

# Development of a fingerprinting approach to determine the geographical origin of vegetable oils: A case study for rapeseed oil from Hesse (Germany)



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

IWAR

vom Fachbereich Bau- und Umweltingenieurwissenschaften  
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades  
Doktor-Ingenieurs (Dr.-Ing.)

genehmigte Dissertation  
von Lili Xia, M.Sc. aus Jianhu

Erstgutachterin: Prof. Dr. rer. nat. Liselotte Schebek

Zweitgutachterin: Prof. Dr.-Ing. Dorota Iwaszczuk

Darmstadt 2023

---

Development of a fingerprinting approach to determine the geographical origin of vegetable oils: A case study for rapeseed oil from Hesse (Germany)

Entwicklung eines Fingerprinting-Ansatzes zur Bestimmung der geografischen Herkunft von Pflanzenölen: Eine Fallstudie für Rapsöl aus Hessen (Deutschland)

Xia, Lili: Development of a fingerprinting approach to determine the geographical origin of vegetable oils: A case study for rapeseed oil from Hesse (Germany).

Darmstadt. Technische Universität Darmstadt.

Jahr der Veröffentlichung der Dissertation auf TUpriints: 2024

URN: urn:nbn:de:tuda-tuprints-280156

Tag der Einreichung: 27.06.2023

Tag der mündlichen Prüfung: 26.09.2023

Veröffentlichung unter CC BY 4.0 International

<https://creativecommons.org/licenses/>

---

---

---

## Acknowledgments (Danksagung)

---

First and foremost, I would like to express my heartfelt gratitude to my advisor, Professor Liselotte Schebek, for her unwavering support and the countless opportunities she provided throughout my research journey. Her positive feedback has been a constant motivator, driving me towards achieving my research goals.

I extend my most profound appreciation to Dr. Kaori Sakaguchi-Söder, who illuminated my path from the very beginning of my research. Throughout this journey, she exhibited remarkable patience in our communications and willingly dedicated time to listen to my thoughts. Every conversation, whether centered on professional knowledge or other topics, has been immensely enriching. She has been like a nurturing gardener, tending to me like a delicate flower, and her dedicated care has played a pivotal role in my personal and academic growth.

My sincere thanks also go to Prof. Dorota Iwaszczuk for providing valuable presentation opportunities within her research group, offering insightful, practical feedback, and guiding me to contemplate my research topic within a different disciplinary context.

Beyond my advisors, I thank the rest of my thesis committee, Prof. Holger Lutze, Prof. Hans-Joachim Linke, and Prof. Boris Lehmann, for their encouragement, insightful comments, and challenging questions.

I extend my heartfelt appreciation to Professor Christoph Schüth for graciously granting me access to his laboratory and offering solutions promptly when issues arose with the lab instruments.

Furthermore, I am grateful to my colleagues at the IAG laboratory and IWAR laboratory: Claudia Cosma, Stefanie Schmidt, Gabriela Schubert, Thomas Schiedek, Zahra Neumann, Nils Michelsen, Nadine Vater, Alexander Pfannenstiel, Sylvia Borsch, and Harald Grund, for their invaluable support and shared laboratory experiences. Completing this thesis would not have been possible without their collaborative efforts. Special thanks to Landesbetrieb Landwirtschaft Hessen, including Mr. Schaumberg, Mr. Deisenroth, Mr. Bickhardt, Ms. Käufler, Mr. Rüdiger, and Mr. Batz, for providing the invaluable rapeseed samples from Hesse.

Acknowledgments are also due to colleagues at FG SuR, Dominik Dörder, Berhane Asfaw, and Michael Gottschling, for engaging in discussions and knowledge exchanges on our research topics. Despite the challenges posed by the pandemic, we found more opportunities to connect and support one another. I appreciate Alessio Campitelli, Steffi Weyand, Karoline Wowra, Alice Lopes, Zoe Miao, Tabea Hagedorn, Vanessa Zeller, Almut Güldemund, and other SuR colleagues for their sunny smiles, kindness, encouragement, and enthusiasm. Also, I thank Lina Emilie Budde for her invaluable data acquisition and Luisa Barkmann for the support of laboratory work.

Finally, I thank my parents for their profound love, akin to the sea's depth and the mountain's

---

steadfastness. Their encouragement has laid a robust foundation for my journey. I thank the members of the Gemeinde Darmstadt for their encouragement and prayers. A special acknowledgment goes to my husband, Yanlong Gao, for his unwavering support and comforting words during moments of self-doubt or difficulty. His presence and love have served as my driving force, instilling in me the determination to overcome challenges.

Ultimately, I thank God for guiding me through the various seasons of life.

Lili Xia  
06.2023

---

## Abstract

---

Vegetable oils play crucial roles as both a food source and a renewable energy resource. Determining their geographical origin is essential for ensuring food safety, health, and energy sustainability. While stable isotopic analysis of light elements combined with chemometric techniques has been effective in tracing the geographical origin of vegetable oils, the potential of stable hydrogen isotopic composition ( $\delta^2\text{H}$ ) of fatty acids (FAs) as analytical markers and the associated fingerprinting approach remains underexplored, particularly for rapeseed oil, which is the third-largest vegetable oil globally.

This dissertation aimed to develop a fingerprinting approach for determining the geographical origin of vegetable oils, using rapeseed oil as a case study. The research objectives are threefold: (1) to optimize the sample preparation method for efficient and high-throughput measurement of  $\delta^2\text{H}$  of FAs in rapeseed; (2) to analyze the spatial distribution of  $\delta^2\text{H}$  of FAs in rapeseeds from different regions and assess its correlation with climatic and soil factors; (3) to establish classification models using  $\delta^2\text{H}$  of FAs and elemental composition of rapeseed to differentiate geographical origins, including distinguishing between Hesse and other regions, as well as various regions within Hesse.

A rapid and straightforward sample preparation method by Garcés & Mancha (1993) enables the direct conversion of rapeseed into fatty acid methyl esters (FAMES). The addition of 2,2-dimethoxypropane (DMP) for effective conversion was evaluated through control experiments to assess its impact on the isotopic composition of individual FAMES using gas chromatography-pyrolysis-isotope ratio mass spectrometry (GC-Py-IRMS).

For the second and third objectives, a total of 121 rapeseed and soil samples were collected from Hesse, Germany, during 2017-2020, and 28 rapeseed samples from Jianhu County, Jiangsu Province, China, in 2019. The  $\delta^2\text{H}$  of major FAs (C18:1, C18:2, C18:3, C16:0) in rapeseed were determined by GC-Py-IRMS, while the elemental composition of rapeseed was analyzed using inductively coupled plasma optical emission spectrometry (ICP-OES) and inductively coupled plasma mass spectrometry (ICP-MS). Climate data and soil parameters were sourced from publicly available databases or measured using standardized laboratory methods. Chemometric techniques were employed for data analysis, including stepwise linear regression analysis, principal component analysis (PCA), and orthogonal partial least squares discriminant analysis (OPLS-DA).

Control experiments confirmed that DMP had no significant influence on the accuracy of  $\delta^2\text{H}$  values of FAs. The precision of hydrogen compound-specific stable isotope analysis (H-CSIA) ranged from  $\pm 0.3$  mUr to  $\pm 3$  mUr, resulting in satisfactory results. This method facilitated the analysis of approximately 150 rapeseed samples, enhancing experimental efficiency.

The spatial distribution of  $\delta^2\text{H}$  values of FAs in Hesse exceeded the instrumental precision and exhibited a spatial pattern similar to the  $\delta^2\text{H}$  values of precipitation on a regional scale. Correlation analysis revealed significant associations between  $\delta^2\text{H}$  of FAs and location-specific factors such as air temperatures in February and May, precipitation amounts in April, and silt content. Climatic conditions exerted a more pronounced influence on  $\delta^2\text{H}$  values of FAs compared to soil properties, indicating the potential of  $\delta^2\text{H}$  of FAs as markers for geographical traceability.

---

Classification models using OPLS-DA were developed to differentiate rapeseed from Hesse and Jianhu. The models achieved notable accuracy, with 82.8% of Hesse samples correctly identified and 93.1% of Jianhu samples accurately classified.  $\delta^2\text{H}$  of FAs and elemental data significantly contributed to the classification.

For differentiating rapeseed within Hesse, the state was divided into eight climatic zones based on three vital climatic conditions—monthly average temperature in February and May, and monthly average precipitation in April—that most influence the  $\delta^2\text{H}$  of FAs, as identified through correlation analysis. Effective discrimination was achieved between two zones by combining  $\delta^2\text{H}$  values of FAs and elemental data using the OPLS-DA method: one zone exhibited low May temperatures and low April precipitation, while the other had high May temperatures and high April precipitation. The classification model was validated, with a  $Q^2$  value of 0.8, indicating good predictive ability. Both  $\delta^2\text{H}$  of FAs and elemental data contributed significantly to the classification. Due to the limited sample size in these zones ( $n=8$  and  $n=7$ , respectively), further verification with external datasets is recommended.

In conclusion, integrating stable hydrogen isotopic composition, elemental data, and OPLS-DA provides a robust approach for determining the origin of rapeseed, both over large distances (e.g., between Germany and China) and within closely situated regions (e.g., within Hesse, Germany), particularly where distinct climatic conditions exist. This approach is more accurate and predictive than using either isotopic or elemental data alone. By combining established soil-related markers with the new analytical markers, such as  $\delta^2\text{H}$  of FAs, this fingerprinting approach could be adapted for other vegetable oils, such as palm oil. This fingerprinting approach has the potential to be a practical tool for monitoring the origin of agricultural commodities, in line with EU sustainability directives and regulations, such as the Regulation on Deforestation-Free Products.

---

## Zusammenfassung

---

Pflanzenöle spielen eine entscheidende Rolle nicht nur als Nahrungsquelle, sondern auch als erneuerbare Energie. Die Rückverfolgbarkeit ihrer geografischen Herkunft ist von großer Bedeutung für Lebensmittelsicherheit, Gesundheit und Energie-Nachhaltigkeit. Die stabile Isotopenanalyse von leichten Elementen in Kombination mit chemometrischen Techniken hat erfolgreich die geografische Herkunft von Pflanzenölen bestimmt. Allerdings bleibt das Potenzial der stabilen Wasserstoff-Isotopenzusammensetzung ( $\delta^2\text{H}$ ) von Fettsäuren (FAs) als analytische Marker und der Fingerprinting-Ansatz unter Verwendung des  $\delta^2\text{H}$  von FAs zur Identifizierung der geografischen Herkunft, insbesondere im Fall von Rapsöl, dem drittgrößten Pflanzenöl weltweit, unerforscht.

Diese Dissertation hatte zum Ziel, einen Fingerabdruckansatz zur Bestimmung des geografischen Ursprungs von Pflanzenölen zu entwickeln, wobei Rapsöl als Fallstudie verwendet wurde. Die Forschung verfolgte drei Hauptziele: (1) die Optimierung der Probenvorbereitungsmethode für eine einfache und hochdurchsatzfähige Messung des  $\delta^2\text{H}$  von FAs in Raps; (2) die Analyse der räumlichen Verteilung des  $\delta^2\text{H}$  von FAs in Raps aus verschiedenen Regionen und deren Korrelation mit klimatischen und bodenbezogenen Faktoren; (3) die Etablierung von Klassifikationsmodellen unter Verwendung des  $\delta^2\text{H}$  von FAs und der Elementzusammensetzung von Raps zur Unterscheidung geografischer Ursprünge, einschließlich der Unterscheidung von Hessen von anderen Regionen und verschiedenen Regionen innerhalb Hessens.

Eine verfügbare schnelle und einfache Probenvorbereitungsmethode von Garcés & Mancha (1993) ermöglicht die direkte Umwandlung von Raps in Fettsäuremethylester (FAME). Für eine effektive Umwandlung wurde 2,2-Dimethoxypropan (DMP) zugesetzt, doch war der Einfluss von DMP auf die Isotopenzusammensetzung der einzelnen FAMEs unklar. Daher wurden Kontrollversuche durchgeführt, um die Eignung von DMP für eine genaue verbindungspezifische Analyse stabiler Isotope (CSIA) mittels Gaschromatographie-Pyrolyse-Isotopenverhältnis-Massenspektrometrie (GC-Py-IRMS) zu bewerten. Für die zweiten und dritten Ziele wurden insgesamt 121 Raps- und Bodenproben aus Hessen, Deutschland, zwischen 2017 und 2020 sowie 28 Rapsproben aus dem Kreis Jianhu, Provinz Jiangsu, China, im Jahr 2019 gesammelt. Das  $\delta^2\text{H}$  der FAs (C18:1, C18:2, C18:3, C16:0) in Raps wurde mittels GC-Py-IRMS bestimmt, während die Elementzusammensetzung von Raps mittels induktiv gekoppelter Plasmaoptischer Emissionsspektrometrie (ICP-OES) und induktiv gekoppelter Plasmamassenspektrometrie (ICP-MS) analysiert wurde. Klimadaten und Bodenparameter wurden aus öffentlich zugänglichen Datensätzen oder mittels standardisierter Methoden im Labor ermittelt. Chemometrische Techniken wurden für die Datenanalyse verwendet, einschließlich schrittweiser linearer Regressionsanalyse, Hauptkomponentenanalyse (PCA) und orthogonaler partieller kleinster Quadrate-Diskriminanzanalyse (OPLS-DA).

Durch Kontrollexperimente wurde bestätigt, dass DMP keinen signifikanten Einfluss auf die Genauigkeit der  $\delta^2\text{H}$ -Werte der FAs hatte. Die Messgenauigkeit der CSIA von Wasserstoff (H-CSIA) reichte von  $\pm 0,3$  mUr bis  $\pm 3$  mUr und lieferte zufriedenstellende Ergebnisse. Diese Methode ermöglichte die Analyse von etwa 150 Rapsproben und steigerte somit die Effizienz der

---

Experimente.

Die räumliche Verteilung der  $\delta^2\text{H}$ -Werte der FAs in Hessen übertraf die instrumentelle Präzision und zeigte ein räumliches Muster, das den  $\delta^2\text{H}$ -Werten der Niederschläge auf regionaler Ebene ähnelt. Korrelationsanalysen offenbarten signifikante Zusammenhänge zwischen dem  $\delta^2\text{H}$  der FAs und ortsspezifischen Faktoren wie Lufttemperaturen im Februar und Mai, Niederschlagsmengen im April und Schluffgehalt. Klimatische Bedingungen hatten einen ausgeprägteren Einfluss auf das  $\delta^2\text{H}$  der FAs als Bodeneigenschaften, was das Potenzial des  $\delta^2\text{H}$  der FAs als Marker zur Bestimmung der geografischen Herkunft von Raps aufzeigt.

Klassifikationsmodelle auf Basis von OPLS-DA wurden entwickelt, um Raps aus Hessen und Jianhu zu unterscheiden. Die Modelle erzielten bemerkenswerte Genauigkeiten: 82,8 % der hessischen Proben wurden korrekt als Hessen identifiziert, und 93,1 % der Jianhu-Proben wurden genau als Jianhu klassifiziert. Sowohl das  $\delta^2\text{H}$  der FAs als auch die Elementdaten trugen signifikant zur Klassifikation bei.

Zur Differenzierung von Raps innerhalb Hessens wurde das Bundesland in acht Klimazonen unterteilt, basierend auf drei wesentlichen klimatischen Bedingungen – der monatlichen Durchschnittstemperatur im Februar und Mai sowie dem monatlichen Durchschnittsniederschlag im April – die den größten Einfluss auf die  $\delta^2\text{H}$ -Werte der FAs haben, wie durch die Korrelationsanalyse festgestellt wurde. Eine effektive Unterscheidung von Raps wurde zwischen zwei Zonen durch Kombination der  $\delta^2\text{H}$ -Werte der FAs und Elementdaten mittels der OPLS-DA-Methode erreicht: Eine Zone wies niedrige Mai-Temperaturen und niedrige April-Niederschläge auf, während die andere hohe Mai-Temperaturen und hohe April-Niederschläge hatte. Das Klassifikationsmodell wurde validiert, mit einem  $Q_2$ -Wert von 0,8, was auf eine gute Vorhersagefähigkeit hinweist. Sowohl das  $\delta^2\text{H}$  der FAs als auch die Elementdaten trugen signifikant zur Klassifikation bei. Aufgrund der begrenzten Stichprobengröße in diesen Zonen ( $n=8$  und  $n=7$ ) wird eine weitere Überprüfung mittels externer Datensätze empfohlen.

Zusammenfassend bietet die Integration der stabilen Wasserstoffisotopenzusammensetzung, der Elementdaten und der OPLS-DA-Methode einen robusten Ansatz zur Bestimmung der Herkunft von Raps, sowohl über große Entfernungen (z. B. zwischen Deutschland und China) als auch innerhalb nahegelegener Regionen (z. B. innerhalb Hessens), insbesondere dort, wo unterschiedliche klimatische Bedingungen vorherrschen. Dieser Ansatz ist genauer und vorhersagestärker als die alleinige Verwendung von Isotopen- oder Elementdaten. Durch die Kombination etablierter bodenbezogener Marker mit den neuen analytischen Markern wie dem  $\delta^2\text{H}$  der FAs könnte dieser Fingerabdruckansatz auch auf andere Pflanzenöle wie Palmöl angepasst werden. Dieser Ansatz hat das Potenzial, ein praktisches Werkzeug zur Überwachung des Ursprungs landwirtschaftlicher Erzeugnisse im Einklang mit den EU-Nachhaltigkeitsrichtlinien und -verordnungen, wie der Verordnung über entwaldungsfreie Produkte, zu sein.



---

---

## Table of Contents

---

Acknowledgments (Danksagung).....	I
Abstract .....	III
Zusammenfassung .....	V
Table of Contents.....	VII
List of Figures .....	X
List of Tables .....	XII
List of Abbreviations.....	XIV
1..... Introduction .....	1
1.1. Background .....	1
1.2. Aims and objectives .....	3
1.3. Structure of the dissertation .....	4
2..... Basics of geographic traceability of vegetable oils.....	5
2.1. Stable isotope analysis .....	5
2.1.1. Definition of stable isotopes .....	5
2.1.2. Stable isotope fractionation .....	6
2.1.3. Isotope ratio and delta notation .....	7
2.2. Rapeseed .....	8
2.2.1. Growth stages of rapeseed .....	8
2.2.2. FA biosynthesis.....	9
2.2.3. Importance of rapeseed oil.....	12
2.3. Analytical approach for geographic traceability of vegetable oil .....	13
2.4. Hydrogen isotopic fractionation of FAs in plants.....	15
2.4.1. Variation in $\delta^2\text{H}$ values of precipitation .....	16
2.4.2. Variation in $\delta^2\text{H}$ values of soil water .....	17
2.4.3. Variation in $\delta^2\text{H}$ values in FA biosynthesis of plant.....	18
2.5. Sample preparation methods for H-CSIA of FAs in vegetable oil .....	19
2.6. Research objectives based on the limitations of the available analytical approach .....	19
3..... Materials and methods .....	22
3.1. Overview of methods based on the research objectives .....	22
3.2. Samples.....	23
3.2.1. Sampling location.....	23
3.2.2. Sampling strategy .....	24

3.3.	Compound-specific stable hydrogen isotope analyses.....	28
3.3.1.	Workflow of GC-Py-IRMS.....	28
3.3.2.	Correction of the $\delta^2\text{H}$ of FAs from FAMES.....	33
3.4.	Elemental analysis.....	34
3.5.	Soil analysis.....	35
3.6.	Acquisition of climate and soil data of sampling sites.....	36
3.7.	Statistical analysis.....	37
4.....	Assessment of the one-step sample preparation method for C- and H-CSIA of FAs in rapeseed oil.....	40
4.1.	Introduction.....	40
4.2.	Basics of the one-step sample preparation method developed by Garcés and Mancha.....	40
4.3.	Experiment.....	42
4.4.	Results and discussion.....	43
4.4.1.	Impact of DMP on transmethylation efficiency of the one-step method.....	43
4.4.2.	Impact of DMP on $\delta^{13}\text{C}$ and $\delta^2\text{H}$ values of FAMES using the one-step method.....	46
4.4.3.	Estimation of $\delta^{13}\text{C}_{\text{FAME}}$ and $\delta^2\text{H}_{\text{FAME}}$ values from $\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}}$ and $\delta^2\text{H}_{\text{H}_3\text{CO-DMP}}$ values with the assumptions.....	49
4.5.	Conclusion.....	50
5.....	Evaluation of the potential of stable hydrogen isotopic composition of FAs for geographical traceability of rapeseed.....	51
5.1.	Introduction.....	51
5.2.	Temporal and spatial variation of the $\delta^2\text{H}$ values of FAs of rapeseed.....	51
5.3.	Effect of rapeseed varieties on $\delta^2\text{H}$ values of FAs.....	57
5.4.	Identifying significant environmental factors affecting the $\delta^2\text{H}$ values of FAs in rapeseed....	59
5.5.	Justifying the correlation between the environmental factors and $\delta^2\text{H}$ values of FAs.....	64
5.6.	Conclusion.....	70
6.....	Development of a fingerprinting approach to determine the geographical origin of rapeseed using H-CSIA of FAs and elemental composition combined with chemometrics.....	72
6.1.	Introduction.....	72
6.2.	Descriptive statistics of rapeseed from JianHu and Hesse.....	72
6.3.	Classification of the geographical origin of rapeseed samples from Hesse and non-Hesse .	74
6.4.	Classification of the geographical origin of rapeseed samples in Hesse.....	81
6.4.1.	Classification based on climatic factors.....	81
6.4.2.	Classification based on the geological substrate.....	85
6.5.	Conclusion.....	88



7..... Conclusion and Outlook ..... 90

7.1. Conclusion ..... 90

7.2. Outlook ..... 92

Literature ..... i

Appendix.....xi

---

---

## List of Figures

---

Figure 1: Hydrogen isotopes in nature. ....	5
Figure 2: Structural feature of triacylglycerol. ....	11
Figure 3: Cycle of fatty acid synthesis in plastids of plant cells.....	11
Figure 4: Composition of major FAs in regular and modified rapeseed oil. ....	12
Figure 5: Simplified diagram showing water fractionation within a cloud through Rayleigh distillation. ....	16
Figure 6: A procedure flow chart for analyzing 2. and 3. Research Objectives (RO). ....	23
Figure 7: Study area: a. Hesse in Germany; b. Topography (elevation between 100 and 950 m a.s.l); c. Sampling sites in Jianhu, China. ....	24
Figure 8: Sampling sites in Hesse from 2017 to 2020. ....	26
Figure 9: a. Schematic diagram of determining soil sampling points; b. Soil sampler. ....	27
Figure 10: Schematic of the GC-C/Py-IRMS system applied in the laboratory. ....	29
Figure 11: Scheme of High Temperature Conversion in the pyrolysis reactor. ....	30
Figure 12: Schematic of the ion source. ....	30
Figure 13: Monitoring the daily H <sub>3</sub> -factors. ....	32
Figure 14: Linearity range for determining the $\delta^2\text{H}$ values of the four FAMEs (C18:1, C18:2, C18:3, and C16:0 FAMEs) with the GC-Py-IRMS system.....	33
Figure 15: a. Structure of the phthalic acid; b. Derivatized phthalic acid, R <sub>1</sub> and R <sub>2</sub> here are acetyl groups.....	34
Figure 16: Flow chart of the one-step sample preparation method. ....	40
Figure 17: Four major reactions during the one-step sample preparation process. ....	41
Figure 18: The weight of four major FAMEs (C16:0, C18:1, C18:2, and C18:3) in 100 g of the rapeseed sample prepared using three transmethylation mixtures.....	44
Figure 19: The $\delta^{13}\text{C}$ (left) and $\delta^2\text{H}$ (right) of C16:0, C18:1, C18:2, and C18:3 FAME in rapeseed using transmethylation mixtures TX-D and TM+D. ....	48
Figure 20: The $\delta^2\text{H}$ values of four FAs of Hessian rapeseed in 2019 and 2020.....	53
Figure 21: Significant correlation of $\delta^2\text{H}$ values of different FAs with latitude and altitude for rapeseed 2019.....	54
Figure 22: Significant correlation of $\delta^2\text{H}$ values of different FAs with latitude and altitude for rapeseed 2020.....	56
Figure 23: Precipitation of sampling sites during the rapeseed growth from 2017 to 2019. ....	60
Figure 24: Precipitation amount variation in April across different years for sampling locations.....	61
Figure 25: Diagnostic plots for the linear regression model of $\delta^2\text{H}$ of C18:1 FA.....	64
Figure 26: Significant correlation between the air temperature in February and $\delta^2\text{H}$ values of C18:1 FA.....	65

---

Figure 27: Correlation of precipitation amount and surface soil moisture in April of sampling sites. .....	66
Figure 28: Significant correlation between the precipitation amount in April and $\delta^2\text{H}$ values of C18:1 FA.....	67
Figure 29: Significant correlation between the air temperature in May and $\delta^2\text{H}$ values of C18:1 FA. .....	67
Figure 30: Significant correlation between the silt content and $\delta^2\text{H}$ values of C18:1 FA. ....	68
Figure 31: Difference in $\delta^2\text{H}$ values between precipitation as source water and FAs of two samples with various soil silt content.....	69
Figure 32: Correlation of air temperature and soil temperature in February of sampling sites. ....	69
Figure 33: Score plots of the three OPLS-DA models and the corresponding permutation plots. ...	76
Figure 34: Clustering of Hesse in 8 climatic zones. ....	82
Figure 35: Samples (blue circles) located in climatic zones 0 and 6+7 .....	83
Figure 36: Score plot of Model 4 and the corresponding permutation plots. ....	84
Figure 37: Score plot of Model 5 and the corresponding permutation plots. ....	84
Figure 38: Spatial Distribution of zechstein (marked in green) and vulcanite in Hesse (marked in blue) and samples (brown circle) from these two substrates. ....	86
Figure 39: Score plot of Model 6 and the corresponding permutation plot. ....	87
Figure 40: Score plot of Model 7 and the corresponding permutation plot. ....	88

---

---

## List of Tables

---

Table 1: Natural isotopic abundances of light elements and their masses in Dalton.....	6
Table 2: Characteristic physical properties of H <sub>2</sub> O. ....	6
Table 3: International isotope reference materials of light elements and their isotopic abundance ratios.....	8
Table 4: Phenological growth stages of oilseed rape in Germany with BBCH code and months. ...	10
Table 5: Overview of studies using stable isotope analysis combined with chemometrics to assess the authenticity of edible oils in recent years (2010–2020). ....	15
Table 6: Stable hydrogen isotopic composition of USGS70 and USGS71.....	31
Table 7: Program of the microwave oven for digestion of rapeseeds. ....	35
Table 8: List of the publicly available datasets. ....	37
Table 9: Interpretation of correlation coefficients. ....	38
Table 10: Composition of transmethylated mixtures for every 20 mg of oil or 50 mg of seeds. ....	43
Table 11: FA composition (mol%) of a rapeseed sample using three transmethylated mixtures... ..	44
Table 12: Molar ratios of each component to oil and FAME yield in the transmethylated mixtures. ....	45
Table 13: Precision of the C- and H-CSIA by analyzing two in-house reference materials by GC-C/Py-IRMS.....	46
Table 14: Comparison of reproducibility of $\delta^{13}\text{C}$ and $\delta^2\text{H}$ values of four FAMES of oil and seeds ..	47
Table 15: The geographic range of the sampling locations for 2019 and 2020. ....	52
Table 16: The mean, minimum, and maximum $\delta^2\text{H}$ values of FAs in different parts of Hesse from 2019 to 2020.....	52
Table 17: Comparison of the $\delta^2\text{H}$ values of FAs among different rapeseed varieties using the ANOVA test.....	58
Table 18: Three types of environmental factors for the stepwise regression analysis.....	59
Table 19: Significant factors affecting stable hydrogen isotopic composition of four FAs using stepwise regression analysis. ....	63
Table 20: Average, minimum, and maximum values of $\delta^2\text{H}$ values of FAs and elements in rapeseeds of 2019 from Hesse and Jianhu. ....	73
Table 21: 16 predictive variables related to the $\delta^2\text{H}$ values of FAs. ....	74
Table 22: Overview of the datasets of three models (Model 1-3).....	75
Table 23: Performance parameter of the three models. ....	78
Table 24: VIP values of variables in three OPLS-DA models (Model 1-3). ....	79
Table 25: Description of the legend for the eight distinct areas in Hesse.....	82
Table 26: Overview of Model 4-5 datasets and their model performance parameters. ....	83
Table 27: VIP values of variables in Model 4-5. ....	85



Table 28: Overview of Model 6-7 datasets and their model performance parameters. .... 87

Table 29: Performance parameter of Model 6-7..... 87

Table 30: Important variables in Models 6-7..... 87

---

## List of Abbreviations

---

acetyl-CoA	Acetyl coenzyme A
ACP	Acyl Carrier Protein
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BBCH	Biologische Bundesanstalt, Bundessortenamt, and Chemical Industry
BSIA	Bulk Stable Isotope Analysis
C-CSIA	Compound-specific Stable Isotope Analysis of Carbon
CSIA	Compound-specific Stable Isotope Analysis
DMP	2,2-Dimethoxypropan
EA	Elemental Analyzer
ER	Endoplasmic Reticulum
EU	European Union
EUDR	Regulation on Deforestation-Free Products
FA(s)	Fatty Acid(s)
FAME	Fatty Acids Methyl Ester
FFAs	Free Fatty Acids
FID	Flame Ionization Detector
GC-C/Py- IRMS	Gas Chromatograph-Combustion Reactor/Pyrolysis-Isotope Ratio Mass Spectrometer
GLCO	Canola Oil with Gamma Linolenic Acid
H-CSIA	Compound-specific stable isotope analysis of hydrogen
HLUNG	Hessisches Landesamt für Umwelt und Geologie
HOCO	High Oleic Acid Canola Oil
HOLLCO	High Oleic Low Linolenic Canola Frying Oil
HSCO	Canola Oil with High Content of Stearic Acid
HTC	High-Temperature Conversion
IAEA	The United Nations International Atomic Energy Agency
ICP-AES	Inductively Coupled Plasma Atomic Emission Spectroscopy
ICP-MS	Inductively Coupled Plasma-Mass Spectroscopy
LDA	Linear discriminant analysis
LICO	Low Linolenic Acid Canola Oil
LLH	Landesbetrieb Landwirtschaft Hessen
NADPH	Nicotinamide Adenine Dinucleotide Phosphate Hydrogen
NADP <sup>+</sup>	Nicotinamide Adenine Dinucleotide Phosphate
NIST	National Institute of Standards and Technology
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PCA	Principal Component Analysis
PGI	Protected Geographical Indication
PLS	Partial least squares
PPS	Probability Proportional to Size



---

PSD	Particle-Size Distribution
RS	Random Start
SD	Standard Deviation
SI	Sampling Interval
SLAP	Standard Light Antarctic Precipitation
TAGs	Triacylglycerols
VCDT	Vienna Canyon Diablo Troilite
VPDB	Vienna Pee Dee Belemnite
VSMOW	Vienna Standard Mean Ocean Water

---

## 1. Introduction

---

### 1.1. Background

Vegetable oils are vital components in agriculture, serving as both a source of nutrition for human consumption and raw materials for various industries, including cosmetics and lubricants. With the rise of renewable energy sources, vegetable oils, such as rapeseed and palm oil, are increasingly used in the production of biofuels. The globalization of commodity trade has emphasized the importance of accurately determining the geographical origin of vegetable oils (Dyer et al., 2008).

The geographical origin of edible oils significantly influences their quality, safety, and taste, providing crucial information for consumers. Products from specific European regions, protected by Protected Geographical Indications (PGI) labeling, exemplify this, as seen in Terra d'Otranto extra-virgin olive oil from Taranto, Italy. Items with PGI labeling or from a specific geographical origin often command higher prices, rendering them susceptible to potential fraud and mislabeling (Camin et al., 2017; Danezis et al., 2016). For instance, Spanish olive oil may be falsely marketed as originating from Italy, or sesame oil grown in India may be erroneously labeled as coming from South Korea (Jeon et al., 2015).

Moreover, the geographical origin of vegetable oils plays a crucial role in determining their sustainability. The rapid expansion of oilseed plant cultivation, driven by the versatile applications of vegetable oils, has raised concerns about illegal land use changes (Fridrihsone et al., 2020; Meijaard et al., 2020). Specifically, certain oils have been grown on farmland obtained through illicit logging of tropical forests, with notable instances in Indonesia, Malaysia, and Brazil (Meijaard et al., 2020). The consequences of such practices have sparked environmental concerns, including loss of biodiversity and a substantial increase in carbon emissions (Friedlingstein et al., 2010; Lapola et al., 2010; Plevin et al., 2010). While measures and regulations, such as certification processes for "sustainable" cultivation products, exist to protect ecologically sensitive areas, they are not foolproof (Sumfleth et al., 2020). For instance, oil palms grown on plantations built on former rainforests rich in carbon are still mislabeled as "sustainable" and exported to other countries (Cazzolla Gatti & Velichevskaya, 2020). To address these issues, the Regulation on Deforestation-Free Products (EUDR) will soon come into force, aiming to minimize the European Union's contribution to global deforestation. One of the objectives of the EUDR is to require companies to provide evidence that their products are not linked to deforestation, thereby increasing transparency and accountability in the supply chain. Therefore, accurately determining the geographical origin of vegetable oils could yield significant economic, social, and environmental benefits.

In recent years, advanced analytical approaches, such as profiling or fingerprinting analysis, coupled with chemometrics, have shown promise in the geographic traceability of vegetable oils and in deterring fraudulent practices (Ballin & Laursen, 2019a). Chemometrics is "the chemical discipline which uses mathematical and statistical methods to achieve the aim of analytical chemistry, namely the obtention in the optimal way of relevant information about material systems" (Kowalski, 1981). Analytical markers, quantitatively determined using modern analytical techniques like mass spectrometry, spectroscopy, and separation methods, are predefined analytical targets linked

---

directly or indirectly to the geographic origin. Examples include chemical compounds, organic compositions, mineral compositions, and stable isotopic compositions of elements. A single marker alone provides limited information and is inadequate for ensuring geographic traceability. Instead, a profiling analysis that quantifies multiple markers can offer a more comprehensive picture and enhance classification accuracy. The nomenclature “fingerprinting analysis” often refers to profiling analysis. Chemometrics evaluates data measured with analytical techniques to establish links to geographic origins (Ballin & Laursen, 2019b; Luykx & van Ruth, 2008).

The stable isotopic composition of light elements, such as carbon, hydrogen, oxygen, nitrogen, and sulfur, is increasingly recognized as a valuable analytical marker for determining the geographical origin of vegetable oils (Camin et al., 2017; Kelly et al., 2018). The stable isotopic composition of these elements reflects the growth environment, including factors like the photosynthetic pathway, latitude, altitude, air temperature, air drought, precipitation, and nitrogen fertilizer source (Camin, Larcher, Nicolini, et al., 2010; Richter et al., 2010; Inácio et al., 2015).

For example, the hydrogen in vegetable oils primarily comes from the hydrogen in the water available to the plants. The stable hydrogen isotopic composition (reported as  $\delta^2\text{H}$  value) in soil water varies due to the source, amount, and movement of the water (Hu et al., 2008; Vargas et al., 2017). A major source of soil water is local precipitation, whose  $\delta^2\text{H}$  value depends on the latitude, longitude, and altitude of the site (Ling et al., 2022). Consequently, the  $\delta^2\text{H}$  of plant-available water in the soil is influenced by the geographical location, local climate, and the physical-chemical properties of the soil that regulate the water movement, such as evaporation and vertical seepage. Suppose the  $\delta^2\text{H}$  of fatty acids (FAs) in vegetable oils correlates well with the  $\delta^2\text{H}$  of plant-available water in the soil, which in turn is influenced by the local climate and soil properties at the site to a certain extent. In this case, the  $\delta^2\text{H}$  can serve as an analytical marker to identify the geographical origin of vegetable oils (Ehtesham et al., 2013). Despite the potential importance of  $\delta^2\text{H}$  of FAs as an analytical marker, its feasibility for determining the geographical origin of vegetable oils requires further examination.

One major gap is the lack of simple and quick analytical methods. Determining  $\delta^2\text{H}$  of FAs involves a complex process that requires converting oil or oilseed samples into a form suitable for instrumental measurement, a time-consuming and labor-intensive task, particularly when dealing with a large number of samples for chemometrics (Paolini et al., 2017). More importantly, the extent of the  $\delta^2\text{H}$  fractionation during the required chemical conversion remains undetermined.

Secondly, the  $\delta^2\text{H}$  of FAs is influenced by multiple factors, including climatic conditions and soil properties in the soil-plant-atmosphere continuum (Luykx & van Ruth, 2008; Camin, Larcher, Nicolini, et al., 2010; Sachse et al., 2012; Eley et al., 2018). However, it remains unclear which of these factors has the most significant impact on  $\delta^2\text{H}$  of FAs, leading to uncertainty in the discriminatory power of  $\delta^2\text{H}$  of FAs in identifying the geographical origin of vegetable oils.

Moreover, the fingerprinting approach that combines the  $\delta^2\text{H}$  of FAs and chemometrics is rarely applied in determining the geographical origin of oilseed plants. The contribution of the stable hydrogen isotopic composition to identifying the geographical origin is unclear and rarely discussed.

---

## 1.2. Aims and objectives

This dissertation aims to develop a fingerprinting approach to determining the geographical origin of vegetable oils, with a specific focus on rapeseed. Rapeseed is chosen for this study due to the limited research on its geographical determination, despite its extensive use as a food ingredient and a raw material for biodiesel production. The research relies on substantial and reliable data, which is facilitated by our collaboration with Landesbetrieb Landwirtschaft Hessen (LLH). LLH provides a comprehensive collection of rapeseed and soil samples with precise sampling locations, essential for obtaining accurate climatic data from publicly available datasets. Should this fingerprinting approach prove successful for rapeseed, it holds potential for application to other significant agricultural commodities mentioned in EU sustainability directives and regulations, such as soybeans and palm oil.

The first objective of this dissertation is to streamline the laboratory work for the compound-specific stable hydrogen isotope analysis (H-CSIA) of rapeseed and its oil. A straightforward and rapid one-step sample preparation method is applied to convert rapeseed into fatty acid methyl esters (FAMES) for instrumental analysis. The feasibility of this one-step method for stable isotope analysis will be thoroughly evaluated.

The second objective is to assess the potential of the  $\delta^2\text{H}$  of FAs as analytical markers for identifying the geographical origin of rapeseed. It involves studying the spatial and temporal variation of  $\delta^2\text{H}$  values of FAs in Hesse and exploring its correlation with selected soil properties and climatic conditions.

The third objective is to determine the geographical origin of rapeseed using  $\delta^2\text{H}$  of FAs and chemometric techniques, such as principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA). Discriminative models will be developed to classify rapeseed samples between Hesse and non-Hesse, where the non-Hesse samples are from Jianhu, China. Additionally, the elemental composition of rapeseed will be utilized to improve classification accuracy.

The following specific objectives are formulated for this dissertation:

1. Evaluate the feasibility of a rapid one-step sample preparation method that converts rapeseed into FAMES for subsequent H-CSIA.
2. Assess the feasibility of  $\delta^2\text{H}$  of FAs as analytical markers for identifying the geographical origin of rapeseed.
  - Investigate spatial and temporal variations in the  $\delta^2\text{H}$  of FAs of rapeseed from Hesse.
  - Explore the correlation of soil and climatic factors with the  $\delta^2\text{H}$  of FAs in rapeseeds.
3. Determine the geographical origin using  $\delta^2\text{H}$  of FAs and elemental composition combined with chemometrics.
  - Distinguish rapeseed samples from Hesse and non-Hesse
  - Distinguish rapeseed samples within Hesse

This study represents the first attempt to apply the  $\delta^2\text{H}$  of FAs as analytical markers for investigating the geographical origin of rapeseed. The research findings are anticipated to significantly contribute to the field by enhancing the accuracy of origin determination, with potential implications for improving the traceability of other vegetable oils.

---

### 1.3. Structure of the dissertation

In the following Chapter, the basics of geographic traceability of vegetable oils are described. The Chapter starts with a brief explanation of the stable isotope ratio of elements. The existing studies on stable isotope analysis employed in the geographic traceability of rapeseed and other vegetable oils are summarized. The physical and biochemical processes that may influence the  $\delta^2\text{H}$  values of FAs in plants and the corresponding environmental factors are outlined. Additionally, a comprehensive overview of the existing sample preparation methods for converting oilseeds into FAMES, specifically applied in the context of geographical traceability of vegetable oil, is presented. The chapter also addresses the research objectives tailored to overcome the limitations of current analytical markers for geographical determination.

Chapter 3 provides a detailed description of the materials and experimental design employed in this dissertation, aligning with the research objectives. The strategy for collecting rapeseed and soil samples is introduced, followed by an explanation of the measurement methods utilized for various chemical and physical parameters of the samples. Protocols for data evaluation, such as the normalization method for isotope analyses and the determination of the isotopic linearity range ( $H_3$ -factor) for hydrogen, are presented. The datasets used for extracting the required climatic and soil data are listed. Moreover, the statistical tools employed for different objectives are described.

The results of the individual research questions are summarized and discussed in Chapters 4, 5, and 6.

In Chapter 4, the feasibility of a one-step sample preparation method to convert oilseeds to FAMES for CSIA is evaluated. In particular, the impact of 2,2-dimethoxypropane (DMP) used in this method on isotopic ratios of carbon and hydrogen is investigated.

Chapter 5 focuses on determining the potential of  $\delta^2\text{H}$  of FAs in identifying the geographical origin of rapeseed. The analysis involves comparing the  $\delta^2\text{H}$  values of FAs in rapeseed collected from Hesse, Germany, between 2017 and 2020. The annual variation in  $\delta^2\text{H}$  values of FAs in Hesse is also analyzed for 2019 and 2020. Environmental conditions, including soil properties and climate factors, that primarily drive or correlate with the variation of  $\delta^2\text{H}$  values of FAs are identified and justified based on available literature.

Chapter 6 involves the construction of discrimination models using  $\delta^2\text{H}$  of FAs and elemental composition, combined with chemometrics, to distinguish rapeseed samples between Hesse and Jianhu. The discrimination power of these models is then tested using an external dataset. In addition, the geographical origin of rapeseed within Hesse is also identified. The samples are initially classified based on the climatic zones and geographical substrates of the sampling sites. Models for distinguishing the samples based on these classifications are constructed and presented with their performances. Further, the contribution of  $\delta^2\text{H}$  of FAs in different geographic ranges is compared and discussed.

Finally, Chapter 7 concludes the dissertation by summarizing the contributions and limitations, providing an outlook for future research.

---

## 2. Basics of geographic traceability of vegetable oils

---

### 2.1. Stable isotope analysis

#### 2.1.1. Definition of stable isotopes

Isotopes refer to atoms whose nuclei have the same number of protons but different numbers of neutrons (Hoefs, 2018). Stable isotopes, resistant to spontaneous disintegration by any known decay mode, differ from unstable isotopes (radioactive isotopes), which emit radiation (Clark & Fritz, 1997). It is conventionally denoted as  ${}^m\text{E}$ , where E is the elemental symbol, and the superscript “m” means the atomic mass number given by the total number of protons and neutrons in the nucleus. For instance, the element hydrogen has two naturally occurring stable isotopes,  ${}^1\text{H}$  and  ${}^2\text{H}$ : the nucleus of the light hydrogen isotope ( ${}^1\text{H}$ ) consists of one proton and no neutron, whereas the heavy hydrogen isotope ( ${}^2\text{H}$ ) contains a proton and a neutron. The third isotope,  ${}^3\text{H}$ , is radioactive (Figure 1). In the subsequent sections, only stable isotopes are relevant to the dissertation.

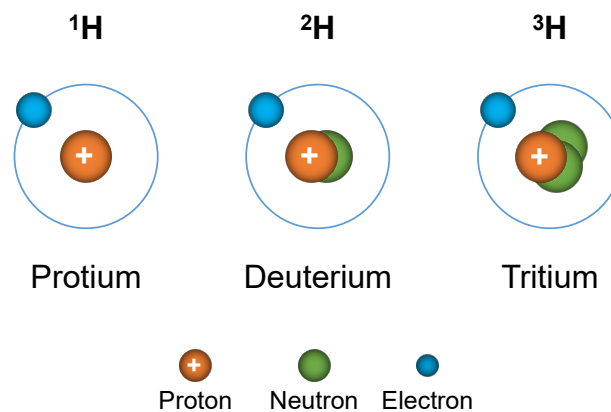


Figure 1: Hydrogen isotopes in nature.

Isotopes of an element have the same proton numbers but different atomic masses. The distinctions in atomic masses among isotopes further vary based on the specific element, as shown in Table 1. For instance, the atomic masses of  ${}^1\text{H}$  and  ${}^2\text{H}$  are 1.0078 Da (Dalton) and 2.014 Da, respectively, while  ${}^{12}\text{C}$  and  ${}^{13}\text{C}$  have masses of 12 Da and 13.003 Da. Isotopic abundances are expressed as the relative proportions (often as atom percentages) of the stable isotopes of an element occurring in nature. For example, the natural abundances of  ${}^1\text{H}$  and  ${}^2\text{H}$  on Earth are 99.985% and 0.015%, respectively (Meija et al., 2016; Sharp, 2017).

The slight differences in atomic masses of isotopes result in subtle changes in the chemical and physical properties of isotopes, a phenomenon termed Isotope effects. For instance, compounds with heavier isotopes typically have higher density, melting points, and boiling points compared to those with lighter isotopes (refer to Table 2). The chemical behavior, such as reaction rate, can also be affected when isotopes replace atoms in molecules. Lighter isotopes, due to their greater mobility and vibrational energy, tend to display higher reactivity (Hoefs, 2018). The velocity of a lighter isotope is proportional to the inverse square root of its mass, resulting in faster movement compared to a heavier isotope (Sharp, 2017). Furthermore, lighter isotopes possess higher vibrational frequencies and require less energy to dissociate from bonds, making them more likely to participate in chemical

processes. This phenomenon is known as isotope fractionation (Hoefs, 2018).

Table 1: Natural isotopic abundances of light elements and their masses in Dalton.

Element	Isotope	Mass	Mass difference	Abundance (%)
Hydrogen	<sup>1</sup> H	1.007825		99.985
	<sup>2</sup> H	2.014102	+1.006277	0.015
Carbon	<sup>12</sup> C	12.0		98.890
	<sup>13</sup> C	13.003355	+1.003355	1.110
Nitrogen	<sup>14</sup> N	14.003074		99.634
	<sup>15</sup> N	15.000109	+0.997035	0.366
Oxygen	<sup>16</sup> O	15.994915		99.762
	<sup>17</sup> O	16.999132	+1.004217	0.038
	<sup>18</sup> O	17.999161	+2.004246	0.200
Phosphor	<sup>31</sup> P	30.973762		100
Sulfur	<sup>32</sup> S	31.972071		95.020
	<sup>33</sup> S	32.971459	+0.999388	0.750
	<sup>34</sup> S	33.967867	+1.995796	4.210
	<sup>36</sup> S	35.967081	+3.995010	0.020

Note: The data in the table is from Meija et al., 2016.

Table 2: Characteristic physical properties of H<sub>2</sub>O.

Property	H <sub>2</sub> <sup>16</sup> O	D <sub>2</sub> <sup>16</sup> O	H <sub>2</sub> <sup>18</sup> O
Density (20 °C, in g cm <sup>-3</sup> )	0.997	1.1051	1.1106
Temperature of greatest density (°C)	3.98	11.24	4.30
Melting point (760 Torr, in °C)	0.00	3.81	0.28
Boiling point (760 Torr, in °C)	100.00	101.42	100.14
Vapor pressure (at 100 °C, in Torr)	760.00	721.60	
Viscosity (at 20 °C, in centipoise)	1.002	1.247	1.056

Note: The data in the table is from Hoefs, 2018.

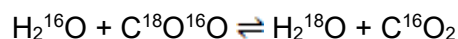
### 2.1.2. Stable isotope fractionation

Isotopic fractionation is defined as partitioning of (stable) isotopes of an element between two compounds or two phases of the same compound. Two different types of processes cause isotope fractionation: isotope exchange reactions and kinetic isotope effects (Hoefs, 2018).

Isotope exchange reactions are equilibrium fractionation processes resulting from the variation in thermodynamic properties of isotopically substituted species or compounds. The isotopes of an element among different species or compounds redistribute until the reaction rates of both forward and reverse reactions are equal. Upon reaching equilibrium, the ratios of different isotopes in each compound remain constant as long as the physicochemical conditions of the system remain unchanged (Sharp, 2017). The oxygen isotope exchange reaction between CO<sub>2</sub> and H<sub>2</sub>O is

---

presented as follows:



The  $^{16}\text{O}$  in  $\text{H}_2\text{O}$  is continuously exchanged with  $^{18}\text{O}$  in  $\text{CO}_2$  during the reaction. At equilibrium, the isotope ratio of  $^{18}\text{O}$  in  $\text{H}_2\text{O}$  and  $\text{CO}_2$  is constant. Water vapor condensation is another example of the isotope exchange process, in which the heavier isotopes ( $^{18}\text{O}$  and  $^2\text{H}$ ) become enriched in the liquid phase, while the lighter isotopes ( $^{16}\text{O}$  and  $^1\text{H}$ ) are favored in the vapor phase (Kendall & McDonnell, 2012). The magnitude of the isotope fractionation effect can be expressed by the fractionation factor  $\alpha$ , which is mainly affected by the temperature (Hoefs, 2018).

Kinetic fractionation refers to continuous variation in partitioning isotopes between phases or compounds with time and reaction progress. It relates to incomplete and unidirectional processes, such as evaporation, diffusion, and biological processes (photosynthesis, root respiration, and bacterial reduction) in nature. Kinetic fractionation is partially accounted for by the different translational velocities of isotopic forms of molecules as they traverse a phase or phase boundary. Generally, kinetic isotope effects are larger than equilibrium isotope effects (Hoefs, 2018).

Many processes in nature involve a combination of equilibrium and kinetic fractionation. For instance, while evaporation can occur under purely equilibrium conditions (e.g., at 100% humidity in still air), more commonly, the products are partially isolated from the reactants (e.g., vapor blown downwind). A variable amount of kinetic isotope fractionation affects the isotopic compositions of water and vapor under these conditions (Kendall & McDonnell, 2012).

### 2.1.3. Isotope ratio and delta notation

As mentioned above, the natural isotopic abundances of elements subtly vary in different compounds or phases during the fractionation processes. The isotope ratio (R), representing the natural abundance of the heavy isotope to the light isotope, expresses the isotopic composition of an element in a compound or a sample:

$$R = \frac{\text{natural abundance of heavy isotope}}{\text{natural abundance of light isotope}}$$

It can be preceded by a superscript representing the mass number of the isotope, such as  $^2R$  for stable hydrogen isotope ratio and  $^{13}R$  for stable carbon isotope ratio.

Absolute isotope ratios are challenging to measure and require sophisticated equipment. Moreover, comparing the difference in isotope ratios among various compounds or samples, rather than the absolute isotope ratio, is commonly significant for research. Therefore, the stable isotopic composition of an element is reported in delta ( $\delta$ ) notation, expressing the deviation of the isotope ratio of this element in samples relative to reference materials:

$$\delta^m\text{E} = (\text{R}_{\text{sample}} - \text{R}_{\text{RM}}) / \text{R}_{\text{RM}}$$

Here,  $\text{R}_{\text{sample}}$  is the isotope ratio of the sample, and  $\text{R}_{\text{RM}}$  is the isotope ratio of the reference material (McKinney et al., 1950).

