Advancements in land-use-specific sediment source apportionment: From concentration-dependent mathematical mixtures to novel lignin derived methoxy isotopes

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Terry Cox

Basel, 2024

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Erstbetreuerin: Prof. Dr. Christine Alewell

Zweitbetreuer: Prof. Dr Ansgar Kahmen

Externe Expertin: Prof. Dr Miriam Glendell

Basel, den 28 Mai 2024

Prof. Dr. Marcel Mayor

Dekan

Summary

Accelerated soil erosion poses a significant global threat to soil health. Sediment source fingerprinting aids in the identification and apportionment of the main erosion sources. Retracing sediments to their sources and respective land-use by sediment fingerprinting offers a semi-empirical, field-based approach to determine the proportional contribution of different soil erosion sources and potentially can provide information to policy makers, land managers and researchers. Given the pivotal role of land-use change in historical and contemporary soil erosion systems, current tracers are often limited by their inability to discriminate amongst land-uses. Thus, the development of tracers which are able to discriminate between land-uses represents a crucial advancement in this field. While compound specific stable isotope (CSSI) tracers are immensely beneficial for land-use-specific sediment source apportionment, they still have several unresolved issues. The aim of this thesis is to enhance land-use-specific sediment source apportionment by addressing these issues and offering advancements in the field to enable more accurate representations of sediment dynamics in the environment.

Initially, I developed a simple approach to evaluate the performance of isotopic mixing models using mathematical mixtures. Isotopic tracers depend not only on their isotopic values and mixing proportions but also on the concentration of each tracer in each source. Therefore, I devised a novel concentration-dependent mathematical mixture tool. Utilizing this tool, I employed a 'brute force' method to investigate how the number of fatty acid (FA) tracers influences the model's performance. Contrary to the previous assumption that the Bayesian framework mixing model (MixSIAR) handles all conservative tracers beneficially, I discovered that tracer redundancy, wherein tracers which share similar mixing spaces and co-linearity (one-dimensional mixing line), negatively impact the model's performance.

Our findings using the concentration-dependent mathematical mixture tool also highlighted the need for an additional land-use-specific tracer to expand the one-dimensional mixing line formed by $\delta^{13}C$ FA tracers with co-linearity (similar mixing spaces). Consequently, I incorporated the $\delta^{13}C$ values of lignin-derived methoxy groups (LMeO) as an additional tracer. This approach was applied to investigate the sediment history of Lake Baldegg over the past 130 years. In particular, I successfully distinguished between the inputs of plant debris (POM_{terr}) and mineral-associated organic matter (MOAM). By assigning POM_{terr} as its own distinct source, I was able to remove the POM_{terr} contribution from the sediment source apportionments and apportion only the MOAM fraction.

Given our focus on Lake Baldegg's historically deposited sediment, Suess corrections are necessary to account for the changing atmospheric $\delta^{13}C$ composition. For more representative Suess corrections, I accounted for multiple tracer turnover times (10, 30, and 100 years) in the Suess correction. This approach facilitated achieving a more representative account of past sediment dynamics within the Baldegg catchment and stands as an important component in using isotopic tracers to estimate historic sediment dynamics. Using both the POM_{terr} removal and Suess correction methods, our estimates for sediment source apportionment aligned well with the policy and land-use change in the catchment.

While our results of the historic apportionment are highly credible, the conservativeness of δ^{13}C LMeO during transport and deposition required further investigation.

To assess the conservativeness of δ^{13}C LMeO values during degradation, I utilized the dual isotopes of LMeO (δ^{2}H LMeO and δ^{13}C LMeO) across the degradation continuum from the litter layer to the MAOM fraction. Since δ^{2}H MeO (lignin and pectin MeO groups) values are known to be stable during degradation, I was able to disentangle isotopic fractionation from source mixing and demonstrated the stability of δ^{13}C LMeO values during degradation. Furthermore, and importantly, I found that the dual isotope approach allowed for the discrimination of the litter layer, above-ground woody material, and root lignin. This method then enabled the apportionment of lignin sources in organic and mineral horizons. By applying this method to two contrasting soil types, a podzol and stagnosol, I elucidated different soil type-specific lignin mixing and sources across the degradation continuum. While the stagnosol demonstrated minimal translocation of above-ground lignin sources to the MOAM fraction, the podzol showed the accumulation of lignin from above-ground sources in the MOAM. Considering the high percentage of lignin in the terrestrial biosphere, this novel, simple, solvent-free, and rapid method of demonstrating lignin dynamics may hold great potential in terms of understanding and modelling carbon sequestration.

Considering the high abundance of LMeO in the terrestrial biosphere and the extremely depleted δ^{13}C LMeO values of leaf litter (~60‰), I assessed how much of the well-known but lesser understood bulk ^{13}C enrichment with organic matter degradation can be explained by the transition of lignin sources from leaf litter to roots. A mass balance approach was used to determine the δ^{13}C values of the non-LMeO fraction. Using the difference in enrichment of the δ^{13}C non-LMeO and δ^{13}C bulk values, up to 14% can be explained by LMeO. Our findings suggest that models using δ^{13}C bulk as a proxy for carbon turnover may overestimate degradation.

By providing a method to test the accuracy of concentration-dependent mixing models and offering an alternative CSSI tracer capable of discriminating between POM_{terr} and MOAM, this thesis makes significant contributions to advancing sediment fingerprinting using CSSI tracers. Additionally, the use of LMeO in determining lignin mixing dynamics in soils represents an important step in understanding carbon sequestration. Simultaneously, this thesis both resolves some of the issues with CSSI sediment source apportionment and opens new questions and tools awaiting exploration by curious and inquisitive researchers in the future.

Contents

Introduction

1.1 The problem of human induced accelerated soil erosion and sedimentation has persisted globally throughout history, with evidence of anthropogenically induced soil erosion dating back to the era of the Palaeolithic-Mesolithic hunter-gatherer (Dotterweich et al. 2013). Soil erosion has then substantially increased in Europe since the Bronze Age (Dotterweich 2008). The rise of population pressure and consequent changes in land-use, leading to heightened susceptibility of soil to extreme weather events, have been proposed as the driving force behind the increase in historical sedimentation rates north of the Alps (Dotterweich 2008, 2013). The deterioration and disappearance of multiple civilisations throughout history has been related to the degradation of resources, accelerated soil erosion and a decline in crop yield (David R. Montgomery 2012 and references within). Throughout history, soil conservation practises were passed down between farmers through generations and have been embedded in traditional practices, while scientific interest in soil erosion is more recent and began at the end of the 18th century (Dotterweich 2013).

Despite more than two centuries of scientific advancements, little has changed regarding soil erosion. Human civilisations are still driven by the effect of depleting soil resources. Soil erosion is estimated to cost around €1.25 billion per year to the EU agricultural sector (Panagos et al. 2018), and costs the EU €2.3 billion annually for sediment removal (Panagos et al. 2024). Furthermore, accelerated soil erosion and sedimentation have negative implications on biogeochemical cycles (Lal 2003; Alewell et al. 2020). Both presently and in the future, we face the additional challenge of global warming and the subsequent increase in extreme weather events. Soil erosion induced carbon emissions have been estimated to be 0.8–1.2 PgC/ year (Lal 2003), potentially generating an accelerated soil erosion feedback loop. The world's population is expected to keep increasing, without effective soil erosion mitigation polices we can expect the problems associated with population pressure and limited resources to enhance the consequences of accelerated soil erosion.

In July 2023, the European Commission introduced the Soil Monitoring Law with the aim of safeguarding both present-day and forthcoming soil quality. This legislation seeks to initiate measures by 2030 that aims to conclude in the promotion of healthy soils by 2050. The policy mandates the assessment of existing land degradation and the evaluation of the potential impact of the proposed legal framework. There is no shortage of models predicting and mapping soil erosion models (Borrelli et al. 2021), however, the validation and testing of these models is lacking, with spatial erosion models not comparing well with independent data (Batista et al. 2019a).

Sediment source apportionment offers a method to test soil erosion models (Fox and Martin 2015; Wiltshire et al. 2023). It is a method of identifying and quantifying the different origins of sediment using the inherent properties (or tracers, characteristics, fingerprint) of source and sediment samples in combination with mixing models. While the application of sediment source apportionment to inform polices and sediment mitigations is omnipresent in literature, there are few documented instances of its implementation in catchment management and policy making (Olley et al. 2009; Mukundan et al. 2012). There is still a disconnect between sediment source apportionment research and application (Collins et

al. 2020; Owens 2022). Potentially this comes from the trouble of upscaling small catchments and the difference in information requirements between scientists, farmers and policy makers (Owens 2022). In general, sediment fingerprinting is not directly applied to inform polices and mitigations strategies but rather a research tool. However, it may be a supportive tool when used as an independent and semi-empirical method to test soil erosion models.

Sediment source apportionment is a semi empirical, field-based approach used to identify erosion hotspots both spatially and temporally. Sediment fingerprinting has been applied to a wide range of fluvial sediment types including; lacustrine sediment cores (Le Gall et al. 2016; Lavrieux et al. 2019), flood plains (Pulley et al. 2015; Kemper et al. 2022), dam reservoirs (Nosrati et al. 2011; Ben Slimane et al. 2016) and riverine systems (Collins et al. 2001; Upadhayay et al. 2018b; Bispo et al. 2020). The method uses various characteristics of potential source soils as a fingerprint. There is a plethora of potential soil characteristics which has been applied, including but not limited to radionuclides (Evrard et al. 2013), elemental geochemistry (Batista et al. 2019b), compound-specific stable isotopes (Upadhayay et al. 2020b; Hirave et al. 2021), colour (Martínez-Carreras et al. 2010), and ultraviolet-visible absorbance (Lake et al. 2022), amongst others. These fingerprints are then compared to the sediment fingerprint and the relative proportion of each source is estimated using either linear mixing equations (Collins et al. 1997), Bayesian models (Stewart et al. 2015; Stock et al. 2018) or frequentist models (Pulley and Collins 2018).

Given the critical role of land-use changes in both historical and contemporary soil erosion, the majority of tracers are limited due to a lack of ability to be land-use specific. Thus, the development of land-use-specific tracers is a crucial advancement in this field. The compound specific stable isotopes (CSSI) of fatty acids (FA) and alkanes have demonstrated to be able to discriminate between different land-uses and have been applied to multiple catchment types and temporal scales (Cooper et al. 2015; Alewell et al. 2016; Brandt et al. 2018; Vale et al. 2020; Gibbs et al. 2020; Hirave et al. 2021). Although CSSI tracers are valuable for sediment source apportionment, they still face several unresolved problems (P):

P1. One dimensional mixing line

Fingerprinting using $\delta^{13}C$ FA values presents limitations, with source values being located along a linear mixing line (Alewell et al. 2016; Lavrieux et al. 2019). This alignment can lead to misclassification of contributions from central sources as originating from sources at either endpoint. The linear mixing line problem is not confined to isotopic tracers. Colour (Barthod et al. 2015) and geochemical tracers (Bouchez et al. 2011) have also presented a linear mixing line. The $\delta^{13}C$ FA one-dimensional mixing line problem can be seen as a product of the elongation during FA biosynthesis. The elongation process occurs with a $\delta^{13}C$ depletion with increasing alkyl length (Chikaraishi et al. 2004). As this process may be the same for very long chain FAs of different plant groups, this could result in the $\delta^{13}C$ FAs of different alkyl lengths demonstrating co-linearity, and subsequently a one-dimensional mixing line. Previous resolutions of this problem have involved grouping sources, leading to the apportionment between two sources, preventing the ability to discriminate between three or more sources (Alewell et al. 2016).

Intriguingly, this effect is not observed in all reported cases of land-use specific sediment source apportionment using $\delta^{13}C$ FAs (Upadhayay et al. 2020b; Lizaga et al. 2021). The absence of a mixing line potentially suggests that the previous land-use or crop type had a higher concentration of specific FAs, which are now more prevalent in the current soil, resulting in their isotopic fingerprint still being present in the soil (Swales and Gibbs 2020; Upadhayay et al. 2020a). The expansion of the mixing line requires an additional tracer. $\delta^{13}C$ FA values have been used in combination with geochemical tracers (Lizaga et al. 2022). However, geochemical tracers are not constrained by land-use type, hence this may give inaccurate source classification. As such, an additional land-use specific tracer is required.

P2. Tracer redundancy

A secondary effect of the mixing line is the potential for tracer redundancy, wherein tracers exhibiting co-linearity possess similar mixing spaces. As mixing models utilize the relative tracer values rather than absolute values, tracers with similar mixing spaces (co-linearity) may essentially provide redundant information. The current approach with MixSIAR is to include all tracers that meet two criteria: optimization of model performance through the Kruskal-Wallis H test and identification of non-conservative tracers using a box and whisker plot. However, the question remains whether tracers with similar mixing spaces offer additional benefits to the model.

