}$ ($p < 0.05$) and $\delta^{18}\text{O}$ ($p < 0.01$) compared to respective malt samples from Munich, with mean offsets of $1.82 \pm 4.51\text{‰}$ for $\delta^2\text{H}$ and $0.21 \pm 0.39\text{‰}$ for $\delta^{18}\text{O}$ (Figure 1). The observed values are consistent with the well-established spatial pattern of mean annual hydrogen and oxygen isotope ratios of precipitation in Germany, which generally decrease from north to south due to the so-called ‘continental effect’ [22,25]. Mean annual precipitation isotope ratios reported for Berlin ($\delta^2\text{H} = -58\text{‰}$; $\delta^{18}\text{O} = -8.1\text{‰}$) are isotopically heavier than those for Munich ($\delta^2\text{H} = -68\text{‰}$; $\delta^{18}\text{O} = -9.6\text{‰}$), corresponding to regional differences of approximately +10‰ in $\delta^2\text{H}$ and +1.5‰ in $\delta^{18}\text{O}$ [35,36]. In Berlin, tap water and thus the water used for malting is derived predominantly from locally recharged groundwater [37] and reflects regional precipitation, whereas Munich’s water supply originates largely from Alpine foreland sources, mainly from the Mangfall valley [38] with precipitation isotope signatures similar to the Munich region. The smaller isotopic offset between Berlin and Munich malts compared to the difference in mean annual precipitation isotope values might result from a bulk-mixing effect. Malting only

leads to a limited formation of new biomass that incorporates the malting water, whereas a substantial fraction of the analysed material remains original barley biomass. As a result, the local signature may be diluted in bulk isotope measurements, and the regional water isotope signal is therefore expected to be preserved in direction but reduced in magnitude. In future studies, this limitation could be circumvented by a technical advancement called compound-specific isotope analyses [39]. For example, compounds of which a large proportion are newly formed during malting, such as the protein fraction, may exhibit a more pronounced local isotope signal than the bulk analysis applied here.

In conclusion, the systematic trends point to a promising applicability of the methodology, which will need to be confirmed in further studies including multiple facilities and repeated sampling across different seasons. Breweries can use retained reference samples of supplied barley lots as internal baselines and directly compare them with the corresponding delivered malt, enabling an analytical plausibility check of the declared malthouse based on relative isotopic shifts. Establishment of reference databases covering malt produced at different contracted malthouses in combination with the discussed technical advancements of the analytical approach could further allow characteristic isotope changes to be identified over time. Overall, stable hydrogen and oxygen isotope abundance analysis of malt provides process-related information linked to malting location and represents a practical tool to strengthen plausibility checks and traceability within brewery supply chains.

Acknowledgements

Special thanks go to our collaboration partner from the brewing industry for providing the sample material. We also thank Carina Bauer, Petra Eckert and Christine Tiroch (BayCEER – Laboratory of Isotope Biogeochemistry, University of Bayreuth) for performing the stable isotope abundance measurements.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This research is a spin-off of the research project AgrOr and was supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support program [Reference number: 2816502414].

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

References

- [1] Everstine K, Hellberg RS, Sklare SA. Introduction to food fraud. In: Hellberg RS, Everstine K, Sklare SA, editors. Food Fraud. Academic Press; 2021. p. 1–7.
- [2] Everstine K, Chin H, Lopes F, et al. Database of food fraud records: Summary of data from 1980 to 2022. *J Food Prot.* 2024;87(3):100227.
- [3] Giannakas K, Yiannaka A. Food fraud: causes, consequences, and deterrence strategies. *Annu Rev Resour Econ.* 2023;15(1):85–104.
- [4] Rao M, Bast A, de Boer A. European private food safety standards in global agri-food supply chains: a systematic review. *Int Food Agribus Manag Rev.* 2021;24(6):739–754.
- [5] Spink JW. The GFSI food fraud prevention compliance development & requirements: a ten-year review. *Trends Food Sci Technol.* 2023;138:766–773.
- [6] Barrere V, Everstine K, Théolier J, et al. Food fraud vulnerability assessment: towards a global consensus on procedures to manage and mitigate food fraud. *Trends Food Sci Technol.* 2020;100:131–137.
- [7] Huisman W, van Ruth S. Risky business: food fraud vulnerability assessments. *Crime Law Soc Change.* 2022;78(5):559–576.
- [8] Haider A, Iqbal SZ, Bhatti IA, et al. Food authentication, current issues, analytical techniques, and future challenges: a comprehensive review. *Compr Rev Food Sci Food Saf.* 2024;23(3):e13360.
- [9] Manning L, MacLeod A, James C, et al. Food fraud prevention strategies: building an effective verification ecosystem. *Compr Rev Food Sci Food Saf.* 2024;23(6):e70036.