In recent decades, internationally recognized reference materials for isotopes have been established to facilitate the comparison of isotopic delta values from different laboratories. These materials are calibrated, cataloged, and distributed by two organizations: The United Nations International Atomic Energy Agency (IAEA) and the National Institute of Standards and Technology



(NIST). For example, the isotopic abundance of hydrogen in Vienna Standard Mean Ocean Water (VSMOW) serves as the reference value for hydrogen isotopic composition. The international isotope reference materials for various light elements are presented in Table 3 (Clark & Fritz, 1997).  $\delta^mE$  values, relative to the corresponding isotope ratio of international isotope reference materials, are also called true  $\delta^mE$  values.

Table 3: International isotope reference materials of light elements and their isotopic abundance ratios.

Isotope	Ratio	(Primary) Reference materials	Reference (Abundance ratio)
$^2\text{H}$	$^2\text{H}/^1\text{H}$	Vienna Standard Mean Ocean Water (VSMOW)	$1.5575 \times 10^{-4}$
$^{13}\text{C}$	$^{13}\text{C}/^{12}\text{C}$	Vienna Pee Dee Belemnite (VPDB)	$1.1237 \times 10^{-2}$
$^{15}\text{N}$	$^{15}\text{N}/^{14}\text{N}$	Air- $\text{N}_2$ (atmospheric nitrogen)	$3.677 \times 10^{-3}$
$^{18}\text{O}$	$^{18}\text{O}/^{16}\text{O}$	VSMOW	$2.0052 \times 10^{-3}$
		VPDB	$2.0672 \times 10^{-3}$
$^{34}\text{S}$	$^{34}\text{S}/^{32}\text{S}$	Vienna Canyon Diablo Troilite (VCDT)	$4.5005 \times 10^{-2}$

Note: The data in the table is from Clark & Fritz, 1997.

These values are mostly small numbers and commonly expressed in units of per mille (‰). This unit is not a true unit as described in the International System of Units but is dimensionless. A new unit, “urey” (Ur), is introduced for the derived quantity isotope  $\delta^mE$  in recognition of Harold C. Urey, who received the Nobel Prize in chemistry for his discovery of deuterium (Cohen et al., 1983). The unit milliurey (mUr), equivalent to the per mil, will be used in the following chapters. The  $\delta^mE$  value can be a negative or a positive number, with a negative number indicating that the heavier isotope abundance of the sample is less than that of the reference material (Meier-Augenstein & Schimmelmann, 2019).

## 2.2. Rapeseed

### 2.2.1. Growth stages of rapeseed

Rapeseed (oilseed rape) is a type of oilseed crop grown for its seeds, and it belongs to the *Cruciferae* family, encompassing several species such as *Brassica napus*, *Brassica rapa*, *Brassica juncea*, and *Brassica carinata*. These species are grown in over thirty countries, primarily in temperate regions, with *Brassica napus* (winter type) being the predominant species in most of Europe (Richter et al., 2010).

Table 4 offers a simplified overview of the winter rape growth cycle based on the BBCH code system (BBCH: Biologische Bundesanstalt, Bundessortenamt, and Chemical Industry), which facilitates the categorization of various phenological growth stages of plants (Meier, 2018). The months corresponding to these stages are extracted from data spanning 2017 to 2020 in Germany (Deutscher Wetterdienst, 2023). In Germany, winter rape is sown towards the end of August and the beginning of September, during which the seeds absorb water and initiate sprouting. By September, the rapeseed plant starts developing its first set of leaves, marking the commencement of leaf development and the appearance of side shoots. From November to the end of February, rapeseed

---

plants experience slower growth and enter dormancy due to colder temperatures and reduced sunlight. With rising temperatures as spring approaches, rapeseed plants enter the stem elongation stage in March, during which the stems grow taller in preparation for further development. In mid to late March, the rapeseed plant initiates the formation and emergence of inflorescences, clusters of flowers marking the transition from vegetative to reproductive growth. April witnesses the flowering stage characterized by the vibrant blossoming of bright yellow flowers, a critical phase for pollination and seed production. By mid to late May, the rapeseed plant focuses on developing seeds. Flowers fade, and seed pods form, gradually filling with rapeseeds. The rapeseed seeds reach the ripening stage in mid-July, with seed pods maturing and seeds attaining optimal quality for harvesting. Toward the end of July, the rapeseed plant enters the senescence stage, with leaves and stems wilting (Cramer, 1990).








### **2.2.2. FA biosynthesis**

Rapeseed oil mainly contains triacylglycerols (TAGs), constituting 94-99% of its lipid amount. In addition, the oil also contains phospholipids, tocopherols, sterols, sulfur-containing compounds, and free FAs. TAGs, illustrated in Figure 2, consist of three FAs, usually straight-chain carboxylic acids with an even number of carbon atoms, esterified on the three hydroxyl groups of glycerol.

Fatty acid synthesis in plants occurs only in plastids, as shown in Figure 3. It begins with acetyl coenzyme A (acetyl-CoA) and CO<sub>2</sub> being converted to malonyl-CoA by acetyl-CoA carboxylase. The reaction of malonyl-CoA with an acyl carrier protein (ACP) produces malonyl-ACP, the substrate for subsequent elongation reactions. In the first fatty acid synthesis cycle, the acetate group from acetyl-CoA combines with malonyl-ACP to form acetoacetyl-ACP via 3-ketoacyl-ACP synthase. Various enzymes remove the keto group at carbon 3 of acetoacetyl-ACP in three steps, yielding butyryl-ACP with a four-carbon acyl chain. Another two-carbon unit from malonyl-ACP is then added to the four-carbon chain by a condensing enzyme. The cycle continues until a 16- or 18-carbon chain is formed, producing saturated palmitic acid (16:0) and stearic acid (18:0). Desaturation and elongation enzymes, located in the endoplasmic reticulum (ER) or plastids, can further modify these FAs by introducing double bonds or additional carbon atoms to the chain (Taiz & Zeiger, 2002). For instance, rapeseeds yield substantial amounts of oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3). Acyltransferases then incorporate these FAs into glycerolipids or TAGs, transferring acyl groups from acyl-CoA or acyl-ACP to glycerol-3-phosphate or diacylglycerol (Bates et al., 2013; Kim, 2020). The resulting TAGs are packaged into oil bodies, which are spherical structures surrounded by a phospholipid monolayer and specific proteins (Taiz & Zeiger, 2002). These oil bodies accumulate in the cytoplasm of seed cells, serving as a reservoir of energy and carbon for seed germination and seedling growth (Miklaszewska et al., 2021).



Table 4: Phenological growth stages of oilseed rape in Germany with BBCH code and months.

BBCH code		Phenological growth stage	Months in Germany
0		Germination	Late August to Early September
10		Leaf development	September
20		Development of side shoots	September
		Vegetation pauses	November to End of February
30		Stem Elongation	March
50		Inflorescence emergence	Mid-March to End of March
60		Flowering	Mid-April
70		Development of seed	Mid-May to End of May
80		Ripening	Mid-July
90		Senescence	End of July

Note: The months in Germany are based on the data from Deutscher Wetterdienst, 2023.

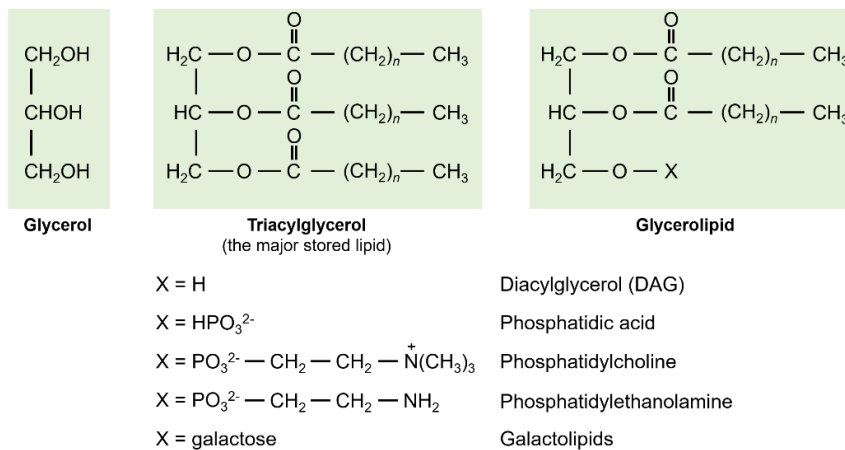


Figure 2: Structural feature of triacylglycerol.

Note: Adapted from *Plant Physiology* (Page 248), by L. Taiz and E. Zeiger, 2002, Sinauer Associates. Copyright 2003 by Annals of Botany Company. Adapted with permission.

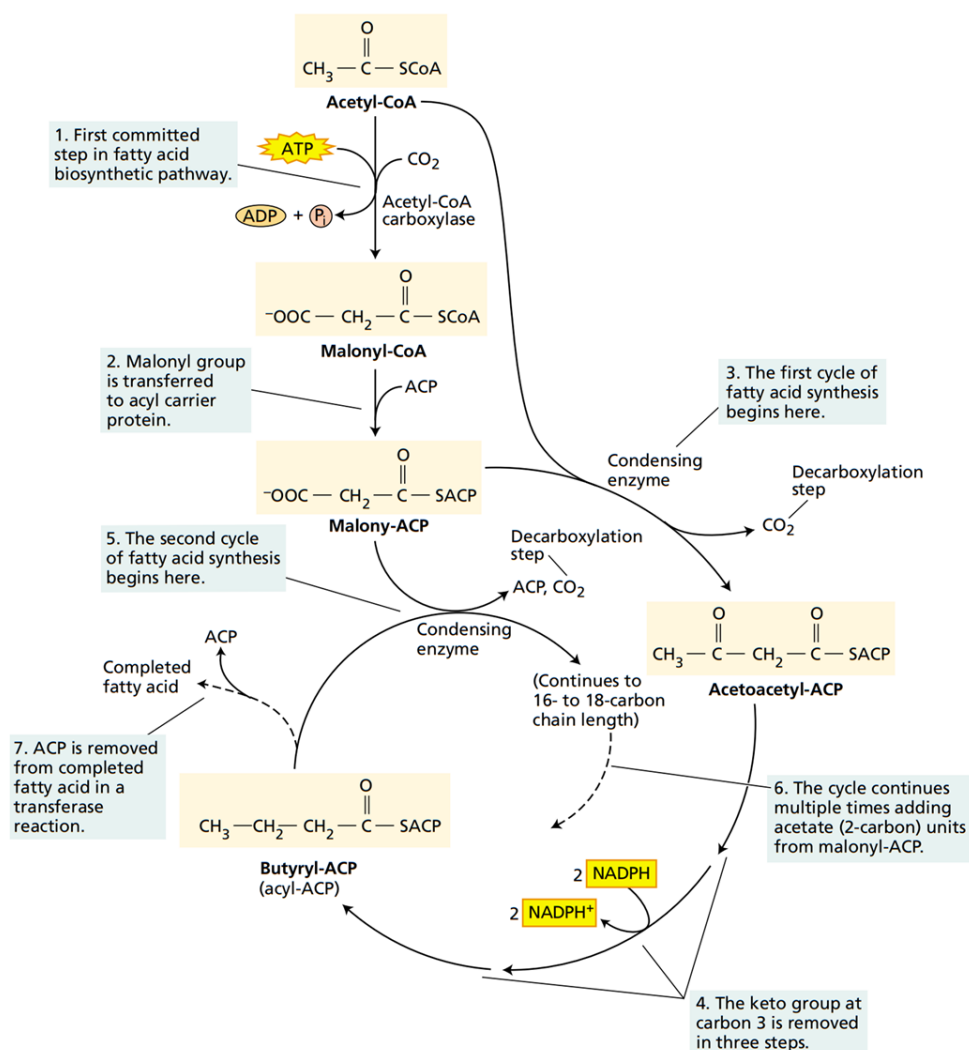


Figure 3: Cycle of fatty acid synthesis in plastids of plant cells.

Note: From *Plant Physiology* (Page 251), by L. Taiz and E. Zeiger, 2002, Sinauer Associates. Copyright 2003 by Annals of Botany Company. Reprinted with permission.

### 2.2.3. Importance of rapeseed oil

Rapeseed, an ancient crop cultivated for centuries, has gained significant importance in various human activities, serving as a valuable source of human food, animal feed, and raw materials for industrial products like lubricants and energy production (Richter et al., 2010). However, rapeseed had limited application before the seventies due to its high erucic acid (C22:1 FA) content in the oil and glucosinolates in the meal, posing potential health risks for humans and animals. With advancements in breeding techniques, the content of these two components was greatly reduced (erucic acid in oil less than 2% and glucosinolates in the meal less than 30µmol/g in the air-dried, oil-free meal). It led to the development of a new rapeseed cultivar, commonly known as double-zero rapeseed or canola in Canada, making rapeseed suitable for human and animal consumption (Przybylski, 2011).

Rapeseed oil, extracted from rapeseed, is recognized as the third most important vegetable oil globally, following palm oil and soybean oil (Jannat et al., 2022). Edible rapeseed oil contains higher levels of unsaturated FAs, such as 20-30% linoleic acid (18:2n-6) and 6-14% linolenic acid (18:3n-3). The health benefits of rapeseed oil are attributed to its high levels of unsaturated FAs, including omega-3 FAs, and low levels of saturated FAs. Breeders have developed various varieties with unique properties by manipulating the fatty acid composition of rapeseed oil. For instance, rapeseed oil with high lauric acid (31%) (C12:0) meets the needs of confectionery coatings, coffee whiteners, whipping toppings, and center-filling fats. Figure 4 shows the fatty acid composition of regular and modified rapeseed oil (Przybylski, 2011).

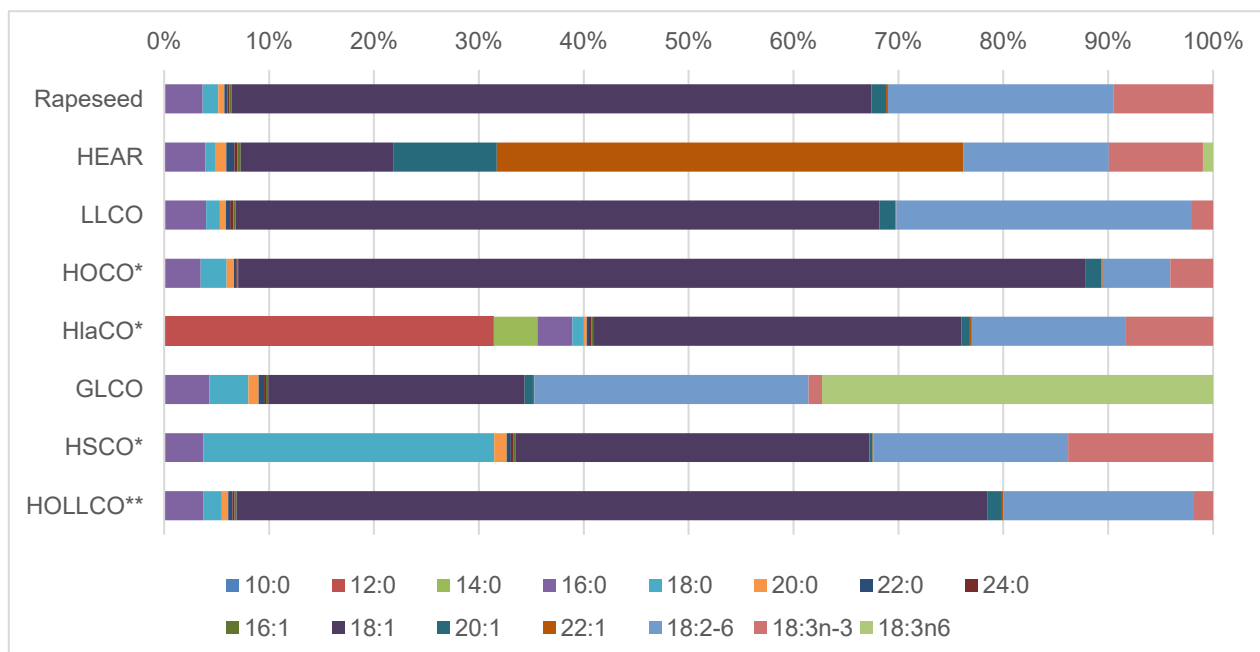


Figure 4: Composition of major FAs in regular and modified rapeseed oil.

HEAR = High Erucic Acid Rapeseed oil; LLCO = Low Linolenic Acid Canola Oil; HOCO = High Oleic Acid Canola Oil; HlaCO = High Lauric Acid Canola Oil; GLCO = Canola Oil with Gamma Linolenic Acid; HSCO = Canola Oil with High Content of Stearic Acid; HOLLCO = High Oleic Low Linolenic Canola Frying Oil.

Note: Adapted based on the data from *Vegetable Oils in Food Technology* (Page 110), by Frank D. Gunstone, 2011, Wiley-Blackwell. Copyright 2011 by Blackwell Publishing Ltd. Adapted with permission.

---

Biodiesel from vegetable oils, including rapeseed oil, is considered a renewable and cleaner-burning alternative to conventional diesel fuel. The European Union (EU) promotes biofuels to reduce greenhouse gas emissions and dependence on fossil fuels. Directive (EU) 2018/2001, amended by Directive (EU) 2023/1791, requires member states to achieve a 42.5% share of renewable energy in gross final energy consumption by 2030 (European Union, 2023). However, concerns have been raised about the sustainability of biofuels, including those derived from rapeseed. Issues such as land-use change, deforestation, and competition with food production have been debated (Fridrihsone et al., 2020; Meijaard et al., 2020). To tackle these concerns, the EU has implemented sustainability criteria for biofuels, ensuring that only sustainable biofuels contribute to meeting the targets. The Regulation on Deforestation-Free Products (EUDR), which will soon come into force, aims to minimize the European Union's contribution to global deforestation. One of the objectives of the EUDR is to require companies to provide evidence that their products are not linked to deforestation, thereby increasing transparency and accountability in the supply chain. Consequently, knowing the geographic origin of rapeseed oil as a biofuel feedstock becomes crucial, playing a role in determining its sustainability.

Additionally, a growing emphasis on food safety and health awareness has recently increased consumer preference for region-specific food products. Researchers like Gatzert et al. have developed methods for geographic identification to determine the origin of crops, such as those from Hesse, Germany (Gatzert et al., 2021). In this context, the geographical origin of rapeseed oil from Hesse, both as a biofuel feedstock and a healthy edible oil, becomes an important aspect.

### **2.3. Analytical approach for geographic traceability of vegetable oil**

In recent years, there has been a significant increase in research activities focused on geographical origin, reflecting a growing interest in the value of food geographical authentication (Fanelli et al., 2021; Katerinopoulou et al., 2020; Wadood et al., 2020). Tahir et al. conducted a study analyzing relevant papers published between 2013 and 2020, revealing that research on the geographical traceability of edible oil accounted for 15% of the total research in this field (Tahir et al., 2022). Combining analytical techniques with chemometric methods has shown promise in geographic traceability and preventing fraudulent practices (Tahir et al., 2022).

Chemometrics involves the mathematical and statistical modeling of analytical data to derive relevant chemical information, facilitating the comparison of numerous responses (dependent variables) against a multitude of independent variables (predictive variables) (Granato et al., 2018). In the context of the geographic traceability of oil, chemometrics enables the analysis of complex data generated by various analytical techniques, allowing for identifying characteristic patterns specific to oils from particular regions (Vandeginste, 2013). The two primary statistical techniques used in chemometrics are PCA and partial least squares (PLS) analysis (Granato et al., 2018). Geographical traceability serves the purpose of either verifying the declared origin of a product or classifying a product into predefined categories. For classification purposes, various discriminating techniques, such as linear discriminant analysis (LDA) and discriminant partial least squares (D-PLS), are employed to build classification models (Vandeginste, 2013).

---

Analytical data obtained through various analytical techniques, referred to as (analytical) markers, plays a crucial role in accurately determining the geographical origin of food products (Ballin & Laursen, 2019a). These techniques provide valuable information serving as predictive variables for models in geographical traceability.

Various chemical components, such as phenolic compounds, fatty acid profiles, sterols, TAGs, and volatile compounds, have been identified as useful markers for distinguishing oils from specific geographical origins. In addition to chemical components, other factors, such as chemical elements and stable isotope ratios, have also demonstrated their potential to trace the geographic origin of food products (Ballin & Laursen, 2019a). A single analytical marker alone provides limited information and is inadequate for ensuring geographic traceability. Instead, a profiling analysis that quantifies multiple markers can offer a more comprehensive picture and enhance classification accuracy. The nomenclature “fingerprinting analysis” often refers to profiling analysis (Tahir et al., 2022). Analytical techniques, including chromatographic methods, inductively coupled plasma-mass spectroscopy (ICP-MS), and inductively coupled plasma atomic emission spectroscopy (ICP-AES), are major methods used for verifying the geographic origin of oils, collectively constituting a significant portion of published articles (Tahir et al., 2022).

Isotope ratio mass spectrometry (IRMS) is a widely used technique for measuring stable isotope ratios of light elements such as C, H, O, N, and S. The isotope ratios of C, H, and O in plants can be influenced by factors like the photosynthesis pathway and environmental conditions, making them valuable markers for geographical traceability (Camin, Larcher, Perini, et al., 2010; Jeon et al., 2015; Luykx & van Ruth, 2008). In the context of tracing the geographical origin of oil, IRMS has great potential that has yet to be fully explored (Tahir et al., 2022).

Two main techniques are used in stable isotope analysis: bulk stable isotope analysis (BSIA) and component-specific stable isotope analysis (CSIA). BSIA involves analyzing the stable isotope ratios of an element in the entire sample, such as  $^2\text{H}/^1\text{H}$  of vegetable oil. The measured isotopic composition is an average value of the interested element, particularly when the sample is a complex mixture of several compounds containing this element. Conversely, CSIA focuses on determining the stable isotope ratios of selected compounds within the sample, for example,  $^2\text{H}/^1\text{H}$  of FAs in vegetable oil (Meier-Augenstein, 2017). For BSIA, a commonly used instrumentation is coupling an elemental analyzer (EA) with an IRMS unit. The sample is first converted to simple gases of the interested isotopes, such as  $\text{CO}_2$ ,  $\text{H}_2$ ,  $\text{CO}$ ,  $\text{N}_2$ , and  $\text{SO}_2$ , immediately after injection into the EA. Subsequently, the IRMS system analyzes the signal intensities of the sample gas with its isotopologues, such as masses 44, 45, and 46 for  $\text{CO}_2$  (Meier-Augenstein, 2017).

In CSIA, a coupled instrumentation known as gas chromatograph-combustion reactor/pyrolysis-isotope ratio mass spectrometer (GC-C/Py-IRMS) is commonly used. The sample compounds are first separated using a GC column and individually converted to the corresponding sample gas in an interface system. For instance,  $\text{H}_2$  gas is generated in a high-temperature microfurnace. The resulting gas is then analyzed using IRMS to determine the stable isotope ratios (van Leeuwen et al., 2014). Table 5 summarizes the studies using isotope analysis and chemometrics to trace the geographical origin of vegetable oils from 2010 to 2020. Most of these studies focused on olive oil,

a widely consumed and traded product. In contrast, rapeseed oil was investigated in only one study. In addition, the main analytical method for vegetable oil traceability was BSIA, measuring the isotopic composition of the whole oil. Only a few studies employed Compound-Specific Isotope Analysis of Hydrogen (H-CSIA), which analyzes the isotopic ratio of a specific compound in vegetable oils.

Table 5: Overview of studies using stable isotope analysis combined with chemometrics to assess the authenticity of edible oils in recent years (2010–2020).

Analytical marker	Detection techniques	Oil types	Sample size	Geographical origin	Chemometrics	Accuracy	References
BSIA-C, -H; <sup>1</sup> H NMR spectra	EA-IRMS; <sup>1</sup> H NMR	Olive oil	125	Italy, Spain, Greece, France, Turkey, Cyprus	PCA, PLS-DA	93%	(Alonso-Salces et al., 2015)
BSIA-C, -O, -H; <sup>1</sup> H NMR spectra; Fatty acid composition	EA-IRMS; <sup>1</sup> H NMR; GC-FID	Olive oil	263	Italy, Tunisia	PCA, RF	100%	(Camin et al., 2016)
BSIA-C, -H; carotenoid content	EA-IRMS; Raman	Olive oil	38	Italy	PCA, LDA	82%	(Portarena et al., 2017)
BSIA-C, -H	EA-IRMS	Palm oil	33	Malaysia	OPLS-DA	71%	(Muhammad et al., 2018)
BSIA-C, -H, -O	EA-IRMS	Olive oil	138	Portugal, France, Turkey	PCA, multi-linear regression	n.a.	(Jiménez-Morillo et al., 2020)
BSIA-C, -H, -O	EA-IRMS	Olive oil	100	Greece	OPLS-DA	91%	(Tarapoulouzi et al., 2021)
BSIA-C; CSIA-C; FA profile	EA-IRMS; GC-C-IRMS; GC-FID	Olive oil	157	Italy	PCA and PLS- DA	93%	(Faber et al., 2014)
CSIA-C, H (n- Alkane)	GC-C/P-IRMS	Olive oil	50	Mediterranean regions	CDA	93%	(Mihailova et al., 2015)
Bulk-C, -H; CSIA-C; FA profile	EA-IRMS; GC/C/IRMS; GC-FID,	Rapeseed oil	25	German, Swiss	PCA	not successful	(Richter et al., 2010)
BSIA-C, -H, -O;	EA-IRMS	Sesame oil	84	Korea, China, India	CDA	89%	(Jeon et al., 2015)
BSIA-C; Elements	EA-IRMS; ICP-MS	Olive oil	49	Turkey	PCA and HCA	n.a.	(Gumus et al., 2017)
BSIA-C, -O; Elements	EA-IRMS; ICP-MS	Olive oil	539	Italy	Kruskall–Wallis multiple bilateral comparison	n.a.	(Camin, Larcher, Perini, et al., 2010)
BSIA-C, -O, -H; Elements	EA-IRMS; ICP-MS	Olive oil	267	8 European sites	CDA	95%	(Camin, Larcher, Nicolini, et al., 2010)

## 2.4. Hydrogen isotopic fractionation of FAs in plants

The  $\delta^2\text{H}$  values of vegetable oils exhibit variations primarily influenced by the composition of their source and the physical and biochemical processes occurring in the soil-water-plant system, both before and after their incorporation into plant tissues. As a result, the  $\delta^2\text{H}$  values of vegetable oils can offer valuable insights into their geographical origin, either directly or indirectly (Richter et al., 2010). Given that FAs are the principal constituents of vegetable oils, predominantly present in the form of TAGs esterified with glycerol, it is reasonable to hypothesize that the  $\delta^2\text{H}$  values of FAs (CSIA-H) could serve as a potential analytical marker for determining the geographical origin of vegetable oils. The following subchapters will summarize the variation sources in  $\delta^2\text{H}$  values of FAs,



from hydrogen source to FA biosynthesis.

### 2.4.1. Variation in $\delta^2\text{H}$ values of precipitation

Plants primarily obtain hydrogen from water sources, including precipitation, soil water, runoff (including snowmelt), and groundwater. All these water sources ultimately trace their origin back to precipitation (Dawson, 1993). The  $\delta^2\text{H}$  value in precipitation varies spatially and temporally, which can be elucidated through Rayleigh distillation, as shown in Figure 5. Briefly, clouds form at the equator through the evaporation of seawater. During this process, the water vapor becomes depleted in the heavy hydrogen  $^2\text{H}$  compared to seawater, since  $\text{H}_2\text{O}$  containing  $^1\text{H}$  has a higher vapor pressure than  $\text{H}_2\text{O}$  containing  $^2\text{H}$ . This causes water molecules with  $^1\text{H}$  to evaporate more readily than those with  $^2\text{H}$ . As the cloud moves towards higher latitudes, it undergoes condensation, resulting in the formation of precipitation that is enriched in  $^2\text{H}$  compared to the original cloud. As a result, the cloud becomes depleted in  $^2\text{H}$  after the removal of the precipitation. Along its poleward movement, the cloud experiences a drop in temperature, leading to increased fractionation and further depletion in  $^2\text{H}$ . This process ultimately yields precipitation with extremely low  $^2\text{H}/^1\text{H}$  values at the poles (Eley, 2014). These temperature and rainout effects result in geographical and seasonal variation in the  $\delta^2\text{H}$  values of precipitation, summarized as follows:

Latitude effect (temperature effect): The  $\delta^2\text{H}$  value in precipitation decreased with increasing geographic latitude, in accordance with the temperature gradient from the equator to the poles.

Seasonal effects: The  $\delta^2\text{H}$  value in precipitation varies seasonally at higher latitudes (over  $30^\circ$ ), peaking in summer and reaching its lowest point in winter.

Continental effects: The  $\delta^2\text{H}$  value in precipitation diminishes as the air mass moves from the coast toward the interior of the continent.

Altitude effect: The  $\delta^2\text{H}$  value in precipitation decreases with increasing altitude, with a more pronounced effect at high latitudes than at low and mid-latitudes.

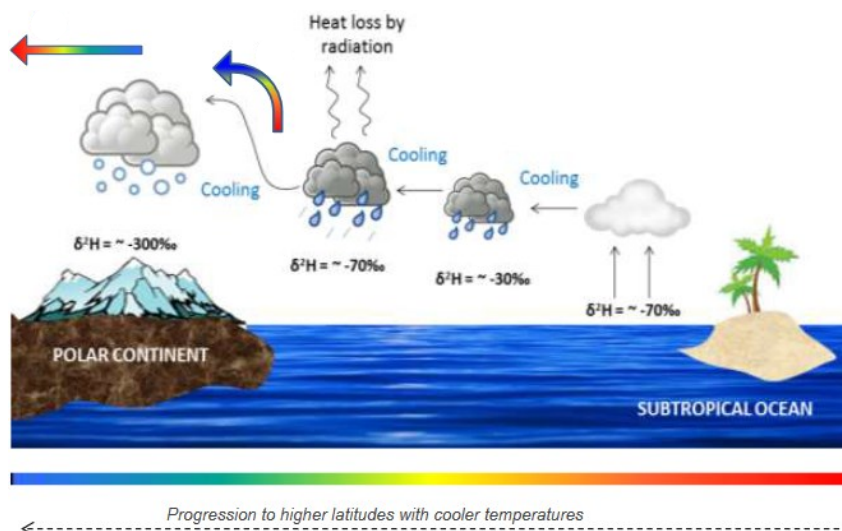


Figure 5: Simplified diagram showing water fractionation within a cloud through Rayleigh distillation.

Note: From Environmental and biochemical controls of the molecular distribution and stable isotope composition of leaf wax biomarkers (Page 9), by Y. Eley, 2014. [Dissertation, School of Environmental Sciences University of East Anglia].

---

Amount effect: The  $\delta^2\text{H}$  value in precipitation decreases with increasing precipitation amounts, a phenomenon particularly notable in tropical regions where seasonal variation is minimal.

Among these factors, temperature effects dominate in non-tropical regions, while in the tropics, amount effects take precedence (Dansgaard, 1964; Marshall et al., 2007).

#### **2.4.2. Variation in $\delta^2\text{H}$ values of soil water**

As precipitation descends, infiltrates into soil, and transforms into soil water, it undergoes a series of processes that influence its isotopic composition:

##### **a) Interception of precipitation by plant leaves**

When precipitation falls, plant leaves can intercept it, altering the isotopic composition through evaporation during interception. Consequently, upon reaching the soil, the isotopic composition of water differs from that of the original precipitation (Marshall et al., 2007).

##### **b) Precipitation mixed with previous soil water**

Hu et al. experimentally demonstrated that as precipitation infiltrates the soil, it mixes with pre-existing soil water, resulting in an isotopic exchange between these two water sources. The resulting  $\delta^2\text{H}$  value of the soil water is influenced by the isotopic compositions of both water sources, precipitation amount, and the timing of mixing (Hu et al., 2008).

##### **c) Evaporation**

Hu et al. summarize the impact of evaporation on isotopes of soil water in two situations: one is in a saturated soil zone, and the other is in an unsaturated zone. Simulation analysis reveals an exponential trend in the  $\delta^2\text{H}$  values in saturated soil layers, depleting with depth and stabilizing at the bottom. Notably, the maximum value is typically observed near the evaporation front, where soil water exhibits enrichment in  $^2\text{H}$ . The water movement within the unsaturated soil zone can be segregated into liquid and vapor phases, primarily defined by the evaporation front. The isotopic composition of the liquid phase beneath the evaporation front exhibits similarities to that of saturated soil columns (Hu et al., 2008).

In natural circumstances, the isotopic patterns observed in the soil profile undergo seasonal variations, which are influenced by water inputs (precipitation) and water losses (evapotranspiration). In winter, reduced evapotranspiration leads to deeper percolation of precipitation into the soil, resulting in  $^2\text{H}$  depletion in deeper soil layers compared to summer precipitation, as mentioned in Chapter 2.4.1. In summer, precipitation is more susceptible to evaporation at the surface soil level. After evaporation, the upper remaining soil water becomes more enriched in  $^2\text{H}$  (Marshall et al., 2007).

##### **d) Water-soil interaction**

Significant isotopic differences are observed among different types of soil water, including bound water and mobile water, with these variations being linked to soil type (Araguás-Araguás et al., 1995; Vargas et al., 2017). Clay minerals are identified as influential factors, as they preferentially adsorb water molecules enriched in the heavier hydrogen isotope over the lighter isotope, resulting in a depletion of  $^2\text{H}$  in the bound water. The precise adsorption mechanism remains incompletely understood, but it is hypothesized that the clay surface exhibits a higher affinity for  $^2\text{H}$ -containing water molecules due to differences in bonding energies (Coplen & Hanshaw, 1973). Moreover,

---

chemical weathering processes can generate cations ( $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{K}^+$ ,  $\text{Na}^+$ ) that exert isotopic effects through their respective hydration shells (Oshun et al., 2016).

The complex interaction of these processes contributes to the overall variations in the  $\delta^2\text{H}$  values of soil water. In particular, evaporation plays a significant role in driving  $\delta^2\text{H}$  value changes in soil water, especially in regions with minimal variation in  $\delta^2\text{H}$  values of precipitation (Dawson, 1993). Considering these processes, it can be hypothesized that various climate factors, such as temperature, precipitation amount, and soil properties, including silt, clay, and others, may influence the  $\delta^2\text{H}$  values of FAs.

### **2.4.3. Variation in $\delta^2\text{H}$ values in FA biosynthesis of plant**

In the complex dynamics of hydrogen isotopes in soil water, it is important to highlight that no fractionation occurs during water uptake by plant roots, preserving the original isotopic composition of soil water (Dawson, 1993). However, hydrogen isotopic fractionation occurs during FA biosynthesis when the hydrogen from root water is involved.

When water molecules are split by light energy, they release hydrogen ions that are used to make Nicotinamide Adenine Dinucleotide Phosphate Hydrogen (NADPH), an important molecule for photosynthesis. NADPH is produced by the combination of  $\text{H}^+$  and electrons photolyzed from soil water with nicotinamide adenine dinucleotide phosphate ( $\text{NADP}^+$ ), a coenzyme in the inner envelope membrane of a plant. This high-energy compound is used together with Adenosine triphosphate (ATP) to synthesize sugars in carbon fixation reactions. The isotope fractionation occurs during light-driven water fission. The hydrogen available for NADPH synthesis is depleted in  $^2\text{H}$  by up to 600 mUr relative to the original water (Luo et al., 1991). However, hydrogen ions from water fission in photosynthesis are not the only source of NADPH. It is possible to reduce  $\text{NADP}^+$  to NADPH by oxidizing sugars, for example, in the pentose-phosphate cycle, in which hydrogen is generated by C-bound hydrogen in the sugars. The different hydrogen sources for NADPH have individual hydrogen isotopic compositions (Schmidt et al., 2003). This discrimination is further enhanced as acetyl-CoA is the substrate for fatty acid synthesis. Similarly to what was observed with  $^{13}\text{C}$ , the pyruvate dehydrogenase complex may discriminate strongly against heavy pyruvate (i.e.,  $^2\text{H}$ -rich) (Ziegler, 1989). In addition, the differences in the  $\delta^2\text{H}$  values in homologous molecules with different degrees of desaturation, such as FAs, are also observed. This variability could result from large enzymatic isotope effects occurring during hydrogenation (saturation) and dehydrogenation (desaturation) (Sachse et al., 2012).

Much literature suggests that environmental factors also affect the  $\delta^2\text{H}$  of FAs in plants. These factors can have direct and indirect effects, either through biosynthesis or the isotope composition of source or cellular water. However, few controlled experiments have been conducted to study it. The environmental influences often mentioned are temperature, humidity, light duration, salinity, and growth stage (Chikaraishi et al., 2004; Sachse et al., 2012; Schmidt et al., 2015). Additionally, environmental stresses such as nutrient, salinity, and water stress have also been identified as potential influencers of material flow within the acetogenic lipid reaction network. These stressors

---

can impact the availability and utilization of key resources involved in fatty acid metabolism, potentially altering the hydrogen isotope composition of the resulting FAs (Eley, 2014).

## **2.5. Sample preparation methods for H-CSIA of FAs in vegetable oil**

Geographical traceability of vegetable oils using the  $\delta^2\text{H}$  values of FAs combined with chemometrics requires measuring large samples. The commonly employed technique to measure  $\delta^2\text{H}$  values of FAs is GC-Py-IRMS. However, directly analyzing oil samples using this technique is not feasible because oils cannot be vaporized directly in the GC inlet. Hence, it is necessary to convert the samples into FAMES, which can be vaporized and separated by the GC column.

The specific sample preparation method depends on the form of the sample, whether it is an oil or a seed. If the sample is an oilseed, the oil must first be extracted from the seed. For instance, Woodbury et al. described the oil extraction from corn and 16 other non-corn seeds/nuts/kernels. They ground the seeds/nuts/kernels and placed them in petroleum ether for 6 hours. Afterward, the residual solvent was removed under a vacuum at 60°C (Woodbury et al., 1998). In the case of sesame oil, Jeon et al. collected the oil by centrifuging roasted and pressed sesame seeds for 10 minutes. The oil was then completely dried under nitrogen-flushing (Jeon et al., 2015).

Once the oil has been extracted, it must be converted into FAMES for GC analysis for the fatty acid profile or stable isotopic analysis of individual FAs. The DIN EN ISO 12966-2 method, which utilizes boron trifluoride ( $\text{BF}_3$ ) as a catalyst, is commonly employed for the production of FAMES for authentication studies (Camin, Larcher, Nicolini, et al., 2010; Richter et al., 2010; Woodbury et al., 1998). However, this method involves multiple steps. Firstly, TAGs are transmethylated in the presence of methanolic sodium hydroxide to produce FAMES. In a subsequent step, free FAs (FFAs) are converted to sodium soaps through an alkaline-catalyzed reaction. The soaps are then transformed into methyl esters by reacting with a  $\text{BF}_3$ -methanol complex. The resulting FAMES are separated into the upper isooctane phase with the addition of iso-octane and saturated sodium chloride solution. Any trace amounts of water are removed by adding anhydrous sodium sulfate, and the purified FAMES are ready for analysis through GC (ISO 12966-2:2017, 2017).

The DIN method has some limitations for producing FAMES in large-scale sample analysis. These limitations arise from the multi-step process involved and the use of boron trifluoride ( $\text{BF}_3$ ) as a catalyst. Multiple steps for sample preparation can hinder efficient sample handling and increase the risk of sample loss, contamination, and isotopic fractionation. Additionally, the use of  $\text{BF}_3$  as a catalyst presents certain concerns.  $\text{BF}_3$  is a toxic substance, and its handling requires strict safety precautions to minimize risks to researchers and the environment. Furthermore, the methanol form of  $\text{BF}_3$  used in the FAME production process is highly reactive and unsuitable for long-term storage (Ichihara & Fukubayashi, 2010).

## **2.6. Research objectives based on the limitations of the available analytical approach**

The combination of stable isotope analyses with chemometrics has successfully determined the geographic origin of vegetable oils. It relies on the understanding that the bulk stable isotopic composition of light elements such as C, H, N, and O in vegetable oils is linked to the environment

---

in which the plants grow. Consequently, these isotopic compositions provide valuable information about their geographical origin, directly or indirectly.

However, the application of  $\delta^2\text{H}$  of FAs (CSIA-H) as a marker for determining geographic origin is still relatively limited despite its potential in this regard. As explained in Chapter 2.4, the hydrogen in FAs originated from precipitation, and the  $\delta^2\text{H}$  of precipitation exhibits spatial distribution patterns influenced by air temperature, precipitation amount, and spatial location. Before plants take in water from precipitation, the rain undergoes evaporation in the soil, mixing with different layers and clay components. These processes can alter its  $\delta^2\text{H}$ , indirectly providing information about geographical origins. Factors such as temperature, soil moisture, silt, and clay content related to soil water mobility can influence the  $\delta^2\text{H}$  of soil water. Furthermore, when the FAs are synthesized in plants and then stored in oilseeds, climate factors and nutrient supply in the soil can also impact  $\delta^2\text{H}$  value, indirectly reflecting geographical sources. These complex natural processes collectively affect the  $\delta^2\text{H}$  value. However, there is limited research on which environmental factors, including climate and soil properties, significantly influence the  $\delta^2\text{H}$  of FA in vegetable oils. This knowledge gap hinders the feasibility of applying  $\delta^2\text{H}$  of FAs as a marker for determining geographic origin.

Additionally, determining the  $\delta^2\text{H}$  values of FAs due to the labor-intensive sample preparation processes with the DIN method is time-consuming, which has practical limitations on applying  $\delta^2\text{H}$  of FAs in geographical determination. Furthermore, while rapeseed oil is the third most important vegetable oil, little research has been conducted on its geographic traceability, emphasizing the need for further investigation.

Based on the mentioned limitations, the primary aim of this dissertation is to develop a fingerprinting approach to determine the geographical origin of vegetable oils. Rapeseed is used as a case study. The LLH provides ample rapeseed and corresponding soil samples with precise sampling locations, which is essential for extracting climatic data from publicly available datasets. Additionally, using rapeseed instead of oil avoids potential interference from the oil extraction process in isotopic measurements. The fundamental knowledge and methods that can serve as a foundation for future research on the geographical traceability of rapeseed oil can be established by studying rapeseed.

For this main aim, three research objectives are defined as follows:

1. Evaluate the feasibility of the rapid one-step sample preparation method that converts rapeseeds into FAMES for subsequent H-CSIA.

The first research objective is to enhance efficiency and streamline laboratory work for the H-CSIA of FAs of rapeseed and its oil. A simple and rapid sample preparation method developed by Garcés & Mancha (1993) allows the conversion of seeds to FAMES in one step, which can be directly analyzed by GC-Py-IRMS. However, it has never been reported if this method is feasible for stable isotope analysis. Therefore, it will be evaluated if this one-step method is feasible for the measurement of  $\delta^2\text{H}$  and  $\delta^{13}\text{C}$  of FAs.

---

2. Assess the feasibility of  $\delta^2\text{H}$  of FAs as analytical markers for identifying the geographical origin of rapeseed.

- Investigate spatial and temporal variations in the  $\delta^2\text{H}$  of FAs of rapeseed from Hesse.
- Explore the correlation of soil and climatic factors with the  $\delta^2\text{H}$  of FAs in rapeseed.

The second objective is to assess the potential of  $\delta^2\text{H}$  of FAs as analytical markers for determining the geographic origin of rapeseed. As mentioned earlier, various environmental factors could potentially influence the  $\delta^2\text{H}$  of FAs in plants. The relative impact of these factors remains unclear, leading to uncertainty regarding the efficacy of H-CSIA of FAs in identifying the geographical origins of vegetable oils. Identifying significant correlations would indicate that the  $\delta^2\text{H}$  of FAs could serve as effective analytical markers. Rapeseed and soil samples from Hesse, Germany, collected between 2017 and 2020, were selected as the primary subjects of investigation. Hesse was chosen due to its distinct regional climatic variations and the availability of comprehensive climatic datasets from Deutscher Wetterdienst (DWD, German Weather Service). This objective is further divided into two parts. First, the spatial and temporal variation of the  $\delta^2\text{H}$  values of FAs in rapeseeds from the Hessian region will be investigated, as there is currently a lack of relevant data in the existing literature. Secondly, the significant environmental factors influencing  $\delta^2\text{H}$  values of FAs will be detected. These factors include climate variables such as temperature, rainfall, sunshine, and drought index during rapeseed growth. Additionally, soil properties, such as moisture, temperature, electrical conductivity, pH, clay content, silt content, and sand content, will be considered. The study will also explore the role of nutrient levels in the soil, specifically potassium, phosphorus, and magnesium content. Furthermore, the geographic location will be considered, including latitude, longitude, and altitude. Finally, the influence of rapeseed variety on the  $\delta^2\text{H}$  values of FAs will also be investigated.

3. Determine the geographical origin using  $\delta^2\text{H}$  of FAs and elemental composition combined with chemometrics.

- Distinguish rapeseed samples from Hesse and non-Hesse.
- Distinguish rapeseed samples within Hesse.

The third objective is to develop classification models that utilize the  $\delta^2\text{H}$  values of FAs in combination with chemometric techniques to determine the origin of rapeseed. PCA and OPLS-DA are used to construct classification models. Two distinct classification targets are considered in this dissertation: Hesse and Jianhu, which are geographically distant locations and different regions within Hesse itself. It enables evaluating the discriminative power of  $\delta^2\text{H}$  values of FAs across varying geographic distances. By considering these multiple classification targets, the performance classification models can be compared, and the contribution of the  $\delta^2\text{H}$  values of FAs in determining the origin of rapeseed across different geographic distances will be evaluated. Additionally, rapeseed elemental composition is incorporated to enhance classification accuracy.

---

## 3. Materials and methods

---

### 3.1. Overview of methods based on the research objectives

These three research objectives necessitate distinct materials and methods involving diverse experimental designs, sample selection across varying periods and geographical ranges, and specific data analysis approaches. Here is an overview of the samples and methods employed for each research objective:

The first objective is to assess the suitability of FAMEs prepared using a selected one-step method for subsequent stable isotope analysis through several comparative experiments conducted on rapeseed in the laboratory. The introduction of the one-step method, the experimental design, and the results are presented in Chapter 4.

For the second objective, rapeseed and related soil samples collected in Hesse between 2017 and 2020 were utilized. Objective 2.1 involves the analysis of samples from 2019 and 2020, while samples from 2017 to 2020 were encompassed for Objective 2.2. Samples distributed across different regions within Hesse were collected to capture the diversity of environmental conditions in the sampling sites. Consequently, the sample collection strategy focused on obtaining representative samples that adequately reflected the overall distribution within Hesse. Given the large number of samples collected, a secondary selection was implemented to reduce the analysis workload. The overarching strategy for sample selection remained unchanged, prioritizing the representativeness of the entire sample set. Further details regarding the sample selection strategy can be found in Chapter 3.2. The data analysis methods for this objective primarily included analysis of variance (ANOVA), correlation analysis (Pearson and Spearman correlations), and stepwise linear regression analysis.

For the third objective, rapeseed samples from Hesse in 2019-2020 were utilized, along with rapeseed samples from Jianhu, China, in 2019 as non-Hesse samples. The data analysis methods PCA and OPLS-DA were employed.

Analytical data acquisition for all research objectives involved two approaches: laboratory measurements and climatic and soil data extraction from publicly available databases. The laboratory measurements included stable hydrogen isotopic analysis of FAs in rapeseed, elemental analysis of rapeseed, and soil property measurements. The methods employed for these measurements are presented in Chapters 3.3 to 3.5. Additionally, data from existing datasets, combined with geographical coordinate data of the sampling points, were utilized. The specific database used is listed in Chapter 3.6. Distinct statistical analysis methods are outlined in Chapter 3.7. Figure 6 presents the flow chart about the methods and samples for the second and third research objectives.

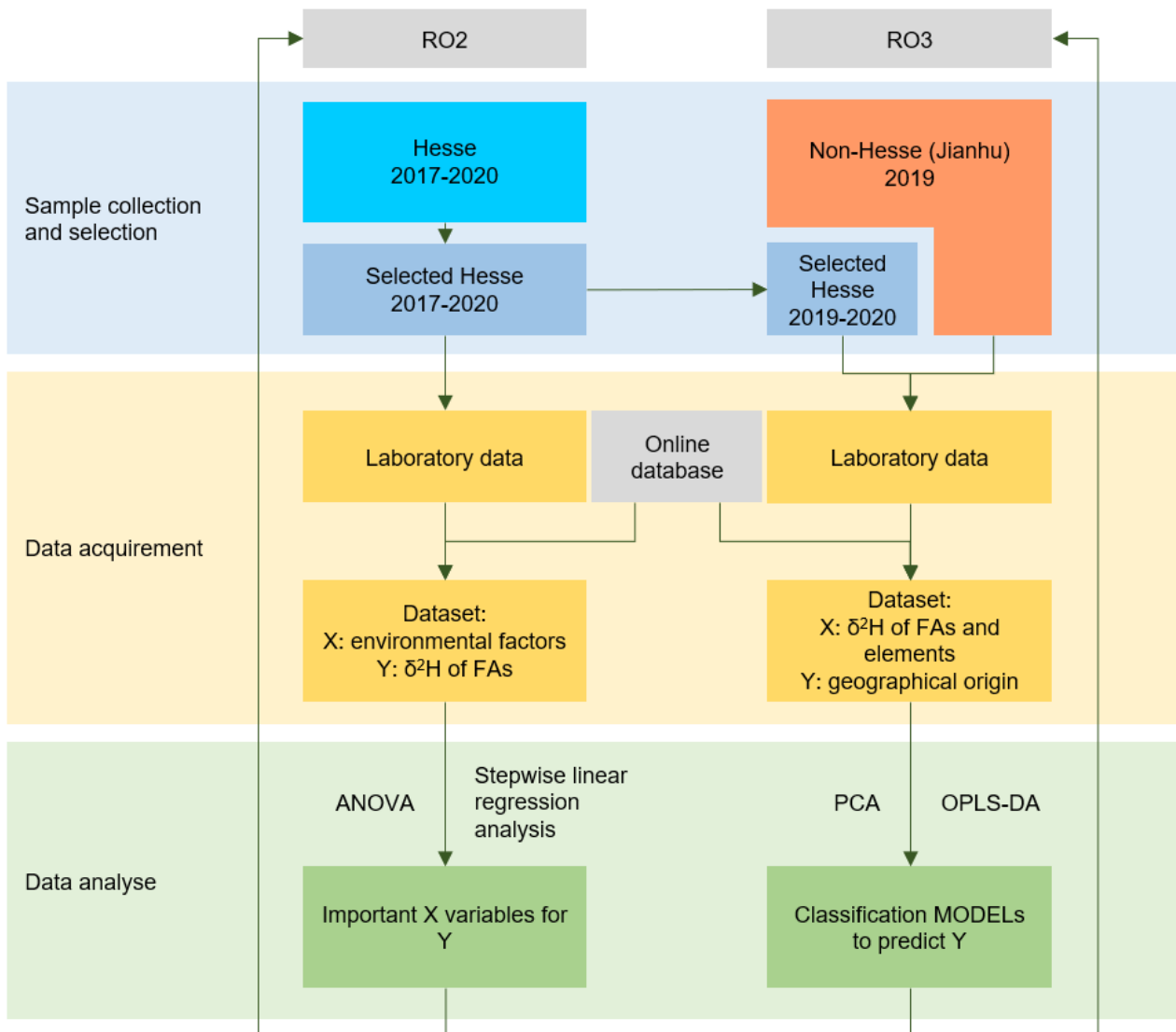


Figure 6: A procedure flow chart for analyzing 2. and 3. Research Objectives (RO).