P3. Evaluating the accuracy of concentration-dependent mixing models

While the precision of sediment fingerprinting models can be assessed through stochastic modelling, determining their accuracy presents challenges (Batista et al. 2022). Artificial mixtures, including laboratory mixtures and mathematical mixtures (also known as 'virtual mixtures'), have been utilized to test model accuracy. Laboratory and mathematical mixtures have demonstrated similar model outputs (Batista et al. 2022). Between 1987 and 2020, only 1.2% of studies used artificial mixtures (based on a Web Of Science word query) (Batista et al. 2022). Although recently, the use of artificial mixtures has increased (29% of studies between 2018 and 2019) (Collins et al. 2020). Considering the simplicity and minimal resource requirements needed for mathematical mixtures, we propose mathematical mixtures should be fundamental step when applying mixing models.

Additionally, when using isotopic tracers, the sediment isotopic value depends on three factors; i) the source mixing proportions, ii) the tracer's isotopic value of source soils and iii) the concentration of the tracer in each source soils (Upadhayay et al. 2018a). The validation of models dealing with concentration-dependent tracers has until recently been dependent on the labour-intensive analysis of a small number of laboratory mixtures (Bravo-Linares et al. 2018) or the oversimplification by assuming isotopic tracers mix linearly (non-concentration dependency) (Collins et al. 2019; Bahadori et al. 2019). While mathematical mixtures may not accurately mirror natural mixing processes due to factors such as signal degradation, tracer alteration, isotope fractionation, and particle size selectivity, they remain essential for assessing model performance (Haddadchi et al. 2014; Batista et al. 2019b; Vale et al. 2022). Considering the increase in using isotopic tracers, the lack of testing models with concentration-dependent mixtures is rather concerning. Such testing is necessary to establish the confidence level applicable to the model.

P4. Non-mineral associated organic matter sources.

Sediment source apportionment aims to utilize the inherent properties of soil and sediment, any additional contribution from non-soil sources resulting in the inaccurate apportionment of soil contributions. The synthesis of FAs is not limited to terrestrial plants; algae and bacteria are known to produce FAs (O'leary 1962; De Carvalho and Caramujo 2018). As the predominant source of saturated long-chain fatty acids (LCFA) is terrestrial plants, the exclusion of short, medium, and non-saturated FAs has been applied to reduce the uncertainty related to input from non-terrestrial plant-derived FAs (Alewell et al. 2016).

While the exclusion of short and medium FAs has been used to reduce contribution from non-soil sources, in Lake Baldegg (Kanton Lucerne, Switzerland) extremely isotopically depleted sediment values were observed for δ^{13}C FA_{24}, and δ^{13}C FA_{26} (Lavrieux et al. 2019). The authors speculated an in-situ source was of FA was missing, most likely algae. Sediment from 1965 was further evidence for this hypothesis. Sediment from 1965 demonstrated most depleted δ^{13}C FA values, and also had the highest TOC and TN content over the 130-year record, suggesting primary production was at its maximum (Lotter et al. 1997). In Lake Baldegg, an additional aquatic source(s) of δ^{13}C FA_{24}, and δ^{13}C FA_{26} is highly probable. Without the identification and measurement of these sources, sediment source apportionment using δ^{13}C FA_{24}, and δ^{13}C FA_{26} cannot be applied accurately.

Contradicting our current understanding of erosion processes and modelling, forests are frequently identified as a major source of sediment (Alewell et al. 2016; Chen et al. 2016; Upadhayay et al. 2020b; Wiltshire et al. 2022). While this may be true in specific circumstances, forests are typically less vulnerable to soil erosion due to their protective canopy cover, understory or ground vegetation, and humus layer (Blanco-Canqui and Lal 2010). Wiltshire et al. (2022) hypothesized this overestimation could be a result of including terrestrial particulate organic matter (POMterr) in the sediment fingerprint. While the physical separation of POM and mineral associated organic matter (MOAM) has been achieved via density centrifugation, it requires relatively large sample amounts (5 g) (Cui et al. 2016). Distinguishing between MAOM and POM_{terr} as separate endmembers using CSSI tracers is challenging due to small isotopic difference but is a necessary step for accurate apportionment of sediment sources rather than sources of organic tracers.

P5. The incorporation of multiple mean residence times when applying the Suess correction to historical sediment

To account for changes in atmospheric $\delta^{13}CO_2$ over the past century (Verburg 2007), the CSSI tracers for historical sediment apportionment requires Suess corrections (Gibbs et al. 2014). However, Suess corrections are often omitted, assuming negligible variability induced by the Suess effect compared to source uncertainty (Brandt et al. 2018), or a single mean residence time (MRT) is used (Bravo-Linares et al. 2020). The mean residence time of FAs varies from decades to millennia (Lützow et al. 2006; Wiesenberg et al. 2010). Therefore, understanding how a range of MRTs affects sediment source apportionment results is crucial for determining the uncertainty associated with modelling sediment source apportionments.

In this thesis, I aim to resolve the problems stated above in Chapter 2 and 3. Additionally, in Chapter 4 I test the conservativeness of a novel tracer whilst simultaneously using a new tracer to apportion lignin sources.

1.2 Objectives of the PhD thesis

Chapter 2:

CSSI tracers often demonstrate a one-dimensional mixing line, in which source values are located on a mixing line. To explore the effect of this on model peformance, I plan to analyse $\delta^{13}C$ FA values in source soils from the Rhine catchment (Basel-Switzerland) and expand the $\delta^{13}C$ FA one-dimensional mixing line into a more suitable multi-dimensional mixing space using $\delta^{15}N$ (P1). As this manuscript solely uses mathematical mixtures, the conservativeness of $\delta^{15}N$ and potential isotopic fractionation of $\delta^{15}N$ during transport and sedimentation are less relevant. The typical approach with mixing models (i.e., MixSIAR) involves maximizing the number of tracers that pass conservativeness tests. However, I hypothesis that co-linear tracers which display a similar mixing space might not improve model output but rather hinder the performance (P2). Using 150 concentration-dependent mathematical mixtures, I will apply a "brute force" method, and evaluate all possible permutations and combinations of the $\delta^{15}N$ and $\delta^{13}C$ FA tracer set and determine the model accuracy by comparing the estimated apportionment to the known proportions (P2). In addition to this rigorous testing of CSSI tracer selection, I plan to investigate the effect of adding non-informative tracers to the model (i.e., identical replicates of $\delta^{13}C$ FA). Using this method, I plan to determine how much information is gained by additional different $\delta^{13}C$ FA tracers compared to adding identical replicate tracers (P3).

Chapter 3:

An additional tracer to $\delta^{13}C$ FAs is required to break up the mixing line. While in Chapter 2 I used $\delta^{15}N$ values with mathematical mixtures, the use of $\delta^{15}N$ values has come under scrutiny due to the potential for isotopic fractionation. In this chapter, I novelly apply the $\delta^{13}C$ values of lignin derived methoxy groups (LMeO) as a sediment tracer (P1). $\delta^{13}C$ LMeO values have demonstrated a considerably large discrimination between C_3, C_4 and CAM plants (Keppler et al. 2004). The conservativeness of $\delta^{13}C$ LMeO is yet to be fully explored, litter bags experiments measuring bulk $\delta^{13}C$ MeO (pectin and lignin) indicated small isotopic fractionation during degradation (Anhäuser et al. 2015). This was suggested to be because of the preferential degradation of pectin. The analysis of the bulk MeO $\delta^{13}C$ values in the degradation continuum from wood to coal demonstrated significant fractionation (Lloyd et al. 2021). However, considering we would not expect LMeO degradation on the same scale as the wood-to-coal degradation continuum (remaining MeO fraction in coal ca. 10^{-4}), we would expect isotopic fractionation to be minimal within LMeO concentrations in soil and sediments.

Leaf and woody material have demonstrated an extremely large difference in $\delta^{13}C$ LMeO (~30‰). I propose that POM_{terr} and MOAM can be differentiated using $\delta^{13}C$ LMeO and utilised to determine the relative proportions of POM_{terr} to the sediment fingerprint (P4). The POM_{terr} signal can then be subsequently removed allowing for the apportionment of the MOAM fraction only.

As stated in P5, historical sediment source apportionments require the correction for the changing $\delta^{13}CO_2$ in the atmosphere using Suess corrections. Here I investigate the effect of 3 residency times (10yr, 30yr, and 100yr) and no Suess corrections on the sediment source apportionment. Considering the MRT of specific compounds is dependent on the soil ecosystem and soil properties, by applying multiple corrections we can more accurately depict the uncertainty when modelling soil erosion using CSSI tracers.

Chapter 4:

In Chapter 3, I plan to use δ^{13}C LMeO as an additional tracer for land-use-specific sediment source apportionment. Here, I further test the conservativeness of δ^{13}C LMeO by the isotopic analysis of LMeO during the degradation continuum from sources of lignin (leaf litter and roots) to soil organic matter. Furthermore, considering the large discrimination between leaf and woody material, I additionally hypothesise that sources of lignin can apportioned using δ^{13}C LMeO. However, the utilisation of only δ^{13}C LMeO values is unable to disentangle whether the change in δ^{13}C LMeO is a result of changing lignin sources or from isotopic fractionation during degradation.

Evidence for the stability of δ^{2}H MeO has been demonstrated by degradation in litter bag experiments (Anhäuser et al. 2015) and cultures of bacteria and fungi with garden waste (Lu et al. 2022). The stability of δ^{2}H MeO has resulted in multiple applications e.g., climate reconstruction (Keppler et al. 2007; Wieland et al. 2024), determining the sources of 20-million-year-old wood deposits (Lee et al. 2019), identifying water sources in salt marshes (Feakins et al. 2013) and locating the sources of potato tubers (Keppler and Hamilton 2008). While the stability of δ^{2}H LMeO has been demonstrated, the stability of δ^{13}C LMeO values is less investigated and requires validation.

In a two-source mixing model, using δ^{2}H LMeO and δ^{13}C LMeO values as tracers, isotopic fractionation can be illustrated as a deviation of a mixture from a 1:1 mixing line, with the mixtures being more enriched in ^{13}C to that predicted by the mixing line (Lutz and Van Breukelen 2014a, b). If no isotopic fractionation occurs, samples will be located on the mixing line. With this in mind, I plan to use the dual isotope (δ^{13}C and δ^{2}H) analysis of LMeO to apportion the sources of lignin during the transformation of fresh material to SOM. Additionally, due to the highly depleted δ^{13}C LMeO values and relatively large percentage in the biosphere (Keppler et al. 2004), I hypothesise that the well-known but less understood bulk δ^{13}C enrichment from the Oi horizon to the mineral soil is partially explained by the changing of lignin sources from leaves to roots.

1.3 References

Alewell C, Birkholz A, Meusburger K, et al (2016) Quantitative sediment source attribution with compound-specific isotope analysis in a C3 plant-dominated catchment (central Switzerland). Biogeosciences 13:1587–1596. https://doi.org/10.5194/bg-13-1587-2016

Alewell C, Ringeval B, Ballabio C, et al (2020) Global phosphorus shortage will be aggravated by soil erosion. Nat Commun 11:. https://doi.org/10.1038/s41467-020-18326-7

Anhäuser T, Greule M, Zech M, et al (2015) Stable hydrogen and carbon isotope ratios of methoxyl groups during plant litter degradation. Isotopes Environ Health Stud 51:143–154. https://doi.org/10.1080/10256016.2015.1013540

Bahadori M, Chen C, Lewis S, et al (2019) A novel approach of combining isotopic and geochemical signatures to differentiate the sources of sediments and particulate nutrients from different land uses. Science of the Total Environment 655:129–140. https://doi.org/10.1016/j.scitotenv.2018.11.084

Barthod LRM, Liu K, Lobb DA, et al (2015) Selecting Color-based Tracers and Classifying Sediment Sources in the Assessment of Sediment Dynamics Using Sediment Source Fingerprinting. J Environ Qual 44:1605–1616. https://doi.org/10.2134/jeq2015.01.0043

Batista PVG, Davies J, Silva MLN, Quinton JN (2019a) On the evaluation of soil erosion models: Are we doing enough? Earth Sci Rev 197

Batista PVG, Laceby JP, Evrard O (2022) How to evaluate sediment fingerprinting source apportionments. J Soils Sediments 22:1315–1328. https://doi.org/10.1007/s11368-022-03157-4

Batista PVG, Laceby JP, Silva MLN, et al (2019b) Using pedological knowledge to improve sediment source apportionment in tropical environments. J Soils Sediments 19:3274–3289. https://doi.org/10.1007/s11368-018-2199-5

Ben Slimane A, Raclot D, Evrard O, et al (2016) Relative Contribution of Rill/Interrill and Gully/Channel Erosion to Small Reservoir Siltation in Mediterranean Environments. Land Degrad Dev 27:785–797. https://doi.org/10.1002/ldr.2387

Bispo DFA, Batista PVG, Guimarães DV, et al (2020) Monitoring land use impacts on sediment production: A case study of the pilot catchment from the brazilian program of payment for environmental services. Rev Bras Cienc Solo 44:1–15. https://doi.org/10.36783/18069657rbcs20190167

Blanco-Canqui H, Lal R (2010) Principles of soil conservation and management. Springer Netherlands

Boerjan W, Ralph J, Baucher M (2003) Lignin Biosynthesis. Annu Rev Plant Biol 54:519–546