- [10] Popping B, Buck N, Bánáti D, et al. Food inauthenticity: authority activities, guidance for food operators, and mitigation tools. *Compr Rev Food Sci Food Saf*. 2022;21(6):4776–4811.
- [11] Rani H, Bhardwaj RD. Quality attributes for barley malt: “The backbone of beer.” *J Food Sci*. 2021;86(8):3322–3340.
- [12] Bamforth CW, Fox GP. Malting and brewing. In: Shewry PR, Koksel H, Taylor JRN, editors. *ICC Handb 21st Century Cereal Sci Technol*. Academic Press; 2023. p. 363–368.
- [13] Yousif AM, Evans DE. Changes in malt quality during production in two commercial malt houses. *J Inst Brew*. 2020;126(3):233–252.
- [14] Fox GP, Bettenhausen HM. Variation in quality of grains used in malting and brewing. *Front Plant Sci*. 2023;14:1172028.
- [15] Kumar V, Chaturvedi SK, Singh GP. Brief review of malting quality and frontier areas in barley. *Cereal Res Commun*. 2023;51(1):45–59.
- [16] Habernicht DK, Blake TK. Application of PCR to detect varietal purity in barley malt. *J Am Soc Brew Chem*. 1999;57(2):64–71.
- [17] Perry DJ, Fernando U, Lee S-J. Simple sequence repeat-based identification of Canadian malting barley varieties. *Can J Plant Sci*. 2014;94(3):485–496.
- [18] Sileoni V, Marconi O, Perretti G. Near-infrared spectroscopy in the brewing industry. *Crit Rev Food Sci Nutr*. 2015;55(12):1771–1791.
- [19] Camin F, Boner M, Bontempo L, et al. Stable isotope techniques for verifying the declared geographical origin of food in legal cases. *Trends Food Sci Technol*. 2017;61:176–187.
- [20] Meier-Augenstein W. *Stable isotope forensics: Methods and forensic applications of stable isotope analysis*. 2nd ed. New York (NY): John Wiley & Sons; 2017.
- [21] Rossmann A. Determination of stable isotope ratios in food analysis. *Food Rev Int*. 2001;17(3):347–381.
- [22] Boner M, Förstel H. Stable isotope variation as a tool to trace the authenticity of beef. *Anal Bioanal Chem*. 2004;378(2):301–310.
- [23] Bhuiyan SA, Jameel Y, Chartrand MMG, et al. Spatial variations in tap water isotopes across Canada: tracing water from precipitation to distribution and assess regional water resources. *PLOS Water*. 2023;2(1):e0000068.
- [24] Bowen GJ, Ehleringer JR, Chesson LA, et al. Stable isotope ratios of tap water in the contiguous United States. *Water Resour Res*. 2007;43(3):W03419.

- [25] Förstel H, Houbé J, Hützen H. Use of tap water samples for monitoring the geographical variation of stable isotopes used in authenticity studies. *Z Lebensm Unters Forsch A*. 1997;204(2):103–108.
- [26] Rani H, Standish A, Walling JG, et al. Development and quality assessment of low-cost benchtop malting protocol for laboratory-scale barley (*Hordeum vulgare* L.) malt quality evaluation. *J Cereal Sci*. 2025;123:104138.
- [27] Gebauer G, Preiss K, Gebauer AC. Partial mycoheterotrophy is more widespread among orchids than previously assumed. *New Phytol*. 2016;211(1):11–15.
- [28] Meier-Augenstein W, Hobson KA, Wassenaar LI. Critique: measuring hydrogen stable isotope abundance of proteins to infer origins of wildlife, food and people. *Bioanalysis*. 2013;5(7):751–767.
- [29] R Core Team. R: A language and environment for statistical computing. Version 4.5.2 [software]. R Foundation for Statistical Computing; 2026. Available from: <https://www.R-project.org/>
- [30] RStudio Team. RStudio: Integrated development for R. Version 2026.01.0+392 [software]. Posit Software, PBC; 2026. Available from: <https://www.rstudio.com/>
- [31] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52(3/4):591–611.
- [32] Mauchly JW. Significance test for sphericity of a normal n-variate distribution. *Ann Math Stat*. 1940;11(2):204–209.
- [33] Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6(2):65–70.
- [34] Hoefs J. Isotope fractionation processes of selected elements. In: Hoefs J, editor. *Stable Isot Geochem*. Cham: Springer International Publishing; 2021. p. 49–265.
- [35] Bowen GJ. The Online Isotopes in Precipitation Calculator [Internet]. 2026 [cited 2026 Feb 17]. Available from: <https://www.waterisotopes.org>.
- [36] Bowen GJ, Revenaugh J. Interpolating the isotopic composition of modern meteoric precipitation. *Water Resour Res*. 2003;39(10):1299.
- [37] Berliner Wasserbetriebe. Wasserqualität [Water quality]. [Internet]. BWB. [cited 2026 Feb 17]. Available from: <https://www.bwb.de/de/wasserqualitaet.php>. German.
- [38] Stadtwerke München. Wasserversorgung Münchens: So gewinnen wir das Münchner Trinkwasser [Munich's water supply: This is how we source Munich's drinking water]. [Internet]. [cited 2026 Feb 17]. Available from: <https://www.swm.de/wasser/trinkwassergewinnung>. German.