## 3.2. Samples

### 3.2.1. Sampling location

#### Hesse, Germany

Hesse, also known as Hessen, is a federal state located in the central part of Germany and positioned between 49°24' to 51°39' north latitude and 7°46' to 10°11' east longitude. It is characterized by a distinctive landscape comprising basins, swales, and elevated areas. The elevation levels vary, with basins and swales ranging below 300 meters above sea level (a.s.l.) and elevated regions exceeding 300 meters a.s.l. The highest point in the region is found in the eastern low mountain range known as the Rhön, reaching an elevation of 950 m a.s.l. Additionally, Hesse has numerous elevations exceeding 550 m a.s.l. Meanwhile, the basin areas lie at altitudes of 100-200 m a.s.l. in the south and 150-250 m a.s.l. in the north (Lüker-Jans et al., 2017). Therefore, the climate in Hesse exhibits distinct regional variations, divided into two main parts. The lowlands experience notably warm conditions, while the highlands are characterized by relatively colder and



more humid climates (Lüker-Jans et al., 2017). There are various main groups of geological substrates in Hesse, including eolian substrates, carbonate substrates, fluvatile substrates, colluvial substrates, pelitic substrates, psephitic substrates, volcanic substrates (HLUNG, 2011).

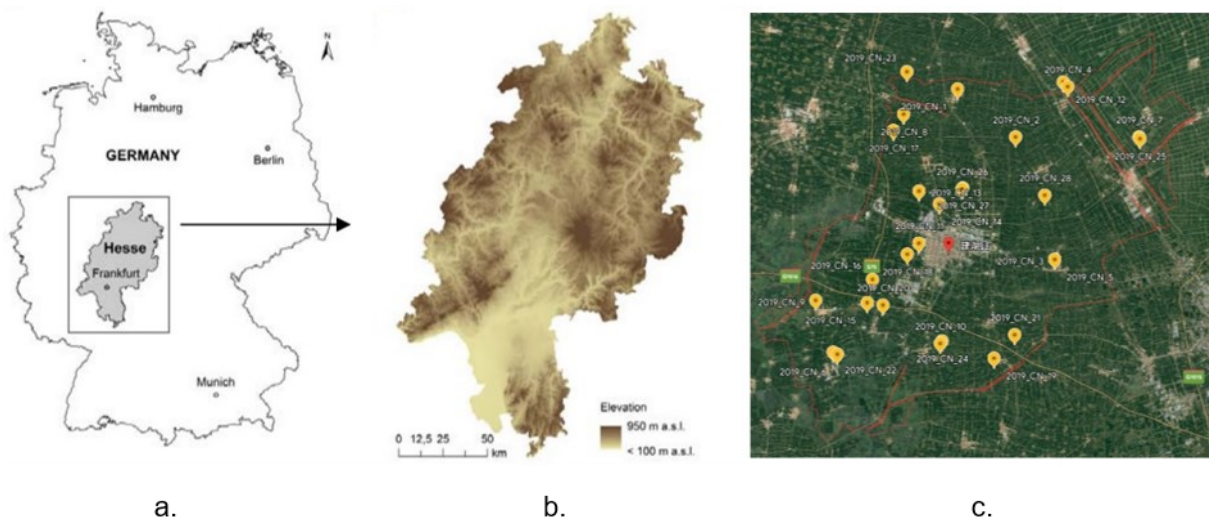


Figure 7: Study area: a. Hesse in Germany; b. Topography (elevation between 100 and 950 m a.s.l.); c. Sampling sites in Jianhu, China.

Note: Figure 7a+b is from *The impact of biogas plants on regional dynamics of permanent grassland and maize area—The example of Hesse, Germany (2005–2010)*, by Lüker-Jans, 2017, *Agriculture, Ecosystems & Environment*, Volume 241. Copyright 2017 Elsevier; Figure 7c: Google Earth.

### Jianhu, China

Jianhu County is situated in the north-central part of Jiangsu Province, China, and falls under the administration of Yancheng City. It is geographically positioned between 33°16' to 33°41' north latitude and 119°32' to 120°05' east longitude. The county stretches approximately 43.7 kilometers from north to south and 48.7 kilometers from east to west, encompassing 1154 square kilometers. Water bodies make up about 19.58% of the total area. Jianhu is located in the Lixiahe Plain, and the topography of the region is notably flat. The average elevation across the county is 1.74 meters, with the highest point reaching 3 meters and the lowest point measuring only 0.7 meters a. s. l. The area falls within the wet zone of the Huaihe River Basin.

The county experiences a distinct monsoon climate. During winter, it is influenced by the cold air mass from Eurasia, resulting in prevailing northerly winds and cold weather. In contrast, summer is affected by the Pacific subtropical high-pressure system, leading to dominant southerly winds and hot weather. The air is warm and humid throughout the year, with abundant rainfall. The climate within the county is relatively uniform, without significant regional variations. Regarding soil types found in the lakeside area, there are four main categories: paddy soil, saline soil, fluvo-aquic soil, and swamp soil. These soil types are specifically associated with the lakes and their surrounding areas in Jianhu County (Jianhu Bureau of Natural Resources and Planning, 2017).

### 3.2.2. Sampling strategy

In Hesse, about 75 rapeseed samples and corresponding soil samples were collected from the entire

---

Hesse each year from 2017 to 2020, respectively. LLH planned and implemented the sampling and collection of samples.

In order to collect the representative samples within Hesse, sampling fields were selected in two stages based on the probability proportional to size (PPS) sampling. PPS is “a method of sampling from a finite population in which a size measure is available for each population unit before sampling and where the probability of selecting a unit is proportional to its size” (Skinner, 2016).

In the first stage, 75 farms that grow rapeseed were selected. In the second stage, one rapeseed field was selected from each of the selected farms. To increase the likelihood of larger farms being sampled in the first stage, PPS sampling was employed using the cumulative method based on plantation area, considering that each farm possesses a distinct rapeseed plantation area. The procedure was that all farms that grow rapeseed in Hesse were arranged one after the other in any order, the rapeseed plantation area of each farm was listed, and a cumulative sum of the plantation area was calculated farm by farm. The sampling Interval (SI) was calculated by dividing the total rapeseed plantation area by the number of farms required for sampling. A Random Start (RS) number was randomly chosen between 1 and the SI value. Each farm had a corresponding cumulative plantation area. In the order in which all farms were listed, the farm that first appeared on the list with a cumulative plantation area greater than the RS number was sampled. The second number for selection was the sum of RS and SI. The farm that first appears in the list with a cumulative plantation area greater than the RS+SI value was sampled. By analogy, the farms were selected. For example, the number of farms growing rapeseed in Hesse reached 4,800 in 2017, with a total area of 55,000 ha (Hessisches Statistisches Landesamt, 2017). The number of farms required for sampling was 75; therefore, the SI value amounted to 733. If an RS number was randomly determined as 600, then the second number for selection was 1333. The farms that first appeared in the list with cumulative plantation areas greater than 600 and 1333 were selected.

In the second stage, one field in each selected farm was sampled again so that farms with larger areas were sampled with a smaller probability, and each rapeseed plantation area had the same probability of being sampled. If the selected farm had only one field for growing rapeseed, it was selected by default. If the selected farm had two or more rapeseed fields that were approximately the same size, then one was randomly selected. If the fields were of unequal size, one rapeseed field was selected from these fields using the same PPS method as above. The sampling fields for the four years were presented in the map of Hesse (Figure 8).

The sample collection sites in the Jianhu region were also selected randomly, but not so strictly as in Hesse. Collecting samples from this area was to test whether they could be distinguished from Hesse. Nevertheless, the sampling sites in Jianhu were roughly scattered across the region, as shown in Figure 7c.

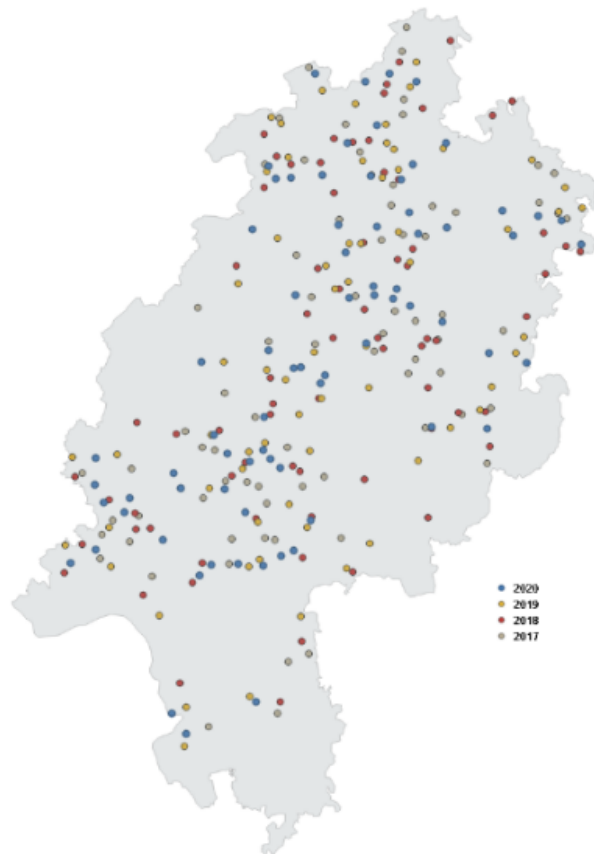


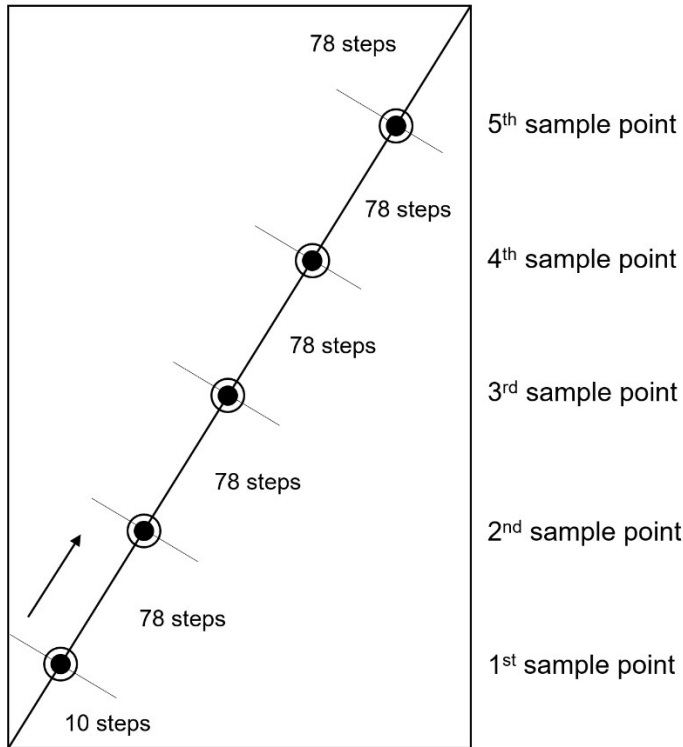
Figure 8: Sampling sites in Hesse from 2017 to 2020.

### Sample collection

To obtain a representative rapeseed sample from the entire selected field, the sampler collected a sample of about 500 g during threshing or when weighing the harvested seeds. The respective soil sample was taken directly when the threshing of rapeseed was carried out.

To collect representative soil samples from the selected field, the sampler followed a standardized method. First, five sampling points in the selected field were determined (Figure 9a). The method was as follows: the staff measured the length of the diagonal of the field with footsteps, for example, 400 steps in total. Then, the first sampling point was determined randomly by the number of footsteps along the diagonal, for example, at step 10. Next, the interval distance between the sampling points was calculated by dividing the difference between the total footsteps of the diagonal and that of the first sampling point by 5, for example, here 78 footsteps. Therefore, the second sampling point was determined by continuing to walk 78 steps along the diagonal; thus, five sampling points were determined.

Subsequent, at each sampling point, three samples were collected from soil depths up to 30 cm with the soil sampler (Figure 9b) so that a total of 15 samples were collected, mixed thoroughly, and packed as one soil sample into a container (Landesbetrieb Landwirtschaft Hessen, 2017). After removing impurities from the rapeseeds through sieving and air classification, the seeds were stored in vacuum-packed bags at room temperature. The air-dried soil samples were ground gently in a mortar and sieved through a 2 mm sieve ready for analysis.



a.

b.

Figure 9: a. Schematic diagram of determining soil sampling points; b. Soil sampler.

### Sample selection

Rapeseed and soil samples were collected for four consecutive years, from 2017 to 2020, with around 70 samples per year in Hesse. In order to complete the work more efficiently without compromising the results, a selection of the collected Hessian samples was performed.

The samples collected in 2019 and 2020 were specifically selected to assess the spatial and temporal distribution of  $\delta^{2}\text{H}$  of FA in rapeseed across Hesse, Germany. For a representative sample selection, locations were randomly distributed throughout the region. Initially, sequential numeric identifiers, spanning from 1 to 70, were allocated according to the south-to-north spatial positioning of the samples. Subsequently, five samples were randomly chosen from each group of ten, delineated by sequential ranges (e.g., 1-10, 11-20), ensuring comprehensive coverage for analysis. Moreover, given the objective to analyze the impact of soil types on the  $\delta^{2}\text{H}$  of FA in rapeseed, it was ensured that at least one sample from each soil type in Hesse was included. Thus, additional samples were selected to fulfill this requirement. The soil types were determined in advance using data obtained from the Bodenviewer.

The sample selection for 2017 and 2018 was less strict than for 2019 and 2020. However, at least one sample was chosen from each range of ten numerical identifiers, indicating that the selected samples were generally distributed throughout Hesse. This approach was acceptable since these samples were not intended for spatial distribution analysis. Their main purpose was to contribute to correlation analysis as long as all the corresponding analytical data for the samples were available.

---

### 3.3. Compound-specific stable hydrogen isotope analyses

This dissertation focuses on the measurement of the hydrogen isotopic composition of FAs in rapeseed oil. In rapeseed, FAs are esterified in the glycerol backbone and exist in the form of TAGs. However, TAGs are too large to be vaporized in the GC Inlet. Therefore, it is necessary to convert TAGs into FAMES before measuring the hydrogen isotope composition using GC-Py-IRMS. During the conversion process from TAGs to FAMES, each mole of FAME incorporates one mole of  $-CH_3$  group from methanol. Consequently, the final measured isotope value requires correction to account for the influence of hydrogen in methanol. To simplify the laboratory work and improve efficiency, a one-step sample preparation method was employed to convert rapeseed into FAMES. The specific details of this method are comprehensively introduced in Chapter 4. In Chapter 3.3, the following main points are covered:

- 1) The workflow for measuring the  $\delta^2H$  values of FAMES using GC-Py-IRMS, including normalizing the raw measurement data and determining the stable operating range of the instrument.
- 2) The correction procedures to calculate the  $\delta^2H$  values of FAs from FAMES.

#### 3.3.1. Workflow of GC-Py-IRMS

The H-CSIA measurement utilized a Thermo Fisher Scientific™ Trace 1310™ GC with a flame ionization detector (FID), coupled with a Delta V Advantage™ IRMS through an IsoLink II™ pyrolysis interface. The continuous helium flow, regulated by a ConFlo IV™ interface, was applied during the analysis (Figure 10).

The GC inlet unit, equipped with a TriPlus RSH™ autosampler, introduced a 1  $\mu$ L liquid FAMES sample into the system in splitless mode, vaporizing it at 250°C. Helium, flowing at a 1.5 mL/min rate, carried the vaporized gas sample into a TG-WAX™ column coated with stationary and mobile phases. This column effectively separated the gas sample mixture into different compounds by distributing them between the two phases. The column temperature was programmatically adjusted during the analytical run to enhance separation sensitivity. Once the vaporized gas sample reached the column, the temperature program initiated at 160°C for 1 min, increased to 190°C at 20°C/min, and then further to 220°C at 7°C/min for 29 min. Subsequently, the GC effluent from the column underwent splitting by a multi-channel splitter: approximately 10% was directed to the FID unit for quantitative analysis, while the remaining 90% was routed to the pyrolysis reactor.

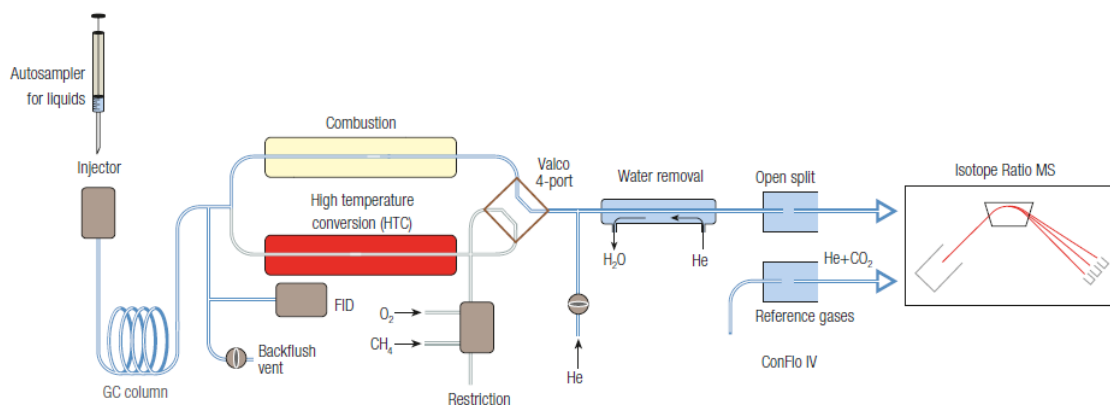


Figure 10: Schematic of the GC-C/Py-IRMS system applied in the laboratory.

Note: From GC-IRMS:  $\delta^{13}\text{C}$  Analysis of PAHs in Soil and Sediment Samples using High Resolution GC Coupled with Isotope Ratio MS, by (Juchelka, 2013) (Thermo Fisher Scientific).

The pyrolysis reactor serves as a quantitative conversion of the GC effluent to  $\text{H}_2$  gas at  $1420^\circ\text{C}$ , thus also called a high-temperature conversion (HTC) reactor (Figure 11). The reactor is a blank ceramic tube and needs to be conditioned to regenerate carbon layers frequently due to oxidation by oxygen and hydrogen in the carrier gas. However, too much carbon generated in the reactor could result in memory effects and influence the accuracy of the succeeding measurement (Sessions, 2006). Before the first measurement, the HTC reactor was preconditioned by  $2\ \mu\text{L}$  n-heptane injections. During the analytical runs, the carbon coating was formed from the highly concentrated FAMES sample containing many carbon atoms. Therefore, it is not necessary to condition the reactor extra. The reactor is used for the pyrolysis of nanogram quantities of the analyte. If the solvent reaches the reactor, a large amount of carbon will be deposited.

Moreover, the pressure can fluctuate, so the protection circuitry inside the IRMS ion source may trip (Sessions, 2006). Therefore, a “Backflushing” system is applied to flush the solvent flow backward out of the system by controlling a backflush vent and a 4-port valve. Once the solvent is vented, the following targeted compounds are allowed to be pyrolyzed to  $\text{H}_2$  gas in the reactor and entered the IRMS unit per the ConFlo IV unit, in which a capillary sniffs the sample gas directly from constant atmospheric pressure into the ion source. As a result, the sample gas that flows into the IRMS is stable, and the pressure in the ion source remains constant, guaranteeing precise IRMS measurement (Sessions, 2006).

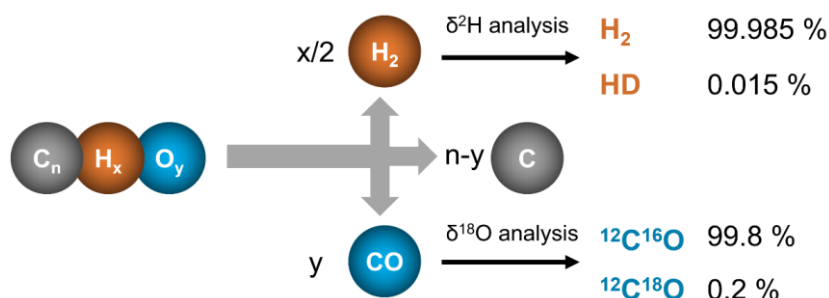


Figure 11: Scheme of High Temperature Conversion in the pyrolysis reactor.

Note: Modified from Thermo Fisher Scientific™ GC Isolink II Operating Manual, Version C (Page 7-15), by Thermo Scientific, 2015.

The gaseous sample enters the IRMS unit, which consists of three parts: an electron-impact ionization source (Figure 12), a single electromagnet analyzer, and a collector assembly. The sample gas is first subjected to electron-impact ionization in the ionization box by the electrons generated by a heated filament. The energy of ionizing electrons is usually 70-124 eV and can be adjusted depending on the sample gas type. A molecule of the sample gas becomes positively charged after an electron is knocked off by the electron impact. The ionized gas leaves the ionization box by an accelerating voltage of up to 3 kV and is focused by electrostatic lenses to form a beam. Subsequently, the ion beam travels to the magnetic field generated by an electromagnet, where the ions are deflected from each other based on the charge/mass ratio of the ion and the kinetic energy. By varying the magnetic field strength and accelerating voltage, the ion beams of different sample gas types are focused and collected by precisely positioned multiple Faraday cups. The multi-cup allows simultaneous measurement of the intensities (ion current) of different ion beams of isotopologues. Each cup is fitted with appropriate resistors to ensure that not only larger but also similar output voltages of different ion beams are generated (Sharp, 2017). Finally, the isotope ratios and delta values are processed by Isodat 3.0. The measured (raw) delta values of the sample (measured  $\delta^m E_{\text{sample}}$ ) are calculated relative to the isotope ratios of the working gas ( $R_{\text{sample}}$ ), which comes from the tank and enters the IRMS unit before and after the sample gas at every single run.

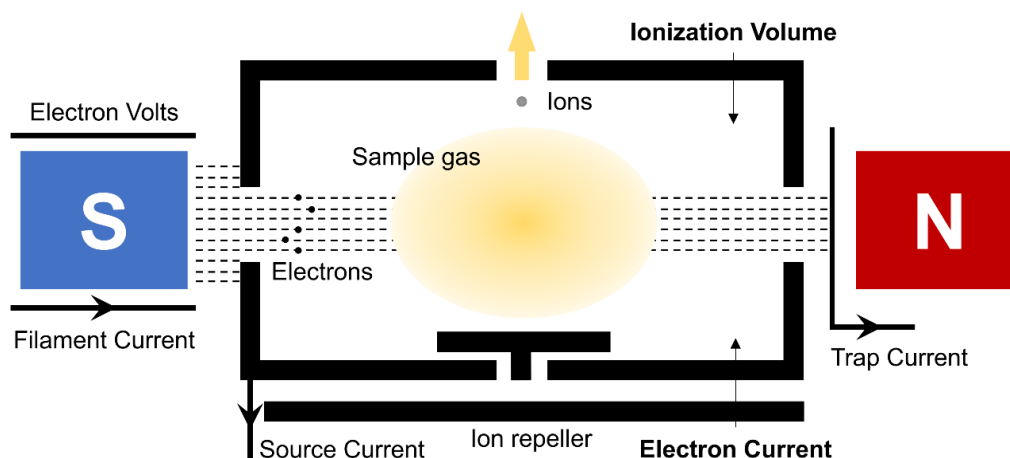


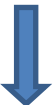
Figure 12: Schematic of the ion source.

Note: Modified from Fundamentals of Isotopic Ratios Mass Spectrometry, by Mounier, 2014 (Thermo Fisher Scientific).

## Two-point isotopic normalization

To compare the delta values between different laboratories, these measured (raw) values are converted and stated on the international isotope reference scale, namely true delta values. This process is called normalization, as shown below.

$$\text{measured } \delta^m E_{\text{sample}} = (R_{\text{sample}} - R_{\text{working gas}}) / R_{\text{working gas}}$$



$$\text{true } \delta^m E_{\text{sample}} = (R_{\text{sample}} - R_{\text{RM}}) / R_{\text{RM}}$$

Normalization

A two-point linear normalization procedure was performed for the  $\delta^2\text{H}$  values of individual FAMEs to correct the effect of the scale compression during the measurement (Meier-Augenstein & Schimmelmann, 2019). Two certified reference materials, USGS70 and USGS71, were purchased from Reston Stable Isotope Laboratory since they meet the following conditions: 1) their chemical structure C20:0 FAME is similar to that of our samples; 2) their isotopic compositions are relative to the international isotope references materials; 3) their true  $\delta^2\text{H}$  values approximately bracket the expected values of the samples (Table 6). The beads of USGS70 and 71 dissolved in n-heptane were used as reference solutions. In each analytical sequence, a linear regression between measured and true delta values was established by analyzing these two reference solutions. The true  $\delta^2\text{H}$  values of the samples were calculated with the regression line.

Table 6: Stable hydrogen isotopic composition of USGS70 and USGS71.

Reference Materials	Compound	$\delta^2\text{H}_{\text{VSMOW-SLAP}}$
USGS70	icosanoic acid methyl esters (C20:0 FAME)	$-183.9 \pm 1.4 \text{ mUr}$
USGS71	icosanoic acid methyl esters (C20:0 FAME)	$-4.9 \pm 1.0 \text{ mUr}$

Note:  $\delta^2\text{H}_{\text{VSMOW-SLAP}}$  values are expressed as delta values relative to VSMOW on a scale normalized such that the  $\delta^2\text{H}$  value of SLAP (Standard Light Antarctic Precipitation) is  $-428 \text{ ‰}$  (Schimmelmann et al., 2016).

## Linearity for H-CSIA

Due to the high ionization efficiency of electron impact sources, there is a large fractionation of isotopes. Providing the ion source conditions (i.e., pressure) remain the same for the sample and working gas, fractionation will not affect the relative isotope ratios. It is called (isotopic) linearity, which is how the fractionation of isotopes changes with source pressure (Sharp, 2017). By tuning the ion source (extraction voltage) and ion optics, the non-linearity can be reduced but not eliminated. Meanwhile, the open split in our continuous-flow IRMS works as a pressure buffering to stabilize the gas flow entering the ion source (Sessions, 2006).

## H<sub>3</sub>-factor

The formation of  $\text{H}_3^+$  ions in the source is the most significant linearity issue in changing  $\delta^2\text{H}$  values with peak intensity.  $\text{H}_3^+$  ions interfere with  $\text{HD}^+$  as they have the same  $m/z$  3, leading to an



overestimation of the HD/HH ratio.  $H_3^+$  ions are produced through the reaction of  $H_2 + H_2^+ \rightarrow H_3^+ + H$ , and their formation is proportional to the pressure of hydrogen gas in the ion source. To correct the contribution of  $H_3^+$  to the m/z of 3 signal, the  $H_3$ -factor, representing the linear correlation between  $H_2$  pressure and the portion of  $H_3^+$ , can be utilized. A low and stable  $H_3$ -factor, normally lower than 10 ppm/nA, is ideal for accurate results (Sharp, 2017). The  $H_3$ -factor was measured daily before every analytical run and presented (Figure 13).

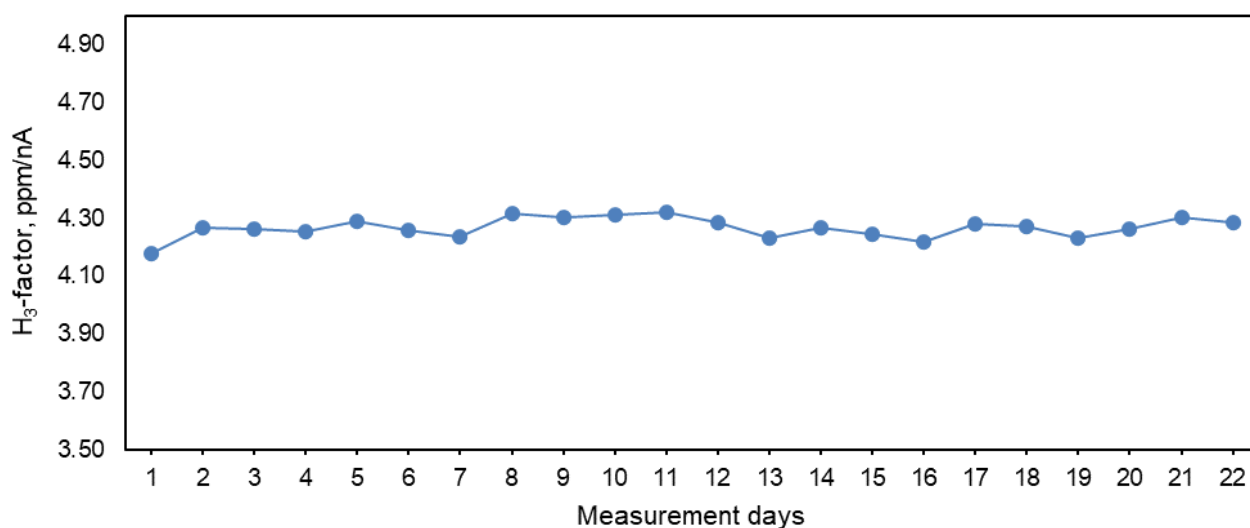


Figure 13: Monitoring the daily  $H_3$ -factors.

### Linearity range for H isotope analysis of FAME

In the GC-Py-IRMS system, non-linearity in H-CSIA can arise from other various sources, such as column temperature gradients and split ratio, due to the non-quantitative transfer of the analyte from the injector to the column (Meier-Augenstein et al., 1996; Schmitt et al., 2003). This effect is called sample size-dependent isotopic fractionation. These sources reduce the useful dynamic range and necessitate linearity checks for system validation to ensure that the measured isotope ratio is within acceptable limits. It can be achieved by injecting varying amounts of the target compounds through the entire analytical path over the desired signal dynamic range (Zhang et al., 2012). For FAMES analysis, different amounts of the reference materials (C16:0, C18:1, C18:2, and C18:3 FAMES) dissolved in n-heptane were injected into the system, and their isotope ratios were measured. The stable ranges for the measurement of the  $\delta^2H$  values of the four FAs were determined, as shown in Figure 14. They present the relationship between  $\delta^2H$  values of FAMES and their amounts ( $H_2$  Intensity).

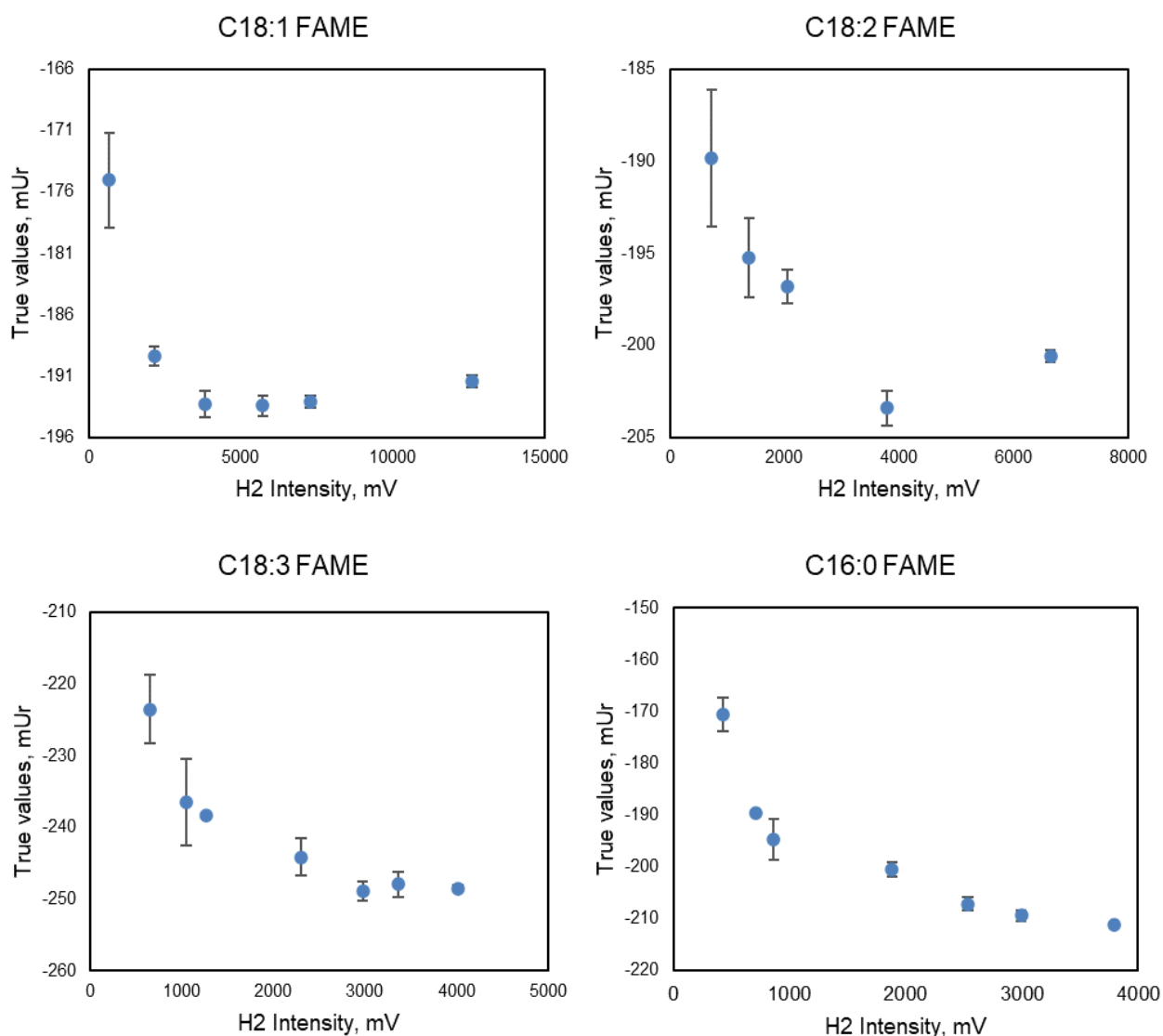


Figure 14: Linearity range for determining the  $\delta^2\text{H}$  values of the four FAMEs (C18:1, C18:2, C18:3, and C16:0 FAMEs) with the GC-Py-IRMS system.

### 3.3.2. Correction of the $\delta^2\text{H}$ of FAs from FAMES

During the production of FAMES from rapeseed, the  $\text{CH}_3$  group of methanol is added to FAs, causing the  $\delta^2\text{H}$  values of FAMES to differ from those of FAs. Hence, mass balance correction equations are used to calculate the  $\delta^2\text{H}$  values of individual FA from its corresponding FAME (Chivall et al., 2012; Eley et al., 2018).

$$(H_n + 3) \times \delta^2\text{H}_{\text{FAME}} = H_n \times \delta^2\text{H}_{\text{FA}} + 3 \times \delta^2\text{H}_{\text{MeOH}}$$

Here,  $H_n$  represents the number of hydrogen atoms in the FA, and  $\delta^2\text{H}_{\text{FAME}}$ ,  $\delta^2\text{H}_{\text{FA}}$ , and  $\delta^2\text{H}_{\text{MeOH}}$  represent the hydrogen isotope compositions of FAME, FA, and the methyl group of methanol, respectively.

To determine the  $\delta^2\text{H}$  of methanol, phthalic acid ( $\text{C}_8\text{H}_6\text{O}_4$ ) with known hydrogen isotope values (-81.9 mUr) was employed. The phthalic acid, which was purchased from Arndt Schimmelmann at the

University of Indiana, Bloomington, IN, USA, was chosen due to its ability to increase the ratio of added methyl hydrogen to analyte hydrogen in the diester product during the derivatization process (Chivall et al., 2012). This acid undergoes the same process as the seed samples for derivatization. The hydrogen isotope ratio of the derivatized phthalic acid, dimethyl phthalate ( $C_{10}H_{10}O_4$ , Figure 15), was then determined with the GC-C-IRMS. The hydrogens on the methyl groups (or the derivatizing agent) that have been added were calculated by comparing their isotopic compositions before and after derivatization. This calculation was performed based on the mass balance correction equation. Two bottles of methanol were employed for the FAMES preparation process from rapeseed. Each bottle of methanol was esterified with phthalic acid to produce two different dimethyl phthalates. The hydrogen isotope values of these dimethyl phthalates, measured by GC-Py-IRMS, were found to be -142 mUr and -101 mUr, respectively. The  $\delta^2H$  values of the two methanol bottles were calculated with the mass balance equation, -182.1 mUr and -113.7 mUr, respectively. The  $\delta^2H$  of FAs of rapeseed was corrected based on the mass balance correction equation and the  $\delta^2H$  values of methanol.

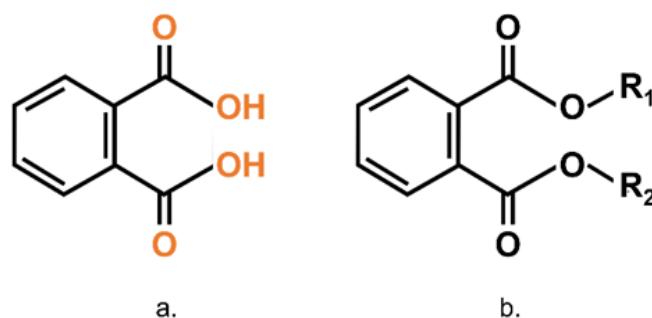


Figure 15: a. Structure of the phthalic acid; b. Derivatised phthalic acid, R<sub>1</sub> and R<sub>2</sub> here are acetyl groups.

### 3.4. Elemental analysis

ICP-OES and ICP-MS techniques were employed to determine elements in rapeseed samples. Ground rapeseed samples (approximately 0.5-1 g) were used for analysis. 8 ml of high purity 60% HNO<sub>3</sub> (EMSURE, Merck) and 2 ml of 30% H<sub>2</sub>O<sub>2</sub> (ROTIPURAN, Roth) were added to digest each ground rapeseed sample. This mixture was allowed to react before the microwave program, enabling the venting of CO<sub>2</sub> gas produced from the reactions due to the organic composition of the rapeseeds. This step was essential to prevent pressure excess during the subsequent heating process. Based on prior experience, it was determined that rapeseed samples weighing more than 0.89 g required additional digestion time to ensure proper gas emission from the chemical reactions. Therefore, overnight digestion at room temperature was conducted before the microwave program for rapeseed samples exceeding this weight. It ensured adequate gas emission and minimized potential pressure-related issues during the subsequent analysis. Then, the mixture was digested with a microwave oven (Multiwave 3000, Anton Paar) program, as shown in Table 7.

Table 7: Program of the microwave oven for digestion of rapeseeds.

Phase	Power (w)	Ramp (w/min)	Holding time (min)	Cooling time (min)
1	250	03:00	05:00	1
2	400	03:00	05:00	1
3	650	03:00	20:00	1
4	0	05:00	10:00	2

Note: Reagent: 2 ml H<sub>2</sub>O<sub>2</sub> + 8 ml HNO<sub>3</sub>.

The elements Ag, As, Ba, Ca, Cd, Co, Cr, Cs, Cu, Fe, K, Mg, Mn, Mo, Nb, Ni, Pb, Rb, Sb, Se, Sn, Sr, Tl, Th, V, U and Zn in the digested diluted solutions were measured. Among these, several elements serve as essential micronutrients and macronutrients for plant growth. Other elements, although not classified as essential nutrients, can still be beneficial for plant growth under certain conditions. Additionally, elements present in soil or rainwater, such as As, Cd, Pb, and U, can be taken up by plants. The concentration of macro elements was determined by the ICP-OES (Spectro Arcos, Spectro Analytical Instruments) working under the following operating conditions: RF power, 1.45 kW; plasma Ar flow rate, 13 L/min; auxiliary Ar flow rate, 0.7 L/min; nebulizer Ar flow, 0.8 L/min. The concentration of trace elements was determined by the ICP-MS (PlasmaQuant MS Elite, Analytik Jena) working under the following operating conditions: RF power, 1.3 kW; plasma Ar flow rate, 9 L/min; auxiliary Ar flow rate, 1.65 L/min; nebulizer Ar flow, 1.03 L/min. Helium, as a collision gas, was used in the integrated collision-reaction cell to minimize potential interferences. The stock solutions of those elements with 1000 mg/L in 2% (v/v) nitric acid (all Merck, Darmstadt, Germany) served for the preparation of calibration solutions.

### 3.5. Soil analysis

#### Soil Particle-size distribution

Particle-size distribution (PSD) is a crucial soil characteristic that significantly influences various properties, including pore distribution, water retention, water conductivity, thermal conductivity, and sorption (Ryżak & Bieganowski, 2011). The equivalent diameter representing particle size in a single soil sample can range from very coarse (>2 mm) to very fine (<2 µm). In the German classification system, the sand fraction, silt fraction, and clay fraction correspond to particles in the range of 2000-63 µm, 63-2 µm, and <2 µm, respectively (Amelung et al., 2018).

Traditional methods for PSD determination involve sieving for the sand fraction and sedimentation with a pipette and hydrometer for the silt and clay fractions. These methods necessitate a combined approach and are time-consuming. In contrast, the laser diffraction technique provides rapid results by measuring the intensity of light scattered as a laser beam passes through a dispersed particulate sample. The particle size is then calculated based on this collected data.

The PSD measurement of the soil samples was performed on the Mastersizer 3000 analyzer, coupled with a Hydro MV unit (Malvern Ltd., UK), capable of determining particle sizes ranging from 0.01 to 3500 µm. A small amount of the finely ground soil sample was placed in the tank, then filled with distilled water as a dispersion medium. Stirring at 2500 rpm and ultrasonication for 5 minutes

---

ensured sufficient dispersion and removal of bubbles. Subsequently, the pump circulated the mixture through the system and the wet cell, facilitating measurement with the optical unit. Red laser light (wavelength: 632.8 nm) and blue light (wavelength: 470 nm) were transmitted through the sample for 30 seconds each. Detectors generated data on the light scattering patterns of the particles, which were then converted to PSD data using the Fraunhofer theory (Ryżak & Bieganowski, 2011).

### **Soil pH**

Soil pH values were determined following the standard method VDLUFA A5.1.1. A 10g/25mL soil-to-water ratio was prepared, and distilled water was added to the soil samples. The mixture was allowed to equilibrate to ensure proper suspension. pH measurements were conducted using a calibrated pH meter.

### **Soil nutrients (P<sub>2</sub>O<sub>5</sub>, K<sub>2</sub>O and Mg)**

The concentration of plant-available nutrients phosphorus (P), potassium (K), and magnesium (Mg) in soil samples was determined according to the procedure stated in A 6.2.1.1 and A 6.2.4.1 of the VDLUFA methods book. The P<sub>2</sub>O<sub>5</sub> and K<sub>2</sub>O of soil samples were extracted with the calcium-acetate-lactate solution. The filtrate was buffered to a pH of 4.2 and determined colorimetrically using a spectrophotometer with a wavelength of 580 nm for P. The K concentration was measured by a flame photometer with a wavelength of 767 nm. The Mg concentration was determined using the calcium chloride (CaCl<sub>2</sub>) extraction method. The extract was determined by an atomic absorption spectrometer (Producer of the instrument) with a wavelength of 285.2 nm.

### **Electrical conductivity**

The electrical conductivity of soil samples was measured according to DIN CEN/TS 15937. An aqueous suspension with a soil-water ratio of 1:5 (g: mL) was prepared. After shaking for 30 minutes and filtration of the suspension, the filtered extract was determined using a conductivity meter.

## **3.6. Acquisition of climate and soil data of sampling sites**

The climatic and soil parameters of the sampling sites from 2017 to 2020 were extracted from various publicly available sources, presented in Table 8. The monthly average values were required. When extracting monthly average data, it is important to consider that the rapeseed data from different years should correspond to the period from September of the previous year to July of the current year, as the growth cycle of rapeseed involves autumn sowing, typically around September, and the harvest takes place approximately in July of the following year.

Table 8: List of the publicly available datasets.

Parameter	Source	Spatial resolution	Temporal resolution	Web
surface soil moisture, %	Copernicus global land service	1 km x 1 km	daily	<a href="#">link</a>
soil temperature in 5 cm, °C	DWD* Climate Center	Data 1 km x 1 km	monthly	<a href="#">link</a>
soil moisture in 60 cm, %nFK	DWD Climate Center	Data 1 km x 1 km	monthly	<a href="#">link</a>
air temperature, °C	Meteostat	1 km x 1 km	monthly	<a href="#">link</a>
precipitation amount, mm	Meteostat	1 km x 1 km	monthly	<a href="#">link</a>
sunlight duration, hours	DWD Climate Center	Data 1 km x 1 km	monthly	<a href="#">link</a>
drought index, mm/°C	DWD Climate Center	Data 1 km x 1 km	monthly	<a href="#">link</a>
geological substrate	HLUNG**	1:50000	n.a.	<a href="#">link</a>
soil types	HLUNG	1:5000	n.a.	<a href="#">link</a>

\* DWD: Deutscher Wetterdienst.

\*\* HLUNG: Hessisches Landesamt für Naturschutz, Umwelt und Geologie.

### 3.7. Statistical analysis

Statistical analysis was performed using SPSS software (version 28.0, SPSS Inc., Chicago, IL) and SIMCA software (version 14.1, Sartorius-Umetric, Umeå, Sweden). Several statistical analysis methods were employed to address different research questions and can be generally categorized into three groups:

#### Comparison of means

ANOVA (Analysis of Variance) is used to compare means between multiple groups, whereas a t-test was employed to compare means between two groups. In addition to requiring observations within each group to be independent, ANOVA assumes that data are normally distributed and that the variances among the groups are approximately equal (McDonald, 2009). These assumptions can be checked using the Shapiro-Wilk and Levene's tests, respectively, which can be conducted using SPSS Software (Shapiro & Wilk, 1965). Meeting these assumptions ensures the validity of the ANOVA test results. The p-value is calculated to assess the significance level of the result. It represents the probability of observing the observed data or more extreme results, assuming the null hypothesis is true (McDonald, 2009). A p-value less than the preset significance level (at 0.05) indicates statistical significance, suggesting that the observed differences are unlikely to have occurred due to chance alone. On the other hand, a p-value greater than the significance level suggests that the observed differences could probably be attributed to random variation (McDonald, 2009).

---

## Correlation analysis

Pearson correlation and Spearman correlation can be used to measure the linear relationship between two variables. Pearson correlation is used for two normally distributed continuous variables, while Spearman correlation is used for the non-normally distributed data variables. The correlation coefficient ( $r$ ) of both tests indicates the strength and direction of the relationship, which ranges from -1 to 1. A negative  $r$  means the variables are inversely related. Table 9 interprets the correlation strength based on the  $r$  values (Newcastle University, n.d.). The  $p$ -value represents the probability that the observed strength of the correlation could happen by random chance (Akoglu, 2018).

Stepwise linear regression analysis was used to determine the environmental factors that significantly impact the hydrogen isotopic composition of FAs of rapeseed in the thesis. This statistical approach is suitable when dealing with many interrelated independent variables (environmental factors) that require investigation (Yan & Su, 2009). By iteratively adding and removing variables, namely environmental factors, based on their contribution to the predictive power of the model, the analysis allowed for the identification of the most important factors influencing the hydrogen isotopic composition of FAs (McDonald, 2009). The regression coefficient is used as a statistical measure to quantify the strength and direction of the relationship between the dependent and independent variables. The standardized regression coefficient (beta coefficient) is used to compare the influence of different independent variables on the dependent variables, especially when the independent variables are measured in different units. The larger the standardized coefficient, the greater the effect of the independent variable on the dependent variable. It means how many standard deviations the dependent variable will change for every standard deviation increase in the independent variable (Baltes-Götz, 2022).

Table 9: Interpretation of correlation coefficients.

Absolute $r$ value	Interpretation
$r = 1$	Perfect linear correlation
$1 > r \geq 0.8$	Strong linear correlation
$0.8 > r \geq 0.4$	Moderate linear correlation
$0.4 > r > 0$	Weak linear correlation
$r = 0$	No correlation

Furthermore, the R-squared value determines the proportion of variance in the dependent variable that could be explained by the independent variables in the model. However, it should be noted that as more independent variables are added, the R-squared value tends to increase, potentially leading to overfitting and inflated R-squared values. In contrast, the adjusted R-squared value adjusts the effect due to the number of predictors in the model and provides a more accurate assessment of the goodness of fit of the model (McDonald, 2009).

## Discriminant analysis (PCA and PLS-DA)

PCA and OPLS-DA are multivariate statistics that cover the simultaneous analysis of one or more dependent variables against two or more independent variables (Granato et al., 2018). Both were

---

employed to analyze the hydrogen isotopic composition of FAs and elements for geographical origin discrimination.

PCA is an unsupervised technique that reduces data dimensionality and identifies principal components representing significant variation. The principal components are linear combinations of the original variables in the dataset, each with specific weights, and these components are uncorrelated with each other. It can explore data structure and detect outliers (Jolliffe & Cadima, 2016).

PLS-DA is a supervised version of PCA that reduces dimensionality with awareness of class labels. It identifies linear combinations of variables that best discriminate between groups (Ruiz-Perez et al., 2020). OPLS-DA is a variation of Partial Least Squares Discriminant Analysis (PLS-DA) incorporating an orthogonal signal correction step. It separates the variation in the data into predictive and orthogonal components. The predictive component captures the variation correlated with the class labels and is used for discrimination and classification. The orthogonal component captures variation unrelated to the class labels (Bylesjö et al., 2006).

However, these methods carry the risk of overfitting, where the model fits the training set well but fails to predict new samples accurately. Therefore, a model diagnostic must assess the model performance based on cross-validation (CV) (Kjeldahl & Bro, 2010). This technique involves partitioning the dataset into multiple subsets or folds. The model is trained on one subset and then evaluated on the remaining data, repeating this process multiple times with different subsets used for training and evaluation each time. By calculating the average results from these iterations, cross-validation provides a more reliable estimate of the performance of the model, enhancing its robustness (Yan & Su, 2009). In SIMCA software, the default cross-validation method is 7-fold cross-validation.



## 4. Assessment of the one-step sample preparation method for C- and H-CSIA of FAs in rapeseed oil

### 4.1. Introduction

This chapter focuses on the first research objective, which is to enhance efficiency and streamline laboratory work in the compound-specific stable hydrogen isotope analysis of rapeseed and its oil. A simplified and rapid sample preparation method developed by Garcés & Mancha (1993) enables the one-step conversion of seeds into FAMES, which can be directly analyzed by GC-Py-IRMS. However, the feasibility of this method for stable isotope analysis has not been previously reported. Therefore, in the following Subchapter 4.2, this one-step method is introduced first, and concerns regarding its potential impact on the accuracy of  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  of FAs are identified. Subsequently, the results of the comparative experiments are presented in Subchapter 4.3 to address these concerns. Experimental results and conclusions are outlined in Chapters 4.4 to 4.5.

### 4.2. Basics of the one-step sample preparation method developed by Garcés and Mancha

Garcés and Mancha have developed a simple way to quickly prepare samples without the drawbacks of the DIN method, as mentioned in Chapter 2.5. Their method involved a one-step process that allowed for the extraction of lipids from seeds, transesterification, methylation, and the subsequent extraction of the produced FAMES, all within a few hours. They applied a transmethylating mixture consisting of methanol, toluene, 2,2-dimethoxypropane (DMP), sulfuric acid, and n-heptane (39:20:5:2:34 vol.%). This mixture was combined with ground seeds and heated at 80°C for 2 hours. The resulting supernatant could be directly used for the GC measurement (Garcés & Mancha, 1993). The sample preparation process is presented in Figure 16.