Borrelli P, Alewell C, Alvarez P, et al (2021) Soil erosion modelling: A global review and statistical analysis. Science of the Total Environment 780

Bouchez J, Gaillardet J, France-Lanord C, et al (2011) Grain size control of river suspended sediment geochemistry: Clues from Amazon River depth profiles. Geochemistry, Geophysics, Geosystems 12:. https://doi.org/10.1029/2010GC003380

Brandt C, Benmansour M, Walz L, et al (2018) Integrating compound-specific δ13C isotopes and fallout radionuclides to retrace land use type-specific net erosion rates in a small tropical catchment exposed to intense land use change. Geoderma 310:53–64. https://doi.org/10.1016/j.geoderma.2017.09.008

Bravo-Linares C, Schuller P, Castillo A, et al (2020) Combining Isotopic Techniques to Assess Historical Sediment Delivery in a Forest Catchment in Central Chile. J Soil Sci Plant Nutr 20:83–94. https://doi.org/10.1007/s42729-019-00103-1

Bravo-Linares C, Schuller P, Castillo A, et al (2018) First use of a compound-specific stable isotope (CSSI) technique to trace sediment transport in upland forest catchments of Chile. Science of the Total Environment 618:1114–1124. https://doi.org/10.1016/j.scitotenv.2017.09.163

Chen F, Fang N, Shi Z (2016) Using biomarkers as fingerprint properties to identify sediment sources in a small catchment. Science of the Total Environment 557–558:123–133. https://doi.org/10.1016/j.scitotenv.2016.03.028

Chikaraishi Y, Naraoka H, Poulson SR (2004) Hydrogen and carbon isotopic fractionations of lipid biosynthesis among terrestrial (C3, C4 and CAM) and aquatic plants. Phytochemistry 65:1369–1381. https://doi.org/10.1016/j.phytochem.2004.03.036

Collins AL, Blackwell M, Boeckx P, et al (2020) Sediment source fingerprinting: benchmarking recent outputs, remaining challenges and emerging themes. J Soils Sediments 20:4160–4193. https://doi.org/10.1007/s11368-020-02755-4

Collins AL, Burak E, Harris P, et al (2019) Field scale temporal and spatial variability of δ13C, δ15N, TC and TN soil properties: Implications for sediment source tracing. Geoderma 333:108–122. https://doi.org/10.1016/j.geoderma.2018.07.019

Collins AL, Walling DE, Leeks GJL (1997) Fingerprinting the origin of fluvial suspended sediment in larger river basins: Combining assessment of spatial provenance and source type. Geografiska Annaler, Series A: Physical Geography 79:239–254. https://doi.org/10.1111/j.0435-3676.1997.00020.x

Collins AL, Walling DE, Sichingabula HM, Leeks GJL (2001) Suspended sediment source fingerprinting in a small tropical catchment and some management implications

Cooper RJ, Pedentchouk N, Hiscock KM, et al (2015) Apportioning sources of organic matter in streambed sediments: An integrated molecular and compound-specific stable isotope approach. Science of the Total Environment 520:187–197. https://doi.org/10.1016/j.scitotenv.2015.03.058

Cui X, Bianchi TS, Hutchings JA, et al (2016) Partitioning of organic carbon among density fractions in surface sediments of Fiordland, New Zealand. J Geophys Res Biogeosci 121:1016–1031. https://doi.org/10.1002/2015JG003225

David R. Montgomery (2012) Dirt: The Erosion of Civilizations. Univ of California press

De Carvalho CCCR, Caramujo MJ (2018) The various roles of fatty acids. Molecules 23

Dotterweich M (2008) The history of soil erosion and fluvial deposits in small catchments of central Europe: Deciphering the long-term interaction between humans and the environment - A review. Geomorphology 101:192–208. https://doi.org/10.1016/j.geomorph.2008.05.023

Dotterweich M (2013) The history of human-induced soil erosion: Geomorphic legacies, early descriptions and research, and the development of soil conservation—A global synopsis. Geomorphology 201:1–34

Dotterweich M, Kühn P, Tolksdorf JF, et al (2013) Late Pleistocene to Early Holocene natural and human influenced sediment dynamics and soil formation in a 0-order catchment in SW-Germany (Palatinate Forest). Quaternary International 306:42–59. https://doi.org/10.1016/j.quaint.2013.04.001

Evrard O, Poulenard J, Némery J, et al (2013) Tracing sediment sources in a tropical highland catchment of central Mexico by using conventional and alternative fingerprinting methods. Hydrol Process 27:911–922. https://doi.org/10.1002/hyp.9421

Feakins SJ, Ellsworth P V., Sternberg L da SL (2013) Lignin methoxyl hydrogen isotope ratios in a coastal ecosystem. Geochim Cosmochim Acta 121:54–66. https://doi.org/10.1016/j.gca.2013.07.012

Fox JF, Martin DK (2015) Sediment Fingerprinting for Calibrating a Soil Erosion and Sediment-Yield Model in Mixed Land-Use Watersheds. J Hydrol Eng 20:. https://doi.org/10.1061/(asce)he.1943-5584.0001011

Gibbs M, Leduc D, Nodder SD, et al (2020) Novel Application of a Compound-Specific Stable Isotope (CSSI) Tracking Technique Demonstrates Connectivity Between Terrestrial and Deep-Sea Ecosystems via Submarine Canyons. Front Mar Sci 7:. https://doi.org/10.3389/fmars.2020.00608

Gibbs M, Swales A, Olsen G (2014) Suess effect on biomarkers used to determine sediment provenance from land-use changes. Proceedings – International Symposium on Managing Soils for Food Security and Climate Change Adaption and Mitigation 371–375

Haddadchi A, Olley J, Laceby P (2014) Accuracy of mixing models in predicting sediment source contributions. Science of the Total Environment 497–498:139–152. https://doi.org/10.1016/j.scitotenv.2014.07.105

Hatfield R, Fukushima RS (2005) Can lignin be accurately measured? Crop Science. pp 832–839

Hirave P, Glendell M, Birkholz A, Alewell C (2021) Compound-specific isotope analysis with nested sampling approach detects spatial and temporal variability in the sources of suspended sediments in a Scottish mesoscale catchment. Science of the Total Environment 755:. https://doi.org/10.1016/j.scitotenv.2020.142916

Kemper JT, Rathburn SL, Friedman JM, et al (2022) Fingerprinting historical tributary contributions to floodplain sediment using bulk geochemistry. Catena (Amst) 214:. https://doi.org/10.1016/j.catena.2022.106231

Keppler F, Hamilton J (2008) Tracing the geographical origin of early potato tubers using stable hydrogen isotope ratios of methoxyl groups. In: Isotopes in Environmental and Health Studies. pp 337–347

Keppler F, Harper DB, Kalin RM, et al (2007) Stable hydrogen isotope ratios of lignin methoxyl groups as a paleoclimate proxy and constraint of the geographical origin of wood. New Phytologist 176:600–609. https://doi.org/10.1111/j.1469-8137.2007.02213.x

Keppler F, Kalin RM, Harper DB, et al (2004) Carbon isotope anomaly in the major plant C1 pool and its global biogeochemical implications. European Geosciences Union 1:123–131

Ko Ègel-Knabner I (2002) The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter

Lake NF, Martínez-Carreras N, Shaw PJ, Collins AL (2022) High frequency un-mixing of soil samples using a submerged spectrophotometer in a laboratory setting—implications for sediment fingerprinting. J Soils Sediments 22:348–364. https://doi.org/10.1007/s11368-021-03107-6

Lal R (2003) Soil erosion and the global carbon budget. Environ Int 29:437–450. https://doi.org/10.1016/S0160-4120(02)00192-7

Lavrieux M, Birkholz A, Meusburger K, et al (2019) Plants or bacteria? 130 years of mixed imprints in Lake Baldegg sediments (Switzerland), as revealed by compound-specific isotope analysis (CSIA) and biomarker analysis. Biogeosciences 16:2131–2146. https://doi.org/10.5194/bg-16-2131-2019

Le Gall M, Evrard O, Foucher A, et al (2016) Quantifying sediment sources in a lowland agricultural catchment pond using 137Cs activities and radiogenic 87Sr/86Sr ratios. Science of the Total Environment 566–567:968–980. https://doi.org/10.1016/j.scitotenv.2016.05.093

Lee H, Galy V, Feng X, et al (2019) Sustained wood burial in the Bengal Fan over the last 19 My. Proc Natl Acad Sci U S A 116:22518–22525. https://doi.org/10.1073/pnas.1913714116

Lizaga I, Bodé S, Gaspar L, et al (2021) Legacy of historic land cover changes on sediment provenance tracked with isotopic tracers in a Mediterranean agroforestry catchment. J Environ Manage 288:. https://doi.org/10.1016/j.jenvman.2021.112291

Lizaga I, Latorre B, Gaspar L, Navas A (2022) Combined use of geochemistry and compound-specific stable isotopes for sediment fingerprinting and tracing. Science of the Total Environment 832:. https://doi.org/10.1016/j.scitotenv.2022.154834

Lloyd MK, Trembath-Reichert E, Dawson KS, et al (2021) Methoxyl stable isotopic constraints on the origins and limits of coal-bed methane. American Association for the Advancement of Science 374:894–897

Lotter AF, Sturm M, Teranes JL, Wehrli B (1997) Varve formation since 1885 and high-resolution varve analyses in hypertrophic Baldeggersee (Switzerland). Aquat Sci 59:304–325. https://doi.org/10.1007/BF02522361

Lu Q, Jia L, Awasthi MK, et al (2022) Variations in lignin monomer contents and stable hydrogen isotope ratios in methoxy groups during the biodegradation of garden biomass. Sci Rep 12:. https://doi.org/10.1038/s41598-022-12689-1

Lutz SR, Van Breukelen BM (2014a) Combined source apportionment and degradation quantification of organic pollutants with CSIA: 1. Model derivation. Environ Sci Technol 48:6220–6228. https://doi.org/10.1021/es405400w

Lutz SR, Van Breukelen BM (2014b) Combined Source Apportionment and Degradation Quantification of Organic Pollutants with CSIA: 2. Model Validation and Application. Environ Sci Technol 48:6220–6228. https://doi.org/10.1021/es405400w

Lützow M V., Kögel-Knabner I, Ekschmitt K, et al (2006) Stabilization of organic matter in temperate soils: Mechanisms and their relevance under different soil conditions - A review. Eur J Soil Sci 57:426–445

Martínez-Carreras N, Krein A, Gallart F, et al (2010) Assessment of different colour parameters for discriminating potential suspended sediment sources and provenance: A multi-scale study in Luxembourg. Geomorphology 118:118–129. https://doi.org/10.1016/j.geomorph.2009.12.013

Mukundan R, Walling DE, Gellis AC, et al (2012) Sediment Source Fingerprinting: Transforming From a Research Tool to a Management Tool. J Am Water Resour Assoc 48:1241–1257. https://doi.org/10.1111/j.1752-1688.2012.00685.x

Nosrati K, Govers G, Ahmadi H, et al (2011) An exploratory study on the use of enzyme activities as sediment tracers: biochemical fingerprints?