- [39] Liu H, Nie J, Liu Y, et al. A review of recent compound-specific isotope analysis studies applied to food authentication. *Food Chem.* 2023;415:135791.

Supplemental material

Table S1. Stable hydrogen ($\delta^2\text{H}$) and oxygen ($\delta^{18}\text{O}$) isotope abundance values of barley and corresponding malt samples. Each barley lot ($n = 34$) was malted in parallel at two geographically separated malthouses (Munich and Berlin).

sample ID	$\delta^{18}\text{O}$	$\delta^2\text{H}$
1_barley	26.96	-79.25
1_malt_munich	26.90	-68.98
1_malt_berlin	27.17	-66.27
2_barley	27.13	-83.11
2_malt_munich	26.52	-71.33
2_malt_berlin	26.96	-70.96
3_barley	26.76	-86.13
3_malt_munich	27.11	-71.92
3_malt_berlin	27.33	-80.29
4_barley	25.76	-79.48
4_malt_munich	26.02	-72.40
4_malt_berlin	26.25	-72.85
5_barley	26.13	-75.89
5_malt_munich	26.07	-75.30
5_malt_berlin	26.04	-69.00
6_barley	25.53	-75.57
6_malt_munich	26.27	-76.05
6_malt_berlin	26.86	-68.93
7_barley	26.31	-75.11
7_malt_munich	25.97	-72.43
7_malt_berlin	25.86	-75.13
8_barley	25.14	-87.80
8_malt_munich	26.01	-70.33
8_malt_berlin	26.29	-71.13

PUBLICATIONS OF THIS THESIS

9_barley	25.51	-76.34
9_malt_munich	25.70	-72.27
9_malt_berlin	25.96	-71.39
10_barley	24.93	-87.21
10_malt_munich	25.30	-79.66
10_malt_berlin	25.53	-73.94
11_barley	24.87	-70.42
11_malt_munich	25.48	-71.44
11_malt_berlin	25.90	-63.36
12_barley	25.58	-70.77
12_malt_munich	25.65	-64.46
12_malt_berlin	25.50	-64.43
13_barley	26.18	-72.92
13_malt_munich	25.98	-64.47
13_malt_berlin	26.50	-62.19
14_barley	25.67	-76.07
14_malt_munich	25.20	-68.92
14_malt_berlin	25.55	-66.55
15_barley	26.35	-80.46
15_malt_munich	26.23	-65.74
15_malt_berlin	25.63	-74.85
16_barley	25.54	-80.65
16_malt_munich	26.27	-75.43
16_malt_berlin	26.06	-75.57
17_barley	25.47	-83.01
17_malt_munich	25.90	-79.01
17_malt_berlin	25.66	-75.84
18_barley	27.03	-79.68
18_malt_munich	27.07	-76.07
18_malt_berlin	27.92	-68.66
19_barley	26.15	-96.45
19_malt_munich	26.93	-76.03

PUBLICATIONS OF THIS THESIS

19_malt_berlin	27.20	-74.66
20_barley	26.62	-74.65
20_malt_munich	27.49	-73.72
20_malt_berlin	27.03	-72.09
21_barley	24.90	-90.82
21_malt_munich	25.26	-74.95
21_malt_berlin	25.01	-73.66
22_barley	24.52	-87.48
22_malt_munich	24.95	-86.66
22_malt_berlin	25.59	-82.95
23_barley	24.60	-89.78
23_malt_munich	24.82	-79.50
23_malt_berlin	25.25	-78.62
24_barley	24.40	-77.34
24_malt_munich	25.43	-79.17
24_malt_berlin	25.67	-76.54
25_barley	24.16	-89.95
25_malt_munich	24.88	-74.25
25_malt_berlin	25.54	-76.00
26_barley	24.57	-81.04
26_malt_munich	24.60	-81.73
26_malt_berlin	25.51	-75.39
27_barley	24.23	-89.43
27_malt_munich	24.30	-90.71
27_malt_berlin	24.46	-75.89
28_barley	24.00	-92.25
28_malt_munich	23.84	-86.71
28_malt_berlin	24.68	-83.49
29_barley	24.17	-92.30
29_malt_munich	24.50	-86.28
29_malt_berlin	24.42	-85.08
30_barley	23.82	-85.73