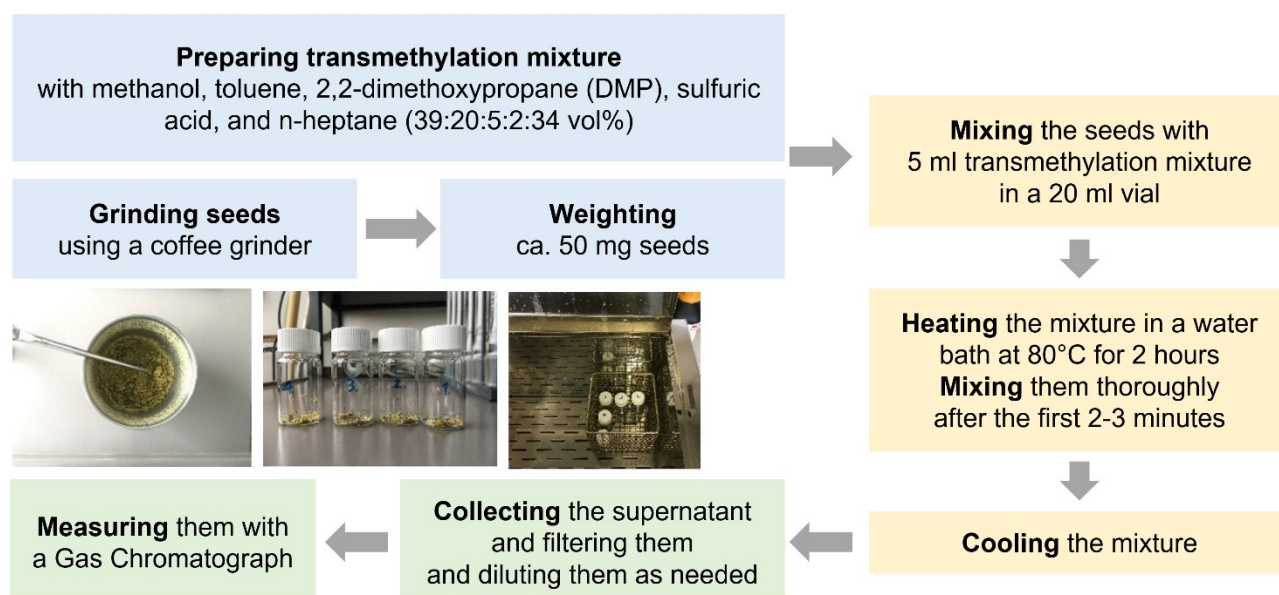
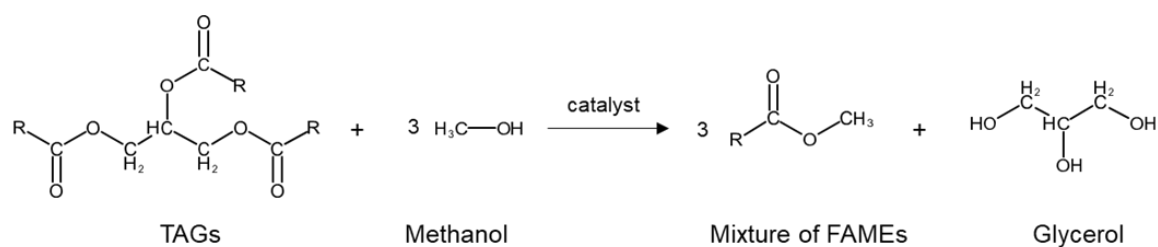
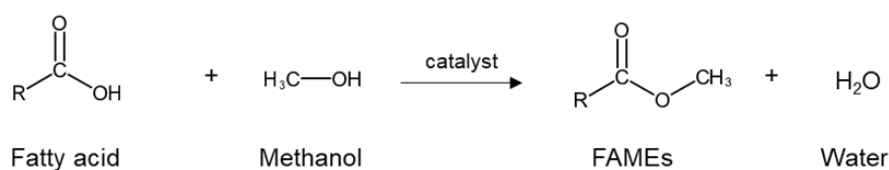


Figure 16: Flow chart of the one-step sample preparation method.

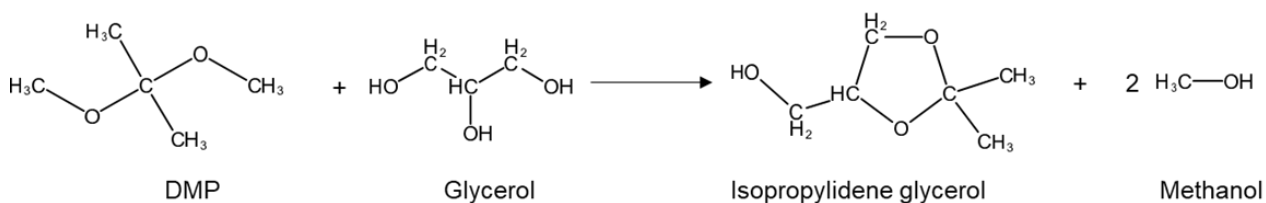
The addition of DMP to the reaction mixture enables the simultaneous occurrence of transesterification and methylation reactions. Glycerol, a byproduct of the transesterification reaction, reacts with DMP to produce isopropylidene glycerol and methanol. Removal of glycerol accelerates the transesterification reaction, and the methanol produced can be utilized as a reactant for the transmethylation/methylation processes. Additionally, DMP reacts with water generated during the methylation process to form acetone and methanol. The presence of water impedes the transmethylation reaction, so removing it from the system helps facilitate the methylation process. Finally, the FAMES that result from these two concurrent reactions can be extracted directly using n-heptane in the same reaction tube. The reaction processes are presented in Figure 17.



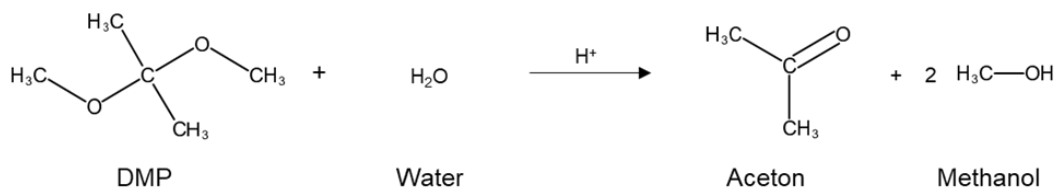
**(1) Transmethylation reaction**



**(2) Methylation reaction**



**(3) Reaction between DMP and glycerol**



**(4) Reaction between DMP and water**

Figure 17: Four major reactions during the one-step sample preparation process.

---

While the one-step method of Garcés and Mancha has been verified for determining the lipid content and FA profile of various oilseeds, its application to determining the isotope composition of individual FAs has not been reported. One concern arises when applying this method for CSIA, specifically for determining the  $\delta^2\text{H}$  and  $\delta^{13}\text{C}$  values of FAs. After preparing the FAMES, the next step involves determining the  $\delta^2\text{H}$  and  $\delta^{13}\text{C}$  of FAs by combining the  $\delta^2\text{H}$  and  $\delta^{13}\text{C}$  of FAMES and those of methanol using the following mass conservation equations. The  $\delta^2\text{H}$  and  $\delta^{13}\text{C}$  values of the FAMES and the initially added methanol can be measured using GC-Py-IRMS. However, due to the chemical reaction presented in Figure 17, the use of DMP in methanol leads to the formation of reaction intermediate methanol, which cannot be directly measured. Therefore, DMP might impact the accuracy of FA isotope measurements through the presence of reaction intermediate methanol.

$$(C_n + 1) \times \delta^{13}\text{C}_{\text{FAME}} = C_n \times \delta^{13}\text{C}_{\text{FA}} + \delta^{13}\text{C}_{\text{MeOH}}$$

$$(H_n + 3) \times \delta^2\text{H}_{\text{FAME}} = H_n \times \delta^2\text{H}_{\text{FA}} + 3 \times \delta^2\text{H}_{\text{MeOH}}$$

Evaluation of the feasibility of using the one-step method of Garcés and Mancha for determining the C- and H-CSIA of rapeseed includes two specific objectives: (1) determining the impact of sample aggregate states on the reproducibility of C- and H-CSIA; (2) evaluating the effect of DMP, particularly the reaction intermediate methanol, on the accuracy of the C- and H-CSIA of FAMES. The methanol produced during the one-step process is presented in Figure 17.

### 4.3. Experiment

#### Materials

A bottle of rapeseed oil was acquired from a local shop in Dottenfelderhof, Germany, while a cultivated rapeseed sample from Hesse, Germany, was provided by the Landesbetrieb Landwirtschaft Hesse. In each experiment, 20 mg of rapeseed oil or 50 mg of rapeseed was employed for the one-step extraction process. All experiments were performed in triplicate. To prepare the rapeseed, 10 g of samples were ground using a coffee grinder for 1 minute to achieve a fine and homogeneous powder.

All solvents and reagents were of analytical grade or higher purity. For assessing the recovery rate of FAMES in the n-heptane phase and determining the FAME yield, Supelco® nonadecanoic acid methyl ester (C19:0 FAME; Merck, Darmstadt, Germany) was used as an internal standard. Individual FAMES were quantified using Supelco® 37 component FAME Mix and ROTICHROM® FO 7 FAMES mixture (Carl Roth, Karlsruhe, Germany), which consists of 11 FAMES. Additionally, ROTICHROM® oleic acid methyl ester (C18:1 FAME) served as an in-house reference material for quality control of the isotope measurements. Two stable isotopic reference materials, USGS 70 and USGS 71, were obtained from the USGS Reston Stable Isotope Laboratory (Reston, VA, USA). These reference materials were utilized to normalize  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  of FAMES. Specifically, the compound used for this purpose was icosanoic acid methyl esters (C20:0 FAME). The instrument GC-Py-IRMS system was introduced in Chapter 3.3.

#### Experimental procedure

Three different transmethylations with and without DMP were prepared for the comparative experiments: TM+D, TM-D, and TX-D. The volume ratio of different components in these mixtures

is presented in Table 10.

TM+D mixtures containing 5% DMP were found to be the most effective for transmethylation by Garcés and Mancha. By substituting methanol for DMP in the TM+D mixture, TM-D was formulated without DMP. A similar formulation was made for TX-D, which does not contain DMP. Methanol was used for TX-D and TM-D in volumes of 3.9 mL (0.096 mol) and 2.2 mL (0.054 mol), but the other components remained unchanged. The transmethylation mixtures were first prepared without n-heptane and stored in the refrigerator.

Table 10: Composition of transmethylation mixtures for every 20 mg of oil or 50 mg of seeds.

	Reagents	Volume (ml)			Amount (mmol)		
		TM+D	TM-D	TX-D	TM+D	TM-D	TX-D
Transmethylation mixture	Methanol	1.95	2.2	3.9	48.2	54.4	96.4
	Toluene	1	1	1	9.4	9.4	9.4
	DMP	0.25	0	0	2	0	0
	Sulphuric acid	0.1	0.1	0.1	1.9	1.9	1.9
<b>N-heptane</b>		1.7	1.7	2	11.6	11.6	13.7
<b>Sum</b>		5	5	7	73.1	77.3	121.3

Nine portions of 50 mg ground rapeseed each, approximately equivalent to 20 mg of rapeseed oil, were taken from the homogenized batch. Each of the three transmethylation mixtures was treated with three portions of samples, using the same conditions for each. To create the reaction, 20 mg of rapeseed oil or 50 mg of ground rapeseed was placed into a 20 mL crimp-top glass vial, also known as a reactor. Before the reaction began, a specific volume of TM+D, TM-D, or TX-D was added to each reactor, along with 1.7 mL of n-heptane for the TM+D and TM-D reactors. The reactors with TX-D received 2 mL of n-heptane once the reaction was finished.

## 4.4. Results and discussion

### 4.4.1. Impact of DMP on transmethylation efficiency of the one-step method

To study the impact of DMP on transmethylation efficiency, the FA profile and absolute mass of FAMES were assessed using three transmethylation mixtures: TM+D (with DMP), TM-D, and TX-D. The TM+D mixture comprised 5 vol% DMP and 39 vol% methanol, while TM-D and TX-D mixtures, prepared without DMP, contained 44 vol% and 56 vol% methanol, respectively.

The FA profile (mol%) was determined by calculating the mass percentage of individual FAs relative to their respective molecular weights. Table 11 summarizes the mean and standard deviation (SD) of each FA composition produced from the three mixtures for comparative purposes. The FA compositions were calculated based on the mol% values of the three mixtures. Additional details can be found in Table A-1 in the Appendix.

Table 11: FA composition (mol%) of a rapeseed sample using three transmethylation mixtures.

	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1	C22:0	C22:1	C24:0	C24:1
<b>TM+D</b>	0.10	5.50	0.40	1.70	61.60	18.20	10.00	0.50	1.20	0.30	0.30	0.10	0.10
<b>TM-D</b>	0.10	5.60	0.40	1.70	61.40	18.40	10.10	0.50	1.20	0.30	0.30	0.10	0.10
<b>TX-D</b>	0.10	5.50	0.30	1.80	62.50	17.70	9.30	0.60	1.30	0.30	0.40	0.10	0.20
<b>Mean</b>	0.10	5.50	0.30	1.70	61.80	18.10	9.80	0.50	1.20	0.30	0.30	0.10	0.10
<b>SD</b>	0.02	0.06	0.05	0.09	0.61	0.36	0.47	0.05	0.04	0.03	0.04	0.03	0.03

The predominant FA components in rapeseed oil were C18:1 (61.8% ± 0.6%), C18:2 (18.1% ± 0.4%), C18:3 (9.8% ± 0.5%), and C16:0 (5.5% ± 0.1%). Remarkably, the FA composition remained consistent regardless of the presence or absence of DMP in the transmethylation mixtures.

To determine the absolute mass of each of the four main FAs, a known quantity of C19:0 FAME internal standard was added to each reactor. When using the TM+D mixture, the absolute mass per 100 g of rapeseed for all major FAMES was higher than that obtained from TM-D and TX-D. Therefore, it can be concluded that adding DMP in the transmethylation process enhances the absolute mass of major FAMES in rapeseed oil, as demonstrated by the results presented in Figure 18 and Table A- 2 in the Appendix.

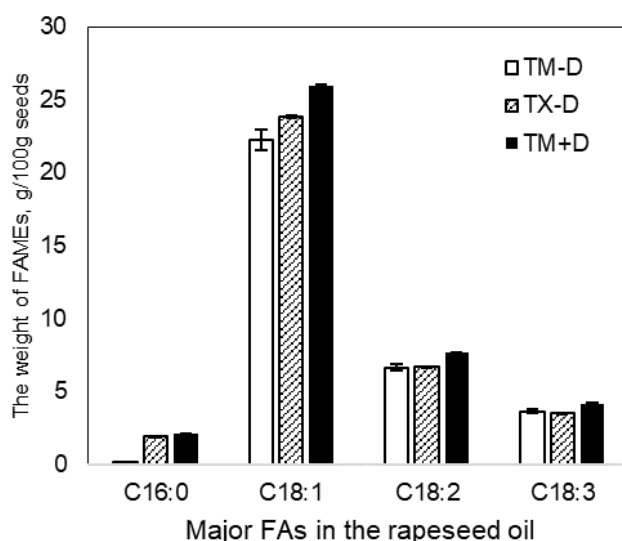


Figure 18: The weight of four major FAMES (C16:0, C18:1, C18:2, and C18:3) in 100 g of the rapeseed sample prepared using three transmethylation mixtures.

Based on the data provided in Table A- 2 in the Appendix, the yield of FAMES obtained from TM-D and TX-D was calculated to be 86% and 91%, respectively, relative to the FAMES yield obtained from TM+D. These results indicate the advantageous effect of DMP on transmethylation efficiency in converting TAGs to FAMES. By replacing 5 vol% methanol in TM-D with 5 vol% DMP from TM+D, the yield increased from 86% to 100%. This finding aligns with Garcés and Mancha's previous research, which demonstrates that the presence of DMP significantly improves transmethylation efficiency when the proportions of methanol, H<sub>2</sub>SO<sub>4</sub>, and oil in the mixtures are maintained.

Additionally, it is noteworthy that the increase in methanol content from 44 vol% (TM-D) to 56 vol% (TX-D) only contributed to a 5% increase in FAME yield.

Figure 17 illustrates the stoichiometric relationship between TAGs and methanol, where 1 mol of TAGs requires 3 moles of methanol to produce 1 mol of glycerol and 3 moles of FAMES. However, an excess amount of methanol ensures a high transmethylation reaction rate. Zheng et al. recommended a molar ratio of methanol to oil at 245:1, along with an acid catalyst, to achieve the maximum FAME yield from frying oil. They observed a 27% reduction in FAME yield when the methanol-to-oil molar ratio was lowered to 106:1 (Zheng et al., 2006).

TAGs consist of FAs esterified to a glycerol backbone (FD Gunstone, 2002). The FA profile in Table 11 allows us to estimate the average molar mass of the TAGs in rapeseed oil to be approximately 881 g/mol (refer to Table A- 3 in the Appendix). Consequently, 50 mg of rapeseed, containing 20 mg of rapeseed oil, can be equated to 0.02265 mmol. Based on this value, the methanol-to-oil molar ratios for the transmethylation mixtures were calculated and are presented in Table 12. The resulting molar ratios for TM+D, TM-D, and TX-D are 2128:1, 2401:1, and 4256:1, respectively. These ratios indicate that the amount of methanol used in the reaction is sufficient, regardless of the presence of DMP. Therefore, the observed increase in yield for TM+D is unlikely to be solely attributed to an increase in methanol content resulting from the use of DMP.

Table 12: Molar ratios of each component to oil and FAME yield in the transmethylation mixtures.

	TM+D	TM-D	TX-D
<b>Rapeseed oil</b>	1	1	1
<b>Methanol</b>	2128	2401	4256
<b>Toluene</b>	415	415	415
<b>DMP</b>	90	0	0
<b>Sulphuric acid</b>	82	82	82
<b>Fame yield</b>	100%	86%	91%

It is assumed that including DMP in the transmethylation reaction has two primary effects: 1) consumption of glycerol, a byproduct of the reaction, and 2) water removal from the system.

The addition of DMP ensures the complete conversion of lipids by consuming the reaction product glycerol (Figure 17). From 20 mg of rapeseed oil (equivalent to 0.02265 mmol), 0.02265 mmol of glycerol was expected to be produced. The amount of DMP (2 mmol) added to the reaction was in excess for the reaction between DMP and glycerol. If all the glycerol reacts with DMP, an additional 0.453 mmol of methanol would be produced. It would increase the methanol to oil molar ratio from 2128:1 to 2130:1. However, considering that the increase in the methanol to oil molar ratio from 2401:1 (TM-D) to 4256:1 (TX-D) only contributed to a 6% increase in yield, it was unlikely that the additional methanol generated through the reaction of DMP and glycerol in TM+D had a significant influence on the observed yield increase.

Furthermore, the presence of DMP exhibits the ability to react with water, thereby promoting the transmethylation reaction (refer to Figure 17). Water in the system can lead to oil hydrolysis, suppressing the transmethylation reaction. Typically, rapeseed contains around 7-10% water and 0.5%-1.8% free fatty acids (FFAs) under optimal storage conditions (FD Gunstone, 2002). It is

important to note that the methylation reaction of FFAs also generates water (as shown in Figure 17). Additionally, solvents often contain small quantities of water. In our case, the methanol used as a reagent contained approximately 0.02% water.

As the amount of methanol increased from 48.2 mmol (TM+D) to 54.4 mmol (TM-D) and further to 96.4 mmol (TX-D), the quantity of water associated with the methanol product also increased from 0.0420 mmol and 0.484 mmol to 0.0858 mmol, respectively. It was generally observed that an increase in methanol content enhanced FAME yield. However, it is worth mentioning that the increase in water associated with the methanol product, if DMP is not used, can inhibit the production of FAMES to some extent (Garcés & Mancha, 1993; Ichihara & Fukubayashi, 2010). It can potentially explain why the transmethylation efficiency of TX-D remained stagnant at 91% despite adding a significant amount of methanol.

#### 4.4.2. Impact of DMP on $\delta^{13}\text{C}$ and $\delta^2\text{H}$ values of FAMES using the one-step method

A preliminary investigation was conducted to ensure the reliability of GC-C/Py-IRMS analysis and to assess the reproducibility of H- and C-CSIA of FAMES. This investigation involved the analysis of two FAME reference materials and evaluating how the aggregate states of samples influence the precision of the isotopic measurements. These steps were carried out before examining the impact of DMP on the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of FAMES generated through different transmethylation mixtures in the rapid one-step sample preparation method.

##### Precision of GC-C/Py-IRMS

The precision of the GC-C/Py-IRMS analysis was evaluated by analyzing two in-house FAME reference materials, specifically oleic acid methyl ester (C18:1 FAME) and nonadecanoic acid methyl ester (C19:0 FAME). The analysis was conducted nine times ( $n = 9$ ) using the GC-C/Py-IRMS unit. The mean and standard deviation of the repetitive measurements were determined to assess the precision of the C- and H-CSIA.

The standard deviation ( $1\sigma$ ) of the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values obtained from the repetitive measurements of the reference materials indicated the precision of the C- and H-CSIA achieved by the GC-C/Py-IRMS unit. The results, presented in Table 13, revealed a precision of  $\pm 0.1$  mUr for C-CSIA and  $\pm 1.7$  mUr for H-CSIA, respectively, for both FAMES. These findings demonstrate the favorable stability of the instrument in performing C- and H-CSIA measurements of FAMES.

Table 13: Precision of the C- and H-CSIA by analyzing two in-house reference materials by GC-C/Py-IRMS.

In-house reference materials	$\delta^{13}\text{C}$ [mUr]			$\delta^2\text{H}$ [mUr]		
	n	M	SD	n	M	SD
FAME						
C18:1	9	-28.8	0.1	9	-165.1	1.7
C19:0	9	-28.7	0.1	9	-244.3	1.7

#### Influence of the aggregate states of samples on the reproducibility of the H- and C-CSIA of

## FAMES

One notable advantage of the one-step method is eliminating the extraction step for obtaining fluid oil from seeds, which can introduce potential isotopic fractionation effects (Chivall et al., 2012). However, it should be noted that solid seeds, in comparison to fluid oil, possess a coarse and heterogeneous nature. This characteristic can potentially influence the reproducibility of transmethylation and methylation reactions. Furthermore, incomplete and inconsistent conversion of TAGs to FAs and subsequently to FAMES can result in poor reproducibility, thereby impacting the precision of the isotope analysis.

FAMES were prepared from rapeseed oil and ground rapeseed samples under identical conditions utilizing the transmethylation mixture TM+D to address this. Three reactors were created using three separate aliquots of each sample. The range of standard deviations for the carbon and hydrogen compositions of the four main FAMES obtained from the oil and rapeseed samples is summarized in Table 14. It is important to note that due to the differing origins of the fluid oil and rapeseed samples, the mean values also exhibit variations (refer to Figure A- 1 in the Appendix).

Table 14: Comparison of reproducibility of  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of four FAMES of oil and seeds

Aggregate states	SD ( $\pm 1\sigma$ ) of $\delta^{13}\text{C}$ values (n = 3)	SD ( $\pm 1\sigma$ ) of $\delta^2\text{H}$ values (n = 3)
Rapeseed oil	$\pm 0.06$ mUr $\sim \pm 0.1$ mUr	$\pm 0.9$ mUr to $\sim 2.4$ mUr
Rapeseed	$\pm 0.1$ mUr $\sim \pm 0.3$ mUr	$\pm 0.7$ mUr to $\sim 1.8$ mUr

The precision of  $\delta^2\text{H}$  values for the different FAMES obtained from the fluid oil and solid seed samples ranged from  $\pm 0.9$  mUr to  $\pm 2.4$  mUr and  $\pm 0.7$  mUr to  $\pm 1.8$  mUr, respectively. These results indicate that the precision of H-CSIA for the four FAMES produced from the fluid oil sample slightly exceeded the precision of the GC-C/Py-IRMS instrument ( $1\sigma = \pm 1.7$  mUr), while the precision of H-CSIA for the solid rapeseed sample fell within the instrument's precision range.

Regarding C-CSIA, the standard deviation of  $\delta^{13}\text{C}$  values for the four FAMES from the fluid oil sample and the solid seeds ranged from  $\pm 0.1$  mUr to  $\pm 0.1$  mUr to  $\pm 0.3$  mUr, respectively. Unlike H-CSIA, the precision of C-CSIA for the fluid oil sample fell within the instrument's precision range ( $1\sigma = \pm 0.1$  mUr), but for the solid seed sample, it was slightly poorer.

The precision of  $\delta^{13}\text{C}$  values for the different FAMES was consistent with the values reported by Richter et al. ( $\pm 0.05$  mUr to  $\pm 0.4$  mUr) in their study, where FAMES were prepared from rapeseed oil using the DIN Norm method involving saponification and methylation with  $\text{BF}_3$  (Richter et al., 2010).

In summary, the precision of both C-CSIA and H-CSIA for the selected FAMES prepared from both fluid oil and solid seeds is comparable. These findings imply that there is no significant difference in the reproducibility of the conversion process from FAs to FAMES using the one-step method, regardless of the aggregate state of the sample. Furthermore, the precision of C-CSIA and H-CSIA for the FAMES produced from solid rapeseed samples using the one-step method and from fluid rapeseed oil using the DIN method are within a similar range, indicating a comparable level of consistency and completeness in FAME production between the one-step method using the transmethylation mixture TM+D under the selected conditions and the DIN method.



### Influence of DMP on the $\delta^{13}\text{C}$ and $\delta^2\text{H}$ values of FAMES

The formation of FAMES involves the addition of 1 mol  $-\text{CH}_3$  from methanol to 1 mol of FA. Consequently, the isotope composition of FAMES can be influenced by the isotope composition of the methanol used in their production.

If the entire 2.0 mmol of DMP added in the TM+D mixture is consumed in reactions with glycerol and water, it theoretically produces 4.1 mmol of methanol as an intermediate product. Thus, the molar ratio of the reaction intermediate methanol to the initially added methanol is approximately 1:12. Since the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the reaction intermediate methanol cannot be measured directly, the influence of DMP on the isotope composition of FAMES needs to be evaluated by comparing the carbon and hydrogen isotope compositions of individual FAMES produced with and without DMP.

The TM+D and TX-D mixtures were used to produce FAMES from the same batch of rapeseed samples. One reason for this selection is that the reactions in both TM+D and TX-D reactors were expected to reach equilibrium, as demonstrated by Garcés and Mancha under comparable conditions (Garcés & Mancha, 1993). Another reason is that the FAME yields of TM+D (100%, as shown in Table 12) and TX-D (91%, as shown in Table 12) are comparable. Previous research by Panetta and Jahren indicated that FAME yields could influence the  $\delta^{13}\text{C}$  values if the reactions have not reached equilibrium. However, according to their results, the difference in  $\delta^{13}\text{C}$  values between FAME yields of 90% and 100% was smaller than 0.1 mUr (Panetta & Jahren, 2011).

Figure 19 shows the isotope compositions of the four major FAMES obtained using the TM+D (with DMP) and TX-D (without DMP) mixtures. A T-test was performed to compare the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of individual FAMES between these two reactors. The results, presented in Table A-4 and Table A-5 of the Appendix, indicate that there is no significant difference in isotope values between TX-D and TM+D, except for the  $\delta^2\text{H}$  value of C18:3. The difference in hydrogen isotope values for C18:3 is 2.6 mUr, slightly higher than the reproducibility of the seed sample preparation (see Table 14), but similar to that of the oil sample preparation.

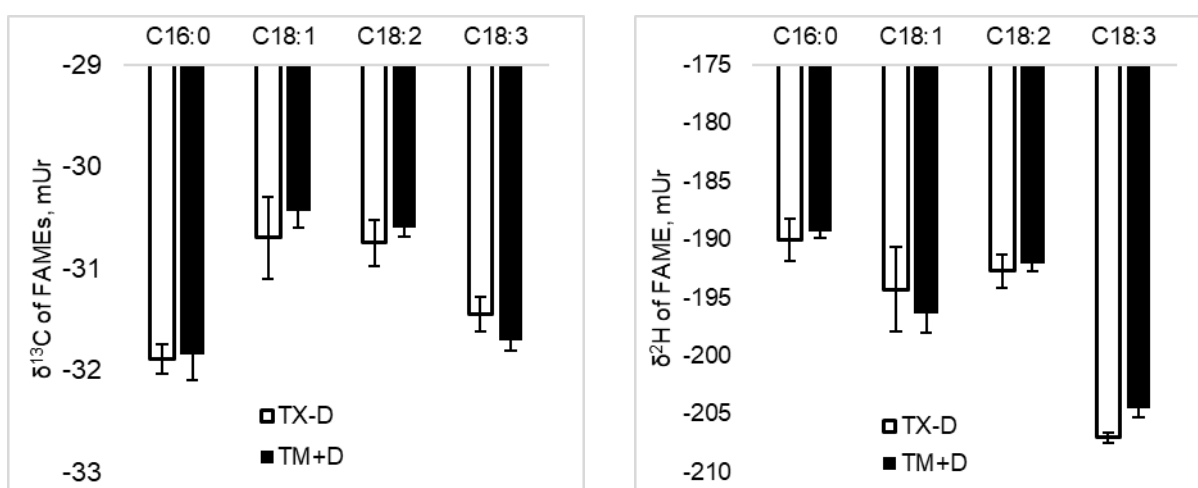


Figure 19: The  $\delta^{13}\text{C}$  (left) and  $\delta^2\text{H}$  (right) of C16:0, C18:1, C18:2, and C18:3 FAME in rapeseed using transmethylation mixtures TX-D and TM+D.

The transmethylation reaction is a complex process involving forming bonds between the carboxyl

---

carbon and the incoming methoxide oxygen (Chivall et al., 2012). In simpler terms, each mol of generated FAME consists of 1 mol  $-\text{CH}_3$  from methanol and 1 mol of FA. According to Paolini et al., the relationship between the  $\delta^2\text{H}$  values of FAs and FAMEs can be estimated using the following mass balance equation (Paolini et al., 2017):

$$(H_n + 3) \times \delta^2\text{H}_{\text{FAME}} = H_n \times \delta^2\text{H}_{\text{FA}} + 3 \times \delta^2\text{H}_{\text{MeOH}}$$

Here,  $H_n$  represents the number of hydrogen atoms in the FA, and  $\delta^2\text{H}_{\text{FAME}}$ ,  $\delta^2\text{H}_{\text{FA}}$ , and  $\delta^2\text{H}_{\text{MeOH}}$  represent the  $\delta^2\text{H}$  values of FAME, FA, and the methyl group of methanol, respectively.

Assuming that the C18:3 FAME produced by TX-D is derived solely from the initially added methanol, while that of TM+D is generated from a mixture of initially added methanol and the reaction intermediate methanol, the difference in  $\delta^2\text{H}$  values ( $\Delta\delta^2\text{H}$ ) of methanol ( $\delta^2\text{H}_{\text{MeOH}}$ ) between TM+D and TX-D can be estimated as 28 mUr. This estimation is based on the measured  $\Delta\delta^2\text{H}$  of FAME 18:3 (2.6 mUr) and the mass balance equation. This assumption suggests that the  $\Delta\delta^2\text{H}$  values of both the initially added methanol and the reaction intermediate methanol should also contribute to the differences in  $\delta^2\text{H}$  values of C16:0, C18:1, and C18:2 FAMEs, with calculated  $\Delta\delta^2\text{H}$  values ranging from 2.3 mUr to 2.4 mUr (see Table A- 6 in the Appendix).

However, the result revealed no significant difference in the measured  $\delta^2\text{H}$  values of C16:0, C18:1, and C18:2 FAMEs between TM+D and TX-D. Consequently, a quantifiable influence of the reaction intermediate methanol produced by DMP on the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of FAMEs was not observed. The results indicate that the isotope composition of FAs can be determined as long as the isotope composition of the initially added methanol is known.

#### **4.4.3. Estimation of $\delta^{13}\text{C}_{\text{FAME}}$ and $\delta^2\text{H}_{\text{FAME}}$ values from $\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}}$ and $\delta^2\text{H}_{\text{H}_3\text{CO-DMP}}$ values with the assumptions**

The effect of DMP on the isotope composition of FAMEs could also be estimated if reasonable assumptions are made for the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of  $\text{H}_3\text{CO}-$  group in DMP structure ( $\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}}$ ,  $\delta^2\text{H}_{\text{H}_3\text{CO-DMP}}$ ). This  $\text{H}_3\text{CO}-$  group in DMP is the source of the  $\text{H}_3\text{CO}-$  group in the reaction intermediate methanol, which is subsequently added during methylation to generate the FAME (Figure 17).

DMP is typically produced through the reaction of methanol and acetone under acidic conditions, implying that the  $\text{H}_3\text{CO}-$  group in DMP originates from the reactant methanol (Hsu & Ellis Jr, 1988). Based on the data reported in the literature, the  $\delta^{13}\text{C}$  value and bulk  $\delta^2\text{H}$  value of methanol can vary between -25 mUr and -46 mUr, and between -18 mUr and -227 mUr, respectively, depending on the source (Ai et al., 2014; Fink et al., 2004; Moussa et al., 1990; Paolini et al., 2017). However, as the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the methanol used for our DMP production are unknown, two assumptions were made: 1) the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the  $\text{H}_3\text{CO}-$  group in DMP are assumed to be similar to the values of methanol found in the literature; and 2) it is assumed that all the DMP present in the TM+D mixture is converted into reaction intermediate methanol, which is then utilized for FAME formation. For the estimation of  $\delta^{13}\text{C}_{\text{FAME}}$  and  $\delta^2\text{H}_{\text{FAME}}$  in FAMEs generated using the TM+D mixture, the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the  $\text{H}_3\text{CO}-$  group in DMP are considered to be in the range of -80 mUr to -20 mUr and -300 mUr to -20 mUr, respectively. This range considers potential isotope fractionation occurring during the production of DMP from methanol, such as separation processes (e.g., distillation) and the complete reaction of methanol. The  $\delta^{13}\text{C}$  and bulk  $\delta^2\text{H}$  values of the originally added methanol

---

in the transmethylation reagent (TX-D and TM+D) were measured to be  $-38.9 \pm 0.08$  mUr and  $-169.5 \pm 2.8$  mUr, respectively, with GC-C/Py-IRMS based on the method described by Ai et al. (Ai et al., 2014). It should be noted that the  $\delta^2\text{H}$  value of the  $\text{H}_3\text{CO}$ - group in methanol is specifically required but not available. In the TM+D mixture, assuming the complete conversion of DMP to methanol, the ratio of originally added methanol to the intermediate reaction methanol is 12:1.

Based on these assumptions, the estimated  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the methanol present in the TM+D mixture during the methylation process were determined to be approximately  $-37.4$  to  $-42.2$  mUr and  $-157.5$  to  $-179.9$  mUr, respectively (refer to Table S-7 in the Appendix). With these estimations, the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the FAMEs were calculated and are presented in Table A-8 and Table A-9 of the Appendix. The absolute differences between the measured and estimated isotope compositions of the individual FAMEs ranged from 0.08 to 0.4 mUr for  $\delta^{13}\text{C}$  values and 1.2 to 3.5 mUr for  $\delta^2\text{H}$  values. These results indicate that our assumptions are reasonable.

Furthermore, the maximum absolute differences between the estimated isotope compositions of the individual FAMEs with and without DMP were approximately 0.2 mUr for  $\delta^{13}\text{C}$  and 1 mUr for  $\delta^2\text{H}$ . This finding suggests that the influence of DMP on the isotope measurements is smaller than the analytical precision, which is consistent with our experimental results. In other words, the difference or the influence of DMP on the carbon and hydrogen stable isotope analysis of FAMEs cannot be detected unless the analytical precision is significantly improved.

#### 4.5. Conclusion

In conclusion, the quick one-step sample preparation method developed by Garcés and Mancha has been proven feasible not only for the FA profiles analysis and lipid content estimation but also for carbon and hydrogen stable isotope analysis of individual FAMEs in both solid rapeseeds and fluid rapeseed oil. The reproducibility of the one-step method in determining the isotope composition of rapeseed samples was found to be within  $\pm 0.3$  mUr for  $\delta^{13}\text{C}$  and  $\pm 3$  mUr for  $\delta^2\text{H}$ . Importantly, the results of the experiments indicate that adding five vol% DMP in the transmethylation mixture does not significantly influence the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of three out of the four major FAMEs. Overall, the performance of the one-step method in determining the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of FAMEs in rapeseed samples is satisfactory. The quick one-step method can be implemented to process many oilseed samples in a significantly shorter timeframe. This advancement facilitates the application of chemometrics for the authentication of various oilseeds.

---

## 5. Evaluation of the potential of stable hydrogen isotopic composition of FAs for geographical traceability of rapeseed

---

### 5.1. Introduction

This Chapter addresses the second research objective, which focuses on evaluating the potential of the stable hydrogen isotopic composition of FAs in rapeseed as an analytical marker for geographic tracking. There are three main considerations:

- To examine its distribution within a specific spatial range, which in this dissertation refers to Hesse, and if the variation can be detected using IRMS technology.
- To determine whether changes in the  $\delta^2\text{H}$  values are related to environmental factors and which factors are related. The potential influencing factors are divided into four categories: climate factors, soil property, rapeseed variety, and geographic location.
- To justify the link between the environmental factors and  $\delta^2\text{H}$  values of the four FAs of rapeseed based on the available literature.

A total of 44 and 38 rapeseed samples from 2019 and 2020, respectively, were analyzed for their  $\delta^2\text{H}$  values of C18:1, C18:2, C18:3, and C16:0 FAs. T-tests and correlation analysis were conducted to study the temporal and spatial variation in the  $\delta^2\text{H}$  values of these FAs. For the effect of rapeseed variety on the  $\delta^2\text{H}$  values of FAs, a comparison was made between four to five different variety groups in 2019 and 2020. The results are presented in Subchapters 5.2 and 5.3.

To increase the sample size and expand the variation of the climatic factors, rapeseed samples from 2017 and 2018 ( $n=19$  and  $n=20$ ) across Hesse were included in the correlation analysis between environmental factors and the  $\delta^2\text{H}$  values of the four FAs. Stepwise linear regression analysis was then applied to identify the important factors for the  $\delta^2\text{H}$  values of the FAs. The results are presented in Subchapter 5.4.

Finally, the significant correlations determined by the stepwise linear regression analysis are justified based on the available literature. The potential of using the stable hydrogen isotopic composition of FAs in rapeseeds as an analytical marker for geographic tracking is evaluated by conducting these analyses and addressing the abovementioned considerations. The results are presented in Subchapter 5.5.

All rapeseeds were prepared with the one-step sample preparation method described in Chapter 4. The  $\delta^2\text{H}$  values of the four FAs of rapeseeds were measured with GC-Py-IRMS based on the method described in Chapter 3.3. The related soil samples were also analyzed for their parameters, as mentioned in Chapter 3.5. The climatic data of the sampling sites were obtained from the available dataset mentioned in Chapter 3.6.

### 5.2. Temporal and spatial variation of the $\delta^2\text{H}$ values of FAs of rapeseed

44 and 38 rapeseed fields were selected in 2019 and 2020, respectively. These sampling locations were distributed throughout the state of Hesse, ranging from a latitude of  $49.68352^\circ\text{N}$  in the south to  $51.44870^\circ\text{N}$  in the north and from a longitude of  $10.14848^\circ\text{E}$  in the east to  $8.01243^\circ\text{E}$  in the west (as shown in Table 15). For better discussion, the entire state of Hesse was divided evenly into three parts according to the latitude of the sampling locations: southern Hesse, central Hesse, and

northern Hesse. Table 15 provides information on the corresponding ranges of longitude and altitude for these sampling locations.

Table 15: The geographic range of the sampling locations for 2019 and 2020.

Year	Hesse-Parts	N	Latitude	Longitude	Elevation* (m, a.s.l.)
2019	Hesse	44	49.68352°N ~ 51.4487°N	8.0125°E ~ 10.14848°E	88 ~ 475
	North Hesse	19	50.92700°N ~ 51.44870°N	8.81906°E ~ 10.14848°E	193 ~ 454
	Central Hesse	19	50.22763°N ~ 50.77646°N	8.0125°E ~ 9.89698°E	117 ~ 475
	South Hesse	6	49.6835°N ~ 50.1842°N	8.4887°E ~ 9.1518°E	88 ~ 238
2020	Hesse	38	49.71674°N ~ 51.34464°N	8.01243°E ~ 10.14375°E	89 ~ 520
	North Hesse	18	50.8634°N ~ 51.34464°N	8.82721°E ~ 10.14375°E	175 ~ 431
	Central Hesse	13	50.29076°N ~ 50.81976°N	8.10829°E ~ 9.90768°E	125 ~ 409
	South Hesse	7	49.71674°N ~ 50.20743°N	8.01243°E ~ 8.80856°E	89 ~ 520

Note: \* Elevation above sea level (a.s.l.) refers to altitude.

Table 16 provides an overview of the  $\delta^2\text{H}$  values of four FAs (C18:1 FA, C18:2 FA, C18:3 FA, and C16:0 FA) in rapeseed samples collected from Hesse in 2019 and 2020. The data includes average values, SD, minimum, and maximum values. The  $\delta^2\text{H}$  values of these rapeseed samples of 2019 ranged from -149 mUr to -163 mUr for C18:1 FA, from -173 mUr to -154 mUr for C18:2 FA, from -194 mUr to -177 mUr for C18:3 FA, and from -176 mUr to -162 mUr for C16:0 FA. Similarly, in 2020, the  $\delta^2\text{H}$  values of the four FAs displayed variations, ranging from -195 mUr to -150 mUr. The differences of  $\delta^2\text{H}$  values of individual FA in these sampling locations in Hesse were in the range of 14 mUr to 30 mUr, which were higher than the IRMS precision for the H-CSIA of FAs ( $\pm 0.7$  to  $\pm 2.4$  mUr), as mentioned in Chapter 4. The  $\delta^2\text{H}$  values of FAs varied more within Hesse in 2020 than in 2019; namely, the difference in 2020 was about 30 mUr, while the difference in 2019 ranged from 14 mUr to 19 mUr.

Table 16: The mean, minimum, and maximum  $\delta^2\text{H}$  values of FAs in different parts of Hesse from 2019 to 2020.

Year	Hesse part	N	$\delta^2\text{H}$ of C18:1 FA			$\delta^2\text{H}$ of C18:2 FA			$\delta^2\text{H}$ of C18:3 FA			$\delta^2\text{H}$ of C16:0 FA		
			M $\pm$ SD	Min	Max	M $\pm$ SD	Min	Max	M $\pm$ SD	Min	Max	M $\pm$ SD	Min	Max
2019	Hesse	44	-156 $\pm$ 3.3	-163	-149	-161 $\pm$ 4.3	-173	-154	-188 $\pm$ 3.8	-194	-177	-169 $\pm$ 3.0	-176	-162
	North Hesse	19	-155 $\pm$ 2.7	-158	-149	-161 $\pm$ 4.5	-173	-154	-188 $\pm$ 2.8	-193	-181	-169 $\pm$ 2.1	-172	-165
	Central Hesse	19	-156 $\pm$ 3.5	-163	-150	-162 $\pm$ 4.3	-169	-154	-187 $\pm$ 4.5	-193	-177	-169 $\pm$ 3.4	-176	-162
	South Hesse	6	-158 $\pm$ 3.6	-162	-152	-161 $\pm$ 4.1	-169	-158	-187 $\pm$ 5.0	-194	-182	-171 $\pm$ 3.6	-176	-165
2020	Hesse	38	-154 $\pm$ 5.5	-164	-135	-156 $\pm$ 5.7	-168	-137	-185 $\pm$ 5.6	-195	-165	-169 $\pm$ 5.7	-179	-150
	North Hesse	18	-156 $\pm$ 4.9	-164	-145	-158 $\pm$ 5.5	-168	-147	-186 $\pm$ 5.3	-195	-175	-170 $\pm$ 5.6	-179	-160
	Central Hesse	13	-153 $\pm$ 4.2	-159	-143	-156 $\pm$ 4.4	-161	-144	-186 $\pm$ 4.2	-193	-175	-169 $\pm$ 4.0	-177	-160
	South Hesse	7	-150 $\pm$ 7.6	-158	-135	-152 $\pm$ 6.8	-157	-137	-182 $\pm$ 7.9	-187	-165	-165 $\pm$ 7.8	-173	-150



A t-test was performed to compare the  $\delta^2\text{H}$  values of four FAs in the rapeseed harvested during these two years. The results showed that the  $\delta^2\text{H}$  values of C18:2 and C18:3 FA in rapeseed had significant differences between 2019 and 2020 ( $p < 0.001$  and  $p < 0.05$ , respectively), and the values in 2019 were slightly lower than those in 2020, while the values of the other two FAs had no difference in years. To ensure a successful crop yield and maintain soil health, it is common practice to rotate crops and avoid planting the same crop on the same farmland for two consecutive years. In Hesse, rapeseed is rotated with wheat or barley as a precursor or successor crop based on the information from LLH. As a result, the rapeseed harvested in 2019 and 2020 came from different fields within Hesse.

Annual variation of  $\delta^2\text{H}$  values of four FAs in 2019 and 2010

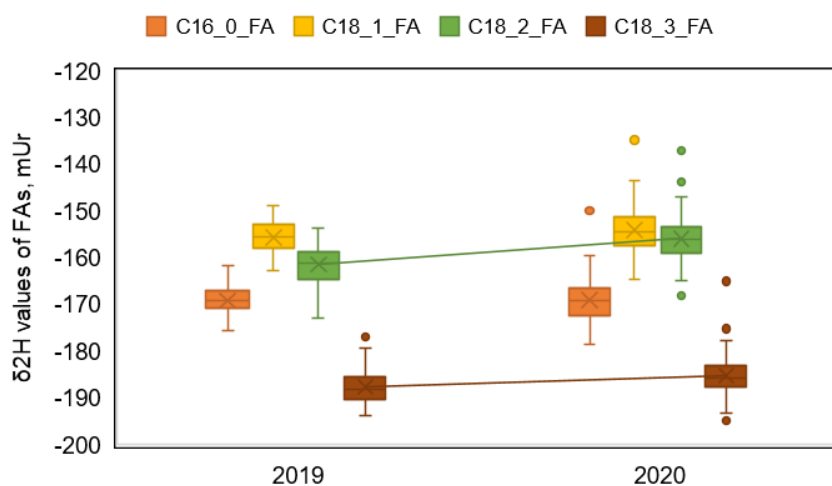


Figure 20: The  $\delta^2\text{H}$  values of four FAs of Hessian rapeseed in 2019 and 2020

Note: Lines indicate significant differences between years.

Regarding the spatial variation in Hesse, the correlation analysis between the  $\delta^2\text{H}$  values of the four FAs of 2019 and 2020 and geographical variables, latitude and altitude, was conducted. Table A- 10 in the Appendix displays the correlation coefficients and p-values from the correlation analysis for the samples of 2019. The statistically significant correlations are presented in Figure 21.

For the entire Hesse, only one significant correlation was observed. A moderately positive correlation existed between the isotopic  $\delta^2\text{H}$  values of C18:1 FA and latitude ( $n=44$ ,  $r=0.32$ ,  $p < 0.05$ ). None of the FAs showed a significant correlation with altitude.

When considering the different parts within Hesse, several significant correlations emerged: a moderate positive correlation was found between the  $\delta^2\text{H}$  values of C18:2 FA and altitude in northern Hesse ( $n=19$ ,  $r=0.51$ ,  $p < 0.05$ ); in central Hesse, the  $\delta^2\text{H}$  values of C18:1, C18:3, and C16:0 FAs all displayed moderate positive correlations with latitude ( $r=0.48$ ,  $p < 0.05$ ); the  $\delta^2\text{H}$  values of C18:3 FA exhibited a strong negative correlation with latitude ( $n=6$ ,  $r=-0.85$ ,  $p < 0.05$ ) in southern Hesse. Compared to the entire Hesse, more correlations were found on the regional scale.

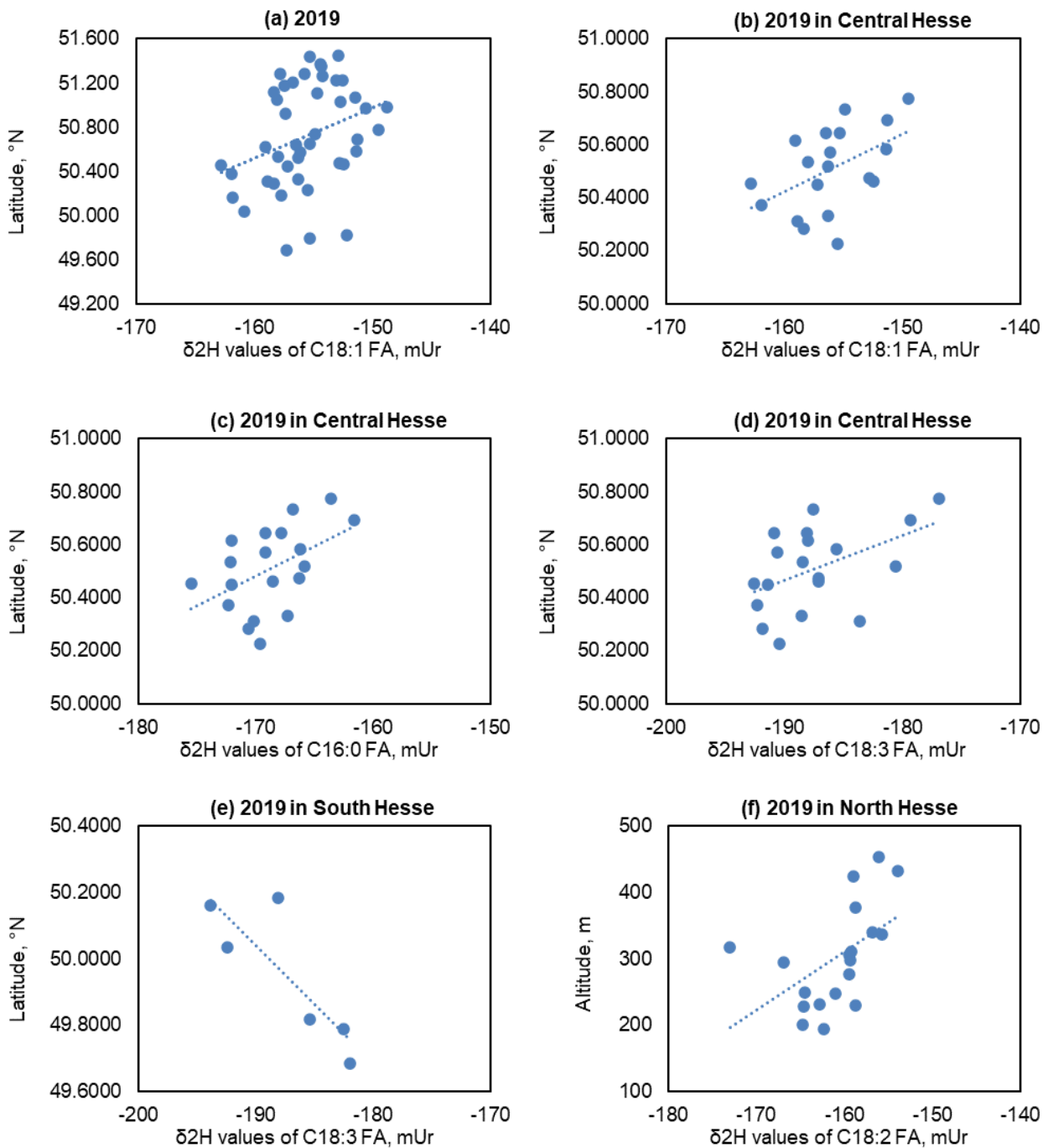


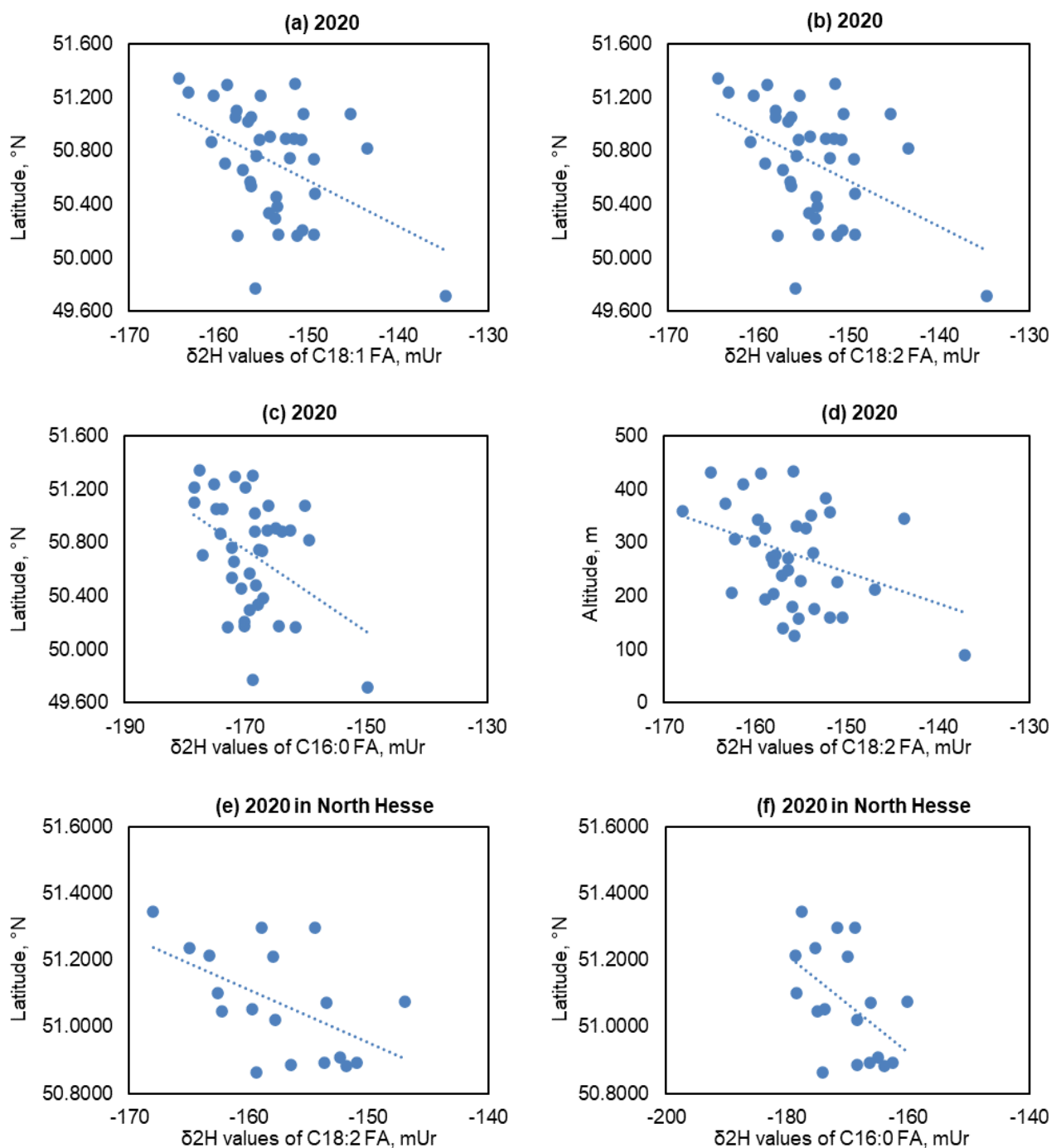
Figure 21: Significant correlation of  $\delta^2\text{H}$  values of different FAs with latitude and altitude for rapeseed 2019.