O'leary WM (1962) THE FATTY ACIDS OF BACTERIA. Bacteriol Rev 26:421–447

Olley J, Saxton NE, Pietsch T, et al (2009) Healthy Country Project Phase 1 Report: Monitoring and evaluation of restoration activities in three SEQ focal catchments

Owens PN (2022) Sediment source fingerprinting: are we going in the right direction? J Soils Sediments 22:1643–1647

Panagos P, Matthews F, Patault E, et al (2024) Understanding the cost of soil erosion: An assessment of the sediment removal costs from the reservoirs of the European Union. J Clean Prod 434:140183. https://doi.org/10.1016/j.jclepro.2023.140183

Panagos P, Standardi G, Borrelli P, et al (2018) Cost of agricultural productivity loss due to soil erosion in the European Union: From direct cost evaluation approaches to the use of macroeconomic models. Land Degrad Dev 29:471–484. https://doi.org/10.1002/ldr.2879

Pulley S, Collins AL (2018) Tracing catchment fine sediment sources using the new SIFT (SedIment Fingerprinting Tool) open source software. Science of the Total Environment 635:838–858. https://doi.org/10.1016/j.scitotenv.2018.04.126

Pulley S, Foster I, Antunes P (2015) The application of sediment fingerprinting to floodplain and lake sediment cores: assumptions and uncertainties evaluated through case studies in the Nene Basin, UK. J Soils Sediments 15:2132–2154. https://doi.org/10.1007/s11368-015-1136-0

Stewart HA, Massoudieh A, Gellis A (2015) Sediment source apportionment in Laurel Hill Creek, PA, using Bayesian chemical mass balance and isotope fingerprinting. Hydrol Process 29:2545–2560. https://doi.org/10.1002/hyp.10364

Stock BC, Jackson AL, Ward EJ, et al (2018) Analyzing mixing systems using a new generation of Bayesian tracer mixing models. PeerJ 2018:. https://doi.org/10.7717/peerj.5096

Swales A, Gibbs MM (2020) Transition in the isotopic signatures of fatty-acid soil biomarkers under changing land use: Insights from a multi-decadal chronosequence. Science of the Total Environment 722:. https://doi.org/10.1016/j.scitotenv.2020.137850

Upadhayay HR, Bodé S, Griepentrog M, et al (2018a) Isotope mixing models require individual isotopic tracer content for correct quantification of sediment source contributions. Hydrol Process 32:981–989. https://doi.org/10.1002/hyp.11467

Upadhayay HR, Griepentrog M, Bodé S, et al (2020a) Catchment-wide variations and biogeochemical time lags in soil fatty acid carbon isotope composition for different land uses: Implications for sediment source classification. Org Geochem 146:. https://doi.org/10.1016/j.orggeochem.2020.104048

Upadhayay HR, Lamichhane S, Bajracharya RM, et al (2020b) Sensitivity of source apportionment predicted by a Bayesian tracer mixing model to the inclusion of a sediment connectivity index as an informative prior: Illustration using the Kharka catchment (Nepal). Science of the Total Environment 713:. https://doi.org/10.1016/j.scitotenv.2020.136703

Upadhayay HR, Smith HG, Griepentrog M, et al (2018b) Community managed forests dominate the catchment sediment cascade in the mid-hills of Nepal: A compound-specific stable isotope analysis. Science of the Total Environment 637–638:306–317. https://doi.org/10.1016/j.scitotenv.2018.04.394

Vale S, Swales A, Smith HG, et al (2022) Impacts of tracer type, tracer selection, and source dominance on source apportionment with sediment fingerprinting. Science of the Total Environment 831:. https://doi.org/10.1016/j.scitotenv.2022.154832

Vale SS, Fuller IC, Procter JN, et al (2020) Storm event sediment fingerprinting for temporal and spatial sediment source tracing. Hydrol Process 34:3370–3386. https://doi.org/10.1002/hyp.13801

Verburg P (2007) The need to correct for the Suess effect in the application of δ13C in sediment of autotrophic Lake Tanganyika, as a productivity proxy in the Anthropocene. J Paleolimnol 37:591–602. https://doi.org/10.1007/s10933-006-9056-z

Wieland A, Römer P, Torbenson M, et al (2024) Tree-ring stable isotopes in cellulose and lignin methoxy groups reveal different age-related behaviour. Quaternary International. https://doi.org/10.1016/j.quaint.2024.02.004

Wiesenberg GLB, Gocke M, Kuzyakov Y (2010) Fast incorporation of root-derived lipids and fatty acids into soil - Evidence from a short term multiple 14CO2 pulse labelling experiment. Org Geochem 41:1049–1055. https://doi.org/10.1016/j.orggeochem.2009.12.007

Wiltshire C, Glendell M, Waine TW, et al (2022) Assessing the source and delivery processes of organic carbon within a mixed land use catchment using a combined n-alkane and carbon loss modelling approach. J Soils Sediments 22:1629–1642. https://doi.org/10.1007/s11368-022-03197-w

Wiltshire C, Waine TW, Grabowski RC, et al (2023) Assessing n-alkane and neutral lipid biomarkers as tracers for land-use specific sediment sources. Geoderma 433:. https://doi.org/10.1016/j.geoderma.2023.116445

Chapter 2

Less is more? A novel method for identifying and evaluating non-informative tracers in sediment source mixing models

Journal of Soils and Sediments (2023) 23:3241–3261

https://doi.org/10.1007/s11368-023-03573-0

Received: 29 November 2022 / Accepted: 1 June 2023 / Published online: 10 July 2023

Less is more? A novel method for identifying and evaluating non-informative tracers in sediment source mixing models

****

Terry Cox[1], J. Patrick Laceby[2], Till Roth[1], Christine Alewell[1]

Corresponding authors email: terry.cox@unibas.ch

[1]Department of Environmental Sciences, University of Basel, Bernoullistrasse 30, 4056 Basel, Switzerland

[2] Environment and Protected Areas, Government of Alberta, 3535 Research Rd NW, Calgary, AB, Canada, T2L 1Y1

****

Abstract

Purpose

Accelerated soil erosion poses a global hazard to soil health. Understanding soil and sediment behaviour through sediment fingerprinting enables the monitoring and identification of areas with high sediment delivery. Land-use specific sediment source apportionment is increasingly determined using the Bayesian mixing model MixSIAR with compound-specific stable isotopes (CSSI). Here, we investigate CSSIs of fatty acid (FA) tracer selection with a novel method to identify and investigate the effect of non-informative tracers on model performance.

Methods

To evaluate CSSI tracer selection, mathematical mixtures were generated using source soils (n=28) from the Rhine catchment upstream of Basel (Switzerland). Using the continuous ranked probability (CRP) skill score, MixSIAR's performance was evaluated for 11 combinations of FAs and 15 combinations of FAs with $\delta^{15}N$ as a mixing line offset tracer. A novel scaling and discrimination analysis (SDA) was also developed to identify tracers with non-unique mixing spaces.

Results

FA only tracer combinations overestimated pasture contributions while underestimating arable contributions. When compared to models with only FA tracers, utilizing $\delta^{15}N$ to offset the mixing line resulted in a 28% improvement in the CRP skill score. $\delta^{15}N$ + $\delta^{13}C$ FA_{26} was the optimal tracer set resulting in a 62% model improvement relative to $\delta^{15}N$ + all $\delta^{13}C$ FAs. The novel SDA method demonstrated how $\delta^{13}C$ FA tracers have a non-unique mixing space and thus behave as non-informative tracers. Importantly, the inclusion of non-informative tracers decreased model performance.

Conclusions

These results indicate that MixSIAR did not handle non-informative CSSI tracers effectively. Accordingly, it may be advantageous to remove non-informative tracers, and where feasible, all combinations and permutations of tracers should be assessed to optimize tracer selection. Application of these tracer selection steps can help improve and advance the performance of sediment fingerprinting models and ultimately aid in improving erosion mitigation and management strategies.

Key Words: Sediment Tracing, Sediment Source Fingerprinting, Tracer selection, MixSIAR, CSSI, Sediment apportionment, Stable isotopes

****

2.1 Introduction

Accelerated soil erosion and sedimentation is a widely recognized global problem that reduces water quality and agricultural output (Bakker et al. 2004; Issaka and Ashraf 2017). Comprehensive and economically feasible mitigation plans are required to reduce accelerated soil erosion. Therefore, effective mitigation plans need to be founded on the accurate identification of erosion sources (Collins and Walling 2004; Walling 2005; Owens et al. 2016).

Sediment source fingerprinting helps identify and apportion the main erosion sources in a catchment. Tracing sediments back to their original sources (e.g., from soils with different land uses), provides a direct, field-based approach that holds the potential to identify the relative contribution of different soil erosion sources to sediment transported downstream in various waterways (Collins et al. 1996; Gibbs 2008; Cooper et al. 2015). The technique uses various properties of the soils and sediments as fingerprints to differentiate the main erosion sources (Collins et al. 1997a, 2001; Walling 2013; Smith et al. 2018). For properties to be able to effectively fingerprint sediment sources, they need to discriminate between sediment sources and remain constant through detachment, transport and deposition processes, or vary in a measurable and predictable way (Motha et al. 2002; Koiter et al. 2013; Belmont et al. 2014; García-Comendador et al. 2023). Essentially, properties of the eroded sediment should remain constant or any variation during these processes must be reproducible.

Sediment fingerprinting has been applied to a wide range of fluvial sediment types including: lacustrine sediment cores (le Gall et al. 2016; Lavrieux et al. 2019), flood plains (Pulley et al. 2015; Kemper et al. 2022), dam reservoir samples (Nosrati et al. 2011; Ben Slimane et al. 2013), and riverine systems (Collins et al. 2001; Bispo et al. 2020; Upadhayay et al. 2018b). A wide range of parameters have been employed to trace sediment sources, including but not limited to radionuclides (Evrard et al. 2013a), elemental geochemistry (Batista et al. 2019), compound-specific stable isotopes (Hirave et al. 2021), colour (Martínez-Carreras et al. 2010), diffuse Reflectance Infrared Fourier Transform Spectroscopy (Evrard et al. 2013b), ultraviolet-visible absorbance (Lake et al. 2022), and eDNA (Evrard et al. 2019)

among others. For more information on the sediment fingerprinting technique, please consult some of the reviews in the literature (Haddadchi et al. 2013; Koiter et al. 2013; Owens et al. 2016; Collins et al. 2017, 2020).

Sediment source apportionment is generally determined by unmixing sediment and soil fingerprints typically using linear equations (Collins et al. 1997b). The reliability of the model outputs is dependent on the mixing model used (Batista et al. 2022), the number of sources, the number of tracers and the dominate source contributing to the sediment load (Vale et al. 2022). Specifically, the accuracy of the apportionment increases when the primary source is well discriminated regardless of the discrimination of the other sources (Vale et al. 2022). Bayesian (e.g., MixSIAR), (Stewart et al., 2015 Stock et al., 2018) and frequentist models (Collins et al. 1997a; Pulley and Collins 2018) have the potential to provide robust and reliable erosion source information fundamental to targeting sediment management practices (Cooper and Krueger 2017; Evrard et al. 2022; Xu et al. 2022).

The compound-specific stable isotopes (CSSI) of plant-derived biomarkers such as fatty acids (FA) and n-alkanes have been used to apportion sediment source contributions from different land uses (Gibbs 2008; Alewell et al. 2016; Upadhayay et al. 2018b; Lavrieux et al. 2019; Hirave et al. 2021). The CO_2 fixation routes (C_3, C_4, or CAM) of plants generate distinct isotopic patterns with the effect being more pronounced in C_4 plants than in C_3 plants (Reiffarth et al. 2016). Even though they are not species specific, CSSI isotopic values can further distinguish between some plant groups, for example, angiosperms and gymnosperms (Collister et al. 1994; Chikaraishi et al. 2004). Additionally, biological and environmental factors (e.g., altitude and rainfall patterns) can induce variation in the isotopic value (Reiffarth et al. 2016).

The exclusion of short, medium, and non-saturated FAs helps reduce the uncertainty related to input from non-terrestrial plant-derived FAs (Alewell et al. 2016; Reiffarth et al. 2016; Lavrieux et al. 2019). Ultimately, the ^{13}C FA fingerprint of the sediment mixture is determined by source mixing proportions and two parameters in each source: FA concentration and $\delta^{13}C$ FA (Upadhayay et al. 2018a). Therefore, the non-linear mixing of the isotopic tracers in the mathematical mixtures requires the incorporation of FA concentration dependency. Additionally, the use of FA isotopes as tracers requires the transformation of unmixing isotopic values to the unmixing sediment. The use of the concentration dependency of FA isotopes incorporates this transformation into the model and therefore requires no additional post organic matter corrections (Alewell et al. 2016).

Fingerprinting with FA CSSIs has limitations, where source values regularly plot along a linear mixing line (Alewell et al. 2016; Lavrieux et al. 2019). Importantly, having sources plot along a mixing line can result in modelled contributions from the central source(s) being misclassified as a contribution from the sources located at the mixing line endpoints. The misclassification has previously been resolved by grouping sources with the subsequent apportionment occurring only between the two grouped sources (Alewell et al. 2016), with the drawback of not being able to distinguish between three or more sources.

The highly correlated $\delta^{13}C$ of FA tracers and the resultant linear mixing line are a product of the biosynthesis of very long chain fatty acids (VLCFA, $FA_{22:0}$ - $FA_{30:0}$). The elongation of long chain fatty

acids (LCFA, $FA_{16:0}$- $FA_{20:0}$) to VLCFA proceeds with a cyclic four-step process of condensation, reduction, dehydration, and reduction (Erdbrügger and Fröhlich 2020). The elongation of FAs occurs with a $\delta^{13}C$ depletion with increasing alkyl length (Chikaraishi et al. 2004). As this relationship is assumed to be similar for VLCFA of different plant groups, this may result in the $\delta^{13}C$ FA tracers of different alkyl lengths having a non-unique mixing space and possibly acting as non-informative tracers.

CSSI tracers have been combined with geochemical tracers in an attempt to improve the discrimination between different land covers using non-land-use specific tracers (Lizaga et al. 2022). As geochemical tracers are not land-use specific, they require significant geological differences between land uses and low variability of lithology within land uses (Blake et al., 2012; Hancock and Revill 2013). $\delta^{15}N$ has been used previously as a tracer for land-use-specific sediment source apportionment (Papanicolaou 2003; Fox and Papanicolaou 2007; Mukundan et al., 2010). However, the conservativeness of $\delta^{15}N$ is questionable (Laceby et al. 2017). Here, we nonetheless use $\delta^{15}N$ to expand the $\delta^{13}C$ FA mixing line to a mixing polygon for mathematical mixtures (also known as: virtual mixtures and artificial mixtures). When investigating model behaviour using mathematical mixtures, the conservativeness of tracers is less relevant as sediment tracer values are calculated from source soil values and are not subject to degradation and possible isotopic fractionation effects.