PUBLICATIONS OF THIS THESIS

30_malt_munich	24.58	-81.96
30_malt_berlin	24.73	-86.61
31_barley	23.96	-90.06
31_malt_munich	23.66	-85.93
31_malt_berlin	24.02	-86.54
32_barley	24.08	-101.84
32_malt_munich	24.34	-89.13
32_malt_berlin	24.86	-88.86
33_barley	25.14	-100.11
33_malt_munich	24.01	-91.31
33_malt_berlin	24.27	-86.92
34_barley	23.48	-83.79
34_malt_munich	24.88	-88.22
34_malt_berlin	24.20	-85.83

7 ACKNOWLEDGEMENTS

I want to express my sincere gratitude to the following individuals and organizations:

Prof. Dr. Gerhard Gebauer (Doctoral Supervisor): For his excellent and professional guidance throughout my research journey. I am grateful for the opportunity to work in a modern laboratory under his supervision, which has allowed me to shape my projects. Furthermore, I am thankful for the opportunity to participate in both national and international scientific events, where I had the chance to present my results and engage in networking. The countless shared team activities will be cherished memories.

Prof. Dr. Stephan Schwarzingger: I am thankful for his outstanding mentorship. Our discussions spanned various areas of science and opened numerous prospects.

Prof. Dr. Johanna Pausch: For her advice and fruitful conversations in a relaxed atmosphere.

Laboratory team consisting of *Carina Bauer, Petra Eckert, Christine Tiroch, Heidi Zier, Ilse Thaufelder, and Isolde Baumann*: Thank you for your invaluable assistance with technical tasks. The pleasant working environment and enjoyable coffee breaks will be fondly remembered.

Industry Collaborators: For providing the foundation for my research projects, offering sample materials, and exciting insights into industry-related challenges that served as the basis for research questions.

Dr. Philipp Giesemann, Dr. Saskia Giesemann, Dr. Julienne Schweiger, Dr. Laura Skates, Dr. Franziska Zahn, Felix Rüll, Peter Kolb, Fabienne Lindner, Felix Brauer, Britta Zimmermann, Sigrid Schwarzingger, Lucas Köberle, Simon Steidele, Hanna Gebhardt, Anne Zeitler, Dr. Jasmin Kniese, Nina Hoffmann: Thank you for the wonderful times we shared. Whether in scientific discussions or private moments, cooking together, exercising, or celebrating, you have been amazing friends, providing much-needed distraction and energy replenishment.

ACKNOWLEDGEMENTS

Stefanie Simon, Christopher Porkert, Nico Rüssel, Kevin Wobedo: For your support as student assistants and the opportunity to gain experience in the field of project management.

Federal Ministry of Food and Agriculture (Germany) and the *Graduate School of the University of Bayreuth:* For the financial support and extensive offerings of education opportunities.

Dr. Merisa Bindereif: I extend my deepest gratitude for her unconditional love and support in all aspects of my life. I couldn't imagine a better wife for the journey ahead.

My family and friends: A heartfelt gratitude goes to my family (*Brigitte, Gerhard, and Florian Bindereif*) as well as my friends, especially *Dinh Liem Nguyen*. From the very beginning, they provided me with unwavering support and stood by me as I embarked on this research journey. I would also like to thank *Jasmina Saric* for her welcome nature and encouragement over the past years. Their constant presence has been a source of strength, even during the most challenging moments.

8 (Eidesstattliche) Versicherungen und Erklärungen

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit versichere ich eidesstattlich, dass ich die Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe (vgl. Art. 97 Abs. 1 Satz 8 BayHIG).

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit erkläre ich, dass ich die Dissertation nicht bereits zur Erlangung eines akademischen Grades eingereicht habe und dass ich nicht bereits diese oder eine gleichartige Doktorprüfung endgültig nicht bestanden habe.

(§ 8 Satz 2 Nr. 4 PromO Fakultät)

Hiermit erkläre ich, dass ich Hilfe von gewerblichen Promotionsberatern bzw. –vermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe noch künftig in Anspruch nehmen werde.

(§ 8 Satz 2 Nr. 7 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass die elektronische Fassung der Dissertation unter Wahrung meiner Urheberrechte und des Datenschutzes einer gesonderten Überprüfung unterzogen werden kann.

(§ 8 Satz 2 Nr. 8 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass bei Verdacht wissenschaftlichen Fehlverhaltens Ermittlungen durch universitätsinterne Organe der wissenschaftlichen Selbstkontrolle stattfinden können.

.....

Ort, Datum, Unterschrift