Note: (a) positive correlation between  $\delta^2\text{H}$  of C18:1 FA and latitude in entire Hesse; (b) positive correlation between  $\delta^2\text{H}$  of C18:1 FA and latitude in central Hesse; (c) positive correlation between  $\delta^2\text{H}$  of C18:3 FA and latitude in central Hesse; (d) positive correlation between  $\delta^2\text{H}$  of C16:0 FA and latitude in central Hesse; (e) negative correlation between  $\delta^2\text{H}$  of C18:3 FA and latitude in southern Hesse; (f) positive correlation between  $\delta^2\text{H}$  of C18:2 FA and altitude in northern Hesse.

For the 2020 rapeseed samples, a different pattern emerged. The statistically significant correlations are presented in Figure 22. For the entire Hesse, significant negative correlations were observed between latitude and the  $\delta^2\text{H}$  values of C18:1 FA ( $n=38$ ,  $r=-0.36$ ,  $p<0.05$ ), C18:2 FA ( $n=38$ ,  $r=-0.47$ ,  $p<0.05$ ), and C16:0 FA ( $n=38$ ,  $r=-0.41$ ,  $p<0.05$ ). The samples from northern Hesse were more

enriched in  $^2\text{H}$  than those from southern Hesse, with a moderate correlation. There were significant correlations of altitude with  $\delta^2\text{H}$  values of C18:2 FA ( $n=38$ ,  $r=-0.33$ ,  $p<0.05$ ) and C16:0 FA ( $n=38$ ,  $r=-0.32$ ,  $p<0.05$ ) in Hesse.

Significant correlations were observed only in northern Hesse for different parts of Hesse. The  $\delta^2\text{H}$  values of C18:2 and C16:0 FAs correlated negatively with latitude ( $n=18$ ,  $r=-0.54$  and  $-0.50$ ,  $p<0.05$ ). A significant negative correlation was identified between the  $\delta^2\text{H}$  value of C18:1 FA and altitude ( $n=18$ ,  $r=-0.51$ ,  $p<0.05$ ). All results are presented in Table A- 11 in the Appendix.





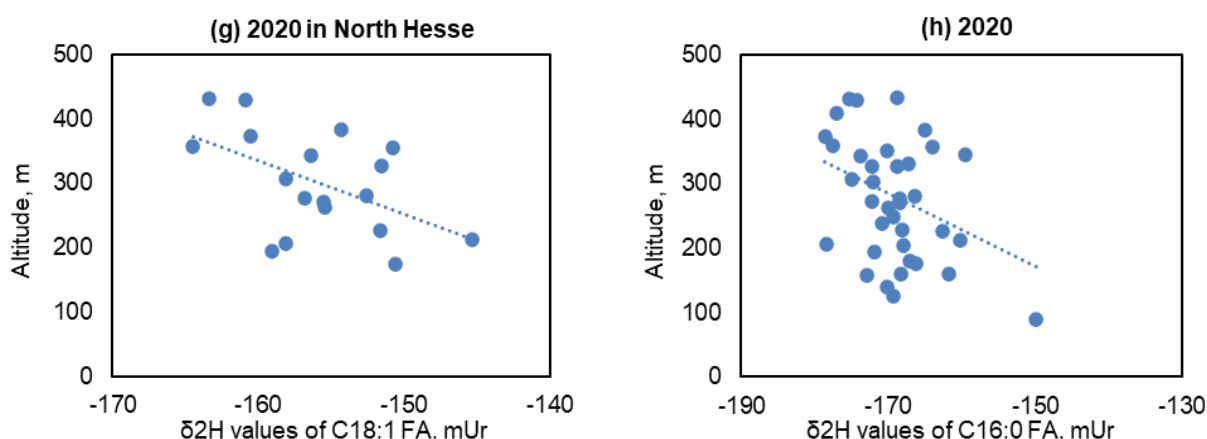


Figure 22: Significant correlation of  $\delta^2\text{H}$  values of different FAs with latitude and altitude for rapeseed 2020.

Note: (a)-(c) negative correlation between latitude and  $\delta^2\text{H}$  values of C18:1, C18:2, and C16:0 FA in entire Hesse; (d) negative correlation between altitude and  $\delta^2\text{H}$  values of C18:2 FA in entire Hesse; (e)-(f) negative correlation between latitude and  $\delta^2\text{H}$  values of C18:2 and C16:0 in northern Hesse; (g)-(h) negative correlation of altitude with  $\delta^2\text{H}$  values of C18:2 FA in northern Hesse, and with C16:0 FA in the entire of Hesse.

These findings are rather consistent with those of Mahindawansha et al., who studied the stable hydrogen and oxygen isotopic composition of precipitation in north Hesse. Their results reveal an altitude effect at the regional scale, with a significant correlation between the  $\delta^2\text{H}$  values of precipitation and altitude that varied seasonally. Specifically, the  $\delta^2\text{H}$  values in spring (March-May) were negatively correlated with altitude, and the correlation was the strongest in March. Very weak positive correlations were observed in autumn and winter, respectively, but they were not statistically significant (Mahindawansha et al., 2022).

The negative correlation observed between altitude and the  $\delta^2\text{H}$  values of FAs in North Hessian rapeseed samples in 2020 is consistent with their conclusion regarding the altitude effect of precipitation in spring. It suggests that the spring precipitation influences the FAs of rapeseeds. Additionally, their results showed the slope of the linear relationship between altitude and the  $\delta^2\text{H}$  values of precipitation was -0.014, meaning that for every hundred-meter increase in altitude, the hydrogen isotopic value of rainwater changed by 1.4 mUr (Mahindawansha et al., 2022). In contrast, for the  $\delta^2\text{H}$  values of FAs in rapeseed, the slope was -0.032, which means that for every hundred-meter change in altitude, the isotopic value of FAs changed by 3.2 mUr. These findings suggest that the differences in  $\delta^2\text{H}$  of FAs resulting from the altitude effect may be amplified during the growth process of rapeseed. However, the FAs of rapeseed of 2019 in northern Hesse showed a positive correlation with altitude, possibly influenced by the precipitation in other seasons.

The  $\delta^2\text{H}$  in precipitation also exhibits the latitude effect, where the  $\delta^2\text{H}$  values decrease with increasing latitude (Dawson, 1993). This pattern was observed in the  $\delta^2\text{H}$  values of C18:1, C18:2, and C16:0 FA in rapeseed samples from the entire Hesse in 2020, with a more pronounced correlation in central Hesse. In contrast, for rapeseed grown in southern Hesse in 2019, the latitude effect was evident in the  $\delta^2\text{H}$  values of C18:3 FA. However, in central Hesse in 2019, opposite correlations were observed for C18:1, C18:3, and C16:0 FAs. This inconsistency cannot be explained by the transfer of the spatial distribution of  $\delta^2\text{H}$  in precipitation. The  $\delta^2\text{H}$  values of FAs may

---

be influenced by additional environmental factors affecting rapeseed growth, which could account for the absence of consistent correlations. While the correlation between different FAs in rapeseed and latitude and altitude often reflected the spatial distribution of isotopes in precipitation, this pattern was sometimes observed across Hesse and at other times more prominently at the regional level. This suggests that rapeseed partially acquires the spatial distribution patterns of  $\delta^2\text{H}$  in precipitation through root absorption. However, the variation in isotopic distribution between 2019 and 2020 indicates that factors beyond rainfall isotopes also influence rapeseed growth throughout its lifecycle.

The spatial variation of  $\delta^2\text{H}$  values of FAs in rapeseed closely resembled that of precipitation, indicating significant correlations of  $\delta^2\text{H}$  values with latitude and altitude. Investigating this correlation is worthwhile, but collecting rainwater samples from numerous sampling points is challenging. Bowen et al. developed a global-scale rain isotopes model using the Global Network of Isotopes in Precipitation (GNIP) data to address this (Bowen & Revenaugh, 2003). However, the accuracy of this model's estimates is limited for our region-specific spatial scope and varying climatic conditions across different years (Mahindawansa et al., 2022). Therefore, instead of relying on precipitation isotope values, directly analyzing the correlation of climatic factors with rapeseed fatty acid isotopes may be a better approach, as air temperature and precipitation amount are known to be the primary driving force for changes in the isotope value of rainwater (Gourcy et al., 2005). The related results are presented in Subchapter 5.4.

### **5.3. Effect of rapeseed varieties on $\delta^2\text{H}$ values of FAs**

When selecting rapeseed varieties, farmers consider several factors, including yield, disease resistance, adaptability, and cost. As a result, the varieties grown in Hesse vary yearly (Bundesministerium für Ernährung und Landwirtschaft, 2020). The effect of rapeseed varieties on the  $\delta^2\text{H}$  values of FAs was investigated by comparing the  $\delta^2\text{H}$  values of FAs between the samples of different varieties of the same year. The analysis focused only on the varieties with sample numbers greater than 3 for each rapeseed variety. This approach allowed us to account for differences in the stable hydrogen isotopic composition of FAs in 7 selected varieties.

The ANOVA test was employed to compare the  $\delta^2\text{H}$  values of FAs among different rapeseed varieties. Details about the test are outlined in Chapter 3. P-values were determined to assess the statistical significance of differences among the varieties.

The ANOVA test was performed on the data from 2019 and 2020. The samples of 2019 were divided into five groups based on the varieties: Architect (n=4), Bender (n=9), Comfort (n=4), DK Exception (n=6), and Penn (n=5). Similarly, in 2020, the samples were categorized into four groups: Avatar (n=3), Architect (n=9), DK Exception (n=7), and DK Expansion (n=3). The results are presented in Table 17. For the isotopic data of C18:1, C18:2, and C16:0 FA in 2019, the Shapiro-Wilk test indicated that the data followed a normal distribution across all five groups of rapeseed varieties. Additionally, the homogeneity of variances of each group was confirmed with Levene's homogeneity test. No significant differences in the  $\delta^2\text{H}$  values of these FAs among the different varieties were observed. However, for the isotopic data of C18:3 FA, one group did not conform to a normal distribution according to the Shapiro-Wilk test, and the variance among groups was not equal based

on Levene's homogeneity test. Therefore, Welch's ANOVA test was conducted to determine differences between the groups, and it revealed that the  $\delta^2\text{H}$  values of C18:3 FA among various varieties did not differ significantly.

For the sample collected in 2020, the Shapiro-Wilk test indicated that the data for C18:1 and C16:0 FAs followed a normal distribution among the four groups of varieties. However, for C18:2 and C18:3 FA, two data groups did not conform to the normal distribution. The homogeneity of variances across all groups for all FAs was confirmed using Levene's chi-square test. Despite the non-normality distribution of the data, the ANOVA test is still considered appropriate (McDonald, 2009). The analysis showed no statistically significant differences in the  $\delta^2\text{H}$  values of the four FAs among the different variety groups.

Table 17: Comparison of the  $\delta^2\text{H}$  values of FAs among different rapeseed varieties using the ANOVA test.

Year	Variety	n	C18:1 FA		C18:2 FA		C18:3 FA		C16:0 FA	
			M	SD	M	SD	M	SD	M	SD
2019	Architect	4	-153.1	2.1	-158.3	4.2	-185.6	4.7	-165.7	3.1
	Bender	9	-155	2.4	-159.5	3.5	-187.4	1.8	-169.4	2.2
	Comfort	4	-158.8	3.6	-161.9	3.2	-187	6.4	-171.9	4.7
	DK Exception	6	-154.2	3.7	-161.1	4.7	-187.3	3.5	-169	2.4
	Penn	5	-157.2	3.2	-163	3.8	-190.1	2.2	-169.2	2.5
	ANOVA p-value		0.06		0.36		0.3 (Welch's ANOVA)		0.077	
2020	Avatar	3	-155.7	4.1	-156.9	5.7	-186.8	7	-170.4	7.8
	Architect	9	-151.7	3.7	-153.4	3.9	-183.6	3.5	-166.6	3.1
	DK Exception	7	-153.6	4.5	-156.3	5.3	-185	6.3	-169	5.9
	DK Expansion	3	-156	5.8	-158	4.6	-187.7	3.5	-172.9	5.7
	ANOVA p-value		0.356		0.41		0.59		0.316	

The statistical results showed no significant differences in the  $\delta^2\text{H}$  values of FAs of rapeseed among the Architect, Bender, Comfort, DK Exception, Penn, Avatar, and DK Expansion varieties. However, it should be noted that using univariate ANOVA does not definitively exclude the potential impact of other factors and their interaction with rapeseed variety on the  $\delta^2\text{H}$  values of FAs. Additionally, due to the limited number of samples available for other rapeseed varieties grown by farmers (less than three samples for each), conducting a statistical analysis on those varieties was not feasible.

Richter et al. compared the bulk  $\delta^2\text{H}$  values of rapeseed oil across various varieties and observed small but significant differences between some varieties. However, the sample size for each variety was limited, ranging from 2 to 5 samples (Richter et al., 2010). Due to the small sample sizes, the statistical power of the findings may be compromised. Therefore, conducting a control experiment with a larger sample size would be a more appropriate approach to identify the specific effects of rapeseed variety. Nevertheless, based on the selected varieties, it can be concluded that the rapeseed varieties do not significantly influence the  $\delta^2\text{H}$  of FAs. In subsequent multivariate analysis, the factor of seed variety was excluded.

## 5.4. Identifying significant environmental factors affecting the $\delta^2\text{H}$ values of FAs in rapeseed

Three types of environmental factors that could potentially influence  $\delta^2\text{H}$  of FAs were collected: climate condition, soil, and location. The specific parameters associated with each type are presented in Table 18. It is important to note that locations and some soil characteristics remain constant and are not subject to change. However, climate conditions, such as temperature and precipitation, are dynamic and vary over time (Figure 23).

Table 18: Three types of environmental factors for the stepwise regression analysis.

Factors type	Factors	Unit	Time resolution
Climate	air temperature in 2 m	°C	monthly, seasonal, and average during the vegetation period
	precipitation amount	mm	monthly, seasonal, and average during the vegetation period
	sunlight duration	hours	monthly, seasonal, and average during the vegetation period
	drought index		monthly, seasonal, and average during the vegetation period
Soil	soil moisture in 60 cm	%nFK	monthly, seasonal, and average during the vegetation period
	soil surface moisture	%	monthly, seasonal, and average during the vegetation period
	soil temperature in 5 cm	°C	monthly, seasonal, and average during the vegetation period
	electrical conductivity	$\mu\text{S}/\text{cm}$	
	pH value		
	phosphor content	$\text{P}_2\text{O}_5$ mg/100g	
	kalium content	$\text{K}_2\text{O}$ mg/100g	
	magnesium content	mg/100g	
	clay content	%	
	silt content	%	
Location	latitude	°N	
	longitude	°E	
	altitude	m	

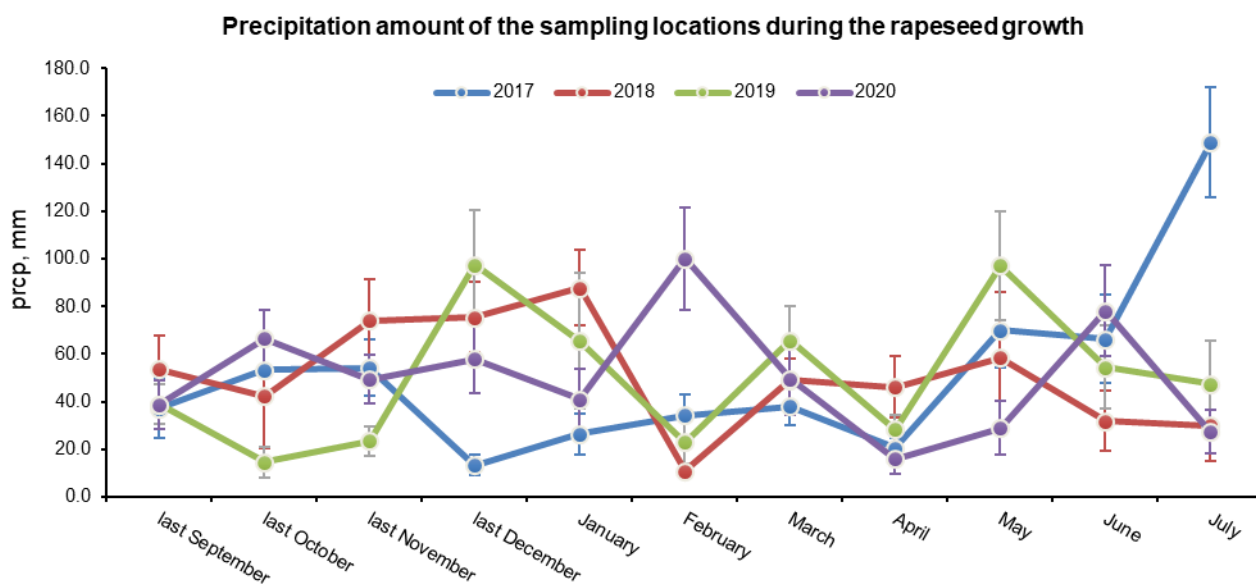


Figure 23: Precipitation of sampling sites during the rapeseed growth from 2017 to 2019.

At the Deutscher Wetterdienst (DWD), the German Weather Service, climate data is accessible in various temporal resolutions. Winter rapeseed is an annual crop sown in late August or early September and harvested from late July to early August of the following year. The growth stages of rapeseed are impacted by the climatic conditions experienced each month. Therefore, climate data were collected monthly, covering the entire growth cycle in the specific sampling locations where the selected rapeseed was cultivated. Furthermore, the analysis incorporates the average values for each season and the vegetation period as separate factors. In this context, the vegetation period refers to the period from March to July, as indicated by the flowering date data provided by DWD and the harvest date information provided by LLH for rapeseed crops in Hesse (Deutscher Wetterdienst, 2023).

These climatic factors are highly variable from year to year, particularly during the four years of our sample collection, and may be due to climate change leading to an increase in extreme weather events (Zscheischler & Fischer, 2020). For instance, in April 2019, the precipitation amounts varied between 20.1 mm and 46.2 mm at each sample collection site, while in 2020, it ranged from 6.3 mm to 30.5 mm, with little intra- and inter-annual variation between the two years. However, when considering the four years from 2017 to 2020, the values ranged from 6.3 mm to 113 mm, showing a more extensive distribution (Figure 24). Therefore, samples from 2017 (n=19) and 2018 (n=20) were included in the statistical analysis to expand the range of climatic factors.

The comprehensive dataset consisted of 89 factors that can potentially influence the  $\delta^2\text{H}$  values of FAs. It is important to note that these factors affect each other and may interact with the  $\delta^2\text{H}$  values of FAs. Therefore, stepwise linear regression analysis was applied to identify the most significant factors that affect the  $\delta^2\text{H}$  values of the FAs of rapeseed in Hesse and eliminate these potential interferences.

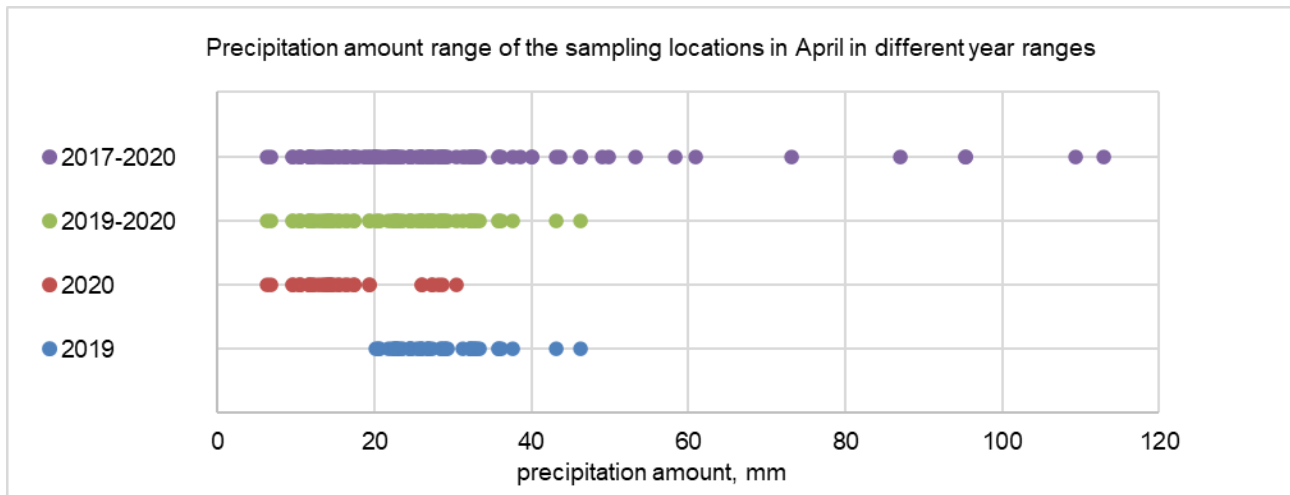


Figure 24: Precipitation amount variation in April across different years for sampling locations.

Before conducting the regression analysis, a thorough screening of all factors was performed to ensure the selection of appropriate factors. The screening process consisted of two steps. In the first step, scatter plots were created to examine the relationship between each factor and the  $\delta^2\text{H}$  values. Factors exhibiting non-linear relationships were excluded from further analysis. This initial screening helped ensure that only factors with a linear relationship to the  $\delta^2\text{H}$  values were considered. The second step involved screening the remaining climate and soil factors that vary with time. Considering these average climates for each month and season as distinct factors, numerous climatic factors may imbalance their influence on the overall model compared to the other conditions. Therefore, Pearson correlation analysis was performed individually between each factor and the  $\delta^2\text{H}$  values to address this issue. The factors with the strongest positive and negative correlations with the  $\delta^2\text{H}$  values were identified and selected. The factors selected through the steps mentioned above, combined with those that remained constant with time during the study period together, were analyzed further with the stepwise regression analysis. The factors considered significant by the analysis are presented in Table 19. Then, the stepwise regression analysis for these variables was conducted using SPSS software.

The final stepwise regression model for C18:1 FA included five factors: air temperature in February and May, soil moisture in 60 cm depth in May, soil surface moisture in April, and silt content. These variables were found to have statistically significant impacts on the  $\delta^2\text{H}$  values of C18:1 FA of rapeseed. One sample was identified as an outlier and subsequently removed from the analysis to ensure the robustness of the results.

The standardized regression coefficients indicate the direction and strength of the relationship between each factor and the  $\delta^2\text{H}$  values of C18:1 FA. A positive coefficient for air temperature in February and soil moisture in 60 cm in spring suggests that warmer weather in February and wetter in deeper soil in spring tended to be associated with higher  $\delta^2\text{H}$  values of C18:1 FA. Similarly, the  $\delta^2\text{H}$  values of C18:1 FA were negatively correlated with air temperature in May, surface soil conditions in April, and silt content in the soil, as indicated by the coefficients for air temperature, soil surface moisture, and silt content. The absolute magnitude of the standardized coefficient offers a standardized metric to compare the relative importance of different factors. They are ranked as

---

follows: air temperature in February > surface soil moisture in April > soil moisture in 60 cm depth in spring > air temperature in May > silt content.

The p-values lower than 0.5 are listed in Table 19, indicating that these environmental variables significantly affected the  $\delta^2\text{H}$  values of C18:1 FA. The adjusted R-squared value of 0.61 suggests that the selected factors could explain approximately 61% of the variance in  $\delta^2\text{H}$  values of C18:1 FA. Furthermore, diagnostic plots, including residual plots, were examined to assess the assumptions of linear regression. The Durbin-Watson statistic assesses the presence of autocorrelation in the residuals of the regression model. The value ranges from 0 to 4, where a value around 2 suggests no significant autocorrelation (Montgomery et al., 2012). The obtained Durbin-Watson value was 2.058, indicating that the assumption of independent and identically distributed errors holds in the model. The residual histogram appeared randomly distributed around zero (Figure 25a), suggesting that the standardized residuals follow a normal distribution. In the scatter plot of standardized residuals against standardized predicted values, the standardized residuals remained relatively stable and showed no significant changes with variations in the standardized predicted values (Figure 25b). This observation suggests that the assumption of homogeneity of variance was essentially met, and the assumptions of linearity, independence, normality, and constant variance of residuals of the model were reasonably met. The F-test is a statistical test used to determine if there are significant differences between the variances of two or more groups (Watkins, 2019). For stepwise regression analysis, the F-test evaluates whether adding or removing an independent variable significantly improves the model fit by comparing the variances explained before and after the change. The p-value of the F-test indicates the probability that the observed improvement in model fit is due to chance.

In conclusion, based on the stepwise regression analysis, air temperature in February and May, soil moisture at 60 cm depth in Spring, soil surface moisture in April, and silt content were identified as the most important factors for the  $\delta^2\text{H}$  values of C18:1 FA in Hesse.

For  $\delta^2\text{H}$  of C18:2 FA, the final model included five factors: precipitation amount in April, phosphorus content in the soil, altitude of the sampling locations, soil temperature in February, and silt content. These factors collectively accounted for approximately 60% of the variation in the  $\delta^2\text{H}$  values of C18:2 FA. Among these five important factors, the soil temperature in May had the most important impact on the  $\delta^2\text{H}$  values of C18:2 FA than the other four factors. The diagnostic plots for the linear regression model are presented in Figure A- 3 of the Appendix.

Table 19: Significant factors affecting stable hydrogen isotopic composition of four FAs using stepwise regression analysis.

Factor type	Important Factors	Units	$\delta^2\text{H}$ of C18:1 FA	$\delta^2\text{H}$ of C18:2 FA	$\delta^2\text{H}$ of C18:3 FA	$\delta^2\text{H}$ of C16:0 FA
Climate	Air temperature in February	°C	0.447	-	0.498	-
	Air temperature in May	°C	-0.179	-	-	-0.476
	Precipitation amounts in April	mm	-	-0.231	-	-
Soil	Soil surface moisture in April	%	-0.396	-	-0.425	-0.404
	Soil temperature in February	°C	-	0.591	-	-
	Soil moisture in spring in 60 cm depth	%nFK	0.247	-	0.197	-
	Silt content	%	-0.12	-0.133	-	-
	Phosphor content	P <sub>2</sub> O <sub>5</sub> mg/100g	-	0.154	-	-
Location	Altitude	m	-	0.17	-	-
Statistics	Adjusted R Square	-	0.608	0.601	0.551	0.552
	Durbin-Watson	-	2.058	2.142	1.974	2.2
	F-test	-	35.781	34.747	46.866	69.913
	P-value of F-test	-	<0.001	<0.001	<0.001	<0.001

Similarly, the regression model for  $\delta^2\text{H}$  of C18:3 FA revealed three important factors: air temperature in February, soil surface moisture in April, and soil moisture at a depth of 60 cm in Spring. These factors could explain about 55% of the variation in the  $\delta^2\text{H}$  values of C18:3 FA. The most important factors for C18:3 FA were air temperature in February and soil surface moisture in April. The diagnostic plots for the linear regression model are presented in Figure A- 4 of the Appendix.

Regarding the  $\delta^2\text{H}$  values of C16:0 FA, the final stepwise regression model included air temperature in May and soil surface moisture in April. These factors accounted for approximately 55% of the variation in the  $\delta^2\text{H}$  values of C16:0 FA. The diagnostic plots for the linear regression model are presented in Figure A- 5 of the Appendix.



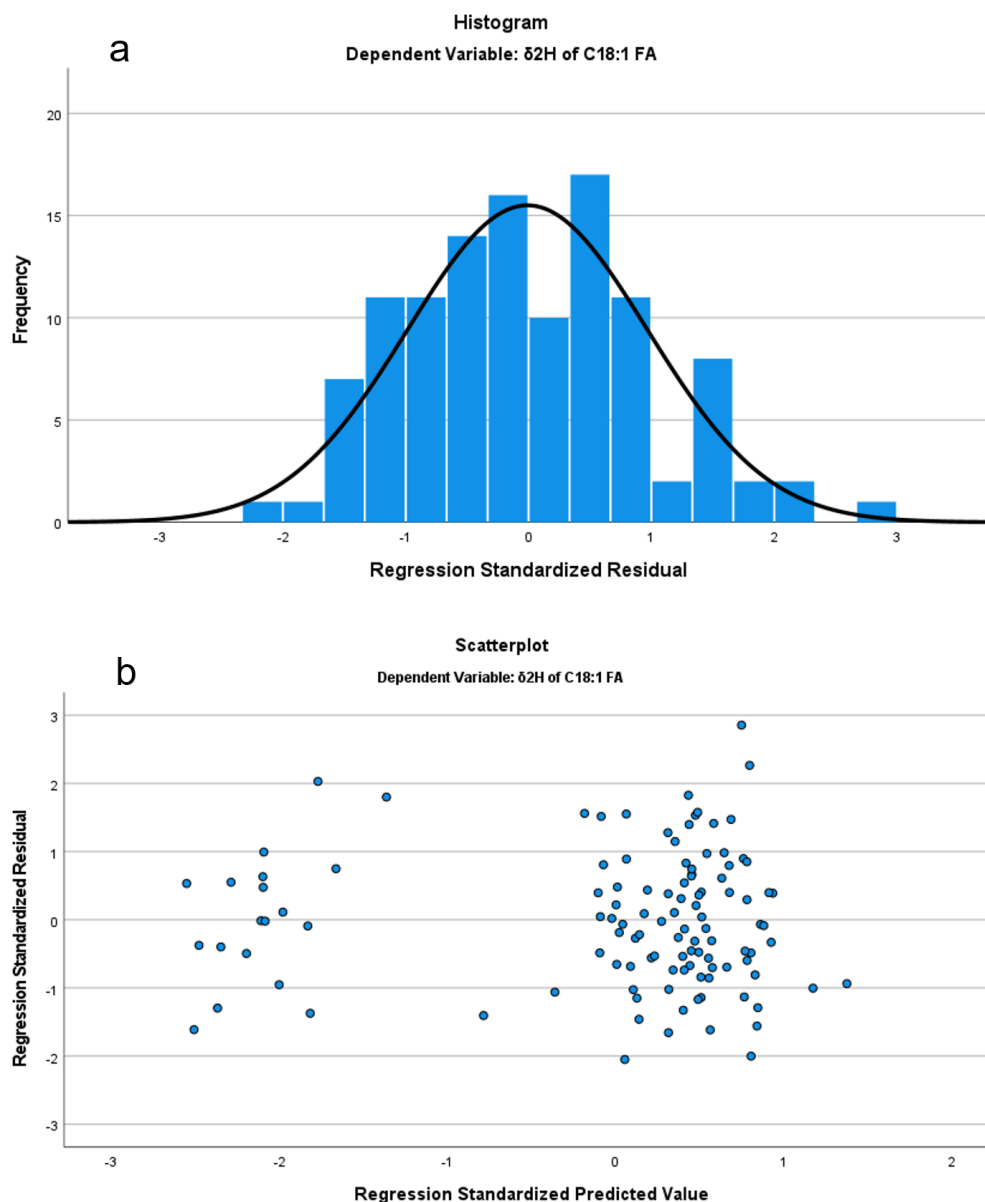


Figure 25: Diagnostic plots for the linear regression model of  $\delta^2\text{H}$  of C18:1 FA.

In conclusion, the factors significantly influencing the  $\delta^2\text{H}$  values of the four FAs in rapeseed were identified by stepwise linear regression analysis. These factors together could explain 55% to 60% of the variation in the  $\delta^2\text{H}$  values of the four FAs in rapeseed. Regarding these important factors, the climatic conditions in February, April, and May significantly affected the  $\delta^2\text{H}$  values of FAs in rapeseed. Among soil factors, soil surface moisture in April and soil moisture at 60 cm depth throughout spring exerted significant effects. Additionally, altitude among geographical locations also impacted these values.

### 5.5. Justifying the correlation between the environmental factors and $\delta^2\text{H}$ values of FAs

Rapeseed synthesizes these four FAs similarly, although their chain lengths and saturation degrees differ (Taiz & Zeiger, 2002). Therefore, the four FAs are discussed together here. These correlations

are discussed from the perspective of temperature effect and altitude effect of the  $\delta^2\text{H}$  value of precipitation and the physiological of plants.

According to the degree of influence on the C18:1 FA, these factors are ranked as follows: air temperature in February > soil surface moisture in April > soil moisture in 60 cm in Spring > air temperature in May > silt content.

Based on the regression analysis, the air temperature in February correlated with the  $\delta^2\text{H}$  values of C18:1 FA of rapeseed positively (shown in Figure 26). In February of the four years, the air temperatures at all sample collection sites ranged from  $-2.8^\circ\text{C}$  to  $6.9^\circ\text{C}$ . During low temperatures, particularly when the average temperature is below  $0^\circ\text{C}$ , precipitation (snow) becomes depleted in  $^2\text{H}$  isotopes, significantly differing from the isotopic composition of rainwater during temperatures above  $0^\circ\text{C}$  (Mahindawansa et al., 2022). After rainfall infiltrates into the soil, a portion of it undergoes evaporation. However, less evaporation in February due to the low temperatures allowed sufficient water to remain in the soil, which can be observed from the soil moisture in 60 cm, 92-112% nFK (Figure 26).

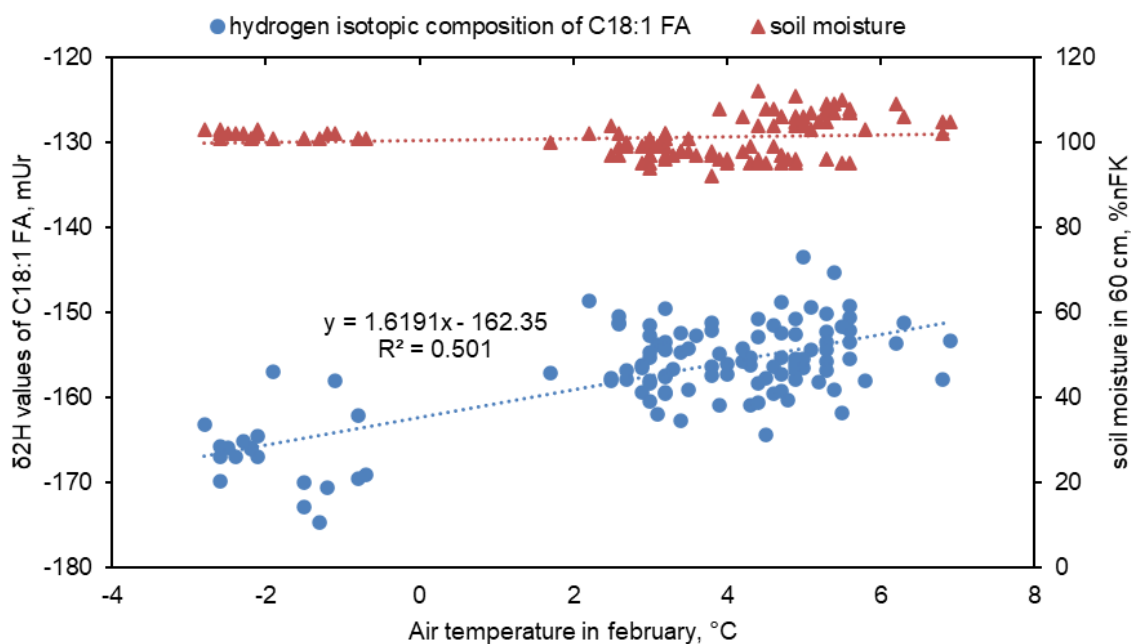


Figure 26: Significant correlation between the air temperature in February and  $\delta^2\text{H}$  values of C18:1 FA.

Once the rapeseed plants emerge from their winter dormancy, they enter a rapid growth phase as temperatures gradually rise. Rapeseeds generally undergo bud formation from mid-March to early April, and their average water demand reaches 4-6 mm/day (Müller et al., 2020). The roots of the plants absorb water stored in the soil, which may include rainfall from February to March. In March, the average temperature of every sampling location ranged from  $1.7^\circ\text{C}$  to  $8.9^\circ\text{C}$  in Hesse. Although the  $\delta^2\text{H}$  value of rainfall during this period also varied due to different temperatures, the difference was supposed to be smaller than the significant difference observed between subzero temperatures and those above freezing in February. That is why the air temperature in February rather than March was determined as an important factor. Consequently, the variation in  $\delta^2\text{H}$  values in rainfall caused by temperature differences at each sampling point in February was likely transferred to the C18:1

FA by water adsorption during the stem elongation phase.

The data-driven regression analysis revealed a notable linear correlation between the  $\delta^2\text{H}$  values of C18:1 FA in rapeseeds and soil surface moisture in April. However, considering the significant positive correlation between precipitation levels and soil surface moisture at each sampling location during April (Figure 27), it is theorized that the influencing factor on the  $\delta^2\text{H}$  values of C18:1 FA might be the April precipitation levels rather than the soil surface moisture itself. This hypothesis is supported by research by Mahindawansha et al., which suggests that in the mid-latitude region of Hesse, the amount effect in  $\delta^2\text{H}$  of precipitation is dominant in spring, meaning that lower amounts of precipitation result in higher  $\delta^2\text{H}$  values (Mahindawansha et al., 2022).

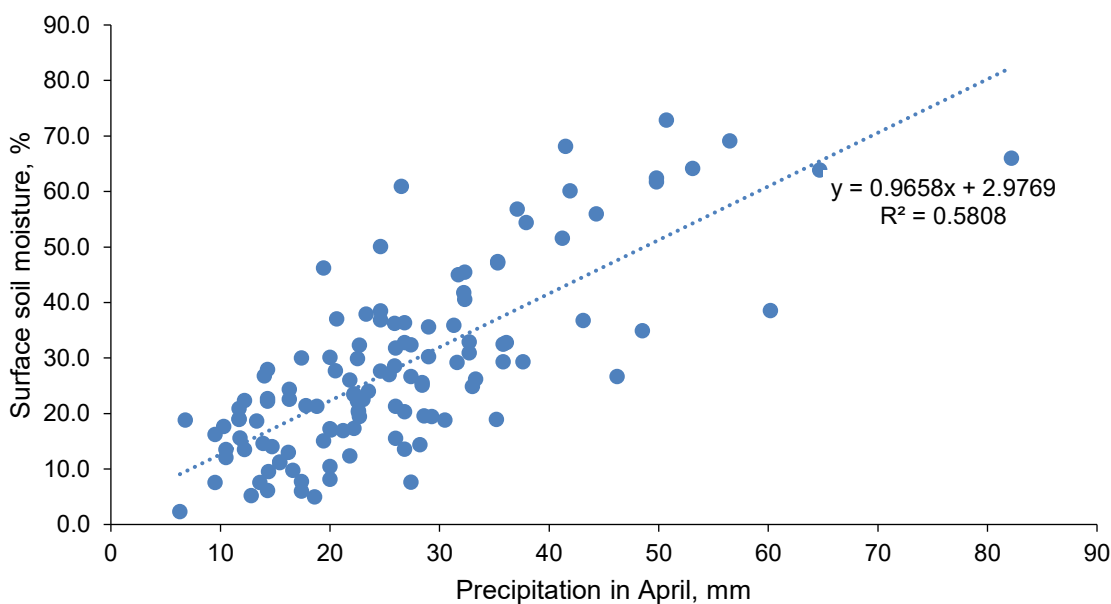


Figure 27: Correlation of precipitation amount and surface soil moisture in April of sampling sites.

During this period, the oilseed rape root system generally undergoes active water and nutrient absorption from the soil, facilitating rapid growth. The monthly average precipitation varied widely in April from 6 to 82 cm at different sampling points in Hesse (Figure 28). By early April, 46% of the rapeseed plants had entered the flowering stage, and by the end of April, all plants were in the flowering stage (Deutscher Wetterdienst, 2023). Rapeseed exhibited a peak water demand of approximately 6.9 mm/day, with a portion of the absorbed water utilized for biosynthesis within the plant (Müller et al., 2020). Therefore, April precipitation had a significant effect on hydrogen in rapeseed oil. That is, the amount effect of precipitation isotopes was likely transferred to the  $\delta^2\text{H}$  values of FAs in seeds during their development.

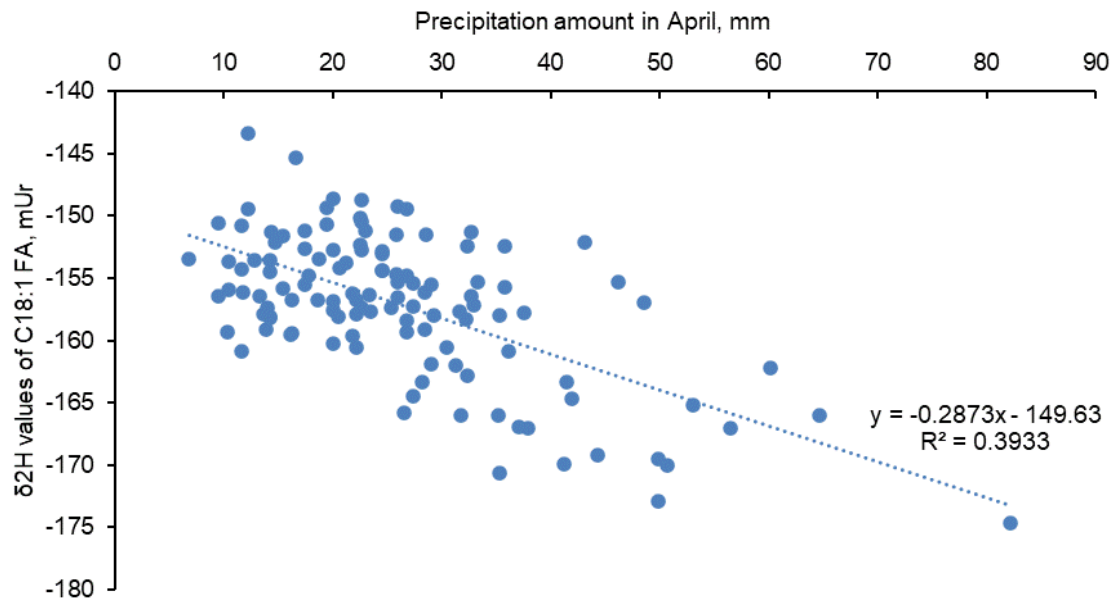


Figure 28: Significant correlation between the precipitation amount in April and  $\delta^2\text{H}$  values of C18:1 FA.

The average air temperature in May significantly negatively correlated with the  $\delta^2\text{H}$  values of C18:1 FA in rapeseed (Figure 29). This correlation is likely due to the temperature effect on the  $\delta^2\text{H}$  of precipitation. As indicated by Mahindawansha et al., the temperature effect was observed in spring rainfall in Hesse, where the  $\delta^2\text{H}$  values in precipitation decreased with higher air temperatures (Mahindawansha et al., 2022). Similarly, rapeseed plants absorb soil water from precipitation during the rapid growth phase in May. The isotopic pattern in precipitation can be transferred to the seed and reflected in  $\delta^2\text{H}$  of FAs in rapeseeds. However, it is noteworthy that the influence of air temperature in May on the  $\delta^2\text{H}$  values of FAs was relatively less pronounced than the impact of the previous three factors mentioned.

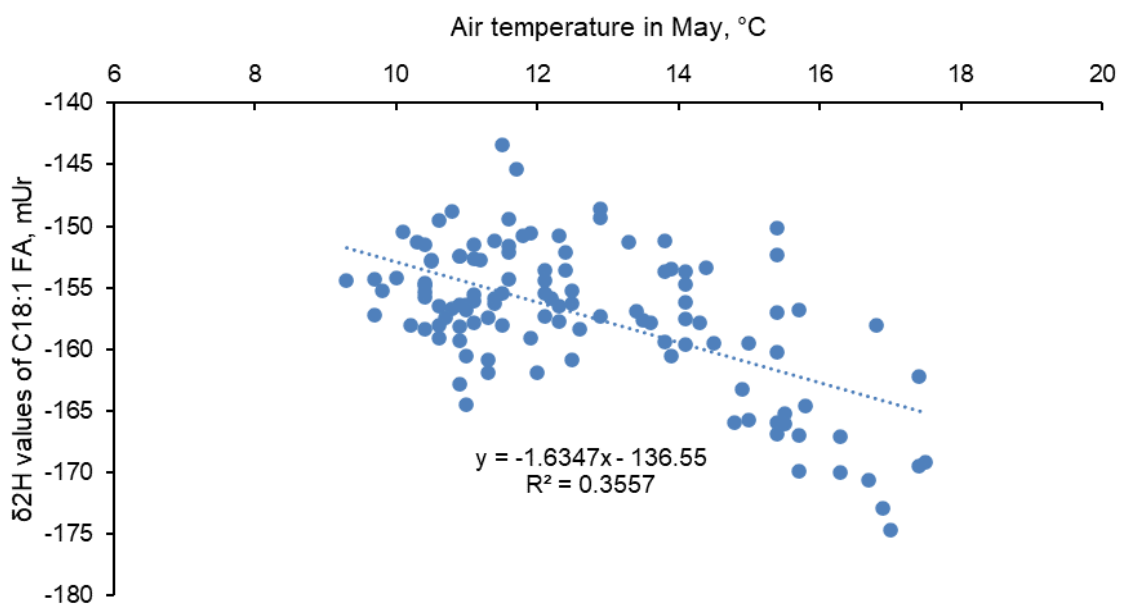


Figure 29: Significant correlation between the air temperature in May and  $\delta^2\text{H}$  values of C18:1 FA.

The regression analysis showed that the silt content in soil negatively correlated with the  $\delta^2\text{H}$  values of C18:1 FA in rapeseed, controlling for other factors. The higher the silt content, the lower the  $\delta^2\text{H}$  values of C18:1 FA.

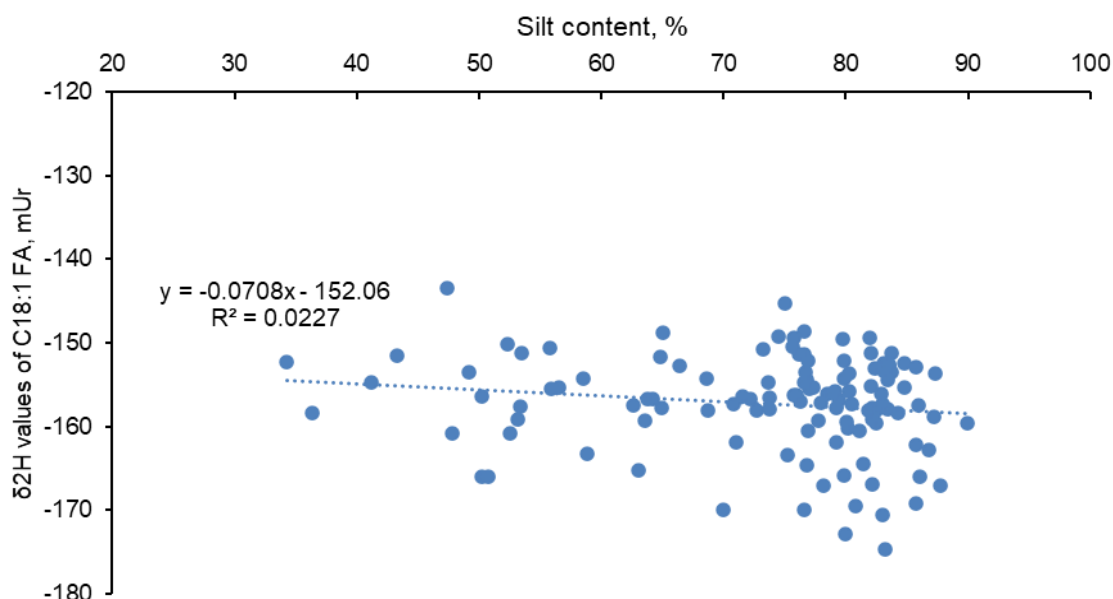


Figure 30: Significant correlation between the silt content and  $\delta^2\text{H}$  values of C18:1 FA.