Mathematical mixtures using non-concentration-dependent tracers (e.g., geochemical tracers) were reported to be equivalent to laboratory mixtures (Batista et al. 2022). Although mathematical mixtures do not fully represent what happens during mixing processes in nature (e.g., signal degradation, tracer alteration in case of non-conservativeness, isotope fractionation, particle size transport selectivity), they are fundamental to understanding and evaluating model performance, and characterizing the uncertainty of the unmixing process (Haddadchi et al. 2014; Batista et al. 2019; Vale et al. 2022).

Currently, there is a limited application of mathematical mixtures to concentration-dependent tracers, in which tracer values of the mixture (e.g., isotopic signatures) are dependent on another parameter in source soils (e.g., the concentration of isotopic tracer). Until recently the validation of concentration-dependent unmixing models has been reliant on the generation of a small number of time-consuming laboratory mixtures (Bravo-Linares et al. 2018) or the over-simplification by the removal of the concentration dependency by assuming that isotopic tracers mix linearly (Collins et al. 2019; Bahadori et al. 2019). However, recently concentration-dependent mathematical mixtures have been explored and utilised (Lizaga et al. 2022; Vale et al. 2022) for investigating model parameters.

The deficiency of suitable evaluation tools and metrics for CSSI tracer selection steps has resulted in the legacy of two commonly used assessments: a Kruskal Wallis test to optimize model performance and a polygon/boxplot range test to identify non-conservative tracers. When using a large number of tracers (e.g., geochemical tracers), linear discrimination analysis (LDA) has also been applied to reduce the suite of tracers to an optimal number with maximum discrimination (Gellis and Noe 2013; Walling 2013; Laceby et al. 2015). The LDA tracer reduction step is not commonly included when using CSSIs due to the limited number of CSSI tracers relative to the number of sources being discriminated. Upadhayay et al. (2018b) used LDA with CSSI tracers to remove bulk ^{13}C from the VLCFA tracer suite

as bulk ^{13}C did not improve source LDA reclassification. Lizaga et al. (2021) also used LDA as a tracer selection step for mixtures from different time points in an attempt to optimise tracer selection for each mixture. An argument for not including the LDA when using MixSIAR is the hypothesis that the covariance structure of MixSIAR (Stock et al. 2018) effectively handles conservative non-informative tracers resulting in a null or a beneficial impact (Smith et al. 2018). The model's output should accurately reflect the real-world scenarios, without being reduced in the interest of enhancing model performance. An additional argument for maximising the number of tracers is to reduce the potential influence of possible non-conservativeness within the tracer set.

Land use-specific sediment source apportionment with CSSIs has been determined with all tracers that pass the two prerequisites without further validation of tracer selection. Consensus ranking (Lizaga et al. 2020) and consistent tracer selection (Latorre et al. 2021) methods have been recently applied to remove non-conservative tracers and tracers which have non-consistent results. Others have argued that tracer selection should be made on a robust bio-physical-chemical foundation (Laceby et al. 2015; Batista et al. 2019), with adjustments to the tracer set aimed at supporting the reliability and accuracy of the model.

We hypothesize that the relationship between $\delta^{13}C$ depletion and increasing alkyl length is similar for all land uses. If this is true, the mixing space for each FA tracer may be seen as a direct isometry translation of each other (i.e., every point of the mixing shape is moved the same distance and in the same direction), resulting in additional FA tracers having non-unique mixing spaces and being non-informative. This effect may result in the inclusion of additional $\delta^{13}C$ FA tracers being seen as essentially comparable to the addition of non-informative clone tracers (i.e., an exact copy of a tracer included as an additional tracer). In this study, clone tracers are used as a standard example of tracers which have identical mixing spaces and therefore can potentially be seen as non-informative tracers. In particular, clone tracers are used to determine the capacity of the mixing model to handle non-informative tracers. The comparison of model performance using additional FA tracers and non-informative clone tracers helps to quantify the information gained by using an additional FA tracer.

Further evaluation and optimization of CSSI tracer selection in sediment source fingerprinting research is critical to increase the reliability of apportionment estimates and as a result the development of appropriate sediment management practices. In this study, we present the results for all combinations of $\delta^{13}C$ FA (n = 11) and FA tracer sets including $\delta^{15}N$ (n = 15) using concentration-dependent mathematical mixtures. Using a novel scaling and discrimination analysis (SDA), non-informative tracers that have a non-unique mixing space are identified. Importantly, we test the hypothesis that the covariance structure of MixSIAR can adequately handle non-informative tracers using clone tracers.

2.2 Materials and methods

Site description and sampling

The study was conducted using source soils from the Rhine catchment upstream of Basel (10,687 km^2) (draining northern Switzerland and parts of southern Baden-Württemberg, Germany) and downstream of the large lakes (i.e., Konstanz, Zürich, Hallwil, Sempach, and Biel) (Fig. 1). Land use within the catchment area was mainly classified as arable land (28%), mixed forest (20%), and pasture (13%).

The Basel Rhine catchment was divided into four sub-catchments: The Birs catchment, the Aare catchment, a Rhine catchment downstream of the Aare entering the Rhine, and a Rhine catchment upstream of the Aare entering the Rhine (Fig. 1). Each sub-catchment contained 3-8 sites of the major land-use classes: arable, pasture, and forest. With the aid of a connectivity model by Borselli et al., (2008), land-use specific sample locations within each sub-catchment were selected based on their high likelihood to contribute suspended sediment to the watercourses. To reduce analytical costs while maintaining the representativeness of the source samples, composite samples were generated from four individual samples located 2 m apart using a soil extraction cylinder (diameter 5.5 cm, length 5 cm). As suggested by Laceby et al. (2017) and Evrard et al. (2022), the size fraction of source soils analysed (<100 µm), was selected based on particle size analysis of flood sediment from wider research project. Information on the sediment collection and size analysis is included in Online Resource 1.

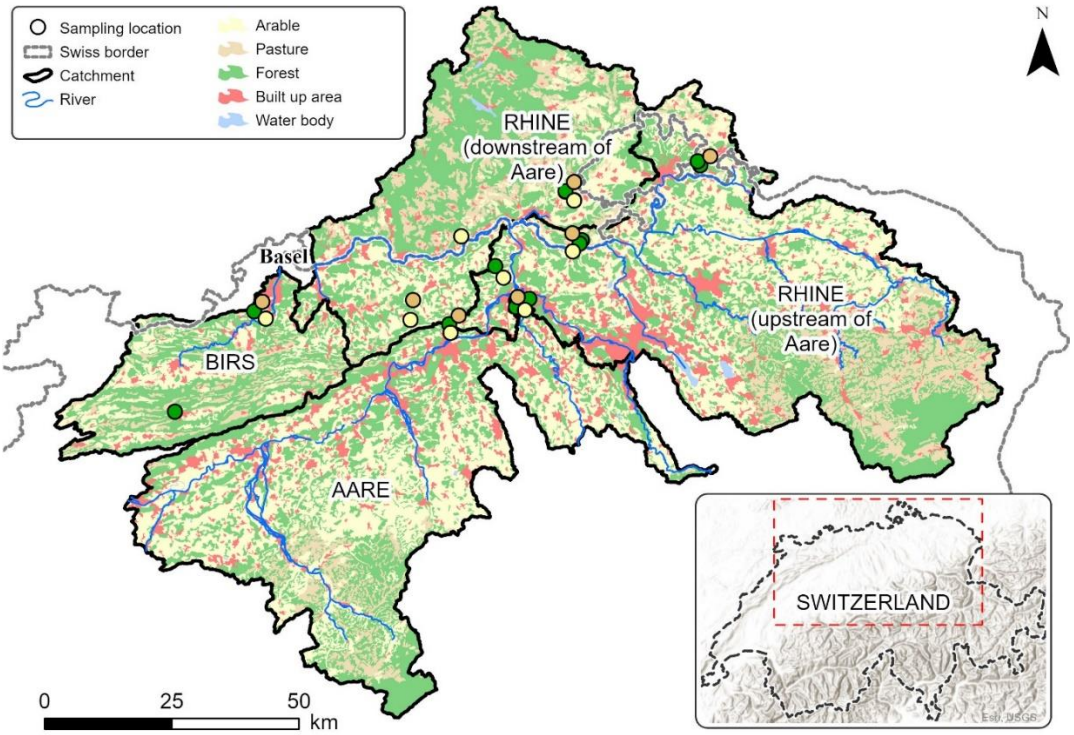


Fig. 1 Land use map of the Basel Rhine catchment area considered showing sampling locations and sub-catchment regions: The Birs catchment, the Aare catchment, a Rhine catchment downstream of the Aare entering the Rhine, and a Rhine catchment upstream of the Aare entering the Rhine

Laboratory Analysis

Lipids were extracted from 0.5-1.5 g of soil using CH_2Cl_2: MeOH (9: 1 *v/v*) in an Accelerated Solvent Extractor (Dionex ASE 350) with the addition of $FA_{19:0}$ as an internal standard. The total lipid extract was separated into polar, neutral, and acidic fractions using solid-phase extraction on aminopropyl bonded silica as described in Jacob et al. (2005). The acidic fractions were eluted using 1% formic acid in diethyl ether on a pre-acidified column. The acidic fraction was subsequently methylated at 60°C for 1 h using 1 ml of 14 % BF_3 in MeOH. Fatty acid methyl esters were extracted from the solution by agitating it four times with 2 mL hexane in the presence of 1 mL of 0.1 M KCl. The $\delta^{13}C$ FA isotopic ratio was measured using a Trace 1310 GC instrument interfaced online through a GC-Isolink II to a Conflo IV and Delta V Plus isotope ratio mass spectrometer (Thermo Fisher Scientific) as described by Lavrieux et al. (2019). Nitrogen concentrations and $\delta^{15}N$ values for source soils were measured by Flash EA (Thermo Finnigan Delta plus XP mass spectrometer coupled with Flash EA 1112 series elemental analyser supplied by Thermo Finnigan, Waltham, MA, USA). Carbon and nitrogen stable isotope ratios were reported in delta notation, per mil deviation from Vienna Pee Dee Belemnite (VPDB) and atmospheric nitrogen (AIR) respectively.

Mathematical mixtures

Mathematical mixtures were generated using the mean stable isotopic ratio and mean concentration of tracers (i.e., bulk N, $FA_{(24, 26, 28, 30)}$) for arable (n = 10), pasture (n = 7), and forest (n = 11) soil samples. Proportions of source contributions were created using a random number generator sampling from a Dirichlet distribution between 0 and 100 with the condition that the sum of source proportions must equal 1. The python script used to generate mathematical mixtures is appended in the Online Resource 2 (an excel version of the mathematical mixture formulation is appended as Online Resource 3). To ensure evenly distributed mixing proportions of each source, 150 mathematical mixtures were generated (the mean of each source proportion of 150 mixtures was ~33%). Concentration-dependent mathematical mixtures were generated as shown in Eq.(1).

$$\frac{\sum_{s}^{so}(C_{s,t} \times P_{s,t} \times V_{s,t})}{\sum_{s}^{so}(C_{s,t} \times P_{s,t})} = V_t \; \forall \; t \in T \qquad \text{Eq.1}$$

Where V is the mean isotopic value of the tracer t, C refers to the mean concentration for all (∀) tracers in a set (∈) of tracers T in source S. S^O refers to the number of sources and P refers to the known proportions of the mathematical mixtures.

End Member Mixing Model

Mathematical mixtures were modelled using the open-source MixSIAR R package (Stock et al. 2018). MixSIAR was run with concentration dependency utilising the concentration of both FAs and N, transforming the unmixing of isotopes to the unmixing sediment/mixtures. Therefore an organic matter correction was not applied post hoc to prevent a secondary transformation. Priors were set to

uninformative and all MixSIAR runs used the same model parameters: chains = 3, chain length= 3,000,000, thin = 500, burn = 2,700,000 with a 'very long' run time. The convergence of the mixing model was assessed by using the Gelman-Rubin diagnostic, with model output being rejected if variables scored >1.0. The R script used for all models is appended in the Online Resource 4.

While the 'residual x process' error structure has been applied appropriately to multiple mixture samples (Cooper et al. 2015; Smith et al. 2018; Upadhayay et al. 2018b; Blake et al. 2018; Vale et al. 2022), likely the high cost and processing time of the analysis of CSSI has led to the predominant use of 'process only' error structure using single mixture samples (Gateuille et al. 2019; Reiffarth et al. 2019; Liu and Han 2021). As such, a single sample of each mixture proportion was unmixed in MixSIAR using the "process only" error structure, in which the variation in the mixtures is assumed to be fully dependent on the variation in the sources (Smith et al. 2018).

Model Evaluation

The probabilistic output of MixSIAR should be evaluated using probabilistic metrics rather than deterministic metrics such as mean absolute error (Batista et al. 2022). The continuously ranked probability score (CRPS) (Matheson and Winkler 1976) is a generalization of the mean absolute error toward a probabilistic perspective and can be thought of as the total displacement needed to move the output distribution density to the observed single outcome or known mixture proportion. CRPS is negatively orientated with smaller values equating to better model performance. A perfect score of 0 represents the entire density of the output placed exactly on the outcome value. Deviation from the perfect score results from a lower density of probability around the observed value. CRPS has provided a useful metric to account for both accuracy and precision of unmixing models and has been suggested to be particularly applicable for model comparison and tracer selection analysis (Batista et al. 2022). CRPS was calculated using the python package 'properscoring' and is used to report on individual model performance.