This result might be related to the hydrogen isotopic fractionation during rapeseed plant growth. Hydrogen isotopic fractionation refers to the phenomenon that light isotopes ( $^1\text{H}$ ) are more likely to participate in biochemical reactions involving H than heavy isotopes ( $^2\text{H}$ ), resulting in different hydrogen isotopic ratios between reaction products and substrates (Ziegler, 1989). FAs in rapeseed are produced by photosynthesis and subsequent synthesis reactions, which all involve hydrogen isotopic fractionation. Therefore, the  $\delta^2\text{H}$  values of FAs in rapeseed are lower (lighter) than that of source water (i.e., water absorbed by plants), which has been confirmed by many studies (Eley et al., 2018; Sachse et al., 2012). Moreover, since FAs are stored in seed coats as TAGs, they are unaffected by evapotranspiration like leaf water. The lack of exposure to evapotranspiration prevents the enrichment of  $^2\text{H}$  isotopes of FA, suggesting that the hydrogen isotopic fractionation may not be offset in this case.

Although the  $\delta^2\text{H}$  values of precipitation, as source water of rapeseed, were not directly measured, they can be estimated by the model developed by Bowen et al. The estimation suggests that the range of  $\delta^2\text{H}$  values in precipitation in Hesse in different months is between -90 mUr and -20 mUr (Bowen & Revenaugh, 2003). Assume that the rapeseed plant absorbed the lightest precipitation, with a  $\delta^2\text{H}$  value of -90 mUr. The measured  $\delta^2\text{H}$  values of FAs in rapeseeds were lighter, between -134 mUr and -195 mUr. The difference in  $\delta^2\text{H}$  values between source water and FAs can reflect the degree of hydrogen isotopic fractionation ( $\epsilon$ ), i.e.,  $\epsilon = \delta^2\text{H}_{\text{water}} - \delta^2\text{H}_{\text{FA}}$ . The smaller the  $\delta^2\text{H}$  value of FAs, the greater the difference from source water (shown in Figure 31) and the greater the degree of fractionation. Silt, i.e., soil particles with a diameter between 2 and 63  $\mu\text{m}$ , has good water-holding and fertilizer-holding capacity, providing sufficient water and nutrients for plants (Amelung et al., 2018). Higher silt content in the soil enhances plant growth conditions and increases the degree of hydrogen isotopic fractionation. This negative correlation with the silt content in soil is similar to the



observation of carbon isotopes in plants from Cornwell et al. They found that high silt content promotes more carbon isotopic fractionation of plants (Cornwell et al., 2018). Therefore, it can be inferred that the higher the silt content, the greater the difference in  $\delta^2\text{H}$  values between FAs and source water, i.e., the lower the  $\delta^2\text{H}$  value of C18:1 FA.

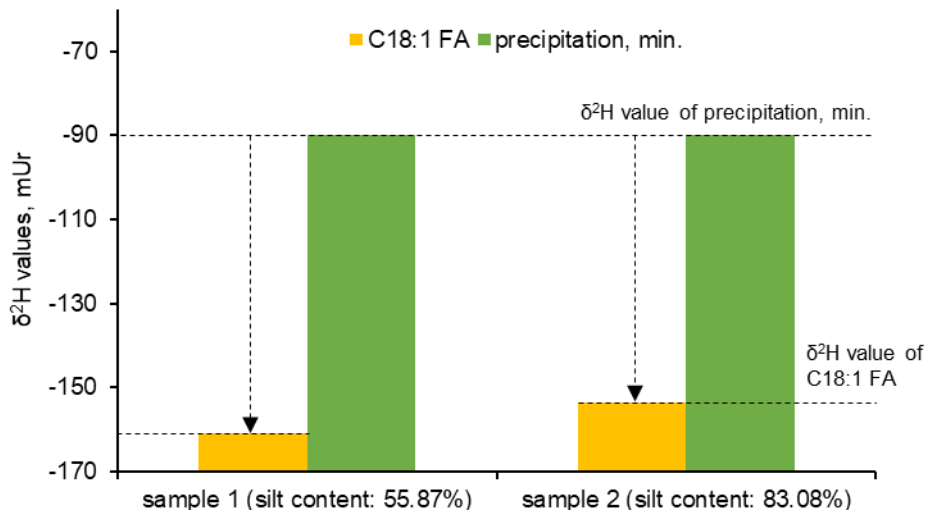


Figure 31: Difference in  $\delta^2\text{H}$  values between precipitation as source water and FAs of two samples with various soil silt content.

The soil temperature in February exhibited a highly significant positive correlation with the February air temperature (Figure 32), although the soil surface temperature was somewhat higher than the air temperature. Its effect on the  $\delta^2\text{H}$  of FAs was discussed earlier.

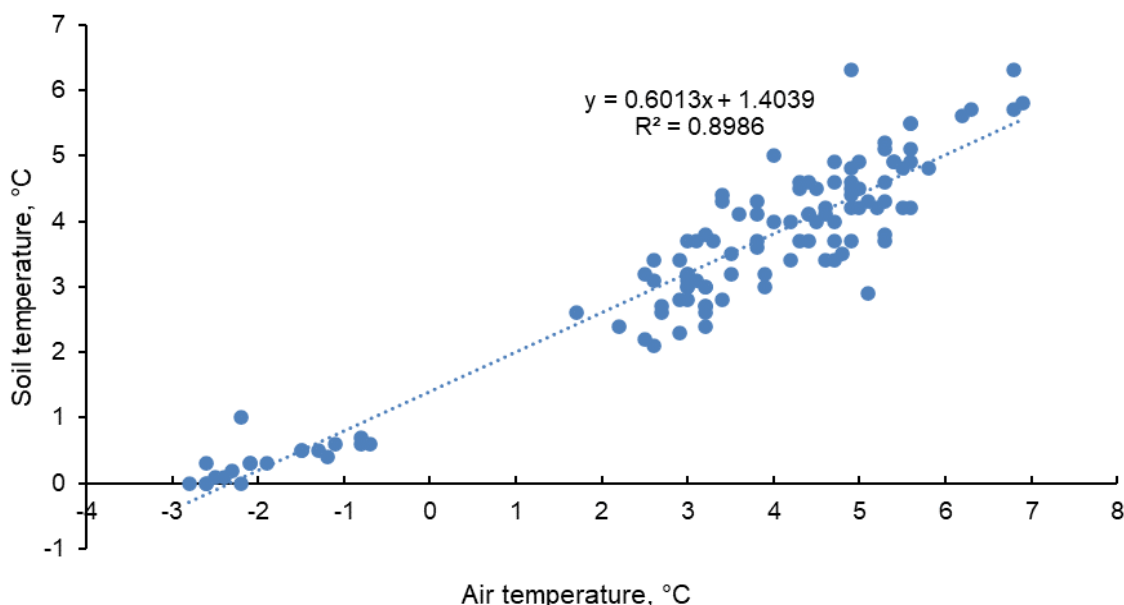


Figure 32: Correlation of air temperature and soil temperature in February of sampling sites.

Altitude was positively correlated with the  $\delta^2\text{H}$  of C18:2 FA of rapeseed, which contradicts the altitude effect of rainfall isotopes, i.e., the isotopic values of rainfall decrease with increasing altitude, so this positive correlation cannot be explained by the effect of source water on the  $\delta^2\text{H}$  values of FAs here.

---

The phosphorus content of the soil is positively correlated with the  $\delta^2\text{H}$  values of C18:2 FA in rapeseed. Phosphorus fertilization in the last autumn and spring of the current year helps rapeseed with photosynthesis, respiration, and fat metabolism, which makes the hydrogen isotopic fractionation in rapeseed more complex. A systematic literature search and experimental studies are required for a comprehensive explanation.

C18:3 and C16:0 FAs were also influenced by the same factors that affect C18:1 and C18:2 FAs, as their biosynthesis pathways were similar. However, there were differences in  $\delta^2\text{H}$  values between these four FAs. The  $\delta^2\text{H}$  values of C18:3 FA were consistently lower than those of the other three FAs. Except for five samples in 2019-2020, almost all rapeseed samples had lower  $\delta^2\text{H}$  values of C18:2 FA values than C18:1 FA, while C16:0 values fell between C18:2 and C18:3, except for one sample (Figure 20). This finding is consistent with the relationship observed between the  $\delta^2\text{H}$  values of FAs in olive oil, where the  $\delta^2\text{H}$  values tend to decrease with an increase in double bonds (Paolini et al., 2017). Fatty acid synthesis in plants involves a series of enzymatic reactions that convert acetyl-CoA to various FAs, as mentioned in Chapter 2.2.2. The initial step occurs in the chloroplasts, where C16:0 FA is formed. Fatty acid synthase then elongates C16:0 FA to stearic acid (C18:0). Subsequently, soluble  $\Delta^9$ -desaturase converts C18:0 FA to C18:1 FA, which can further undergo desaturation by  $\Delta^{12}$ -desaturase to produce C18:2 FA. Finally,  $\Delta^{15}$ -desaturases catalyze the desaturation of linoleic acid to yield C18:3 FA (Duan et al., 2002; Richter et al., 2010). The observed magnitude of the enrichment in  $^2\text{H}$  between C18:1, C18:2, and C18:3 FAs is consistent with the findings reported by Duan et al. As the seed matures,  $\Delta^{12}$ -desaturase tends to remove  $^1\text{H}$  rather than  $^2\text{H}$  atoms from oleic acid during the conversion to linoleic acid. Consequently, this process leads to an enrichment in  $^2\text{H}$  in the residual pool of oleic acid (Duan et al., 2002; Porokhvinova et al., 2022). During chain elongation, there is a slight preference for  $^1\text{H}$  atoms due to the less favorable kinetics of reactions involving molecules carrying heavy isotopes. However, it should be noted that the source of the added hydrogen atoms is the methyl group of acetate, and the isotopic ratio of this methyl group is unknown. This uncertainty may explain why C16:0 showed a depletion in  $^2\text{H}$  rather than C18:1 in our samples (Duan et al., 2002).

In summary, regression analysis is a data-driven exploratory tool to identify environmental factors that display significant linear associations with fatty acid isotopes. While some of these correlations may have a theoretical basis, others may not. Nevertheless, these findings offer valuable insights into comprehending the underlying factors contributing to the variability in  $\delta^2\text{H}$  of FAs in rapeseed. Further experiments with rigorous controls are necessary to establish a theoretical causal relationship and provide conclusive evidence.

## 5.6. Conclusion

The stable hydrogen isotopic composition of FAs has demonstrated potential as analytical markers for identifying the geographical origins of rapeseed. The detailed approach is described in Chapter 5. First, an analysis of its spatial distribution in Hesse revealed that the difference in the  $\delta^2\text{H}$  of the four FAs among different sampling sites ranged from 14 mUr to 30 mUr, exceeding the measurement

---

precision of  $\pm 3$  mUr of the instrument. This finding indicates that the spatial variations of rapeseed in Hesse can be accurately measured using existing techniques, which is a crucial prerequisite for further analysis. Furthermore, a similarity in the distribution pattern of  $\delta^2\text{H}$  values between the FAs in rapeseed and precipitation in 2020 was observed. For instance, the  $\delta^2\text{H}$  values of FAs negatively correlated with the latitude of the sampling sites, which was more pronounced in certain local areas. This correlation was also observed in  $\delta^2\text{H}$  of the precipitation in Hesse. It suggests that the  $\delta^2\text{H}$  of precipitation was partially transferred to rapeseed. Notably, the slope of this correlation in rapeseed was higher than that observed in precipitation, which suggests that the differences in  $\delta^2\text{H}$  values of FAs in rapeseed were amplified during the plant growth process. The correlation with the geographical variables, including latitude and altitude, further implies that the  $\delta^2\text{H}$  of FAs in rapeseed might be analytical markers for identifying geographical origin. However, this similarity between the  $\delta^2\text{H}$  of FAs and precipitation in 2020 was not observed in 2019, indicating that factors other than precipitation also influence the  $\delta^2\text{H}$  values of FAs in rapeseed.

Therefore, a total of 89 climatic and soil factors were collected using publicly available databases and laboratory measurements. Stepwise linear regression analysis was applied to identify and prioritize the factors affecting the  $\delta^2\text{H}$  values of FAs. The approach to determine these factors was summarized in three steps: 1) linear regression to exclude factors that do not have a linear relationship with  $\delta^2\text{H}$  of FAs; 2) Pearson correlation analysis to determine the most correlated factors with time resolution in Table 18; 3) stepwise regression analysis to prioritize these factors by iteratively adding and removing environmental factors based on their contribution to the predictive power of the model. The factors that could be rationally explained with available literature were considered effective. The analysis demonstrated that the most significant factors were the air temperature in February, precipitation amount in April, air temperature in May, soil silt content, altitude, and  $\text{P}_2\text{O}_5$  content in the soil. Some of these factors may affect the  $\delta^2\text{H}$  values of FAs in rapeseed through their influence on the precipitation, while others may potentially impact through isotopic fractionation during plant growth. The influence of rapeseed variety was not observed. Moreover, the climate conditions had a more significant impact on rapeseed than the soil properties. These factors are site-specific and vary spatially and temporally. Therefore, the significant correlations between them suggest that the  $\delta^2\text{H}$  of FAs in rapeseed could be used to identify the geographical origin.

In summary, the  $\delta^2\text{H}$  values of FAs in rapeseeds in Hesse can be accurately measured using the existing GC-Py-IRMS technique. The spatial distribution pattern of the  $\delta^2\text{H}$  of FAs in rapeseed from 2020 showed similarities to the  $\delta^2\text{H}$  of precipitation. However, this similarity was not observed in 2019, which means more factors influence the  $\delta^2\text{H}$  values of FAs. Therefore, the linear stepwise regression analysis was performed between various environmental factors and  $\delta^2\text{H}$  values of FAs in rapeseed. The analysis reveals that environmental factors, including some climatic and soil conditions, significantly influence the  $\delta^2\text{H}$  values of FAs. The finding highlights the potential of  $\delta^2\text{H}$  of FAs as markers for geographical identification, considering the variations resulting from different factors in distinct geographic locations.



---

## 6. Development of a fingerprinting approach to determine the geographical origin of rapeseed using H-CSIA of FAs and elemental composition combined with chemometrics

---

### 6.1. Introduction

This chapter addresses the third research objective, which focuses on determining the geographical origin of rapeseed using stable hydrogen isotopic composition of FAs and elemental composition of rapeseed combined with chemometrics. The specific objectives are:

1. To distinguish rapeseed samples from Hesse and non-Hesse;
2. To identify the geographical origin of rapeseed within Hesse.

A diverse set of samples was employed for the first sub-objective, including those selected in Hesse during 2019 and 2020, as well as additional rapeseed samples (*Brassica napus*) from Jianhu, China, collected in 2019, serving as representatives of the non-Hesse region. For the second sub-objective, the samples in Hesse during 2019 were applied.

These samples were analyzed for the  $\delta^2\text{H}$  values of the four FAs and elemental composition, including macro- and microelements. Chemometrics tools, PCA and OPLS-DA, were applied to build discriminative models to classify the rapeseed samples based on their geographical origins. In the context of chemometrics, the  $\delta^2\text{H}$  values of FAs and elementals composition are also called predictive variables, which are used to differentiate the geographic origin. The labels for the classification, such as geographic origin or other groups, are called response variables. The software SIMCA was employed for these data analyses.

The following sections present the constructed discriminative models with different predictive variables and their discrimination performance. The role of the  $\delta^2\text{H}$  values of FAs and elemental composition in differentiating the geographical origin of rapeseed is discussed.

### 6.2. Descriptive statistics of rapeseed from JianHu and Hesse

Table 20 shows the  $\delta^2\text{H}$  values of the four main FAs and elements of rapeseed samples obtained from JianHu and Hesse in 2019. The range of  $\delta^2\text{H}$  values from Jianhu was -179 mUr to -159 mUr for C16:0 FA, -158 mUr to -141 mUr for C18:1 FA, -165 mUr to -150 mUr for C18:2 FA, and -195 mUr to -178 mUr for C18:3 FA. The t-test showed a statistically significant difference in  $\delta^2\text{H}$  values of C18:1 FA and C18:2 FA between Hesse and Jianhu. However, no significant difference was observed for C18:3 FA and C16:0 FA.

For the rapeseed samples from Jianhu, the relationship of the  $\delta^2\text{H}$  values between the four FAs was similar to that observed in Hesse. Specifically, the C18:1 FA had higher values compared to C18:2 and C18:3 FA, while the  $\delta^2\text{H}$  values of C16:0 FA fell between C18:2 and C18:3. The observed isotopic fractionation in hydrogen between C18:1, C18:2, and C18:3 FAs is consistent with the findings of other vegetable oil such as olive oil and peanut seed oil (Duan et al., 2002; Paolini et al., 2017). This can be explained by the enzymes' preference for light hydrogen isotope  $^1\text{H}$  during elongation and desaturation (Duan et al., 2002).

Additionally, Table 20 presents the content of 20 elements in rapeseed from both regions, measured in milligrams per kilogram (mg/kg) or micrograms per kilogram ( $\mu\text{g}/\text{kg}$ ). However, the measured values of certain trace elements (V, As, Se, Nb, Sb, Th, and U) fell below the detection limit of the measuring instrument, precluding their use as valid markers. Consequently, their results are not presented in Table 20. Among the 20 elements, only K, Cu, Rb, Ni, Mo, Cd, Sn, and Pb exhibited no significant differences in content between the two regions. The content of the remaining elements showed significant variations between the two regions.

Table 20: Average, minimum, and maximum values of  $\delta^2\text{H}$  values of FAs and elements in rapeseeds of 2019 from Hesse and Jianhu.

Analytical markers	Unit	Hesse, Germany (n = 37)				JianHu, China (n = 28)			
		Min	Max	M	SD	Min	Max	M	SD
C16_0_FA	mUr	-168.6	-173.1	-161.6	2.7	-179.1	-158.8	-167.5	5.1
C18_1_FA*	mUr	-161.9	-148.8	-154.9	2.9	-158.1	-141.2	-150.9	3.8
C18_2_FA*	mUr	-173.0	-153.7	-161.1	4.4	-164.8	-149.6	-156.0	3.6
C18_3_FA	mUr	-193.9	-176.9	-187.6	3.5	-194.7	-177.7	-186.5	4.5
Ca**	mg/kg	3869.5	6271.6	5064.1	532.5	3783.8	5197.5	4394.7	353.1
Fe*	mg/kg	48.8	99.4	65.2	9.3	45.9	365.6	112.4	80.6
K	mg/kg	6393.0	10314.0	8404.3	926.2	5895.4	9739.9	7607.2	988.3
Mg*	mg/kg	2389.3	3308.0	2866.5	256.2	2573.0	3712.0	3098.9	300.7
Mn**	mg/kg	18.11	42.68	29.83	5.33	26.7	54.0	35.7	6.3
Sr*	mg/kg	7.01	25.51	12.03	4.69	10.9	19.0	14.9	2.2
Zn**	mg/kg	31.40	53.88	39.45	4.98	25.1	42.3	33.3	4.2
Cu	mg/kg	1.88	3.99	2.86	0.44	1.2	3.7	2.9	0.6
Rb	mg/kg	1.17	71.61	6.42	11.55	1.0	34.1	8.8	7.8
Ba*	mg/kg	0.889	10.195	4.022	2.349	0.7	7.9	2.8	1.7
Ni	mg/kg	0.12	2.18	0.61	0.47	0.2	2.2	0.8	0.5
Mo	mg/kg	0.3	1.2	0.5	0.2	0.2	0.8	0.5	0.2
Cr**	$\mu\text{g}/\text{kg}$	24.2	442.5	88.6	67.1	44.2	1994.7	442.3	479.0
Co*	$\mu\text{g}/\text{kg}$	10.9	106.6	25.6	17.3	9.1	132.6	44.8	27.3
Ag**	$\mu\text{g}/\text{kg}$	1.8	18.6	6.5	4.9	3.5	17.6	11.1	3.8
Cd	$\mu\text{g}/\text{kg}$	14.2	71.7	33.1	13.6	5.5	92.4	30.7	16.5
Sn	$\mu\text{g}/\text{kg}$	10.3	39.4	18.8	7.4	5.3	45.6	16.8	7.7
Cs*	$\mu\text{g}/\text{kg}$	0.6	129.1	14.8	26.5	2.8	107.8	35.9	28.2
Tl*	$\mu\text{g}/\text{kg}$	45.2	473.2	159.1	91.9	40.1	199.0	105.3	45.3
Pb	$\mu\text{g}/\text{kg}$	9.8	756.5	53.2	124.6	7.5	209.5	60.6	43.7

Note: Analytical markers followed by one and two asterisks in the table indicate significant differences between two locations at  $p < 0.05$  and  $p < 0.001$  by T-Test.

### 6.3. Classification of the geographical origin of rapeseed samples from Hesse and non-Hesse

This subchapter presents three distinct models developed based on a combination of the  $\delta^2\text{H}$  values of FAs and elemental composition (predictive variables) coupled with chemometrics techniques. They were built to classify the rapeseeds between Hesse and Jianhu (an example of non-Hesse). The contribution of different types of predictive variables for the classification by comparing these three models is discussed. Further, the temporal stability of the model is assessed.

Model 1 focused solely on the  $\delta^2\text{H}$  values of the four FAs in rapeseed. Since the hydrogen isotopic fractionation occurs between FAs with different chain lengths and saturations, which is influenced by environmental factors, additional isotopic variables (16 in total, shown in Table 21) related to differences and ratios between the four main hydrogen isotopic composition were included as predictive variables. The samples included 37 rapeseed samples from Hesse and 28 samples from Jianhu, China, in 2019. About 70% of the samples were randomly selected as training samples, while the remaining 30% were used as prediction samples to test the discrimination performance of the models. Hence, the training sample sizes for Hesse and Jianhu were 28 and 19, respectively, with nine predicted samples in both regions.

Table 21: 16 predictive variables related to the  $\delta^2\text{H}$  values of FAs.

$\delta^2\text{H}$ of FAs	$\delta^2\text{H}$ difference between FAs	$\delta^2\text{H}$ ratio between FAs
$\delta^2\text{H}$ of C18:1 FA	$\delta^2\text{H}$ difference: C16:0-C18:1 FA	$\delta^2\text{H}$ ratio: C16:0/C18:1 FA
$\delta^2\text{H}$ of C18:2 FA	$\delta^2\text{H}$ difference: C16:0-C18:2 FA	$\delta^2\text{H}$ ratio: C16:0/C18:2 FA
$\delta^2\text{H}$ of C18:3 FA	$\delta^2\text{H}$ difference: C16:0-C18:3 FA	$\delta^2\text{H}$ ratio: C16:0/C18:3 FA
$\delta^2\text{H}$ of C16:0 FA	$\delta^2\text{H}$ difference: C18:1-C18:2 FA	$\delta^2\text{H}$ ratio: C18:1/C18:2 FA
	$\delta^2\text{H}$ difference: C18:1-C18:3 FA	$\delta^2\text{H}$ ratio: C18:1/C18:3 FA
	$\delta^2\text{H}$ difference: C18:2-C18:3 FA	$\delta^2\text{H}$ ratio: C18:2/C18:3 FA

Model 2 extended the analysis by combining the elemental composition of rapeseeds (listed in Table 20) and the  $\delta^2\text{H}$  values. The samples used for this model were the same as Model 1. Model 2 aimed to assess the combined effect of isotopic and elemental composition in distinguishing geographical origins.

Model 3 incorporated the samples from multiple years, including samples from 2019 and 2020 in Hesse and samples of 2019 in Jianhu, using the same predictive variables as Model 2. This model aimed to assess the temporal stability of the model in identifying geographic origins. 36 Hessian samples from 2020 were added, resulting in 73 Hessian samples (52 as training samples and 21 as prediction samples). The datasets (samples and variables) for the three models are listed in Table 22.

Table 22: Overview of the datasets of three models (Model 1-3).

Model	Origin	Year	n		Predictive variables	
			Training set	Prediction set	Hydrogen isotopic composition	Elements
1	Hesse	2019	28	9	16	n.a.
	Jianhu		19	9		
2	Hesse	2019	28	9	16	20
	Jianhu		19	9		
3	Hesse	2019 - 2020	52	21	16	20
	Jianhu	2019	20	8		

The model-building process involves the following steps:

First, PCA was performed to detect outliers and to reduce the dimension by converting the predictive variables into principal components, which retain the most information while reducing the number of variables. PCA is an unsupervised method that does not require any labeled or target variable during the analysis. In this case, the geographic origins of the training samples were unknown. A few outliers were identified for Models 1, 2, and 3 (n=1,2,3, respectively), but they were not significant measurement errors and were retained for the model building. The analysis revealed that most samples from the two geographical origins were distributed in different areas of the score plot (Figure A- 6). It indicates that the used variables can reflect geographic differences and enable geographic discrimination.

Second, discriminant analysis using orthogonal least squares was conducted to identify key variables driving separation between the groups and to assess their importance in class discrimination. This supervised approach utilizes training samples of known origin to build a discriminative model that maximizes the difference between samples from different origins. The score plot visually demonstrates the discriminatory power of the models and provides insights into group separation. Three OPLS-DA models are presented in Figure 33a, c, and e. All of them revealed two distinct groups, rapeseed from Hesse and Jianhu.

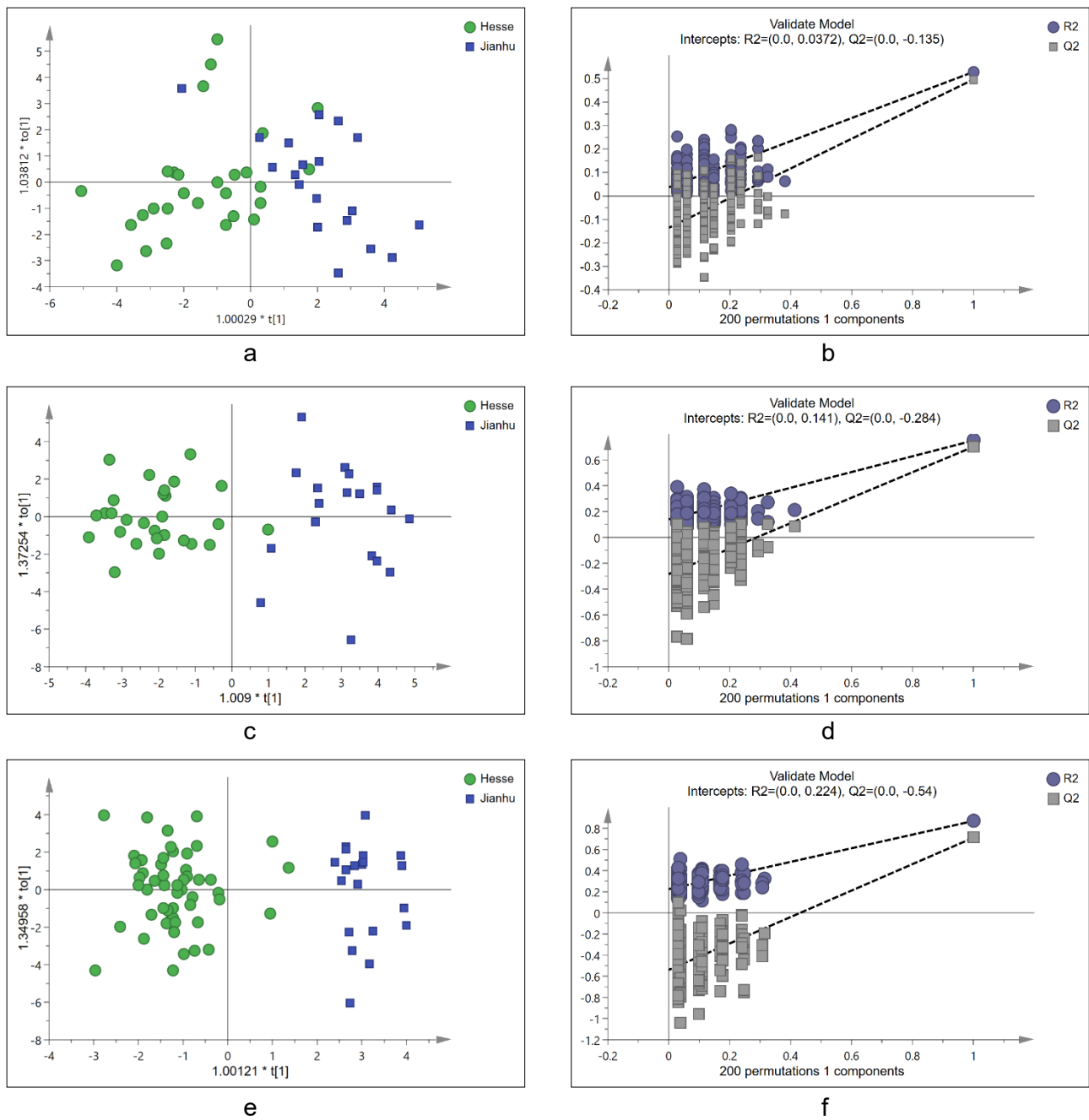


Figure 33: Score plots of the three OPLS-DA models and the corresponding permutation plots.

Note: a and b for Model 1, c and d for Model 2, e and f for Model 3

However, the OPLS-DA algorithm carries the risk of overfitting, where the model fits the training set well but fails to predict new samples accurately. Therefore, a model diagnostic test was performed to assess the model's performance. The combination of the cross-validation and permutation tests can provide a reliable estimation of the model performance (Kjeldahl & Bro, 2010).

The following parameters and tests were applied:

$R^2$  (coefficient of determination) and  $Q^2$  (cross-validated coefficient of determination) are indicators to evaluate the robustness and predictive power of the models.  $R^2$  quantifies the percentage of the variance of the reference values that the fitted line can explain. The value ranges from 0 to 1, with

---

higher values indicating better explanatory power.  $Q^2$  assesses the predictive ability, with values ranging from  $-\infty$  to 1. The closer the value is to 1, the better the predictive ability. Low  $Q^2$  values indicate an overfitted model that may not accurately predict new samples (MKS Umetrics, 2012).

CV-ANOVA (Cross-Validated Analysis of Variance testing of predictive residuals) is a significance test to assess the reliability of the model. It compares the residuals of a global model containing all OPLS-DA components with a purely random model that lacks any relationship between the response variable and the predictors. The null hypothesis is that the residuals of these two models are equal (MKS Umetrics, 2012). A p-value less than 0.05 indicates a significant model. In this case, the response variables were the geographical origin, while the predictors are  $\delta^2H$  values of FAs and elemental composition, shown in Table 20 and Table 21.

Furthermore, a permutation test aims to assess the significance of a classification result. It involves comparing the goodness of fit ( $R^2$  and  $Q^2$ ) of the original model with those of models based on randomly permuted sample groups while keeping the variables-matrix intact. By conducting multiple permutations, the test determines if the fit of the original model is significantly better than that of the random permutation model. If yes, it suggests that the observed class separation in the original data is unlikely due to random factors. The plot's slopes of the regression lines (shown in Figure 33 b, d, f) were greater than 0, indicating that the original model is statistically significant (Rubingh et al., 2006). In this case, the class assignment was permuted 200 times.

Table 23 and Figure 33 present the model diagnostic results. Model 1, Model 2, and Model 3 explained 52.9%, 75.1%, and 87% of the variation, respectively. The cross-validation test indicated that these models had 49.7%, 70.3%, and 71.2% predictive abilities for samples of unknown origin based on the training datasets. The p-values of CV-ANOVA were less than 0.05. The slopes of the permutation lines of the three models were more significant than 0 (Figure 33 b, d, f). All three models demonstrated statistical significance and did not exhibit signs of overfitting.

Finally, three external datasets (prediction set shown in Table 22), which were not applied during the model training phase, were employed to assess the predictive performance of each model, respectively. Sensitivity, specificity, and accuracy were applied to evaluate the model's predictive ability.

Sensitivity, also known as true positive rate, refers to the proportion of samples belonging to the modeled classes correctly identified and accepted by the classification model. Specificity represents the percentage of samples from different classes that are accurately recognized as not belonging to the modeled classes and are consequently rejected by the classification model (Ruiz-Samblás et al., 2013). Accuracy represents the percentage of samples correctly classified (Bajoub et al., 2017). The results are listed in Table 23.

Once validated, important variables were identified using VIP (Variable Importance for the Projection) values. VIP values greater than 1 indicate the most relevant variables for discriminating between the two groups (shown in Table 24). They are arranged in descending order based on their VIP values. Comparing VIP values of the same variable across different models is not meaningful. However, a valid comparison can be made between the VIP values of different variables within the same model.

Table 23: Performance parameter of the three models.

Performance parameter		Model 1	Model 2	Model 3
R <sup>2</sup>		0.529	0.751	0.87
Q <sup>2</sup>		0.497	0.703	0.712
p-value of CV-ANOVA		< 0.001	< 0.001	< 0.001
sensitivity	Hesse	66.70%	55.60%	80.95%
	Jianhu	88.90%	100%	75%
specificity	Hesse	44.40%	88.90%	87.50%
	Jianhu	77.80%	100%	100%
accuracy	Hesse	55.60%	72%	82.75%
	Jianhu	83.30%	100%	93.10%

In Model 1, out of the 16 variables, nine contributed significantly to the discriminative ability, as indicated in Table 24. Interestingly, the ratio and difference of  $\delta^2\text{H}$  values in different FAs played more critical roles in the classification than the  $\delta^2\text{H}$  values of individual FAs. It suggests that the hydrogen isotopic fractionation occurring during the FA biosynthesis process in rapeseed exhibited significant regional variation between the two regions. Consequently, this finding implies that employing CSIA, specifically considering the  $\delta^2\text{H}$  values of different FAs, could provide more informative data for distinguishing the geographical origin of the rapeseed.

In Model 2, as seen in Table 24, 19 of 36 variables contributed significantly to the discriminative ability. The  $\delta^2\text{H}$  values of FAs and their differences and ratios were more effective in distinguishing the two origins than the elements. Nevertheless, elements such as Cr, Fe, Ag, Ca, Zn, Mn, and K contributed to distinguishing the two regions. The differences in these element contents could be attributed to various factors, including differing fertilization practices, varying degrees of soil pollution, and natural variations in soil composition.

Regarding the predictive power of the models, Model 2, combining both isotopic composition and elements as predictive variables, exhibited superior predictive performance compared to Model 1. The prediction accuracy of Model 2 for external Hessian rapeseed reached 72%, while for external Jianhu's rapeseed, it reached 100%. The variables in Model 2 represented different types, which offers more perspectives on the differences between geographical origins. The  $\delta^2\text{H}$  values of FAs primarily reflect climate variations, while the element composition predominantly reflects soil conditions, including human activities like fertilizer use. These differences ultimately contributed to the differentiation between the two regions.

Table 24: VIP values of variables in three OPLS-DA models (Model 1-3).

Model 1		Model 2		Model 3	
Variables	VIP	Variables	VIP	Variables	VIP
$\delta^2\text{H}$ ratio: C18:2/C18:3 FA	1.4	$\delta^2\text{H}$ ratio: C18:2/C18:3 FA	1.6	Cs	1.6
$\delta^2\text{H}$ difference: C18:2-C18:3 FA	1.3	$\delta^2\text{H}$ difference: C18:2-C18:3 FA	1.6	Cr	1.5
$\delta^2\text{H}$ ratio: C18:1/C18:3 FA	1.3	$\delta^2\text{H}$ ratio: C16:0/C18:2 FA	1.5	$\delta^2\text{H}$ ratio: C18:1/C18:3 FA	1.4
$\delta^2\text{H}$ ratio: C16:0/C18:2 FA	1.3	Cr	1.5	$\delta^2\text{H}$ ratio: C18:2/C18:3 FA	1.3
$\delta^2\text{H}$ difference: C16:0-C18:1 FA	1.3	$\delta^2\text{H}$ ratio: C18:1/C18:3 FA	1.5	Mg	1.3
$\delta^2\text{H}$ ratio: C16:0/C18:1 FA	1.3	$\delta^2\text{H}$ difference: C16:0-C18:2 FA	1.4	$\delta^2\text{H}$ difference: C18:2-C18:3 FA	1.3
$\delta^2\text{H}$ difference: C16:0-C18:2 FA	1.3	$\delta^2\text{H}$ difference: C16:0-C18:1 FA	1.4	$\delta^2\text{H}$ difference: C18:1-C18:3 FA	1.2
$\delta^2\text{H}$ difference: C18:1-C18:3 FA	1.2	$\delta^2\text{H}$ ratio: C16:0/C18:1 FA	1.4	Fe	1.2
$\delta^2\text{H}$ of C18:1 FA	1.1	$\delta^2\text{H}$ difference: C18:1-C18:3 FA	1.4	Sn	1.2
		$\delta^2\text{H}$ value of C18:1 FA	1.3	Rb	1.2
		Fe	1.2	Sr	1.1
		Ag	1.2	Ti	1.1
		Ca	1.2	$\delta^2\text{H}$ value of C18:1 FA	1.1
		Zn	1.2	$\delta^2\text{H}$ value of C18:2 FA	1
		$\delta^2\text{H}$ of C18:2 FA	1.1	$\delta^2\text{H}$ difference: C16:0-C18:3 FA	1
		Mn	1.1	Zn	1
		K	1.1		
		Cs	1		
		Co	1		

Model 3, including the additional dataset of Hessian samples 2020, considered the temporal stability of the model in identifying geographic origins. The same variables as those in Model 2 were applied in Model 3. The predictive accuracy for external rapeseed, regardless of whether it originated from 2019 or 2020, reached 82.75% for Hesse and 93.1% for Jianhu (Table 23). The high predictive accuracy suggests that the variables applied in the model exhibited a degree of resilience against temporal fluctuations. In other words, these variables remained informative and can effectively distinguish the geographic origins of Hessian rapeseed samples consistently across different years. Compared to Model 2, there was a decrease in the number of variables related to the  $\delta^2\text{H}$  values with VIP values greater than 1, and their importance rankings also decreased (shown in Table 24). On the other hand, the importance of certain elements increased, mainly Cs and Cr, in Model 3. The elements such as Cs, Cr, Mg, Fe, Sn, Rb, Sr, Ti, and Zn were identified as important variables for the model. The changes in rankings highlighted the impact of incorporating data from multiple years on the discriminative ability of the model. It implies that the variables associated with the  $\delta^2\text{H}$  values may exhibit less stability and consistency across different years, while the importance of certain



---

elements in determining geographical origin remains robust.

Through modeling, key elements crucial for classification were identified, as shown in Table 24. These elements were selected purely through mathematical methods. However, within the context of soil-plant interactions, many of these plant elements originate from their environments, such as soil and water. Therefore, the significant role these elements play in classifying rapeseed from the two regions can likely be attributed to differences in the soil environment of these regions.

For instance, Chromium (Cr) in seeds from both regions was a significant variable in Model 2 and Model 3. The total Cr content in plant tissues does not directly correlate with the total Cr content in the soil, as plants can only absorb bioavailable Cr. Nonetheless, if the total soil Cr increases, the amount of bioavailable Cr also tends to increase (Adriano, 2001). The total Cr content in soil is largely determined by the parent material; for example, soils derived from ultramafic igneous rocks, shales, and clays tend to have high levels of Cr (Adriano, 2001). In general, the total Cr content in agricultural soils ranges from 1.4 to 1300 mg/kg (Alloway, 2013). The forms of Cr in the soil that are available for plant absorption include the most readily absorbed forms, typically existing as trivalent chromium ( $\text{Cr}^{3+}$ ) and hexavalent chromium ( $\text{CrO}_4^{2-}$  or  $\text{Cr}_2\text{O}_7^{2-}$ ) (Adriano, 2001); organically bound Cr; exchangeable Cr, which is adsorbed onto the surfaces of soil particles, especially clay minerals and organic matter; and weakly bound Cr, which is physically adsorbed or weakly chemically bound to soil particles (Alloway, 2013). Several factors affect the bioavailability of Cr to plants, such as soil pH, the oxidation state of Cr, the presence of reducing and oxidizing materials, Fe and Mn oxides, and redox potential. These factors influence the mobility and solubility of Cr, thereby affecting its bioavailability (Adriano, 2001).

Additionally, trace amounts of Cr in fertilizers can be absorbed by plants, with levels ranging from 30 to 3000 mg/kg in P fertilizers and 60 to 250 mg/kg in superphosphates (Adriano, 2001). Ultimately, the typical range of Cr in food crops is between 0.013 and 4.2 mg/kg (Alloway, 2013). The measurements show that Cr content in seeds from Hesse ranged from 0.024 to 0.44 mg/kg, while in Jianhu, it ranged from 0.044 to 1.99 mg/kg, both within the typical range. Table 20 shows a significant difference in Cr content between seeds from the two regions. This discrepancy may be due to differences in the total soil Cr content or soil properties such as pH and Fe and Mn content. Additionally, as previously mentioned, variations in the fertilizers used in the two regions, which contain different amounts of bioavailable Cr, could also contribute to the differences in seed Cr content. Further analysis of specific factors would require additional measurement data.

In summary, this fingerprinting approach, using the  $\delta^2\text{H}$  values of FAs and elemental composition, with the aid of OPLS-DA, proves highly effective in classifying the geographical origin of rapeseed samples from Hesse and non-Hesse (Jianhu). Upon comparing Model 1 and Model 2, it becomes clear that the integration of both hydrogen isotopic composition and elemental composition as predictive variables significantly enhances the discriminative capability of the model, surpassing the efficacy observed when solely relying on hydrogen isotopic composition. Further, contrasting Model 2 with Model 3, using a dataset spanning two years rather than just one, results in notable improvement in accuracy. The predictive accuracy reaches 82.75% for external Hessian rapeseed

---

and 93.1% for external Jianhu rapeseed. This suggests that despite the temporal variation, the spatial difference in geographical origin can still be detected using these variables.

## **6.4. Classification of the geographical origin of rapeseed samples in Hesse**

The same chemometric techniques were employed to distinguish rapeseed samples from various locations within Hesse, Germany. The objective was to explore the classification potential in Hesse and to assess the discriminative power of different models, including different types of predictive variables. PCA was performed on the dataset of rapeseed samples from Hesse in 2019, consisting of 37 samples with 36 variables (Table 22). However, no clear grouping pattern was observed based on northern, central, and southern Hesse (Figure A- 7 in the Appendix).

To further explore the classification potential of rapeseed samples in Hesse, the  $\delta^2\text{H}$  values of FAs and elemental composition were analyzed concerning climatic and geological conditions, respectively, rather than relying on administrative boundaries.

### **6.4.1. Classification based on climatic factors**

The analysis conducted in Chapter 5 highlighted the significant influence of climatic factors, specifically the air temperature in February and May and the precipitation amount in April, on the  $\delta^2\text{H}$  values of FAs in rapeseed. Therefore, the classification of samples based on climatic factors included two key aspects:

- 1) Dividing Hesse into distinct areas based on these climatic factors, termed as climatic zones.
- 2) Classifying the rapeseed samples according to these climatic zones.

Hesse was first divided based on climatic factors. The numerical values of the climatic conditions divided Hesse into high and low regions. For instance, considering the air temperature in February, with a distribution ranging from 1.3 to 6.6 °C and a median value of 4.2 °C, all regions below 4.2 °C were classified as low-temperature areas, while those above 4.2 °C were classified as high-temperature areas. Similar categorization methods were applied to the other two climate factors. Consequently, these three categorizations resulted in three different sets of regions within Hesse. The regions defined by each climatic factor were overlaid, creating a composite map with eight unique combinations. As a result, Hesse was effectively differentiated into eight climatic zones based on these combinations of climatic factors. The clustering was implemented using ArcGIS Pro software. The eight climatic zones of Hesse and the description of the legend are presented in Figure 34 and Table 25.

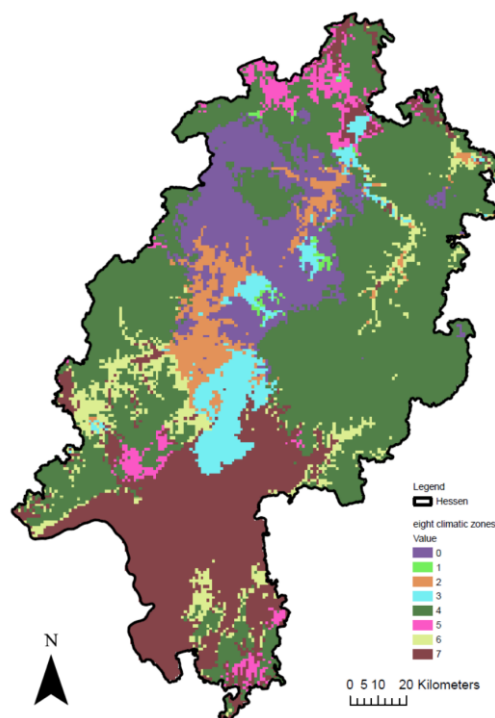


Figure 34: Clustering of Hesse in 8 climatic zones.

Table 25: Description of the legend for the eight distinct areas in Hesse.

Climatic zone	Description	n
0	low temperature in February, low temperature in May, and low precipitation in April	8
1	high temperature in February, low temperature in May, and low precipitation in April	0
2	low temperature in February, high temperature in May, and low precipitation in April	1
3	high temperature in February, high temperature in May, and low precipitation in April	3
4	low temperature in February, low temperature in May, and high precipitation in April	15
5	high temperature in February, low temperature in May, and high precipitation in April	3
6	low temperature in February, high temperature in May, and high precipitation in April	3
7	high temperature in February, high temperature in May, and high precipitation in April	4

Note: Low temperature in February < 4.2 °C; low temperature in May < 11 °C; low precipitation amount < 27 mm.

A total of 37 samples from 2019 were assigned to 7 groups based on the climatic zones. No sample originated from Area 1, as indicated in Table 25.

PCA was performed on the dataset of rapeseed samples from Hesse in 2019, consisting of 37 samples with 16 variables related to the  $\delta^2\text{H}$  values of FAs in rapeseed. However, the resulting score plot did not exhibit any clear groupings, as depicted in Figure A- 8 of the Appendix. A further classification between any two climatic zones was attempted. It was found that the samples from Climatic zone 0 and Climatic zones 6+7 can be differentiated from each other, as shown in Figure 35. Climatic zone 0 comprised eight samples, while Climatic zones 6+7 consisted of seven samples. Climatic zone 0 represents an area characterized by low temperatures in February and May and low precipitation in April. Conversely, Climatic zones 6+7 refer to regions with high temperatures in May and a high precipitation amount in April, regardless of whether the temperature in February was high

or low.

The discriminatory model (Model 4), developed using OPLS-DA, is illustrated in Figure 36. The same parameters outlined in Section 6.3 were employed to validate the model, with the results summarized in Table 26. The obtained model exhibited an  $R^2$  value of 0.49 and a predictive  $Q^2$  value of 0.35. The positive slope of the permutation test with 200 permutations indicates that the model could effectively discriminate between the two groups, and this classification was unlikely to be driven by random factors. However, the p-value of CV-ANOVA exceeded 0.05, indicating that this model was not statistically significant. This discrepancy between CV-ANOVA and the permutation test might stem from the small sample size. The important predictive variables for the classification of Model 4 are presented in Table 27.

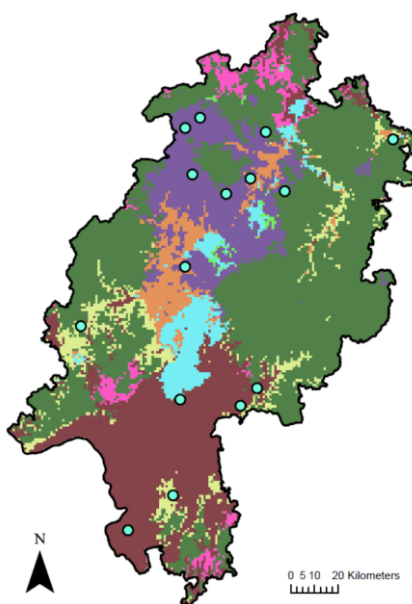


Figure 35: Samples (blue circles) located in climatic zones 0 and 6+7

Model 5, expanding upon Model 4, included elemental composition as additional predictive variables to further perform discriminant analysis on the two sample classes. The performance parameters of this discriminant model are presented in Table 26. The resulting Model 5 illustrates an enhanced ability to distinguish between the sample classes, as evidenced by an impressive  $R^2$  value of 0.95 and a robust  $Q^2$  value of 0.8 (Figure 37). Additionally, both the permutation test (200 permutations) and the p-value of CV-ANOVA confirm the statistical significance of the model. Key predictive variables for the separation of Model 5 are outlined in Table 27.

Table 26: Overview of Model 4-5 datasets and their model performance parameters.

Model	Year	Climatic zone (response variables)	Sample size	Predictive variables	$R^2$	$Q^2$	P-value of CV-ANOVA
4	2019	0	8	$\delta^2H$ values of FAs	0.49	0.35	> 0.05
		6+7	7				
5	2019	0	8	$\delta^2H$ values of elements FAs	0.95	0.8	< 0.05
		6+7	7				

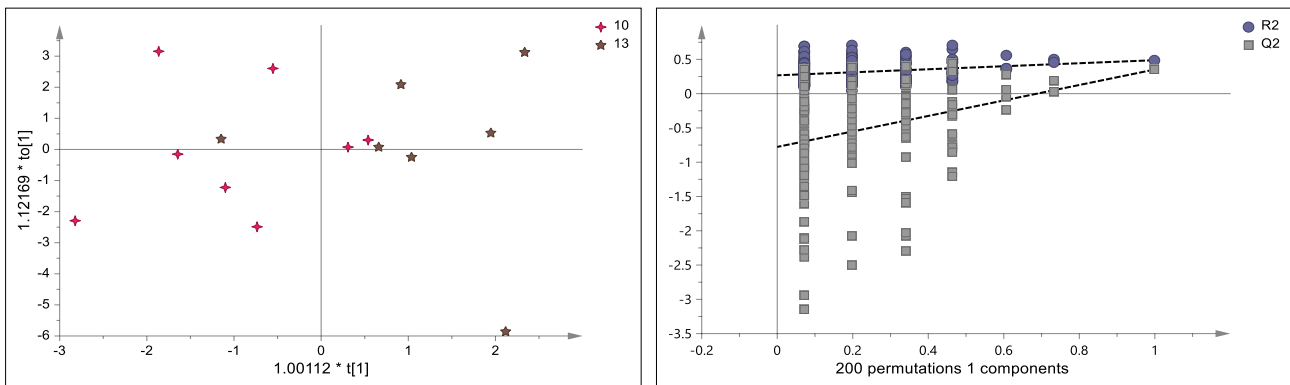


Figure 36: Score plot of Model 4 and the corresponding permutation plots.

Note: Numbers 10 and 13 in the left diagram represent samples from climatic zones of 0 and 6+7.

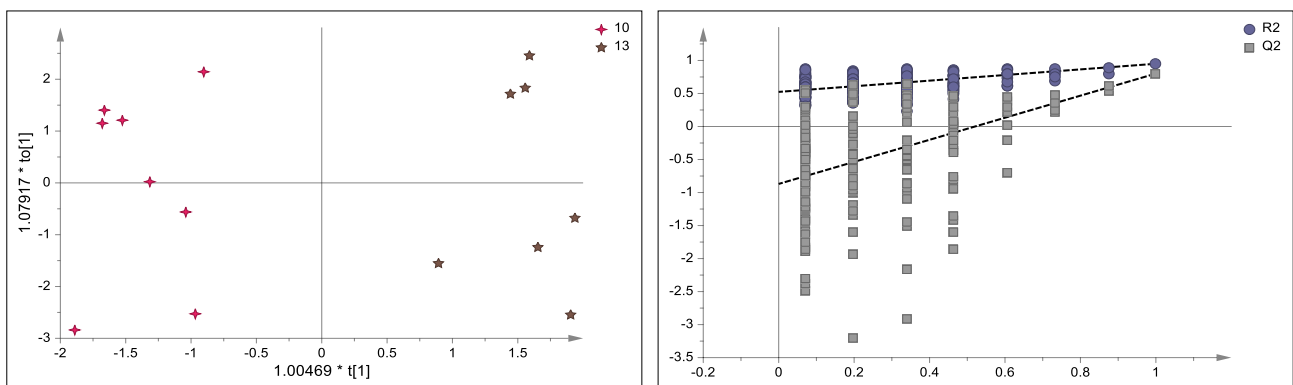


Figure 37: Score plot of Model 5 and the corresponding permutation plots.

Note: Numbers 10 and 13 in the left diagram represent samples from climatic zones of 0 and 6+7.

The significant improvement in discriminant ability indicates slight differences in the element content of rapeseed between the two climatic zones. Cs and TI contributed significantly to the differentiation between the two climatic zones, as shown in Table 27. Cs and TI are not essential elements for plants. TI can particularly inhibit germination, plant growth, and chlorophyll content when its concentration in roots reaches about 2 mg/kg (Kabata-Pendias & Mukherjee, 2007). However, both elements are relatively active in soils and are easily taken up by plants. For example, plants readily absorb Cs from sandy acid soils. Based on limited data, the common range of Cs in various plants is estimated at <math>0.1\text{--}3\text{ mg/kg}</math> (Kabata-Pendias & Mukherjee, 2007). Most Cs are concentrated in the non-edible parts of plants. The measurements indicate that the Cs content in rapeseeds from Hesse in 2019 ranged between 0.001-1.374 mg/kg, which aligns with this estimated range. The concentrations of these two elements in plants are related to factors such as the parent materials of their source area, soil conditions, as previously mentioned in section 6.3, and the pollution level (Kabata-Pendias & Mukherjee, 2007). The reason these elements can distinguish between the two climatic zones (i.e., the zone with low temperatures in May and low precipitation in April from the

zone with high temperatures in May and high precipitation in April) is likely because temperature and precipitation affect the solubility and mobility of elements in soil solutions, thereby influencing plant uptake. For instance, the amount of precipitation can affect the concentration of Cs<sup>+</sup> in soil solution, influencing plant root absorption, which has been observed (Bange & Overstreet, 1960). The impact of temperature on Cs uptake has also been noted, although the mechanisms are not well understood (Bange & Overstreet, 1960). Further research is needed to explore this subject matter more comprehensively.

In summary, rapeseed samples from Hesse were divided into eight groups based on the climatic zones. Distinguishing between the samples from two climatic zones - specifically, the climatic zone characterized by low temperature in May and low precipitation amount in April and the climatic zone marked by high temperature in May and high precipitation amount in April - was achieved using the  $\delta^2\text{H}$  values as predictive variables. Incorporating seed elements into predictive variables improved the ability to differentiate between the two sample classes. Ultimately, differentiation among rapeseed from various climatic zones in Hesse was feasible through a blend of  $\delta^2\text{H}$  values of fatty acids and elemental composition in rapeseed. Notably, these zones were delineated not by direct orientation or administrative boundaries but by climatic factors.

Table 27: VIP values of variables in Model 4-5.

Model 4		Model 5	
Variables	VIP	Variables	VIP
$\delta^2\text{H}$ of C18:1 FA	1.4	TI	1.5
$\delta^2\text{H}$ ratio: C18:1/C18:2 FA	1.3	$\delta^2\text{H}$ ratio: C18:1/C18:3 FA	1.3
$\delta^2\text{H}$ of C16:0 FA	1.3	Cs	1.3
$\delta^2\text{H}$ difference: C18:1-C18:2 FA	1.3	$\delta^2\text{H}$ of C18:1 FA	1.3
$\delta^2\text{H}$ ratio: C18:1/C18:3 FA	1	$\delta^2\text{H}$ difference: C18:1-C18:3 FA	1.2
$\delta^2\text{H}$ difference: C16:0-C18:2 FA	1	Fe	1.1
		$\delta^2\text{H}$ of C16:0 FA	1.1
		Mg	1.1
		Cd	1.1
		Cu	1

#### 6.4.2. Classification based on the geological substrate

An attempt was made to utilize elemental content as predictive variables to differentiate rapeseed based on the geological substrate, which served as the response variable in Hesse. As per data from HLUNG, the geological substrates in Hesse can be roughly categorized into seven main types: aeolian substrate, carbonate substrate, fluvatile substrate, colluvial substrate, pelitic substrate, psephitic substrate, and volcanic substrate (HLUNG, 2011). Hesse boasts about 200 substrate types, comprising the substrates of three soil layers (topsoil, subsoil, and underground) (Friedrich et al., 2011). Accordingly, 37 samples were classified into 17 classes.

PCA was conducted on the elemental composition of rapeseed samples from 2019. There was no obvious grouping pattern based on the main substrates (Figure A- 9). However, samples grown on volcanic substrate (n=7) and substrate with zechstein (n=5) could be differentiated, as illustrated in Figure 38. The discriminatory model, Model 6, constructed by OPLS-DA, is depicted in Figure 39. The  $R^2$  and  $Q^2$  of Model 5 were 0.77 and 0.51, indicating a reasonable goodness fit of the model and a limited predictive ability for discrimination, probably due to the small sample size. The slope of the permutation plot was greater than 0, and the p-value from CV-ANOVA was 0.04, both confirming the significance of this classification. The key elements for separation are listed in Table 30. Thus, the 20 elements can differentiate between these two substrates: vulcanite and zechstein.

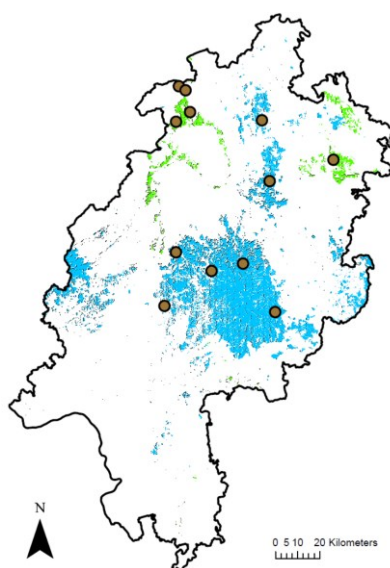


Figure 38: Spatial Distribution of zechstein (marked in green) and vulcanite in Hesse (marked in blue) and samples (brown circle) from these two substrates.

Model 7, building upon Model 6, incorporated additional  $\delta^2\text{H}$  values of FAs as predictive variables. The  $R^2$  value was 0.78, indicating a reasonable goodness fit, while the  $Q^2$  value was 0.60. Its discrimination performance slightly improved compared to Model 6. The p-value of the CV-ANOVA and the permutation test indicate that the model was statistically significant. By comparing the VIP values of variables in Model 7,  $\delta^2\text{H}$  values were deemed more important than elements. However, there was almost no difference in their discriminant ability, indicating that  $\delta^2\text{H}$  values contributed little to the separation. It might be because, according to the classification based on the climatic factors (shown in Figure 34), these 12 samples were located in climatic zones 0 and 4, with different precipitation amounts in April. However, as mentioned earlier, samples from these two climatic zones could not be distinguished by the  $\delta^2\text{H}$  values of FAs. Samples grown on the same geological substrate did not necessarily belong to the same climatic zone, which either interfered with or, at the very least, did not increase the discriminative power of the model. Therefore, in this case, the  $\delta^2\text{H}$  values of FAs did not play a role in distinguishing geology.

Table 28: Overview of Model 6-7 datasets and their model performance parameters.

Model	Year	Group (response variables)	Sample size	Predictive variables	R <sup>2</sup>	Q <sup>2</sup>	P-value of CV-ANOVA	
6	2019	Vulcanite	7	n.a.	Elements	0.77	0.51	< 0.05
		Zechstein	5					
7	2019	Vulcanite	7	$\delta^2\text{H}$ values of FAs	Elements	0.78	0.6	<0.05
		Zechstein	5					

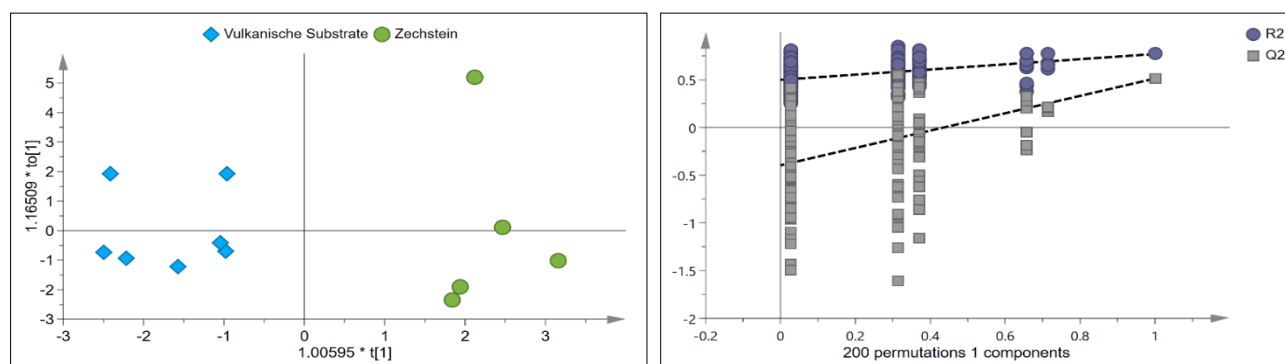


Figure 39: Score plot of Model 6 and the corresponding permutation plot.

Table 29: Performance parameter of Model 6-7.

Performance parameter	Model 6	Model 7
R <sup>2</sup>	0.77	0.78
Q <sup>2</sup>	0.51	0.6
P-value of CV-ANOVA	< 0.05	< 0.05

Table 30: Important variables in Models 6-7.

Model 6		Model 7	
Variables	VIP	Variables	VIP
Ag	1.7	$\delta^2\text{H}$ difference: C16:0-C18:2 FA	1.8
Mg	1.6	$\delta^2\text{H}$ ratio: C16:0/C18:2 FA	1.7
Mn	1.5	$\delta^2\text{H}$ difference: C18:2-C18:3 FA	1.7
K	1.5	$\delta^2\text{H}$ ratio: C18:2/C18:3 FA	1.6
Ni	1.3	Ag	1.6
		Mg	1.6
		Ni	1.5
		K	1.4
		Mn	1.2



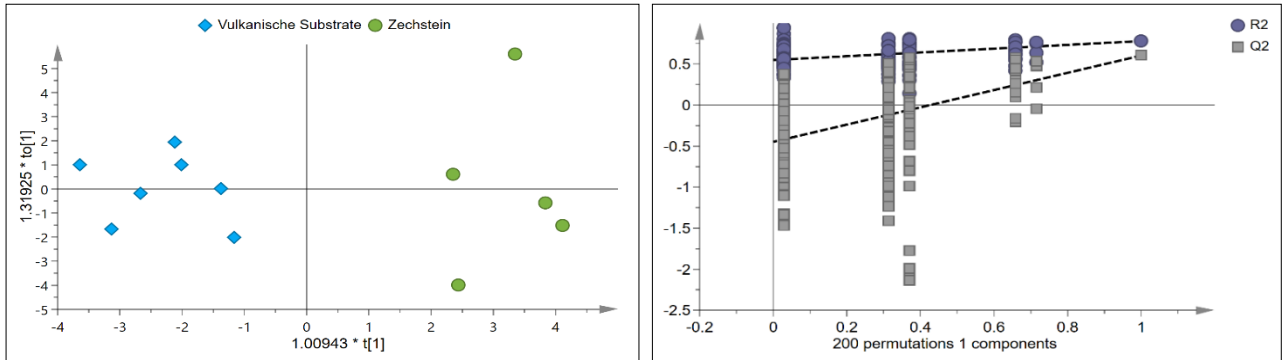


Figure 40: Score plot of Model 7 and the corresponding permutation plot.

In summary, the samples were categorized into different groups based on the geological substrates of the sampling sites. Although employing the 20 elements combined with OPLS-DA did not enable differentiation of all the samples, it effectively distinguished between two substrate types, vulcanite and zechstein, as demonstrated in Model 6. The discriminant model exhibited moderate discriminant ability, suggesting the need to incorporate additional elements and other predictive variables to enhance the modeling. Furthermore, increasing the number of samples to train the discriminant model effectively is essential, ultimately resulting in improved performance.

Moreover, utilizing elements and hydrogen isotopic composition as predictive variables did not significantly enhance the discriminant capability. It could be attributed to the limited differences in climatic conditions between the two substrate types, which may not effectively influence the  $\delta^2\text{H}$  of FAs.

## 6.5. Conclusion

In this chapter, the geographic origin of rapeseed was determined using the  $\delta^2\text{H}$  values of the four FAs and elemental composition. The classification models were developed using chemometrics techniques PCA and OPLS-DA.

Three models were constructed to differentiate rapeseed between Hesse and Jianhu (an example of non-Hesse), with each model using different predictive variables and sample sets.

Model 1 applied the  $\delta^2\text{H}$  values of the four FAs as predictive variables, while Model 2 additionally incorporated the elemental composition. Further, Model 3 extended the analysis to include samples from multiple years (2019-2020). These models demonstrate significant discriminative ability and temporal stability in classifying the selected rapeseed samples from Hesse, Germany, and Jianhu, China. The results indicate that combining  $\delta^2\text{H}$  values of FAs and elemental composition can improve the discriminative ability compared to using the  $\delta^2\text{H}$  values of the four FAs alone. Model 3, trained by data from multiple years, shows temporal stability by achieving higher accuracy rates. That is, 82.75% of the Hessian samples were correctly identified as Hessian, and 93.1% of the Jianhu samples were accurately classified as Jianhu.

Next, the same approach was attempted to distinguish rapeseed samples from various locations within Hesse, Germany. As simple geographic orientations like northern and southern Hesse were

---

insufficient for classification, two approaches were employed: classification based on combined climatic factors and classification based on geological substrate.

Rapeseed samples were divided into eight climatic zones within Hesse based on significant climatic factors such as temperature and precipitation. However, the model only with the  $\delta^2\text{H}$  values of FAs as predictive variables exhibits limited discriminative ability, likely due to the small sample size (only 15 samples). It can still distinguish rapeseed samples grown in areas with high temperatures and high precipitation from areas with low temperatures and low precipitation. More samples from these two climatic zones are needed to validate the model. Adding the elemental composition as predictive variables could significantly improve the distinction between two sample classes grown in these two climatic zones. It might be because climatic conditions may influence elemental composition in rapeseed through root adsorption from the soil materials or the soil elements themselves. Further investigation is required to explore this relationship. Nevertheless, more samples are needed from these two climatic zones for the model training. Therefore, a fingerprinting approach was developed to determine the origins in a small-scale region. Instead of classifying the samples directly by their administrative regions or geographic directions, they were classified by clusters based on significant environmental factors.

Regarding the classification based on geological substrate, seven main types of geological substrates in Hesse were identified. While the initial analysis did not reveal clear grouping patterns based on the main substrate types, a classification model using elements as predictive variables successfully differentiated rapeseed samples grown on the volcanic substrate and substrate with zechstein. The addition of the  $\delta^2\text{H}$  values of FAs did not contribute significantly to the classification, likely due to limited differences in climatic conditions and their influence on the  $\delta^2\text{H}$  values of FAs.

In conclusion, combining the elemental composition, hydrogen isotopic composition of rapeseed, and the OPLS-DA can effectively differentiate the rapeseed samples from Hesse and Jianhu. The proposed model correctly classified 82.75% of the Hessian samples and 93.1% of the Jianhu's samples. Even for the samples within Hesse, this fingerprinting has proven its potential for distinguishing the origins based on climatic factors. In addition, rapeseed samples from the selected two geological substrates in Hesse were also well classified using their elemental composition. However, the discriminative ability of the latter two models can be influenced by factors such as sample size, climatic variations, and the addition of other predictive variables. Therefore, further research, including a larger sample size and combining additional variables, is necessary to enhance the accuracy and robustness of the classification models.

---

## 7. Conclusion and Outlook

---

### 7.1. Conclusion

In this work, a fingerprinting approach to determine the geographical origin of vegetable oils was developed using rapeseed as an example. This dissertation details the procedure for this approach, including the analysis of various parameters of vegetable oils, the acquisition of environmental factors from publicly available databases, and the selection of relevant factors to build models that differentiate vegetable oils based on their geographical origin. Additionally, it presents procedures to test the robustness of these models to ensure their reliability and accuracy. The three research objectives achieved and detailed in Chapters 4-6 are as follows:

1. Evaluate the feasibility of a rapid one-step sample preparation method that converts rapeseed into FAMES for subsequent H-CSIA.
2. Assess the feasibility of  $\delta^2\text{H}$  of FAs as analytical markers for identifying the geographical origin of rapeseed:
  - Investigating spatial and temporal variations in the  $\delta^2\text{H}$  of FAs of rapeseed from Hesse.
  - Exploring the correlation of soil and climatic factors with the  $\delta^2\text{H}$  of FAs in rapeseeds.
3. Determine the geographical origin using  $\delta^2\text{H}$  of FAs and elemental composition combined with chemometrics, including:
  - Distinguishing rapeseed samples from Hesse and non-Hesse.
  - Distinguishing rapeseed samples within Hesse.

Chapter 4 addressed the first objective, enhancing efficiency and streamlining laboratory work in the compound-specific stable hydrogen isotope analysis of rapeseed and its oil. A rapid and simple sample preparation method from Garcés & Mancha (1993) was employed to convert rapeseed into FAMES directly. The influence of 2,2-dimethoxypropane (DMP), added for effective conversion, on the isotopic composition of individual FAMES was evaluated through control experiments. The method demonstrated satisfactory reproducibility, with  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values determined within a precision of  $\pm 0.3$  mUr and  $\pm 3$  mUr, respectively. Notably, the addition of 5 vol% DMP to the transmethylation mixture did not significantly affect the  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of individual FAMES. This result indicates that the rapid one-step method allows for processing multiple oilseed samples in a shorter timeframe, facilitating the application of chemometrics to authenticate various oilseeds.

Chapter 5 addressed the second research objective, evaluating the applicability of  $\delta^2\text{H}$  of FAs in rapeseed as markers for identifying its geographical origin. This evaluation focused on two aspects: (1) determining the spatial variation of  $\delta^2\text{H}$  of FAs in rapeseed samples from Hesse, Germany, and (2) analyzing the correlation of  $\delta^2\text{H}$  of FAs with various climatic and soil factors. The analysis included rapeseed samples collected from Hesse between 2017 and 2020. The results revealed significant spatial variations of  $\delta^2\text{H}$  in rapeseed (14 mUr to 30 mUr for the four major FAs), surpassing the measurement precision of the instrument ( $\pm 3$  mUr). The four major FAs in rapeseed were C18:1, C18:2, C18:3, and C16:0 FA. In 2020,  $\delta^2\text{H}$  values of these FAs in rapeseed showed a negative correlation with the altitude of the sampling sites, more pronounced in certain local areas. This spatial distribution pattern resembled that of precipitation in northern Hesse, indicating that  $\delta^2\text{H}$  values in

---

precipitation were transferred to rapeseed. However,  $\delta^2\text{H}$  values of FAs in rapeseed in 2019 exhibited an inconsistent spatial distribution pattern compared to 2020. For instance, in northern rapeseed,  $\delta^2\text{H}$  values correlated positively with altitude, suggesting that additional environmental factors beyond precipitation influence the  $\delta^2\text{H}$  of FAs. Therefore, stepwise linear regression analysis was applied to determine and prioritize the factors affecting the  $\delta^2\text{H}$  values of FAs. A total of 89 environmental factors were collected. The process for identifying these factors involved three key steps: (1) linear regression to exclude factors that do not exhibit a linear relationship with  $\delta^2\text{H}$  of FAs; (2) Pearson correlation analysis to determine the most correlated factors with time resolution in Table 18; and (3) stepwise regression analysis to prioritize these factors by iteratively adding and removing environmental factors based on their contribution to the predictive power of the model. Factors that could be rationally explained by existing literature were considered effective. It was found that environmental factors such as temperature in February and May, precipitation amount in April, silt content, and altitude influence the  $\delta^2\text{H}$  of the FAs, with climate conditions showing a greater impact than soil properties. These findings suggest that the  $\delta^2\text{H}$  of FAs could serve as markers for identifying geographical origins based on these environmental influences.

Chapter 6 addressed the third research objective, focusing on determining the geographical origin of rapeseed using stable hydrogen isotopic composition of FAs and elemental composition, combined with chemometrics. Classification models were developed to distinguish rapeseed samples between Hesse and non-Hesse, as well as to differentiate among rapeseed within Hesse. First, three models were constructed to differentiate rapeseed between Hesse and Jianhu (an example of non-Hesse), with each model using different predictive variables and sample sets. It was found that the combination of  $\delta^2\text{H}$  of FAs and the elemental composition of rapeseed has an improved classification ability compared to the  $\delta^2\text{H}$  values of FAs alone. Specifically, 82.8% of the Hessian samples were correctly identified as Hesse, and 93.1% of the Jianhu samples were accurately classified as Jianhu. Both  $\delta^2\text{H}$  of FAs and elemental data contributed significantly to the classification.

In distinguishing rapeseed samples from various locations within Hesse, classification based on climatic factors and geological substrate was conducted. Regarding the classification based on climatic factors, the state of Hesse was divided into eight climatic zones based on the three key climatic conditions—monthly average temperature in February and May, monthly average precipitation amount in April—that most influence the  $\delta^2\text{H}$  of FAs, as identified through correlation analysis. Using the combination of  $\delta^2\text{H}$  of FAs and elemental composition, a classification model differentiating between two climatic zones was developed and validated, with a  $Q^2$  value of 0.8, indicating good predictive ability. One climatic zone exhibited low May temperatures and low April precipitation, while the other had high May temperatures and high April precipitation. The classification model only using  $\delta^2\text{H}$  of FAs showed limitations. The significant enhancement in discriminant ability suggests subtle variations in element content in rapeseed between the two climatic zones, potentially attributable to differences in the geological condition itself or the impact of varying climatic factors on rapeseed root absorption. Further research is needed to explore this topic in more detail.

---

Regarding the classification based on geological substrate, the classification model using elemental composition was achieved with the  $Q^2$  value of 0.5 between rapeseed samples grown on the volcanic substrate and substrate with Zechstein. Adding  $\delta^2\text{H}$  of FAs as predictive variables did not significantly improve the classification performance, and the  $Q^2$  value was raised to 0.6. It may imply that no evident correlation existed between the geological substrate and the  $\delta^2\text{H}$  of FAs. Thus, the variations in the geological substrate cannot be reflected in the  $\delta^2\text{H}$  values of FAs.

Overall, in this doctoral work, a fingerprinting approach to determine the geographical origin of vegetable oils was developed. The analysis demonstrated that  $\delta^2\text{H}$  values of FAs in rapeseed are influenced by the environmental factors of their growth sites, including air temperature, precipitation amount, silt content, and soil nutrients. These significant correlations indicate the potential of  $\delta^2\text{H}$  of FAs as analytical markers for geographical traceability. Combining  $\delta^2\text{H}$  of major FAs and elemental composition could be applied to differentiate samples not only between regions with substantial spatial distances (e.g., Germany and China) but also between closely situated regions (e.g., within Hesse, Germany), provided they are clustered by significant environmental factors. This approach could also be adapted for other vegetable oils like soybeans and palm oil with some modifications and adjustments. It holds the potential to be a practical tool for monitoring the origin of agricultural commodities, aiding compliance with EU directives and regulations on sustainability, such as the EUDR. By ensuring the traceability and authenticity of vegetable oils, this approach could help enforce regulatory compliance and promote sustainable practices in the agricultural sector.

## 7.2. Outlook

Several key areas of future research are outlined to improve the field of geographical origin determination using stable hydrogen isotopic composition of FAs in vegetable oils, including rapeseed oil:

### Expanding the sample numbers from target locations

The elemental composition and hydrogen isotopic composition in rapeseed show their potential for distinguishing geographic origins from different climatic zones in Hesse. Further research, with a larger sample size in these two climatic zones, is necessary to enhance the accuracy and robustness of the classification models.

### Increasing the geographic range of sample locations

The current study primarily focused on differentiating rapeseed samples between Hesse and Jianhu, China, and within Hesse. In terms of geographical distance, these are two extreme situations. To gain a more comprehensive understanding of stable hydrogen isotopic composition for identifying geographical origins, future research should expand the range of sampling sites. For instance, including vegetable oils from different states within Germany can provide a broader dataset for analysis.

### Integrating more predictive variable types

While this dissertation investigated the influence of climatic and soil factors on the  $\delta^2\text{H}$  values of FAs,

---

other predictive variables may also contribute to geographical variations. To improve the discriminative ability of the models, additional types of variables should be considered, such as  $\delta^{13}\text{C}$  values and  $^{87}\text{Sr}/^{86}\text{Sr}$ . Incorporating a wider range of variables could result in more accurate and robust geographical origin determination.

#### Exploring non-linear modeling techniques

Although the current research employed linear statistical techniques, there is potential for further improvement by exploring non-linear modeling approaches. Advanced techniques such as machine learning algorithms, artificial neural networks, or other non-linear statistical methods can capture complex relationships between variables and potentially yield superior classification performance.

#### Adaptation of the approach to the other vegetable oil

When analyzing other types of vegetable oils, it is important to consider the unique characteristics of the plants themselves. For example, palm trees have different growth rates and fruiting frequencies compared to rapeseed plants. Therefore, climate data collection and selection should also be adjusted accordingly. Additionally, the elemental composition of other vegetable oils can vary significantly from that of rapeseed oil. The choice of elements to be measured should take into account the accuracy of the laboratory instruments, the specific nature of the oil, and other relevant factors.

By addressing these areas, future research can further refine the approaches for determining the geographical origin of vegetable oils, thereby enhancing the reliability and applicability of this approach across different oil types and geographical regions.

---

---

## Literature

---

- Adriano, D. C. (2001). *Trace Elements in Terrestrial Environments*. Springer New York. <https://doi.org/10.1007/978-0-387-21510-5>
- Ai, G., Sun, T., & Dong, X. (2014). Gas chromatography/isotope ratio mass spectrometry: Analysis of methanol, ethanol and acetic acid by direct injection of aqueous alcoholic and acetic acid samples: GC/IRMS analysis of directly diluted MeOH, EtOH and HOAc. *Rapid Communications in Mass Spectrometry*, 28(15), 1674–1682. <https://doi.org/10.1002/rcm.6948>
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Alloway, B. J. (2013). Bioavailability of Elements in Soil. In O. Selinus (Ed.), *Essentials of Medical Geology* (pp. 351–373). Springer Netherlands. [https://doi.org/10.1007/978-94-007-4375-5\\_15](https://doi.org/10.1007/978-94-007-4375-5_15)
- Alonso-Salces, R. M., Segebarth, N., Garmón-Lobato, S., Holland, M. V., Moreno-Rojas, J. M., Fernández-Pierna, J. A., Baeten, V., Fuselli, S. R., Gallo, B., Berrueta, L. A., Reniero, F., Guillou, C., & Héberger, K. (2015). <sup>1</sup>H-NMR and isotopic fingerprinting of olive oil and its unsaponifiable fraction: Geographical origin of virgin olive oils by pattern recognition. *European Journal of Lipid Science and Technology*, 117(12), 1991–2006. <https://doi.org/10.1002/ejlt.201400243>
- Amelung, W., Blume, H.-P., Fleige, H., Horn, R., Kandeler, E., Kögel-Knabner, I., Kretschmar, R., Stahr, K., & Wilke, B.-M. (with Gaiser, T., Gauer, J., Stoppe, N., Thiele-Bruhn, S., & Welp, G.). (2018). *Scheffer/Schachtschabel Lehrbuch der Bodenkunde* (F. Scheffer & P. Schachtschabel, Eds.; 17., überarbeitete und ergänzte Auflage). Springer Spektrum Berlin, Heidelberg. <https://doi.org/10.1007/978-3-662-55871-3>
- Araguás-Araguás, L., Rozanski, K., Gonfiantini, R., & Louvat, D. (1995). Isotope effects accompanying vacuum extraction of soil water for stable isotope analyses. *Journal of Hydrology*, 168(1), 159–171. [https://doi.org/10.1016/0022-1694\(94\)02636-P](https://doi.org/10.1016/0022-1694(94)02636-P)
- Bajoub, A., Medina-Rodríguez, S., Gómez-Romero, M., Ajal, E. A., Bagur-González, M. G., Fernández-Gutiérrez, A., & Carrasco-Pancorbo, A. (2017). Assessing the varietal origin of extra-virgin olive oil using liquid chromatography fingerprints of phenolic compound, data fusion and chemometrics. *Food Chemistry*, 215, 245–255. <https://doi.org/10.1016/j.foodchem.2016.07.140>
- Ballin, N. Z., & Laursen, K. H. (2019a). To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication. *Trends in Food Science & Technology*, 86, 537–543. <https://doi.org/10.1016/j.tifs.2018.09.025>
- Ballin, N. Z., & Laursen, K. H. (2019b). To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication. *Trends in Food Science & Technology*, 86, 537–543. <https://doi.org/10.1016/j.tifs.2018.09.025>
- Baltes-Götz, B. (2022). *Lineare Regressionsanalyse mit SPSS*. Zentrum für Informations-, Medien- und Kommunikationstechnologie (ZIMK). <https://www.uni-trier.de/fileadmin/urt/doku/linreg/linreg.pdf>
- Bange, G. G. J., & Overstreet, R. (1960). Some Observations on Absorption of Cesium by Excised Barley Roots. *Plant Physiology*, 35(5), 605–608. <https://doi.org/10.1104/pp.35.5.605>

- Bates, P. D., Stymne, S., & Ohlrogge, J. (2013). Biochemical pathways in seed oil synthesis. *Current Opinion in Plant Biology*, 16(3), 358–364. <https://doi.org/10.1016/j.pbi.2013.02.015>
- Bowen, G. J., & Revenaugh, J. (2003). Interpolating the isotopic composition of modern meteoric precipitation: ISOTOPIC COMPOSITION OF MODERN PRECIPITATION. *Water Resources Research*, 39(10). <https://doi.org/10.1029/2003WR002086>
- Bundesministerium für Ernährung und Landwirtschaft. (2020). *Besondere Ernte- und Qualitätsermittlung (BEE) 2019: Reihe Daten-Analysen*. 35.
- Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. (2006). OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20(8–10), 341–351. <https://doi.org/10.1002/cem.1006>
- Camin, F., Boner, M., Bontempo, L., Fauhl-Hassek, C., Kelly, S. D., Riedl, J., & Rossmann, A. (2017). Stable isotope techniques for verifying the declared geographical origin of food in legal cases. *Trends in Food Science & Technology*, 61, 176–187. <https://doi.org/10.1016/j.tifs.2016.12.007>
- 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>
- Camin, F., Larcher, R., Perini, M., Bontempo, L., Bertoldi, D., Gagliano, G., Nicolini, G., & Versini, G. (2010). Characterisation of authentic Italian extra-virgin olive oils by stable isotope ratios of C, O and H and mineral composition. *Food Chemistry*, 118(4), 901–909. <https://doi.org/10.1016/j.foodchem.2008.04.059>
- Camin, F., Pavone, A., Bontempo, L., Wehrens, R., Paolini, M., Faberi, A., Marianella, R. M., Capitani, D., Vista, S., & Mannina, L. (2016). The use of IRMS, <sup>1</sup>H NMR and chemical analysis to characterise Italian and imported Tunisian olive oils. *Food Chemistry*, 196, 98–105. <https://doi.org/10.1016/j.foodchem.2015.08.132>
- Cazzolla Gatti, R., & Velichevskaya, A. (2020). Certified “sustainable” palm oil took the place of endangered Bornean and Sumatran large mammals habitat and tropical forests in the last 30 years. *Science of The Total Environment*, 742, 140712. <https://doi.org/10.1016/j.scitotenv.2020.140712>
- Chikaraishi, Y., Naraoka, H., & Poulson, S. R. (2004). Hydrogen and carbon isotopic fractionations of lipid biosynthesis among terrestrial (C<sub>3</sub>, C<sub>4</sub> and CAM) and aquatic plants. *Phytochemistry*, 65(10), 1369–1381. <https://doi.org/10.1016/j.phytochem.2004.03.036>
- Chivall, D., Berstan, R., Bull, I. D., & Evershed, R. P. (2012). Isotope effects associated with the preparation and methylation of fatty acids by boron trifluoride in methanol for compound-specific stable hydrogen isotope analysis via gas chromatography/thermal conversion/isotope ratio mass spectrometry: Hydrogen stable isotope analysis of FAMES via GC/TC/IRMS. *Rapid Communications in Mass Spectrometry*, 26(10), 1232–1240. <https://doi.org/10.1002/rcm.6188>
- Clark, I. D., & Fritz, P. (1997). *Environmental isotopes in hydrogeology*. CRC Press/Lewis Publishers.
- Cohen, K. P., Runcorn, S. K., Suess, H. E., & Thode, H. G. (1983). Harold Clayton Urey, 29 April



- 
- 1893—5 January 1981. *Biographical Memoirs of Fellows of the Royal Society*, 29, 622–659. <https://doi.org/10.1098/rsbm.1983.0022>
- Coplen, T. B., & Hanshaw, B. B. (1973). Ultrafiltration by a compacted clay membrane—I. Oxygen and hydrogen isotopic fractionation. *Geochimica et Cosmochimica Acta*, 37(10), 2295–2310. [https://doi.org/10.1016/0016-7037\(73\)90105-1](https://doi.org/10.1016/0016-7037(73)90105-1)
- Cornwell, W. K., Wright, I. J., Turner, J., Maire, V., Barbour, M. M., Cernusak, L. A., Dawson, T., Ellsworth, D., Farquhar, G. D., Griffiths, H., Keitel, C., Knohl, A., Reich, P. B., Williams, D. G., Bhaskar, R., Cornelissen, J. H. C., Richards, A., Schmidt, S., Valladares, F., ... Santiago, L. S. (2018). Climate and soils together regulate photosynthetic carbon isotope discrimination within C<sub>3</sub> plants worldwide: XXXX. *Global Ecology and Biogeography*, 27(9), 1056–1067. <https://doi.org/10.1111/geb.12764>
- Cramer, N. (1990). *Raps: Anbau und Verwertung*. Ulmer. <https://books.google.de/books?id=iPHFAAAACAAJ>
- Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *TrAC Trends in Analytical Chemistry*, 85, 123–132. <https://doi.org/10.1016/j.trac.2016.02.026>
- Dansgaard, W. (1964). Stable isotopes in precipitation. *Tellus*, 16(4), 436–468. <https://doi.org/10.1111/j.2153-3490.1964.tb00181.x>
- Dawson, T. E. (1993). Water Sources of Plants as Determined from Xylem-Water Isotopic Composition: Perspectives on Plant Competition, Distribution, and Water Relations. In *Stable Isotopes and Plant Carbon-water Relations* (pp. 465–496). Elsevier. <https://doi.org/10.1016/B978-0-08-091801-3.50040-4>
- Deutscher Wetterdienst. (2023). *Phänologische Meldestatistik*. [https://www.dwd.de/DE/leistungen/phaeno\\_sta/meldestatistik.html?nn=575800&lsbld=522616](https://www.dwd.de/DE/leistungen/phaeno_sta/meldestatistik.html?nn=575800&lsbld=522616)
- Duan, J.-R., Billault, I., Mabon, F., & Robins, R. (2002). Natural Deuterium Distribution in Fatty Acids Isolated from Peanut Seed Oil: A Site-Specific Study by Quantitative <sup>2</sup>H NMR Spectroscopy. *ChemBioChem*, 3(8), 752–759. [https://doi.org/10.1002/1439-7633\(20020802\)3:8<752::AID-CBIC752>3.0.CO;2-G](https://doi.org/10.1002/1439-7633(20020802)3:8<752::AID-CBIC752>3.0.CO;2-G)
- Dyer, J. M., Stymne, S., Green, A. G., & Carlsson, A. S. (2008). High-value oils from plants. *The Plant Journal*, 54(4), 640–655. <https://doi.org/10.1111/j.1365-313X.2008.03430.x>
- Ehtesham, E., Baisden, W. T., Keller, E. D., Hayman, A. R., Van Hale, R., & Frew, R. D. (2013). Correlation between precipitation and geographical location of the δ<sup>2</sup>H values of the fatty acids in milk and bulk milk powder. *Geochimica et Cosmochimica Acta*, 111, 105–116. <https://doi.org/10.1016/j.gca.2012.10.026>
- Eley, Y. (2014). *ENVIRONMENTAL AND BIOCHEMICAL CONTROLS ON THE MOLECULAR DISTRIBUTION AND STABLE ISOTOPE COMPOSITION OF LEAF WAX BIOMARKERS* [School of Environmental Sciences University of East Anglia]. [https://ueaeprints.uea.ac.uk/id/eprint/52165/1/Yvette\\_Eley\\_2014.pdf](https://ueaeprints.uea.ac.uk/id/eprint/52165/1/Yvette_Eley_2014.pdf)
- Eley, Y., White, J., Dawson, L., Hren, M., & Pedentchouk, N. (2018). Variation in Hydrogen Isotope Composition Among Salt Marsh Plant Organic Compounds Highlights Biochemical Mechanisms Controlling Biosynthetic Fractionation. *Journal of Geophysical Research: Biogeosciences*, 123(9), 2645–2660. <https://doi.org/10.1029/2018JG004403>

- 
- European Union. (2023). *Renewable energy targets*. [https://energy.ec.europa.eu/topics/renewable-energy/renewable-energy-directive-targets-and-rules/renewable-energy-targets\\_en#:~:text=The%20revised%20Renewable%20Energy%20Directive,to%20a%20minimum%20of%2042.5%25.&text=The%20energy%20sector%20is%20responsible,the%20EU's%20greenhouse%20gas%20emissions](https://energy.ec.europa.eu/topics/renewable-energy/renewable-energy-directive-targets-and-rules/renewable-energy-targets_en#:~:text=The%20revised%20Renewable%20Energy%20Directive,to%20a%20minimum%20of%2042.5%25.&text=The%20energy%20sector%20is%20responsible,the%20EU's%20greenhouse%20gas%20emissions).
- Faberi, A., Marianella, R. M., Fuselli, F., Mantia, A. L., Ciardiello, F., Montesano, C., Mascini, M., Sergi, M., & Compagnone, D. (2014). Fatty acid composition and  $\delta^{13}\text{C}$  of bulk and individual fatty acids as marker for authenticating Italian PDO/PGI extra virgin olive oils by means of isotopic ratio mass spectrometry. *Journal of Mass Spectrometry*, 49(9), 840–849. <https://doi.org/10.1002/jms.3399>
- Fanelli, V., Mascio, I., Miazzi, M. M., Savoia, M. A., De Giovanni, C., & Montemurro, C. (2021). Molecular Approaches to Agri-Food Traceability and Authentication: An Updated Review. *Foods*, 10(7), Article 7. <https://doi.org/10.3390/foods10071644>
- FD Gunstone. (2002). Vegetable oils in food technology: Composition, properties and uses. *Canola/Rapeseed Oil*, 99–127.
- Fink, K., Richling, E., Heckel, F., & Schreier, P. (2004). Determination of  $2\text{H}/1\text{H}$  and  $13\text{C}/12\text{C}$  Isotope Ratios of (E)-Methyl Cinnamate from Different Sources Using Isotope Ratio Mass Spectrometry. *Journal of Agricultural and Food Chemistry*, 52(10), 3065–3068. <https://doi.org/10.1021/jf040018j>
- Fridrihsone, A., Romagnoli, F., & Cabulis, U. (2020). Environmental Life Cycle Assessment of Rapeseed and Rapeseed Oil Produced in Northern Europe: A Latvian Case Study. *Sustainability*, 12(14), Article 14. <https://doi.org/10.3390/su12145699>
- Friedlingstein, P., Houghton, R. A., Marland, G., Hackler, J., Boden, T. A., Conway, T. J., Canadell, J. G., Raupach, M. R., Ciais, P., & Le Quéré, C. (2010). Update on CO<sub>2</sub> emissions. *Nature Geoscience*, 3(12), Article 12. <https://doi.org/10.1038/ngeo1022>
- Friedrich, K., Lügger, K., Wolfgang, J., & Sehr, J. (2011). *Hintergrundwerte von Spurenstoffen in hessischen Böden*. Hessisches Landesamt für Umwelt und Geologie.
- Garcés, R. F., & Mancha, M. (1993). One-step lipid extraction and fatty acid methyl esters preparation from fresh plant tissues. *Analytical Biochemistry*. <https://doi.org/10.1006/abio.1993.1244>
- Gatzert, X., Chun, K. P., Boner, M., Hermanowski, R., Mäder, R., Breuer, L., Gattinger, A., & Orłowski, N. (2021). Assessment of multiple stable isotopes for tracking regional and organic authenticity of plant products in Hesse, Germany. *Isotopes in Environmental and Health Studies*, 57(3), 281–300. <https://doi.org/10.1080/10256016.2021.1905635>
- Granato, D., Putnik, P., Kovačević, D. B., Santos, J. S., Calado, V., Rocha, R. S., Cruz, A. G. D., Jarvis, B., Rodionova, O. Y., & Pomerantsev, A. (2018). Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing. *Comprehensive Reviews in Food Science and Food Safety*, 17(3), 663–677. <https://doi.org/10.1111/1541-4337.12341>
- Gumus, Z. P., Celenk, V. U., Tekin, S., Yurdakul, O., & Ertas, H. (2017). Determination of trace elements and stable carbon isotope ratios in virgin olive oils from Western Turkey to authenticate geographical origin with a chemometric approach. *European Food Research and Technology*, 243(10), 1719–1727. <https://doi.org/10.1007/s00217-017-2876-4>
- Hessisches Statistisches Landesamt. (2017). *Bodennutzung in Hessen 2017*. Hessisches

---

Statistisches Landesamt.

- HLUNG. (2011). *Substratgruppen zur Kennzeichnung der Hintergrundwerte in Oberböden*. <https://www.hlnug.de/themen/boden/information/bodenflaechenkataster-und-kartenwerke/bfd50/themen/hgw-oberboden>
- Hoefs, J. (2018). *Stable Isotope Geochemistry*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-78527-1>
- Hsu, C.-Y., & Ellis Jr, P. E. (1988). *Process for the production of 2, 2-dimethoxypropane*. <https://www.freepatentsonline.com/4775447.html>
- Hu, H., Bao, W., Wang, T., & Qu, S. (2008). Study on the Variation of Hydrogen and Oxygen Isotopes in Soil Water. *2008 2nd International Conference on Bioinformatics and Biomedical Engineering*, 2978–2981. <https://doi.org/10.1109/ICBBE.2008.1077>
- Ichihara, K., & Fukubayashi, Y. (2010). Preparation of fatty acid methyl esters for gas-liquid chromatography. *Journal of Lipid Research*, 51(3), 635–640. <https://doi.org/10.1194/jlr.D001065>
- 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>
- ISO 12966-2:2017. (2017). Animal and vegetable fats and oils—Gas chromatography of fatty acid methyl esters—Part 2: Preparation of methyl esters of fatty acids. *Beuth, Berlin*.
- Jannat, A., Ishikawa-Ishiwata, Y., & Furuya, J. (2022). Does Climate Change Affect Rapeseed Production in Exporting and Importing Countries? Evidence from Market Dynamics Syntheses. *Sustainability*, 14(10), Article 10. <https://doi.org/10.3390/su14106051>
- Jeon, H., Lee, S.-C., Cho, Y.-J., Oh, J.-H., Kwon, K., & Kim, B. H. (2015). A triple-isotope approach for discriminating the geographic origin of Asian sesame oils. *Food Chemistry*, 167, 363–369. <https://doi.org/10.1016/j.foodchem.2014.07.032>
- Jianhu Bureau of Natural Resources and Planning. (2017). *建湖县地质地貌土壤\_科学普及*. [http://zrzy.jiangsu.gov.cn/ycjh/gtzx/ztlz/kxpj/201702/t20170216\\_54953.htm](http://zrzy.jiangsu.gov.cn/ycjh/gtzx/ztlz/kxpj/201702/t20170216_54953.htm)
- Jiménez-Morillo, N. T., Palma, V., Garcia, R., Dias, C. B., & Cabrita, M. J. (2020). Combination of Stable Isotope Analysis and Chemometrics to Discriminate Geoclimatically and Temporally the Virgin Olive Oils from Three Mediterranean Countries. *Foods*, 9(12), Article 12. <https://doi.org/10.3390/foods9121855>
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Juchelka, D. (2013). *GC-IRMS:  $\delta^{13}C$  Analysis of PAHs in Soil and Sediment Samples using High Resolution GC Coupled with Isotope Ratio MS* (Application Note 30268). Thermo Fisher Scientific. <https://assets.thermofisher.com/TFS-Assets/CMD/Application-Notes/AN-30268-IRMS-PAHs-Soil-Sediment-AN30268-EN.pdf#page=2>
- Kabata-Pendias, A., & Mukherjee, A. B. (2007). *Trace Elements from Soil to Human*. Springer. <https://doi.org/10.1007/978-3-540-32714-1>
- Katerinopoulou, K., Kontogeorgos, A., Salmas, C. E., Patakas, A., & Ladavos, A. (2020). Geographical Origin Authentication of Agri-Food Products: A Review. *Foods*, 9(4), Article 4.

- 
- <https://doi.org/10.3390/foods9040489>
- Kelly, S., Brodie, C., & Hilkert, A. (2018). Isotopic-Spectroscopic Technique: Stable Isotope-Ratio Mass Spectrometry (IRMS). In *Modern Techniques for Food Authentication* (pp. 349–413). Elsevier. <https://doi.org/10.1016/B978-0-12-814264-6.00011-6>
- Kendall, C., & McDonnell, J. J. (2012). *Isotope tracers in catchment hydrology*. Elsevier. <https://wwwrcamnl.wr.usgs.gov/isoig/isopubs/itchch2.html#2.3.2>
- Kim, H. U. (2020). Lipid Metabolism in Plants. *Plants*, 9(7), Article 7. <https://doi.org/10.3390/plants9070871>
- 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>
- Kowalski, B. R. (1981). Analytical chemistry as an information science. *TrAC Trends in Analytical Chemistry*, 1(3), 71–74. [https://doi.org/10.1016/0165-9936\(81\)80030-1](https://doi.org/10.1016/0165-9936(81)80030-1)
- Landesbetrieb Landwirtschaft Hessen. (2017, April). *Besondere Ernte- und Qualitätsermittlung (BEE) 2017-Arbeitsanweisung für die Beauftragten*.
- Lapola, D. M., Schaldach, R., Alcamo, J., Bondeau, A., Koch, J., Koelking, C., & Priess, J. A. (2010). Indirect land-use changes can overcome carbon savings from biofuels in Brazil. *Proceedings of the National Academy of Sciences*, 107(8), 3388–3393. <https://doi.org/10.1073/pnas.0907318107>
- Ling, Y., Zheng, M., Sun, Q., Zhang, C., Wang, Y., & Xie, M. (2022). Impact of climatic and environmental factors on  $\delta D$  of n-alkanes in the lake surface sediments of arid Tibet. *Quaternary International*, 637, 44–56. <https://doi.org/10.1016/j.quaint.2022.08.014>
- Lüker-Jans, N. (2017). *Analysing regional differences of agricultural land use and land-use change (2005-2010) in Hesse, Germany*.
- Lüker-Jans, N., Simmering, D., & Otte, A. (2017). The impact of biogas plants on regional dynamics of permanent grassland and maize area—The example of Hesse, Germany (2005–2010). *Agriculture, Ecosystems & Environment*, 241, 24–38. <https://doi.org/10.1016/j.agee.2017.02.023>
- Luo, Y.-H., Steinberg, L., Suda, S., Kumazawa, S., & Mitsui, A. (1991). Extremely Low D/H Ratios of Photoproduced Hydrogen by Cyanobacteria. *Plant and Cell Physiology*, 32(6), 897–900. <https://doi.org/10.1093/oxfordjournals.pcp.a078158>
- Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, 107(2), 897–911. <https://doi.org/10.1016/j.foodchem.2007.09.038>
- Mahindawansa, A., Jost, M., & Gassmann, M. (2022). Spatial and Temporal Variations of Stable Isotopes in Precipitation in the Mountainous Region, North Hesse. *Water*, 14(23), Article 23. <https://doi.org/10.3390/w14233910>
- Marshall, J. D., Brooks, J. R., & Lajtha, K. (2007). Sources of Variation in the Stable Isotopic Composition of Plants. In R. Michener & K. Lajtha (Eds.), *Stable Isotopes in Ecology and Environmental Science* (pp. 22–60). Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470691854.ch2>
- Mcdonald, J. H. (2009). *HANDBOOK OF BIOLOGICAL STATISTICS* (2. Edition). SPARKY HOUSE PUBLISHING.
- McKinney, C. R., McCrea, J. M., Epstein, S., Allen, H. A., & Urey, H. C. (1950). Improvements in

- 
- Mass Spectrometers for the Measurement of Small Differences in Isotope Abundance Ratios. *Review of Scientific Instruments*, 21(8), 724–730. <https://doi.org/10.1063/1.1745698>
- Meier, U. (2018). *Growth stages of mono- and dicotyledonous plants: BBCH Monograph*. <https://doi.org/10.5073/20180906-074619>
- Meier-Augenstein, W. (2017). *Stable Isotope Forensics* (1st ed.). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119080190>
- Meier-Augenstein, W., & Schimmelmann, A. (2019). A guide for proper utilisation of stable isotope reference materials. *Isotopes in Environmental and Health Studies*, 55(2), 113–128. <https://doi.org/10.1080/10256016.2018.1538137>
- Meier-Augenstein, W., Watt, P. W., & Langhans, C.-D. (1996). Influence of gas chromatographic parameters on measurement of  $^{13}\text{C}/^{12}\text{C}$  isotope ratios by gas-liquid chromatography-combustion isotope ratio mass spectrometry. I. *Journal of Chromatography A*, 752(1), 233–241. [https://doi.org/10.1016/S0021-9673\(96\)00498-0](https://doi.org/10.1016/S0021-9673(96)00498-0)
- Meija, J., Coplen, T. B., Berglund, M., Brand, W. A., Bièvre, P. D., Gröning, M., Holden, N. E., Irrgeher, J., Loss, R. D., Walczyk, T., & Prohaska, T. (2016). Isotopic compositions of the elements 2013 (IUPAC Technical Report). *Pure and Applied Chemistry*, 88(3), 293–306. <https://doi.org/10.1515/pac-2015-0503>
- Meijaard, E., Brooks, T. M., Carlson, K. M., Slade, E. M., Garcia-Ulloa, J., Gaveau, D. L. A., Lee, J. S. H., Santika, T., Juffe-Bignoli, D., Struebig, M. J., Wich, S. A., Ancrenaz, M., Koh, L. P., Zamira, N., Abrams, J. F., Prins, H. H. T., Sendashonga, C. N., Murdiyarso, D., Furumo, P. R., ... Sheil, D. (2020). The environmental impacts of palm oil in context. *Nature Plants*, 6(12), 1418–1426. <https://doi.org/10.1038/s41477-020-00813-w>
- Mihailova, A., Abbado, D., Kelly, S. D., & Pedentchouk, N. (2015). The impact of environmental factors on molecular and stable isotope compositions of n-alkanes in Mediterranean extra virgin olive oils. *Food Chemistry*, 173, 114–121. <https://doi.org/10.1016/j.foodchem.2014.10.003>
- Miklaszewska, M., Zienkiewicz, K., Inchana, P., & Zienkiewicz, A. (2021). Lipid metabolism and accumulation in oilseed crops. *OCL*, 28, 50. <https://doi.org/10.1051/ocl/2021039>
- MKS Umetrics. (2012). *User Guide to SIMCA 13*. <http://chemsrv0.pph.univie.ac.at/scripten/EDV/Software/Simca13/User%20Guide%20to%20SIMCA%2013.pdf>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed). Wiley.
- Mounier, L. (2014). *Fundamentals of Isotopic Ratios Mass Spectrometry*. [http://www.ebd.csic.es/IsotopeCourse/Conferences/2014%20Fundamentals%20of%20IRM%20-%20EBD%20course%20-%20Sevilla%20\[Lecture%20seule\].pdf](http://www.ebd.csic.es/IsotopeCourse/Conferences/2014%20Fundamentals%20of%20IRM%20-%20EBD%20course%20-%20Sevilla%20[Lecture%20seule].pdf)
- Moussa, Issam., Naulet, Norbert., Martin, M. L., & Martin, G. J. (1990). A site-specific and multielement approach to the determination of liquid-vapor isotope fractionation parameters: The case of alcohols. *The Journal of Physical Chemistry*, 94(21), 8303–8309. <https://doi.org/10.1021/j100384a056>
- Muhammad, S. A., Seow, E.-K., Mohd Omar, A., Rodhi, A. M., Mat Hassan, H., Lalung, J., Lee, S.-C., & Ibrahim, B. (2018). Variation of  $\delta^2\text{H}$ ,  $\delta^{18}\text{O}$  &  $\delta^{13}\text{C}$  in crude palm oil from different regions in Malaysia: Potential of stable isotope signatures as a key traceability parameter.