The CRPS of all tracer combinations are further used to evaluate the accuracy of using LDA for tracer selection. Using the R package 'KlaR', the model performance of the tracer section by a stepwise forward variable selection using the Wilk's Lambda criterion (niveau =0.1) is compared to the empirically selected optimal tracer combination with the lowest CRPS.

Model comparisons are then evaluated by the continuously ranked probability skill score (CRP skill score) shown in Eq. (2) (Pedro et al. 2018). The CRP skill score is a comparative metric of the accuracy and precision between two mode outputs. Where $CRPS_m$ and $CRPS_{ref}$ are the CRPS of the new model (the model compared, e.g., $\delta^{15}N+ \delta^{13}C$ all FA's) and the reference model (the model compared against, e.g., only $\delta^{13}C$ all FA's.) respectively.

$$\text{CRP Skill score} = 1 - (CRPS_m / CRPS_{ref}) \quad \text{Eq.2}$$

Negative CRP skill score values indicate the new model does not outperform the reference model as the newer model requires more displacement of output distribution density to be shifted onto the known

value than the reference model. A value of one indicates that the newer model has a perfect skill score compared to the reference model (Pedro et al. 2018).

Tracer selection and prediction bias analysis (PBA)

The unmixing performance of ideal tracers should be independent of the source contribution, as contribution-dependent model performance is an indication of prediction bias. Predictive bias and dominate source effects on model output have been previously recognised and suggested to be an effect of poor tracer source discrimination (Vale et al. 2022). The hypothesis that FA tracers have similar and non-informative mixing spaces infers that additional FA tracers provide minimal additional source discrimination information. If the hypothesis is true, predictive bias will be decreased by reducing the number of tracers with a non-unique mixing spaces (e.g., FA tracers) as any added source uncertainty is removed. To assess if apportionment estimates occur with predictive bias, known source proportions of mathematical mixtures are plotted against the model performance (CRPS) for each source (predictive bias analysis-PBA). Two tracer sets ($\delta^{15}N$ + all $\delta^{13}C$ FAs and $\delta^{15}N$ + $\delta^{13}C$ FA_{26}) were used to illustrate the effect of reducing the number of tracers on predictive bias.

Non-informative tracers – Scaling and discrimination Analysis (SDA)

MixSIAR uses the relative source-sediment-source positions for un-mixing. Therefore, tracers that exhibit differences in their mixing space by only direct isometry translation, can be seen mathematically, as being identical and potentially non-informative. To evaluate if $\delta^{13}C$ FA tracers have non-unique mixing spaces and are direct isometry translations of each other, a novel scaling and discrimination analysis (SDA) was developed. Scikit-learn's MinMaxScaler package (Pedregosa et al. 2011) was used to scale tracer values between 0 and 1 across all sources. Scaling retains the relative location, shape, and distribution of the sources for each tracer, enabling comparison of relative source locations between tracers. A Kruskal Wallis H-test ($p<0.05$) was used to evaluate the similarity between the relative source locations of the scaled tracers. Scaled tracers depicting a non-significant difference in source locations will consequently have mixing spaces which are direct isometry translations and can therefore be seen as non-informative.

Non-informative tracers – Clone Tracer Analysis

To evaluate MixSIAR's effectiveness to model tracers with non-unique mixing spaces, $\delta^{13}C$ FA_{26} was utilized as a non-informative clone tracer (an exact copy of a tracer used as an additional tracer). This clone tracer was then added three times to the $\delta^{15}N$+ $\delta^{13}C$ FA_{26} tracer set. Each addition of FA_{26} was evaluated individually by CRPS to identify the effect of additional non-informative tracers. The comparison of model performance using additional FA tracers and non-informative clone tracers is then used to quantify the information gained by using an additional FA tracer.

2.3 Results and Discussion

Source tracer values

To reduce errors associated with input from non-terrestrial plant-derived FAs, only VLCFAs ($FA_{22:0}$-$FA_{30:0}$) (hereby referred to as FAs) were considered (Alewell et al. 2016; Reiffarth et al. 2016; Upadhayay et al. 2017; Lavrieux et al. 2019). Forest sources contained the highest concentration of FAs (mean: 19.4 µg g^{-1}, SD: 5.0 µg g^{-1}) and the most $\delta^{13}C$ enriched isotopic values for all FA tracers (mean $\delta^{13}C$: -33.4 ‰, SD: 1.3 ‰), (Tables 1 and 2, Fig. 2). Pasture sources had the most $\delta^{13}C$ depleted isotopic values for all FAs (mean $\delta^{13}C$: -36.2 ‰, SD: 1.4 ‰) and mid-ranged FA concentrations (mean: 11.9 µg g^{-1}, SD: 1.9 µg g^{-1}). Arable sources contained the lowest concentration of FAs (mean: 7.7 µg g^{-1}, SD: 1.1 µg g^{-1}) and mid-ranged FA isotopic values (mean $\delta^{13}C$: -35.0 ‰, SD: 1.7 ‰). The $\delta^{13}C$ FA values for these land uses are similar to previous findings in fresh biomass (Chikaraishi et al. 2004) and soils from the same land use classification in similar geographic and climate regions (Alewell et al. 2016; Lavrieux et al. 2019; Hirave et al. 2021).

The $\delta^{15}N$ value of soil reflects the isotopic signature of nitrogen inputs, outputs and internal processes of the system (Amundson et al. 2003). $\delta^{15}N$ values ranged from a mean of 6.3 ‰ (SD 0.9 ‰) in arable land to 4.0 ‰ (SD 0.9 ‰) in pasture and 0.0 ‰ (SD 1.6 ‰) in forest soil. Nitrogen concentrations ranged from a mean 0.5 mg g^{-1} (SD 0.1 mg g^{-1}) in pasture to a 0.4 mg g^{-1} (SD 0.1 mg g^{-1}) in forest and 0.3 mg g^{-1} (SD 0.1 mg g^{-1}) in arable soil (Tables 1 and 2, Fig. 2). The $\delta^{15}N$ values are comparable to previous results of similar land uses (Fox and Papanicolaou, 2008; Mukundan et al. 2010). Source tracer distribution are similar to those reported by (Fox and Papanicolaou, 2008; Mukundan et al., 2010; Alewell et al. 2016; Lavrieux et al. 2019; Hirave et al. 2021). As such, we found the samples representative of their land use classification and therefore are suitable for the mathematical mixture analysis of this study. However, we suggest that further source soil sampling should be done for the reliable unmixing of real suspended sediment. The full source value data set is appended in the Online Resource 5.

Table 1 Summary of the isotopic tracer values for each land-use ($\delta^{15}N$ ‰, $\delta^{13}C$ ‰ FA) (see the supplementary data set for the full dataset)

Land-use	$\delta^{15}N$ ‰		$\delta^{13}C$ ‰ FA_{24}		$\delta^{13}C$ ‰ FA_{26}		$\delta^{13}C$ ‰ FA_{28}		$\delta^{13}C$ ‰ FA_{30}	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Forest	0.0	1.6	-31.9	1.2	-32.9	0.4	-33.9	0.4	-35.0	1.2
Arable	6.3	0.9	-32.7	1.0	-34.7	0.7	-35.8	0.7	-36.6	0.8
Pasture	4.0	0.9	-34.5	0.6	-35.8	0.4	-36.8	0.3	-37.7	0.5

Table 2 Summary of tracer concentrations for each land-use (see supplementary for full dataset).

Land-use	N mg g^{-1}		FA_{24} µg g^{-1}		FA_{26} µg g^{-1}		FA_{28} µg g^{-1}		FA_{30} µg g^{-1}	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Forest	0.4	0.1	19.6	11.2	14.7	5.9	26.2	14.9	17.1	6.6
Arable	0.3	0.1	6.1	2.0	8.6	3.5	8.1	3.1	8.0	3.7
Pasture	0.5	0.1	9.2	1.9	13.8	3.9	12.4	3.3	12.1	2.9

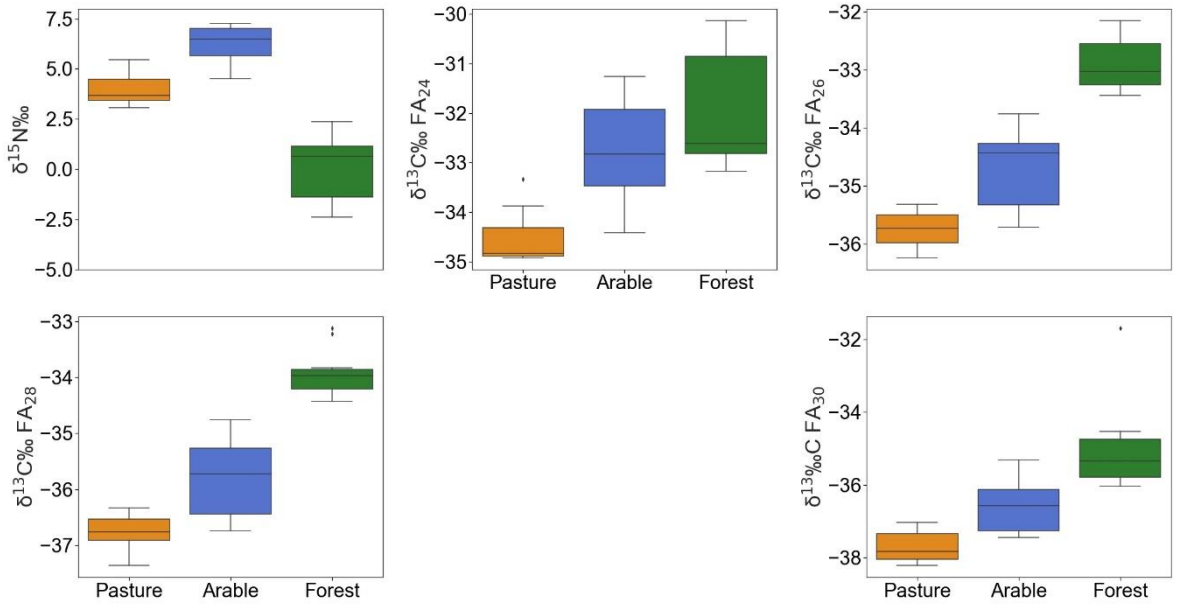


Fig. 2 Isotopic ratios of $\delta^{13}C$ FA and $\delta^{15}N$ tracers for each source group. The boxes represent 25, 50 and 75% quantiles with whiskers showing a 1.5 interquartile range (IQR)

Source discrimination and mixing line origins

The discriminative power of the isotopic tracers between each possible source pair was tested before MixSIAR modelling. 93% of all tracers significantly discriminated between all pairs of sources (Kruskal Wallis, $p<0.05$). Only $\delta^{13}C$ $FA_{24:0}$ did not discriminate between arable and forest sources (Fig. 2). All land uses displayed similar $\delta^{13}C$ depletion with increasing alkyl chain length ($\bar{x}$ = -1.1 ‰ $\delta^{13}C$ per two additional carbon atoms, SD=0.13 ‰).

The results are consistent with the literature that suggests a depletion of up to -2.7 ‰ in C_3 plants from $FA_{24:0}$ to $FA_{32:0}$ (Agrawal et al., 2014 and references within). The small variation of $\delta^{13}C$ enrichment and alkyl chain length (SD=0.13) between all sources suggests the $\delta^{13}C$ enrichment during FA elongation is not land-use dependent. The similar enrichment of the mean $\delta^{13}C$ FAs with increasing alkyl length (forest: $R^2 = 0.999$, arable: $R^2 = 0.952$ and pasture: $R^2 = 0.995$) (Fig. 3) results in a mixing line for all FA tracers (Fig. 4) with the isotopic value of forest and pasture located at either end of the mixing line.

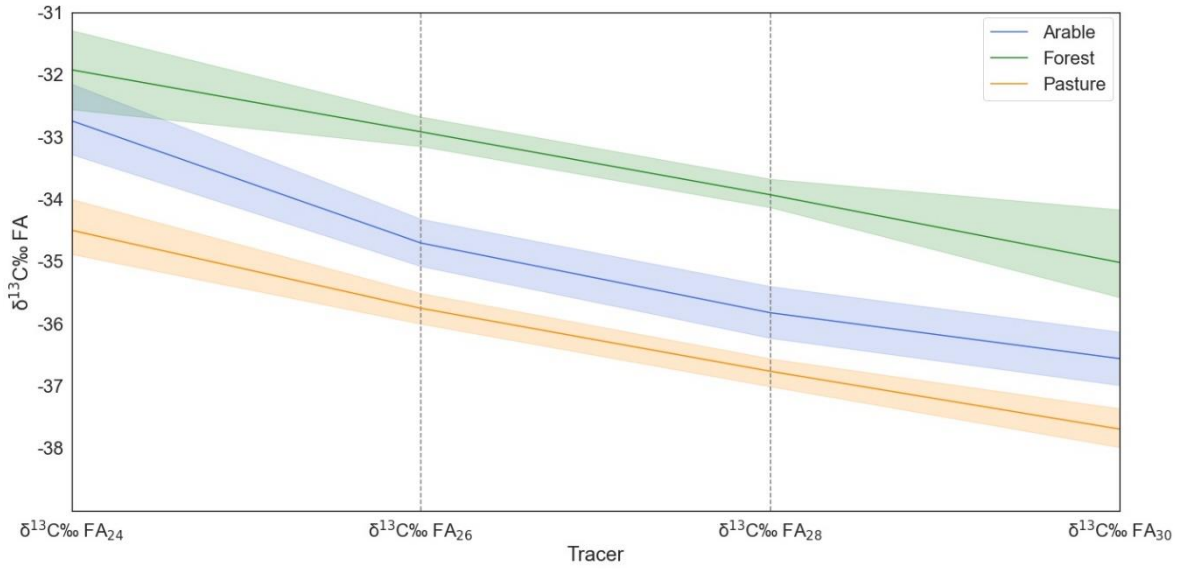


Fig. 3 All land uses displayed a similar $\delta^{13}C$ depletion with increasing alkyl chain length (mean = -1.1 ‰ δ^{13} C per two additional carbon atoms) indicating the $\delta^{13}C$ enrichment during FA elongation is not land-use dependent. Uncertainty is depicted with 95% confidence intervals

The mixing line problem

The FA mixing line illustrated in Fig. 4 is present in $\delta^{13}C$ FA sediment fingerprinting studies with a similar land-use classification of sediment sources (Alewell et al. 2016; Lavrieux et al. 2019). The linear mixing line problem is not confined to isotopic tracers. Colour (Barthod et al. 2015) and geochemical tracers (Bouchez et al. 2011) have also presented a linear mixing line. The similar alkyl length - $\delta^{13}C$ relationship of the different land uses is a result of the same mechanistic process of FA elongation for all land uses. Interestingly, this effect is not observed in all reported cases of arable, pasture and forest land uses (Upadhayay et al. 2020; Lizaga et al. 2021). Deviation from this relationship and the absence of a mixing line could indicate that the previous land use or crop type contained a higher concentration of a specific FA, which is now more present in the current soil compared to other legacy FAs. Conservative tracers, such as FAs, can persist in the soil after a change in land management (Upadhayay et al. 2020). Swales and Gibbs (2020) demonstrated that isotopic shifts occur during a land use transition, and therefore, past land use management should be considered when grouping source soils. This legacy effect has the potential to increases the uncertainty in source distributions and reduce source discrimination and subsequent unmixing performance. However, the legacy effect can potentially be used beneficially for fingerprinting if sources are grouped by their crop cycle rather than the current crop.

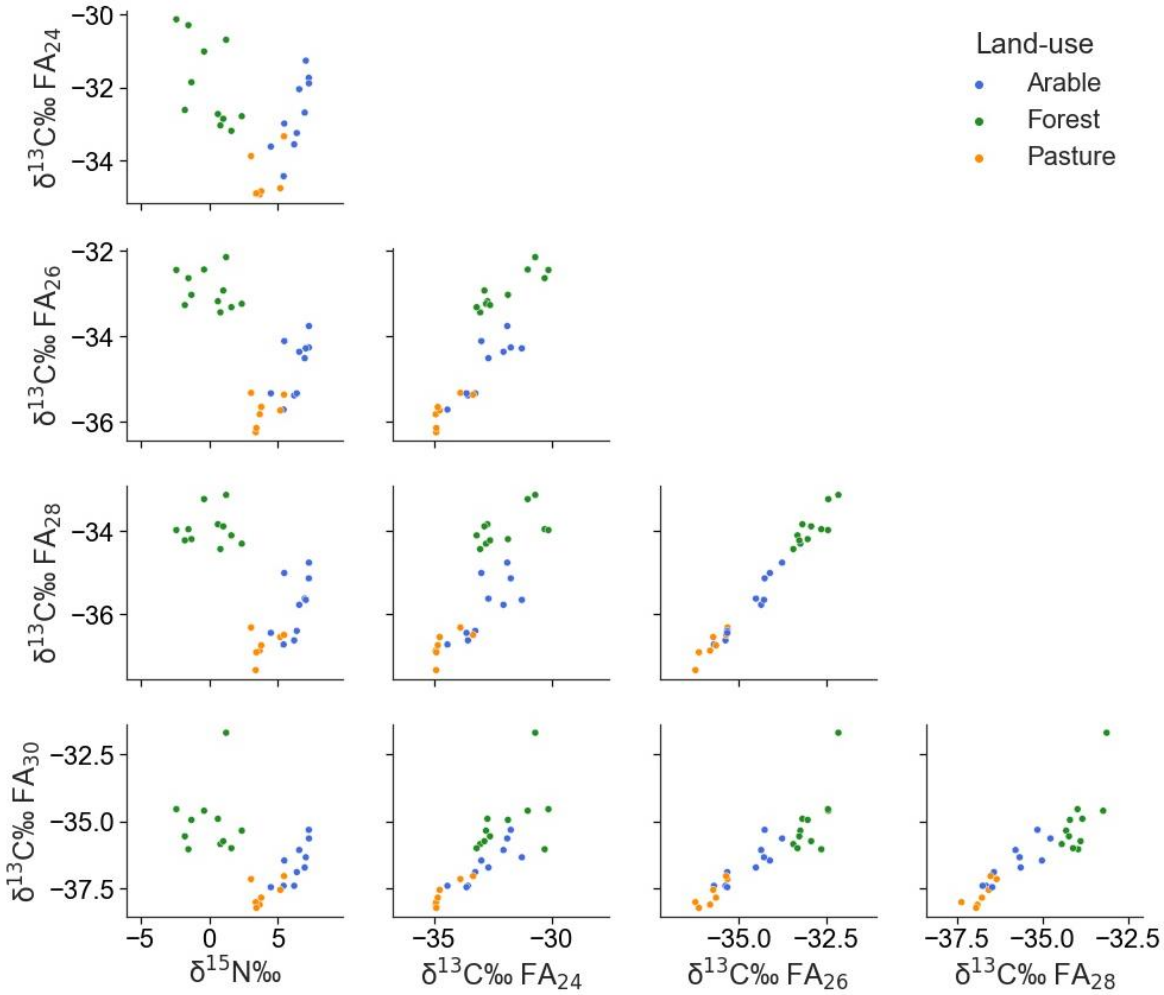


Fig. 4 Iso plots of δ^{15}N and δ^{13}C FA with colours indicating land use type. δ^{13}C FA tracers present the mixing line problem that occurred using FA tracers only. It is unlikely that there will be a perfect 1:1 mixing line when there are multiple samples for each source. Nonetheless, the central location of one source consistently between two other source end members will create challenges (i.e. central source(s) being misclassified as a contribution from endpoint sources) during the modelling process. The addition of δ^{15}N expands the mixing line/space to more of a mixing polygon that provides the source discrimination necessary for more accurate and less uncertain model results

Evaluation of mathematical mixtures

All possible permutations and combinations of δ^{13}C FA tracers (n=11) were evaluated using 150 concentration-dependent mathematical mixtures (Fig. 5, A). Results demonstrated a general increase in CRPS of all sources as the number of δ^{13}C FA tracers is increased (2 FAs mean CRPS: 0.165, 3 FAs: 0.188, 4 FAs: 0.195). A summary of all tracer combination CRPS is appended in the Online Resource 1.

The elevated errors for the arable source contributions (mean CRPS: 0.260) are probably directly related to the location of arable FAs in the mixing space, resulting in the underreporting of arable sources with their contributions likely being misclassified as pasture. Misclassification of the arable source contribution potentially results in the over-estimation of pasture contributions, as suggested by pasture having the second highest CRPS value (mean CRPS: 0.209) for all δ^{13}C FA tracers sets. Strong discrimination of forest sources resulted in a relatively low CRPS value (mean CRPS: 0.058) for all sets of tracers. Forest apportionment estimates were relatively independent of the number of δ^{13}C FA tracers, suggesting that any additional source-based uncertainty induced by additional δ^{13}C FAs was out-weighed by the beneficial source discrimination gained. This collaborates with the iso plots that display low source uncertainty of the forest using any δ^{13}C FA tracer (Fig. 4). Although, our findings differ from those of Vale et al. (2022), who reported that the forest source apportionment had the highest mean absolute error (MAE) among all sources. Both studies demonstrate that the sources with a higher source discrimination have increased model performance. This indicates the ability to discriminate between sources is likely a crucial factor in model performance. Nonetheless, there is a necessity for catchment specific apportionment validation as source discrimination is highly variable even with similar land use groups. Overall, the performance of the model is more dependent on the number of δ^{13}C FA tracers rather than the selection of individual tracers due to mixing space similarities.

Including δ^{15}N to offset the mixing line for all combinations of tracers (n=15) increased the performance of the model of all sources (FA combinations mean CRPS: 0.175, δ^{15}N +FA combinations mean CRPS: 0.091) (Fig. 5, B, Table S2). Importantly, pasture and arable source apportionment estimates decreased in performance with additional δ^{13}C FA tracers (Pasture mean CRPS: δ^{15}N +1 FA 0.073,2 FAs CRPS: 0.102, 3 FAs: 0.161, 4 FAs: 0.184 and arable mean CRPS: δ^{15}N +1 FA 0.079,2 FAs CRPS: 0.109, 3 FAs: 0.172, 4 FAs: 0.184) suggesting any beneficial source discrimination by additional δ^{13}C FA tracers is out-weighted by the increase in source-based uncertainty. Further evidence supporting these results is the iso plots that depict a large intersection between the arable and pasture source groups for all δ^{13}C FA tracers (Fig. 4). Consequently, the mixing space shifts from a mixing line to a mixing polygon with the inclusion of δ^{15}N, reducing pasture-arable misclassification.

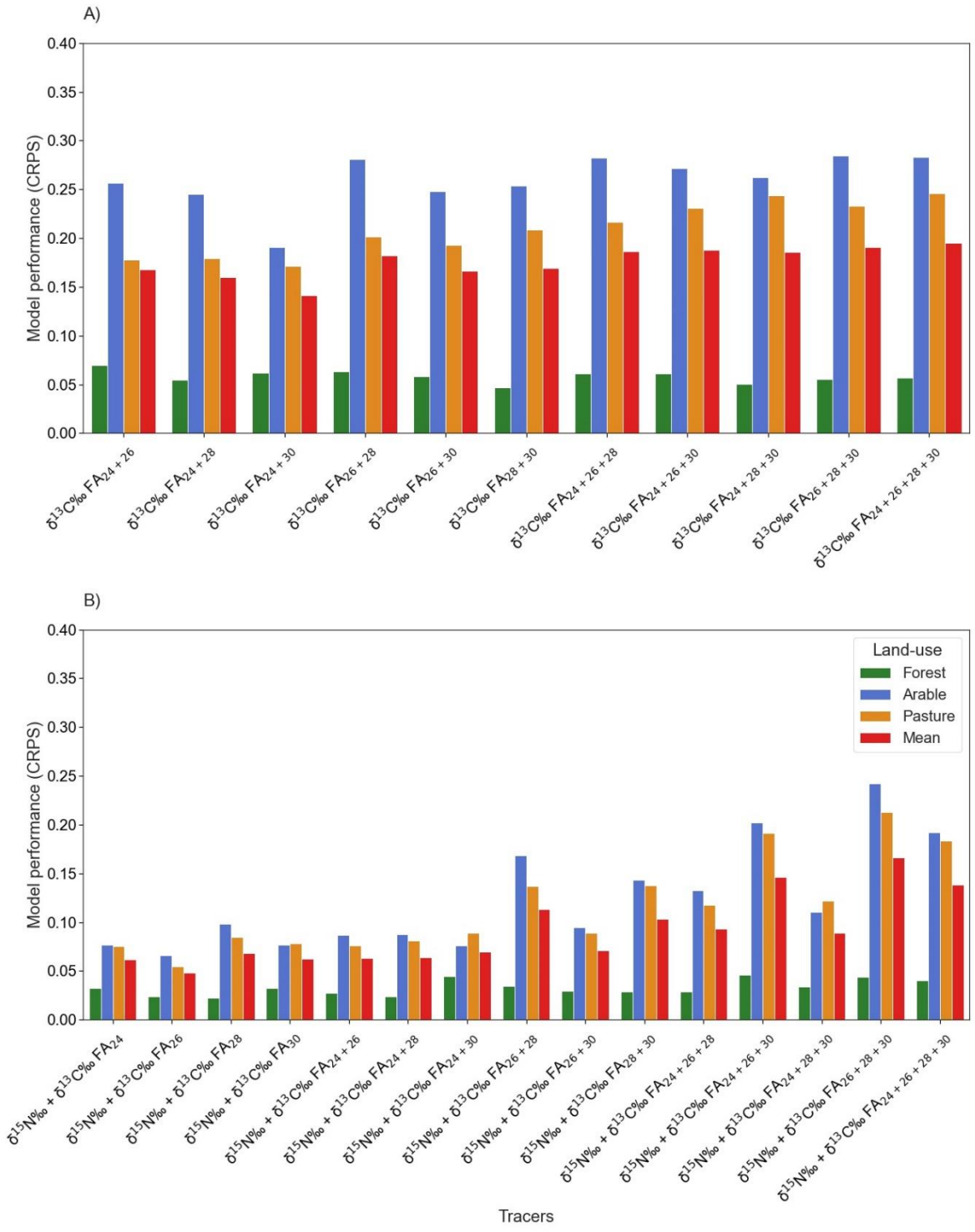


Fig. 5 Mean CRPS of all possible permutations and combinations of A) $\delta^{13}C$ FA tracers (n=11) and B) $\delta^{15}N$ and all $\delta^{13}C$ FA tracers (n=15). Tracer sets were evaluated using 150 concentration-dependent mathematical mixtures with CRPS (a higher CRPS indicates lower performance)

Tracer selection by the analysis of all combinations

The benefit of using $\delta^{15}N$ as a mixing line offset tracer is presented in Fig. 6 with the solid line indicating perfect fit (i.e., estimated proportion equals the known proportion). Fig. 6 (A) highlights the only $\delta^{13}C$ FA tracer set's inaccurate and underestimated apportionment of arable contribution and the overestimation of pasture contribution. Again, the inaccuracy can be attributed to the central location of arable in the mixing space for all $\delta^{13}C$ FA tracers, causing an underestimation of arable contributions as they are misclassified as pasture contributions. This centralised source location challenge and misclassification has been presented previously by Alewell et al. (2016) and Lavrieux et al. (2019).