- 
- Science & Justice*, 58(1), 59–66. <https://doi.org/10.1016/j.scijus.2017.05.008>
- Müller, D. M., Zinkernagel, D. J., Kleber, J., Fricke, E., Beck, D. M., & Göttl, M. (2020). *Kulturartbezogene Kennzahlen*. Arbeitsgemeinschaft Landtechnik und Landwirtschaftliches Bauwesen in Bayern e.V. (ALB),. <https://www.alb-bayern.de/media/files/0004/bb.bef2-kennzahlenkulturen-20200706.pdf>
- Newcastle University. (n.d.). *Numeracy, Maths and Statistics—Academic Skills Kit*. Retrieved June 27, 2023, from <https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/regression-and-correlation/strength-of-correlation.html>
- Oshun, J., Dietrich, W. E., Dawson, T. E., & Fung, I. (2016). Dynamic, structured heterogeneity of water isotopes inside hillslopes. *Water Resources Research*, 52(1), 164–189. <https://doi.org/10.1002/2015WR017485>
- Panetta, R. J., & Jahren, A. H. (2011). Single-step transesterification with simultaneous concentration and stable isotope analysis of fatty acid methyl esters by gas chromatography-combustion-isotope ratio mass spectrometry: Analysis of fatty acid methyl esters by GC-C-IRMS. *Rapid Communications in Mass Spectrometry*, 25(10), 1373–1381. <https://doi.org/10.1002/rcm.5001>
- Paolini, M., Bontempo, L., & Camin, F. (2017). Compound-specific  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  analysis of olive oil fatty acids. *Talanta*, 174, 38–43. <https://doi.org/10.1016/j.talanta.2017.05.080>
- Plevin, R. J., Michael O'Hare, Jones, A. D., Torn, M. S., & Gibbs, H. K. (2010). Greenhouse Gas Emissions from Biofuels' Indirect Land Use Change Are Uncertain but May Be Much Greater than Previously Estimated. *Environmental Science & Technology*, 44(21), 8015–8021. <https://doi.org/10.1021/es101946t>
- Portarena, S., Baldacchini, C., & Brugnoli, E. (2017). Geographical discrimination of extra-virgin olive oils from the Italian coasts by combining stable isotope data and carotenoid content within a multivariate analysis. *Food Chemistry*, 215, 1–6. <https://doi.org/10.1016/j.foodchem.2016.07.135>
- Przybylski, R. (2011). Canola/Rapeseed Oil. In *Vegetable Oils in Food Technology* (pp. 107–136). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781444339925.ch4>
- Richter, E. K., Spangenberg, J. E., Kreuzer, M., & Leiber, F. (2010). Characterization of Rapeseed (*Brassica napus*) Oils by Bulk C, O, H, and Fatty Acid C Stable Isotope Analyses. *Journal of Agricultural and Food Chemistry*, 58(13), 8048–8055. <https://doi.org/10.1021/jf101128f>
- Rubingh, C. M., Bijlsma, S., Derks, E. P. P. A., Bobeldijk, I., Verheij, E. R., Kochhar, S., & Smilde, A. K. (2006). Assessing the performance of statistical validation tools for megavariate metabolomics data. *Metabolomics*, 2(2), 53–61. <https://doi.org/10.1007/s11306-006-0022-6>
- Ruiz-Perez, D., Guan, H., Madhivanan, P., Mathee, K., & Narasimhan, G. (2020). So you think you can PLS-DA? *BMC Bioinformatics*, 21(1), 2. <https://doi.org/10.1186/s12859-019-3310-7>
- Ruiz-Samblás, C., Arrebola-Pascual, C., Tres, A., van Ruth, S., & Cuadros-Rodríguez, L. (2013). Authentication of geographical origin of palm oil by chromatographic fingerprinting of triacylglycerols and partial least square-discriminant analysis. *Talanta*, 116, 788–793. <https://doi.org/10.1016/j.talanta.2013.07.054>
- Ryżak, M., & Bieganowski, A. (2011). Methodological aspects of determining soil particle-size distribution using the laser diffraction method. *Journal of Plant Nutrition and Soil Science*, 174(4), 624–633. <https://doi.org/10.1002/jpln.201000255>

- 
- Sachse, D., Billault, I., Bowen, G. J., Chikaraishi, Y., Dawson, T. E., Feakins, S. J., Freeman, K. H., Magill, C. R., McInerney, F. A., van der Meer, M. T. J., Polissar, P., Robins, R. J., Sachs, J. P., Schmidt, H.-L., Sessions, A. L., White, J. W. C., West, J. B., & Kahmen, A. (2012). Molecular Paleohydrology: Interpreting the Hydrogen-Isotopic Composition of Lipid Biomarkers from Photosynthesizing Organisms. *Annual Review of Earth and Planetary Sciences*, 40(1), 221–249. <https://doi.org/10.1146/annurev-earth-042711-105535>
- Schimmelmann, A., Qi, H., Coplen, T. B., Brand, W. A., Fong, J., Meier-Augenstein, W., Kemp, H. F., Toman, B., Ackermann, A., Assonov, S., Aerts-Bijma, A. T., Brejcha, R., Chikaraishi, Y., Darwish, T., Elsner, M., Gehre, M., Geilmann, H., Gröning, M., Hélie, J.-F., ... Werner, R. A. (2016). Organic Reference Materials for Hydrogen, Carbon, and Nitrogen Stable Isotope-Ratio Measurements: Caffeines, *n*-Alkanes, Fatty Acid Methyl Esters, Glycines, L-Valines, Polyethylenes, and Oils. *Analytical Chemistry*, 88(8), 4294–4302. <https://doi.org/10.1021/acs.analchem.5b04392>
- Schmidt, H.-L., Robins, R. J., & Werner, R. A. (2015). Multi-factorial *in vivo* stable isotope fractionation: Causes, correlations, consequences and applications. *Isotopes in Environmental and Health Studies*, 51(1), 155–199. <https://doi.org/10.1080/10256016.2015.1014355>
- Schmidt, H.-L., Werner, R. A., & Eisenreich, W. (2003). Systematics of 2H patterns in natural compounds and its importance for the elucidation of biosynthetic pathways. *Phytochemistry Reviews*, 2(1–2), 61–85. <https://doi.org/10.1023/B:PHYT.0000004185.92648.ae>
- Schmitt, J., Glaser, B., & Zech, W. (2003). Amount-dependent isotopic fractionation during compound-specific isotope analysis. *Rapid Communications in Mass Spectrometry*, 17(9), 970–977. <https://doi.org/10.1002/rcm.1009>
- Sessions, A. L. (2006). Isotope-ratio detection for gas chromatography. *Journal of Separation Science*, 29(12), 1946–1961. <https://doi.org/10.1002/jssc.200600002>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3–4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Sharp, Z. (2017). *Principles of Stable Isotope Geochemistry, 2nd Edition*. [https://digitalrepository.unm.edu/unm\\_oer/1](https://digitalrepository.unm.edu/unm_oer/1)
- Skinner, C. J. (2016). Probability Proportional to Size (PPS) Sampling. In *Wiley StatsRef: Statistics Reference Online* (pp. 1–5). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118445112.stat03346.pub2>
- Sumfleth, B., Majer, S., & Thrän, D. (2020). Recent Developments in Low iLUC Policies and Certification in the EU Biobased Economy. *Sustainability*, 12(19), Article 19. <https://doi.org/10.3390/su12198147>
- Tahir, H. E., Arslan, M., Komla Mahunu, G., Adam Mariod, A., B.H. Hashim, S., Xiaobo, Z., Jiyong, S., El-Seedi, H. R., & Musa, T. H. (2022). The use of analytical techniques coupled with chemometrics for tracing the geographical origin of oils: A systematic review (2013–2020). *Food Chemistry*, 366, 130633. <https://doi.org/10.1016/j.foodchem.2021.130633>
- Taiz, L., & Zeiger, E. (2002). *Plant physiology. 3rd edn.* (3rd ed., Vol. 91). Sinauer Associates. <https://academic.oup.com/aob/article-lookup/doi/10.1093/aob/mcg079>
- Tarapoulouzi, M., Skiada, V., Agriopoulou, S., Psomiadis, D., Rébufa, C., Roussos, S., Theocharis, C. R., Katsaris, P., & Varzakas, T. (2021). Chemometric Discrimination of the Geographical

- 
- Origin of Three Greek Cultivars of Olive Oils by Stable Isotope Ratio Analysis. *Foods*, 10(2), Article 2. <https://doi.org/10.3390/foods10020336>
- Thermo Scientific. (2015). *GC\_IsoLink\_II\_Operating\_Manual\_Rev\_C.pdf*.
- van Leeuwen, K. A., Prenzler, P. D., Ryan, D., & Camin, F. (2014). Gas Chromatography-Combustion-Isotope Ratio Mass Spectrometry for Traceability and Authenticity in Foods and Beverages: GC-C-IRMS for authentication in foods.... *Comprehensive Reviews in Food Science and Food Safety*, 13(5), 814–837. <https://doi.org/10.1111/1541-4337.12096>
- Vandeginste, B. (2013). Chemometrics in studies of food origin. In *New Analytical Approaches for Verifying the Origin of Food* (pp. 117–145). Elsevier. <https://doi.org/10.1533/9780857097590.2.117>
- Vargas, A. I., Schaffer, B., Yuhong, L., & Sternberg, L. da S. L. (2017). Testing plant use of mobile vs immobile soil water sources using stable isotope experiments. *New Phytologist*, 215(2), 582–594. <https://doi.org/10.1111/nph.14616>
- Wadood, S. A., Boli, G., Xiaowen, Z., Hussain, I., & Yimin, W. (2020). Recent development in the application of analytical techniques for the traceability and authenticity of food of plant origin. *Microchemical Journal*, 152, 104295. <https://doi.org/10.1016/j.microc.2019.104295>
- Watkins, J. C. (2019). *An introduction to the science of statistics: From theory to implementation*. Joseph C. Watkins.
- Woodbury, S. E., Evershed, R. P., & Barry Rossell, J. (1998). Purity assessments of major vegetable oils based on  $\delta^{13}\text{C}$  values of individual fatty acids. *Journal of the American Oil Chemists' Society*, 75(3), 371–379. <https://doi.org/10.1007/s11746-998-0055-2>
- Yan, X., & Su, X. (2009). *Linear regression analysis: Theory and computing*. World Scientific.
- Zhang, Y., Tobias, H. J., Sacks, G. L., & Thomas Brenna, J. (2012). Calibration and data processing in gas chromatography combustion isotope ratio mass spectrometry. *Drug Testing and Analysis*, 4(12), 912–922. <https://doi.org/10.1002/dta.394>
- Zheng, S., Kates, M., Dubé, M. A., & McLean, D. D. (2006). Acid-catalyzed production of biodiesel from waste frying oil. *Biomass and Bioenergy*, 30(3), 267–272. <https://doi.org/10.1016/j.biombioe.2005.10.004>
- 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](https://doi.org/10.1007/978-1-4612-3498-2_8)
- Zscheischler, J., & Fischer, E. M. (2020). The record-breaking compound hot and dry 2018 growing season in Germany. *Weather and Climate Extremes*, 29, 100270. <https://doi.org/10.1016/j.wace.2020.100270>



## Appendix

Table A-1: Fatty acid profile of rapeseed (mol%) using three different transmethylation mixtures.

transmethylation mixture		TM+D, n=3			TM-D, n=3			TX-D, n=3			mean	SD
FAME composition		Fi <sup>a</sup>	mean	SD	Fi <sup>a</sup>	mean	SD	Fi <sup>a</sup>	mean	SD		
Fatty acid composition (mol %)	<b>C14:0</b>	1.02	0.1%	0.01%	1.02	0.1%	0.01%	1.02	0.1%	0.00%	<b>0.1%</b>	0.02%
	<b>C16:0</b>	1.00	5.5%	0.04%	1.00	5.6%	0.02%	1.00	5.5%	0.07%	<b>5.5%</b>	0.06%
	<b>C16:1</b>	1.00	0.4%	0.02%	1.00	0.4%	0.02%	1.00	0.3%	0.00%	<b>0.3%</b>	0.05%
	<b>C18:0</b>	0.99	1.7%	0.01%	0.99	1.7%	0.03%	0.99	1.8%	0.09%	<b>1.7%</b>	0.09%
	<b>C18:1</b>	1.00	61.6%	0.06%	1.00	61.4%	0.04%	1.00	62.5%	0.12%	<b>61.8%</b>	0.61%
	<b>C18:2</b>	0.98	18.2%	0.05%	0.98	18.4%	0.02%	0.98	17.7%	0.02%	<b>18.1%</b>	0.36%
	<b>C18:3</b>	1.00	10.0%	0.01%	1.00	10.1%	0.02%	1.00	9.3%	0.02%	<b>9.8%</b>	0.47%
	<b>C20:0</b>	0.99	0.5%	0.00%	0.99	0.5%	0.01%	0.99	0.6%	0.00%	<b>0.5%</b>	0.05%
	<b>C20:1</b>	1.00	1.2%	0.01%	1.00	1.2%	0.01%	0.96	1.3%	0.01%	<b>1.2%</b>	0.04%
	<b>C22:0</b>	1.02	0.3%	0.01%	1.02	0.3%	0.02%	1.01	0.3%	0.00%	<b>0.3%</b>	0.03%
	<b>C22:1</b>	1.01	0.3%	0.02%	1.01	0.3%	0.02%	1.01	0.4%	0.02%	<b>0.3%</b>	0.04%
	<b>C24:0</b>	1.05	0.1%	0.01%	1.05	0.1%	0.01%	1.21	0.1%	0.00%	<b>0.1%</b>	0.03%
	<b>C24:1</b>	1.05	0.1%	0.00%	1.05	0.1%	0.01%	1.21	0.2%	0.01%	<b>0.1%</b>	0.03%

a: correction factor

Table A- 2: Calculation of weight of individual FAMES (FAME yield).

FAME Composition	TM-D, g/100 g seeds		TX-D, g/100 g seeds		TM+D, g/100 g seeds	
	mean	SD	mean	SD	mean	SD
C14:0	0.03	0.001	0.02	0.000	0.03	0.000
C16:0	1.85	0.062	1.92	0.005	2.11	0.003
C16:1	0.12	0.004	0.10	0.000	0.15	0.000
C18:0	0.61	0.020	0.70	0.002	0.70	0.001
C18:1	22.24	0.738	23.86	0.065	25.98	0.039
C18:2	6.61	0.219	6.70	0.018	7.64	0.011
C18:3	3.61	0.120	3.48	0.010	4.17	0.006
C20:0	0.19	0.006	0.24	0.001	0.23	0.000
C20:1	0.47	0.016	0.53	0.001	0.55	0.001
C22:0	0.11	0.004	0.14	0.000	0.14	0.000
C22:1	0.13	0.004	0.17	0.000	0.16	0.000
C24:0	0.04	0.001	0.07	0.000	0.05	0.000
C24:1	0.05	0.002	0.08	0.000	0.06	0.000
sum	36.07	1.197	38.01	0.104	41.98	0.062
	86%		91%		100%	

Calculation:

The internal standard C19:0 FAME with known amounts was added into each reactor to determine individual FAME yields. The yields of FAMES in three transmethylation mixtures were calculated with the peak area ratio between C19:0 FAME and total FAMES.

Table A- 3: Calculation of molecular weight of TAG of rapeseed oil.

Fatty acid	MWi, Molar weight of FAi	fi, FA weight fraction	fi/MWi
C14:0	228.4	0.1%	3.5679E-06
C16:0	256.4	5.0%	1.9503E-04
C16:1	254.4	0.3%	1.3731E-05
C18:0	284.5	1.7%	5.9021E-05
C18:1	282.5	61.9%	2.1914E-03
C18:2	280.5	18.2%	6.4917E-04
C18:3	278.4	9.9%	3.5680E-04
C20:0	312.5	0.5%	1.7563E-05
C20:1	310.5	1.3%	4.2723E-05
C22:0	340.6	0.3%	9.5666E-06
C22:1	338.6	0.4%	1.1354E-05
C24:0	368.6	0.1%	3.3698E-06
C24:1	366.6	0.1%	3.7931E-06
Summe		1	3.5570E-03
Average molecular weight of the fatty acid mixture, g/mol		281.1	
Molar Weight of TAG in rapeseed oil, g/mol		881.4	

Calculation:

The average molecular weight of fatty acids =  $\sum fi / \sum (fi/MWi)$ , where fi is the FA weight fraction and MWi is the molar weight of FAi.

The molecular weight of the TAG is  $MW=3 \times \text{Average molecular weight of fatty acids} + 38.049$ , where 3 means TAG contains three fatty acids and 38.049 is the weight of the glycerol backbone.

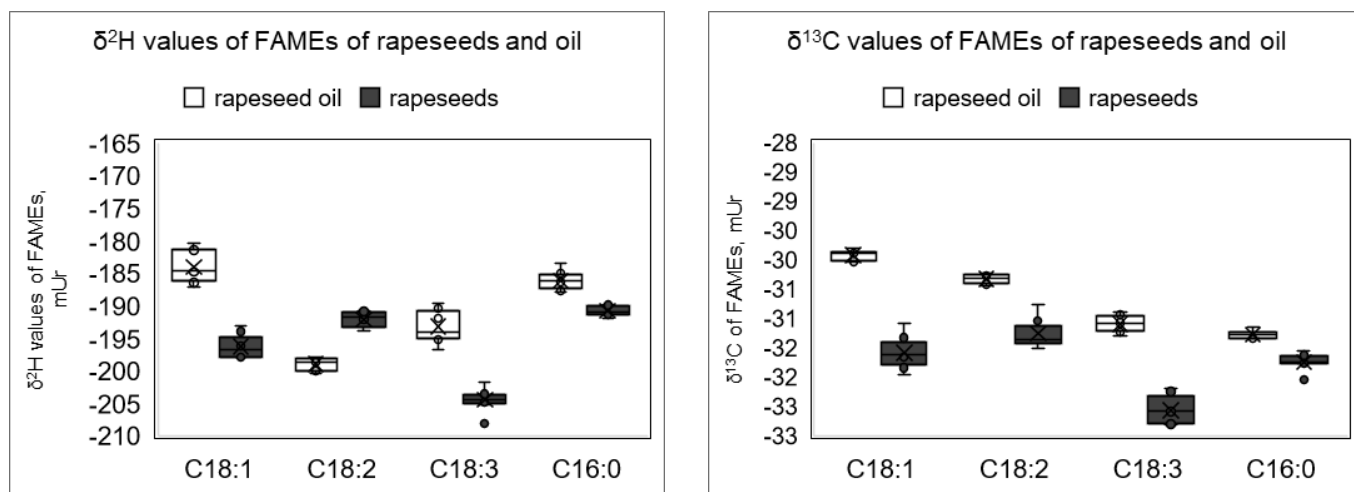


Figure A- 2: The results of carbon and hydrogen isotope compositions of the four major FAMES of rapeseeds and rapeseed oil (C16:0, C18:1, C18:2, and C18:3)

Table A- 4: The  $\delta^{13}\text{C}$  values of C16:0, C18:1, C18:2 and C18:3 FAME in rapeseed using transmethylation mixtures TX-D and TM+D (F-Test and T-Test)

transmethylation mixture		$\delta^{13}\text{C}$ of C16:0 FAME		$\delta^{13}\text{C}$ of C18:1 FAME		$\delta^{13}\text{C}$ of C18:2 FAME		$\delta^{13}\text{C}$ of C18:3 FAME	
		mean, n=3	SD, n=3	mean, n=3	SD, n=3	mean, n=3	SD, n=3	mean, n=3	SD, n=3
TX-D	sample 1	-31.72	0.12	-31.09	0.07	-31.00	0.14	-31.59	0.21
	sample 2	-31.96	0.02	-30.71	0.20	-30.56	0.13	-31.26	0.11
	sample 3	-31.98	0.03	-30.29	0.04	-30.68	0.03	-31.50	0.20
	mean, n=3	-31.88	-	-30.70	-	-30.75	-	-31.45	-
	SD, n=3	-	0.14	-	0.40	-	0.23	-	0.17
TM+D	sample 4	-31.56	0.02	-30.60	0.02	-30.61	0.17	-31.58	0.08
	sample 5	-32.00	0.03	-30.45	0.05	-30.53	0.05	-31.75	0.10
	sample 6	-31.98	0.05	-30.27	0.06	-30.68	0.13	-31.78	0.08
	mean, n=3	-31.85	-	-30.44	-	-30.60	-	-31.70	-
	SD, n=3	-	0.25	-	0.16	-	0.08	-	0.11
p-value of F-Test		0.49		0.29		0.20		0.57	
p-value of T-Test		0.82		0.36		0.36		0.09	

Table A- 5: The  $\delta^2\text{H}$  values of C16:0, C18:1, C18:2 and C18:3 FAME in rapeseed using transmethylation mixtures TX-D and TM+D (F-Test and T-Test)

transmethylation mixture		$\delta^2\text{H}$ of C16:0 FAME		$\delta^2\text{H}$ of C18:1 FAME		$\delta^2\text{H}$ of C18:2 FAME		$\delta^2\text{H}$ of C18:3 FAME	
		mean n=3	SD n=3	mean n=3	SD n=3	mean n=3	SD n=3	mean n=3	SD n=3
TX-D	sample 1	-188.16	2.35	-190.15	2.58	-191.04	2.26	-206.49	2.71
	sample 2	-191.78	1.08	-196.71	1.31	-193.49	0.96	-207.21	1.08
	sample 3	-190.12	1.09	-195.99	0.74	-193.54	0.94	-207.36	1.15
	mean, n=3	-190.02	-	-194.28	-	-192.69	-	-207.02	-
	SD, n=3	-	1.81	-	3.60	-	1.43	-	0.46
TM+D	sample 4	-189.52	0.07	-194.44	1.62	-192.39	1.21	-203.67	1.67
	sample 5	-189.69	0.83	-197.02	1.20	-192.61	1.56	-205.37	2.44
	sample 6	-188.70	0.54	-197.51	0.56	-191.37	0.66	-204.32	0.79
	mean, n=3	-189.30	-	-196.32	-	-192.13	-	-204.45	-
	SD, n=3	-	0.53	-	1.65	-	0.66	-	0.86
p-value of F-Test		0.16		0.35		0.35		0.45	
p-value of T-Test		0.55		0.42		0.57		0.01 <sup>a</sup>	

a: p-value < 0.05

Table A- 6: Calculation of  $\Delta\delta^2\text{H}$  of FAME between TM+D and TX-D of C16:0, C18:1 and C18:2

FAME	H number of FA	H number of FAME	$\Delta\delta^2\text{H}$ of FAME between TX-D and TM+D (measured)	$\Delta\delta^2\text{H}$ of -CH <sub>3</sub> of methanol between TX-D and TM+D (calculated)
C18:3	29	32	2.6	27.7

Calculation:

The relationship of  $\delta^2$  values between FAs and FAMEs can be estimated using the following mass balance equation (6)

$$(H_n + 3) \times \delta^2\text{H}_{\text{FAME}} = H_n \times \delta^2\text{H}_{\text{FA}} + 3 \times \delta^2\text{H}_{\text{MeOH}} \quad \text{equation (6)}$$

Where  $H_n$  is the number of H atoms in fatty acids and  $\delta^2\text{H}_{\text{FAME}}$ ,  $\delta^2\text{H}_{\text{FA}}$ , and  $\delta^2\text{H}_{\text{MeOH}}$  are the hydrogen isotope composition of FAME, FA, and methyl group of methanol.

Assuming that C18:3 FAME by the TX-D is generated from the initially added methanol, whereas that of TM+D is generated from the mixture of initially added. The reaction intermediate methanol, then the difference of  $\delta^2\text{H}$  values ( $\Delta\delta^2\text{H}$ ) of methanol ( $\delta^2\text{H}_{\text{MeOH}}$ ) between TM+D and TX-D is 28 mUr, derived based on the measured  $\Delta\delta^2\text{H}$  of FAME 18:3 of 2.6 mUr and the equation.

$$\text{TX-D: } (H_n + 3) \times \delta^2\text{H}_{\text{FAME}} = H_n \times \delta^2\text{H}_{\text{FA}} + 3 \times \delta^2\text{H}_{\text{MeOH}} \quad \text{equation (6-1)}$$

$$\text{TM+D: } (H_n + 3) \times \delta^2\text{H}_{\text{FAME}} = H_n \times \delta^2\text{H}_{\text{FA}} + 3 \times \delta^2\text{H}_{\text{MeOH}} \quad \text{equation (6-2)}$$

Equation (1-1) – equation (1-2):

$$\delta^2\text{H}_{\text{MeOH in TX-D}} - \delta^2\text{H}_{\text{MeOH in TM+D}} = (H_n + 3) \times (\delta^2\text{H}_{\text{FAME in TX-D}} - \delta^2\text{H}_{\text{FAME in TM+D}}) / 3 \quad \text{equation (6-3)}$$

$\Delta\delta^2\text{H}_{\text{MeOH}}$  between TX-D and TM+D is 27.7 mUr, derived from the  $\Delta\delta^2\text{H}_{\text{FAME}}$  between TX-D and TM+D of 2.6 mUr and equation (6-3).

FAME	H number of FA	H number of FAME	$\Delta\delta^2\text{H}$ of -CH <sub>3</sub> of methanol between TX-D and TM+D	$\Delta\delta^2\text{H}$ of FAME between TX-D and TM+D
C16:0	31	34	27.7	2.4
C18:1	33	36	27.7	2.3
C18:2	31	34	27.7	2.4

Calculation:

$$\delta^2\text{H}_{\text{FAME in TX-D}} - \delta^2\text{H}_{\text{FAME in TM+D}} = (\delta^2\text{H}_{\text{MeOH in TX-D}} - \delta^2\text{H}_{\text{MeOH in TM+D}}) \times 3 / (H_n + 3) \quad \text{equation (6-4)}$$

For the other FAMEs, the  $\Delta\delta^2\text{H}_{\text{FAME}}$  between TX-D and TM+D should be about 2.3-2.4 mUr, derived based on  $\Delta\delta^2\text{H}_{\text{MeOH}}$  between TX-D and TM+D of 28 mUr and equation (6-4).

Table A- 7: Estimation of  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the mixed methanol in TM+D mixture

$\delta^{13}\text{C}$ value of methoxy group of DMP	$\delta^{13}\text{C}$ value of originally added Methanol	$\delta^{13}\text{C}$ value of mixture Methanol
assumed	measured	estimated
$\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}}$ , mUr	$\delta^{13}\text{C}_{\text{MeOH}}$ , mUr	$\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP+MeOH}}$ , mUr
-80	-38.9	-42.2
-20		-37.4

$\delta^2\text{H}$ value of methoxy group of DMP	$\delta^2\text{H}$ value of originally added Methanol (bulk)*	$\delta^2\text{H}$ value of mixture Methanol
assumed	measured	estimated
$\delta^2\text{H}_{\text{H}_3\text{CO-DMP}}$ , mUr	$\delta^2\text{H}_{\text{MeOH}}$ , mUr	$\delta^2\text{H}_{\text{H}_3\text{CO-DMP+MeOH}}$ , mUr
-300	-169.5	-179.9
-20		-157.5

\*The  $\delta^2\text{H}$  value used for the calculation should be the  $\delta^2\text{H}$  value of  $\text{H}_3\text{CO-}$  from methanol. It cannot be measured by GC-C/Py-IRMS. So, for the rough estimation, the bulk  $\delta^2\text{H}$  value was used.

Calculation:

The  $\delta^{13}\text{C}$  and  $\delta^2\text{H}$  values of the mixed methanol, i.e., the originally added methanol + reaction intermediate methanol, in TM+D mixture, were calculated using equations (7-1) and (7-2), respectively:

$$\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP+MeOH}} = a * \delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}} + b * \delta^{13}\text{C}_{\text{MeOH}} \quad \text{equation (7-1)}$$

$$\delta^2\text{H}_{\text{H}_3\text{CO-DMP+MeOH}} = a * \delta^2\text{H}_{\text{H}_3\text{CO-DMP}} + b * \delta^2\text{H}_{\text{MeOH}} \quad \text{equation (7-2)}$$

a: Fraction of MeOH from DMP (0.08)

b: Fraction of MeOH from MeOH (0.92)

Table A- 8: Estimation <sup>i</sup> of  $\delta^{13}\text{C}_{\text{FAME}}$  in TM+D mixture and the difference ( $\Delta\delta^{13}\text{C}_{\text{FAME}}$ ) between the measured and estimated values

$\delta^{13}\text{C}$ value of $\text{H}_3\text{CO}$ -group of DMP	Comp.	Number of C in FA	Number of C in FAME	in TX-D, without DMP			in TM+D, with DMP			$\Delta\delta^{13}\text{C}_{\text{FAME}}$ , mUr	
				$\delta^{13}\text{C}$ value of FAMEs	$\delta^{13}\text{C}$ value of methanol	$\delta^{13}\text{C}$ value of FAs	$\delta^{13}\text{C}$ value of FAMEs	$\delta^{13}\text{C}$ value of methanol	$\delta^{13}\text{C}$ value of FAs		
assumed				measured (a)	measured (b)	estimated (c)	measured (d)	estimated (e)	estimated (f)	A: (f)-(a)	B: (f)-(d)
$\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}}$ , mUr				$\delta^{13}\text{C}_{\text{FAME}}$ , mUr	$\delta^{13}\text{C}_{\text{MeOH}}$ , mUr	$\delta^{13}\text{C}_{\text{FA}}$ , mUr	$\delta^{13}\text{C}_{\text{FAME}}$ , mUr	$\delta^{13}\text{C}_{\text{MeOH}}^*$ , mUr	$\delta^{13}\text{C}_{\text{FAME}}$ , mUr		
-80	C16:0	16	17	-31.88	-38.9	-31.4	-31.9	-42.2	-32.1	-0.2	-0.2
	C18:1	18	19	-30.7	-38.9	-30.2	-30.4	-42.2	-30.9	-0.2	-0.4
	C18:2	18	19	-30.75	-38.9	-30.3	-30.6	-42.2	-30.9	-0.2	-0.3
	C18:3	18	19	-31.45	-38.9	-31	-31.7	-42.2	-31.6	-0.2	0.1
-20	C16:0	16	17	-31.88	-38.9	-31.4	-31.9	-37.4	-31.8	0.1	0.1
	C18:1	18	19	-30.7	-38.9	-30.2	-30.4	-37.4	-30.6	0.1	-0.2
	C18:2	18	19	-30.75	-38.9	-30.3	-30.6	-37.4	-30.7	0.1	-0.1
	C18:3	18	19	-31.45	-38.9	-31	-31.7	-37.4	-31.4	0.1	0.3

<sup>i</sup> Estimation was made with an assumption that  $\delta^{13}\text{C}_{\text{H}_3\text{CO-DMP}}$  ranges from -80 to -20 mUr

\* Mixture of originally added methanol and the reaction intermediate methanol

(c): calculated based on the (a) and (b) values and the mass balance equation (6) in Supporting Information

(e): calculated based on the equation (7) in Supporting Information

(f): calculated based on the (d) and (e) values and the mass balance equation (6) in Supporting Information

Table A- 9: Estimation <sup>i</sup> of  $\delta^2\text{H}_{\text{FAME}}$  in TM+D mixture and the difference ( $\Delta\delta^2\text{H}_{\text{FAME}}$ ) between the measured and estimated values

$\delta^2\text{H}$ value of $\text{H}_3\text{CO}$ -group of DMP	Comp.	Number of C in FA	Number of C in FAME	in TX-D, without DMP			in TM+D, with DMP			$\Delta\delta^2\text{H}_{\text{FAME}}$ , mUr	
				$\delta^2\text{H}$ value of FAMEs	$\delta^2\text{H}$ value of methanol	$\delta^2\text{H}$ value of FAs	$\delta^2\text{H}$ value of FAMEs	$\delta^2\text{H}$ value of methanol	$\delta^2\text{H}$ value of FAs		
				assumed	measured (a)	measured (b)	estimated (c)	measured (d)	estimated (e)	estimated (f)	A: (f)-(a)
$\delta^2\text{H}_{\text{H}_3\text{CO-DMP}}$ , mUr				$\delta^2\text{H}_{\text{FAME}}$ , mUr	$\delta^2\text{H}_{\text{MeOH}}$ , mUr	$\delta^2\text{H}_{\text{FA}}$ , mUr	$\delta^2\text{H}_{\text{FAME}}$ , mUr	$\delta^2\text{H}_{\text{MeOH}}^*$ , mUr	$\delta^2\text{H}_{\text{FAME}}$ , mUr		
-300	C16:0	31	34	-190	-169.5	-192	-189.3	-179.9	-190.9	-0.9	-1.6
	C18:1	33	36	-194.3	-169.5	-196.6	-196.32	-179.9	-195.2	-0.9	1.2
	C18:2	31	34	-192.7	-169.5	-194.9	-192.13	-179.9	-193.6	-0.9	-1.5
	C18:3	29	32	-207	-169.5	-210.9	-204.45	-179.9	-208	-1	-3.5
-20	C16:0	31	34	-190	-169.5	-192	-189.3	-157.5	-188.9	1.1	0.4
	C18:1	33	36	-194.3	-169.5	-196.6	-196.32	-157.5	-193.3	1	3
	C18:2	31	34	-192.7	-169.5	-194.9	-192.13	-157.5	-191.6	1.1	0.5
	C18:3	29	32	-207	-169.5	-210.9	-204.45	-157.5	-205.9	1.1	-1.4

<sup>i</sup> Estimation was made assuming that  $\delta^2\text{H}_{\text{H}_3\text{CO-DMP}}$  ranges from -300 to -20 mUr.

\* Mixture of originally added methanol and the reaction intermediate methanol

(c): calculated based on the (a) and (b) values and the mass balance equation (6) in Supporting Information

(e): calculated based on the equation (7) in Supporting Information

(f): calculated based on the (d) and (e) values and the mass balance equation (6) in Supporting Information

Table A- 10: Correlation analysis of the  $\delta^2\text{H}$  values of the four FAs with latitude and altitude in 2019

2019	sample size	C18:1 vs. latitude			C18:1 vs. altitude		
		correlation test	r	p values	correlation test	r	p values
Hesse	44	Pearson test	0.316*	0.04	Pearson test	0.29	0.05
North Hesse	19	Pearson test	-0.16	0.52	Pearson test	0.34	0.16
Central Hesse	19	Pearson test	.486*	0.03	Pearson test	0.05	0.85
South Hesse	6	Pearson test	-0.63	0.18	Pearson test	0.36	0.49
		C18:2 vs. latitude			C18:2 vs. altitude		
Hesse	44	Pearson test	0.11	0.49	Pearson test	0.09	0.58
North Hesse	19	Pearson test	0.10	0.69	Pearson test	0.51*	0.03
Central Hesse	19	Pearson test	0.32	0.18	Pearson test	-0.24	0.32
South Hesse	6	Pearson test	-0.75	0.09	Pearson test	-0.13	0.81
		C18:3 vs. latitude			C18:3 vs. altitude		
Hesse	44	Pearson test	-0.06	0.70	Pearson test	-0.01	0.98
North Hesse	19	Pearson test	-0.15	0.53	Pearson test	0.22	0.36
Central Hesse	19	Spearman test	.475*	0.04	Spearman test	0.06	0.81
South Hesse	6	Pearson test	-.846*	0.03	Pearson test	-0.27	0.60
		C16:0 vs. latitude			C16:0 vs. altitude		
Hesse	44	Pearson test	0.22	0.16	Pearson test	0.13	0.39
North Hesse	19	Pearson test	-0.33	0.16	Pearson test	0.00	1.00
Central Hesse	19	Pearson test	.480*	0.04	Pearson test	-0.12	0.63
South Hesse	6	Pearson test	-0.20	0.70	Pearson test	0.65	0.16

Note: significant correlation is marked with \* ( $p < 0.05$ ).



Table A- 11: Correlation analysis of the  $\delta^2\text{H}$  values of the four FAs with latitude and altitude in 2020

2020	sample size	C18:1 vs. latitude			C18:1 vs. altitude		
		correlation test	r	p values	correlation test	r	p values
Hesse	38	Spearman test	-0.36*	0.03	Spearman test	-0.31	0.06
North Hesse	18	Pearson test	-0.40	0.10	Pearson test	-0.511*	0.03
central Hesse	13	Pearson test	0.21	0.50	Pearson test	-0.13	0.68
south Hesse	7	Spearman test	-0.07	0.88	Pearson test	-0.28	0.54
		C18:2 vs. latitude			C18:2 vs. altitude		
Hesse	38	Pearson test	-0.47*	0.00	Pearson test	-0.33*	0.04
North Hesse	18	Pearson test	-0.54*	0.02	Pearson test	-0.39	0.11
Central Hesse	13	Spearman test	-0.19	0.54	Spearman test	-0.34	0.26
South Hesse	7	Spearman test	-0.43	0.34	Spearman test	-0.11	0.82
		C18:3 vs. latitude			C18:3 vs. altitude		
Hesse	38	Spearman test	-0.28	0.09	Spearman test	-0.29	0.08
North Hesse	18	Pearson test	-0.44	0.07	Pearson test	-0.42	0.09
Central Hesse	13	Spearman test	0.02	0.95	Spearman test	-0.15	0.62
South Hesse	7	Spearman test	-0.68	0.09	Spearman test	0.04	0.94
		C16:0 vs. latitude			C16:0 vs. altitude		
Hesse	38	Pearson test	-0.41*	0.01	Pearson test	-0.32*	0.05
North Hesse	18	Pearson test	-0.50*	0.04	Pearson test	-0.36	0.14
Central Hesse	13	Pearson test	0.08	0.81	Pearson test	-0.23	0.46
South Hesse	7	Spearman test	-0.61	0.15	Pearson test	-0.31	0.49

Note: significant correlation is marked with \* ( $p < 0.05$ ).

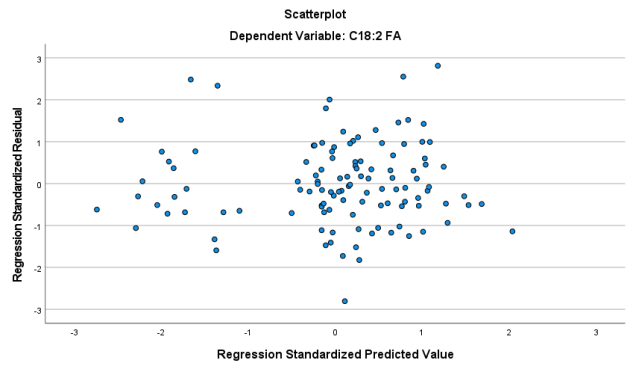
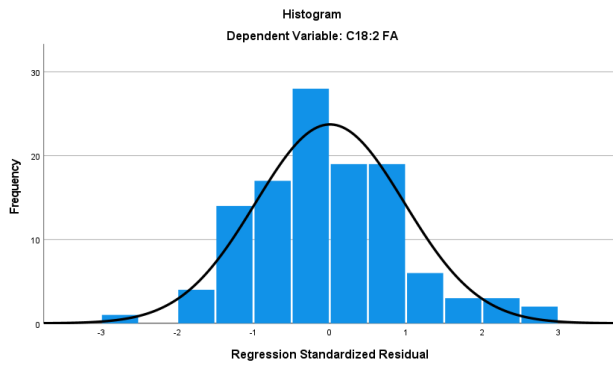


Figure A- 3: Diagnostic plots for the linear regression model of  $\delta^2\text{H}$  of C18:2 FA.

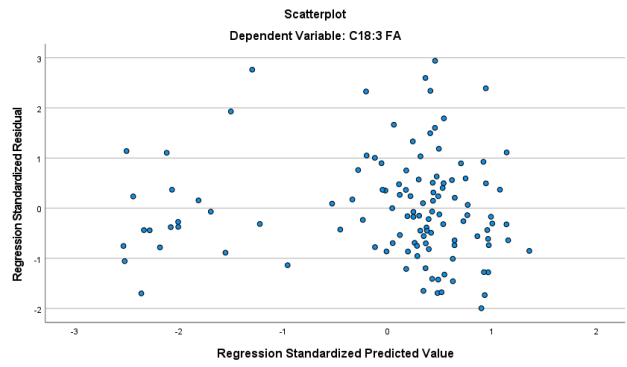
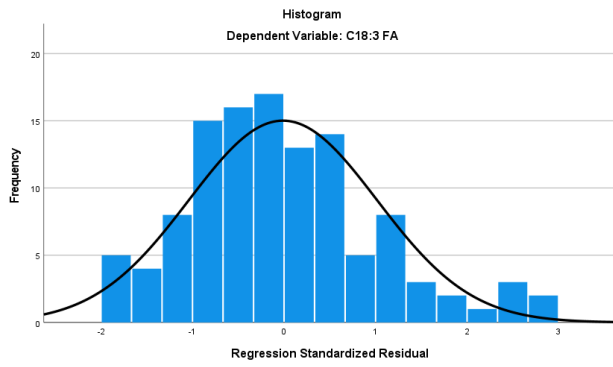


Figure A- 4: Diagnostic plots for the linear regression model of  $\delta^2\text{H}$  of C18:3 FA.

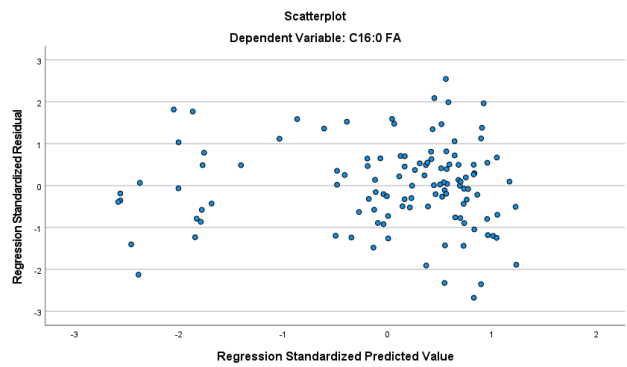
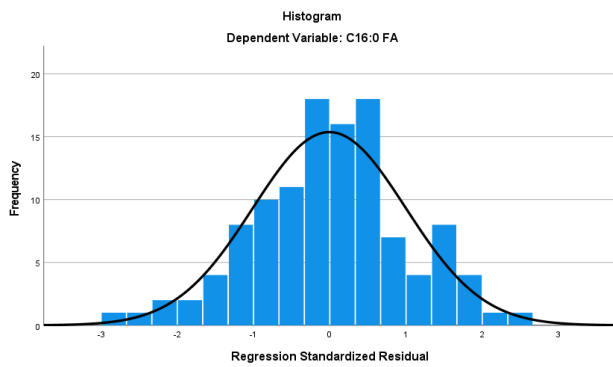


Figure A- 5: Diagnostic plots for the linear regression model of  $\delta^2\text{H}$  of C16:0 FA.

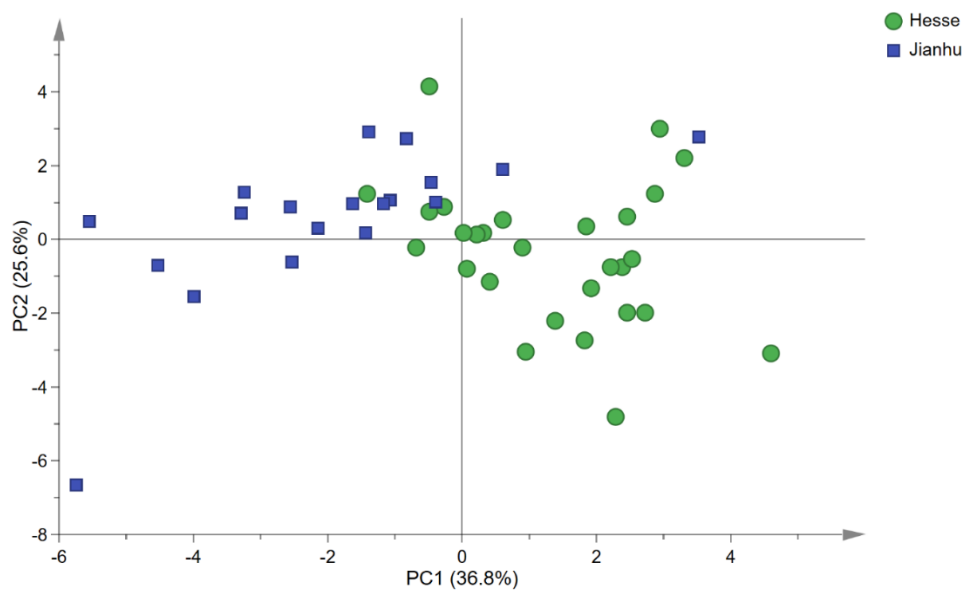


Figure A- 6: PCA score scatter plot for rapeseed samples from Hesse and Jianhu in 2019.

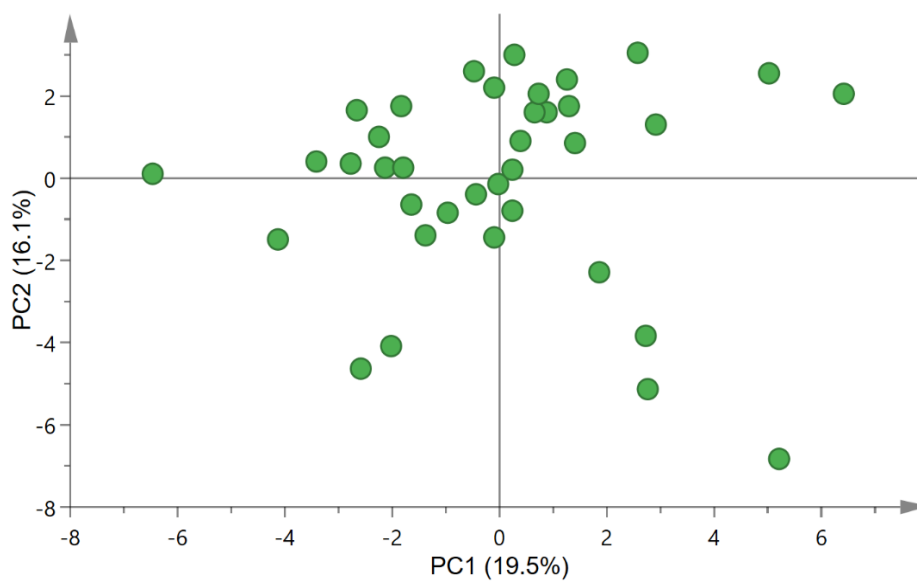


Figure A- 7: PCA score scatter plot for rapeseed samples from Hesse in 2019 with 36 variables (16 variables related to the  $\delta^2\text{H}$  values of FAs and 20 variables related to elemental composition).

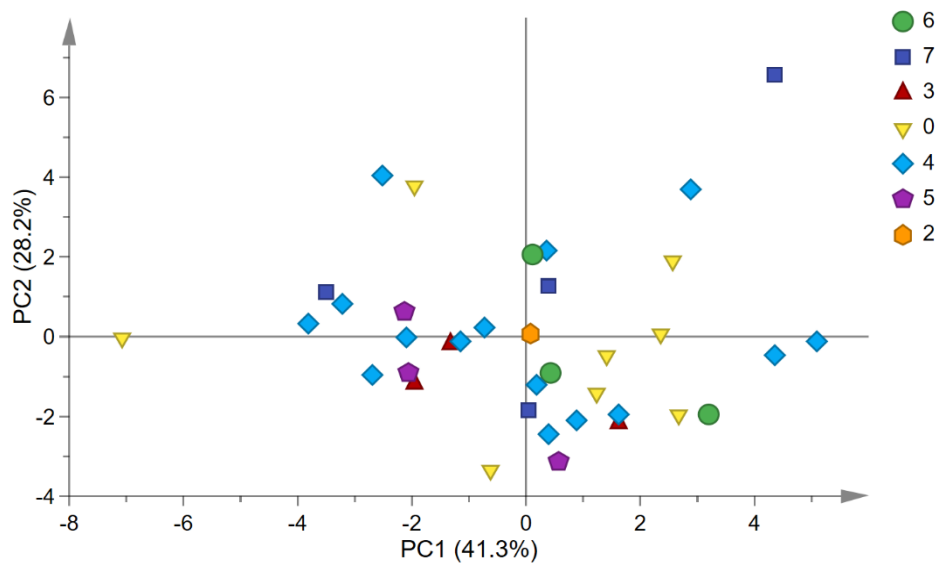


Figure A- 8: PCA score scatter plot for rapeseed samples from Hesse in 2019 with 16 variables related to the  $\delta^2\text{H}$  values

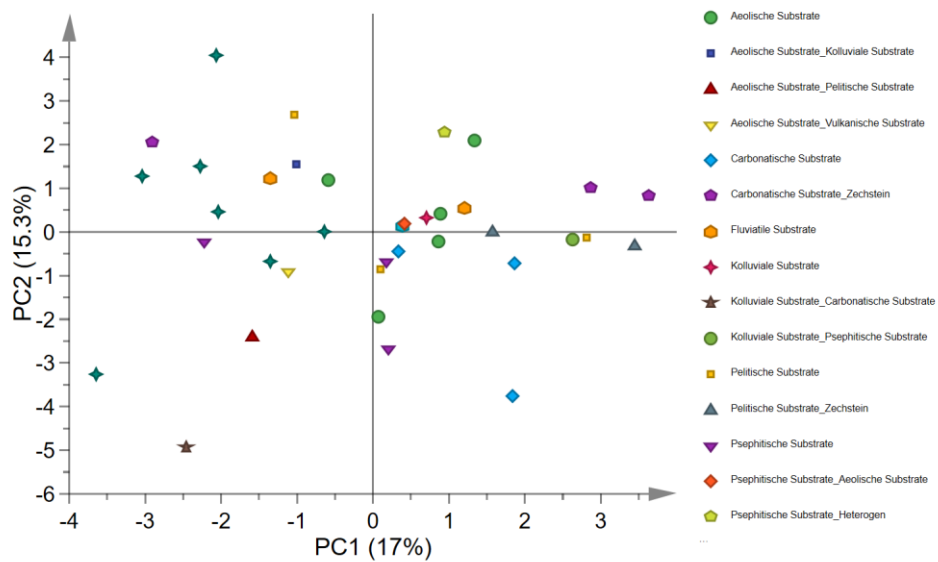


Figure A- 9: PCA score scatter plot for rapeseed samples from Hesse in 2019 with 20 variables related to the elemental composition.