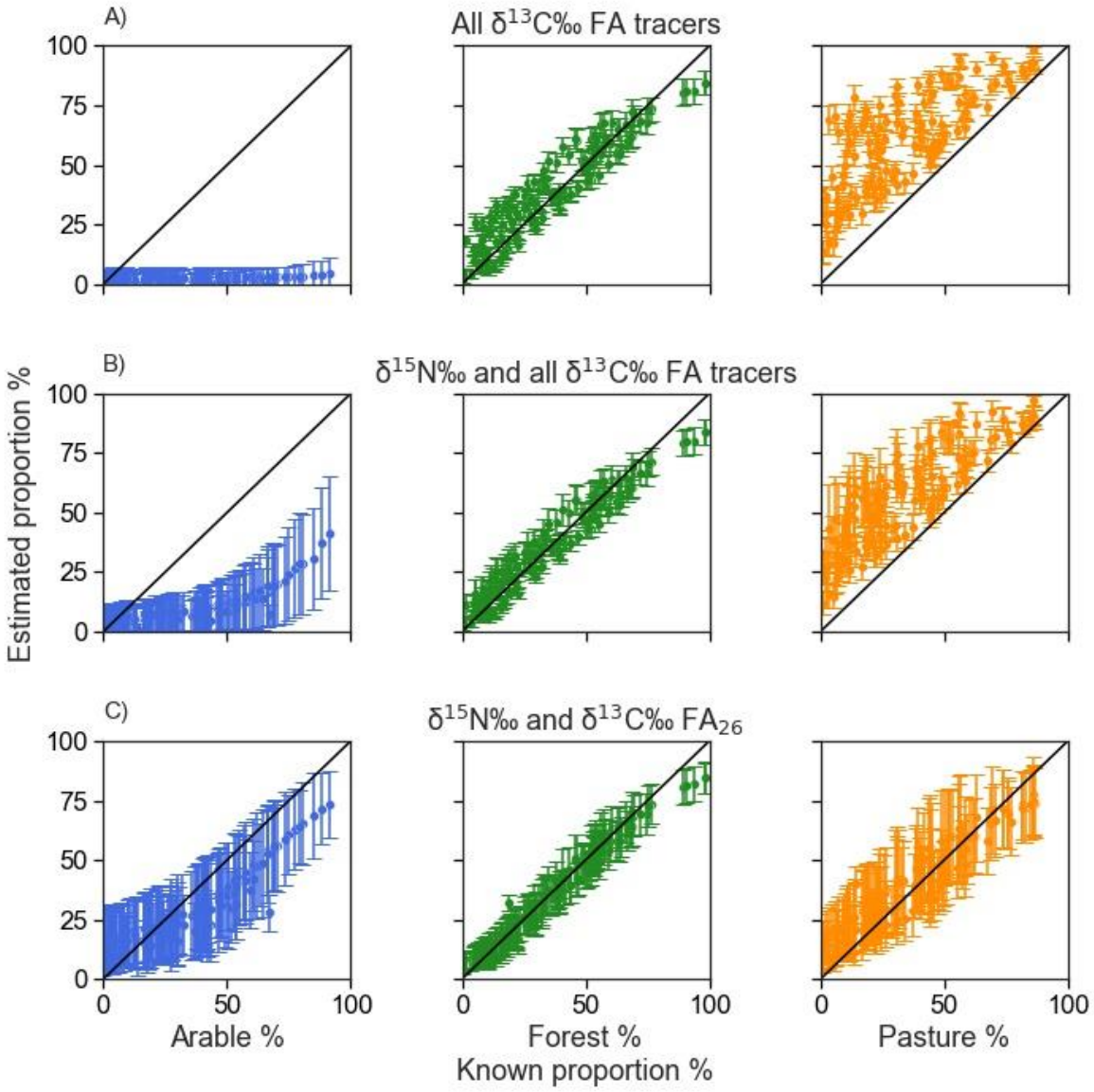


Fig. 6 Estimated proportions vs known mixture proportion for different tracer sets, with the solid line indicating perfect fit (estimated proportion= known proportion). A) all FA tracers, B) $\delta^{15}N$+all $\delta^{13}C$ FA tracers, C) $\delta^{15}N$+ $\delta^{13}C$ FA_{26} tracer set

The inclusion of the $\delta^{15}N$ tracer reduced the overestimation of pasture and the underestimation of the contribution of arable sources (Fig. 6, B). Using the CRP skill score as a comparative model performance metric, the expansion of the mixing space using $\delta^{15}N$ had a mean 18% (median 22%) CRP skill score model improvement compared to using only $\delta^{13}C$ FAs. The improvement of the model output using $\delta^{15}N$ as an additional tracer was expected by the expansion of the $\delta^{13}C$ FA linear mixing line into a more suitable mixing polygon.

When examining all potential tracer combinations, $\delta^{15}N + \delta^{13}C\ FA_{26}$ (Fig. 5, C) had the best model performance for all permutations with a mean 16% (median 62%) improvement compared to $\delta^{15}N$ + all $\delta^{13}C$ FAs (Fig. 5, B). The offset between mean and median is a result of the model predicting low contributions for all arable mixture proportions, at these low arable contributions, the model is likely correct for the wrong reasons. The increase in the accuracy and uncertainty of estimated source apportionment using $\delta^{15}N + \delta^{13}C\ FA_{26}$ (Fig. 6, C) results in the known source proportions being bracketed by the estimated values. The increase in uncertainty of the model compared to using $\delta^{15}N$ +all $\delta^{13}C$ FA tracers suggests that a reduction in the number of $\delta^{13}C$ FA tracers increases the number of possible solutions to the unmixing equation.

Although LDA is commonly used to optimise the power of discrimination when handling a large number of tracers, it is used irregularly for CSSI tracer selection. The accuracy of using LDA for tracer selection was investigated with a stepwise forward variable selection using the Wilk's Lambda criterion (niveau =0.1), which selected $\delta^{15}N + \delta^{13}C\ FA_{24} + \delta^{13}C\ FA_{26}$ as the optimal tracer set (LDA reclassification score 89%). Interestingly, the model performance of the LDA selected tracers decreased by 24% compared to $\delta^{15}N + \delta^{13}C\ FA_{26}$. The poor performance of the LDA selected tracers may be attributable to the mixing model's inclusion of concentration dependency, which is ignored by the LDA.

Tracer Selection and prediction bias analysis (PBA)

Predictive bias and the impact of the dominant source on model output has been identified previously in sediment fingerprinting and been described as a product of the source discrimination (Vale et al. 2022). Ideal tracers should contain enough discrimination power for null predictive bias; however, this is not the case with real tracers. To assess if predictive bias effects are reduced by the removal of tracers which have non-unique mixing spaces, known source proportions of mathematical mixtures are plotted against the model performance (CRPS) for each source (predictive bias analysis-PBA) (Fig. 7, Fig. 8). PBA of $\delta^{15}N$ + all $\delta^{13}C$ FA and $\delta^{15}N + \delta^{13}C\ FA_{26}$ was used to illustrate the effect of reducing source uncertainty by removing non-informative tracers.

PBA of the $\delta^{15}N$ + all $\delta^{13}C$ FA tracer set illustrates the decrease in arable and pasture performance with increasing arable contribution (Fig. 7). The extremely similar and linear relationship between arable and pasture CRPS is strong evidence for the misclassification of arable and pasture as the model underestimates and overestimates contributions from arable and pasture respectively. The clear discrimination of the forest source for all tracers (Fig. 4) resulted in the performance of the forest estimates being not affected by different source contributions. The PBA of the $\delta^{15}N + \delta^{13}C\ FA_{26}$ tracer

set depicts a reduction in the linear regression slope indicating a reduction in predictive bias effects (Fig. 8). It can be assumed that this is the result of a reduction in source uncertainty, when using a single FA tracer. The PBA highlights the balance between the source uncertainty error and the discriminative information gained by additional tracers.

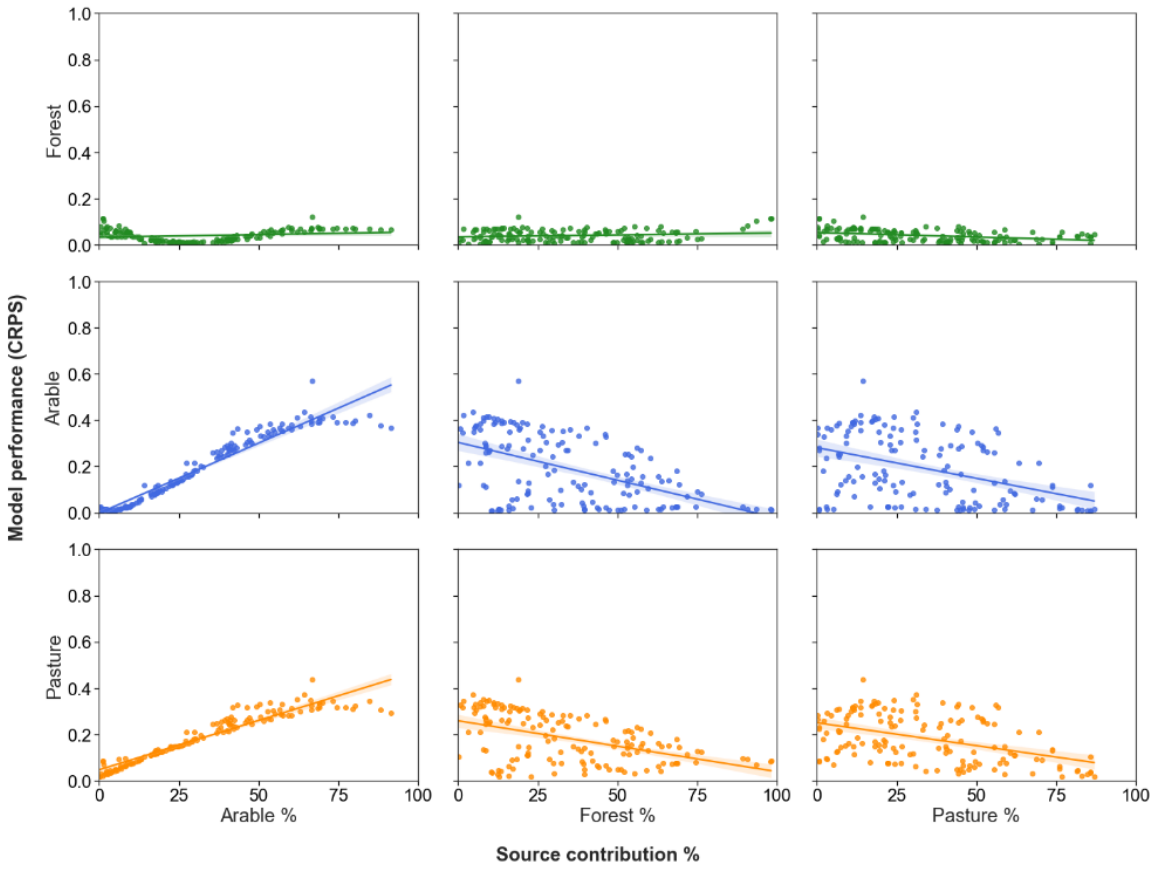


Fig. 7 The prediction bias analysis of $\delta^{15}N$ + all $\delta^{13}C$ FA illustrates how the model's performance for each source is influenced by varying source proportions, with each source's contribution plotted against the mean model's performance (where higher CRPS values indicate lower performance)

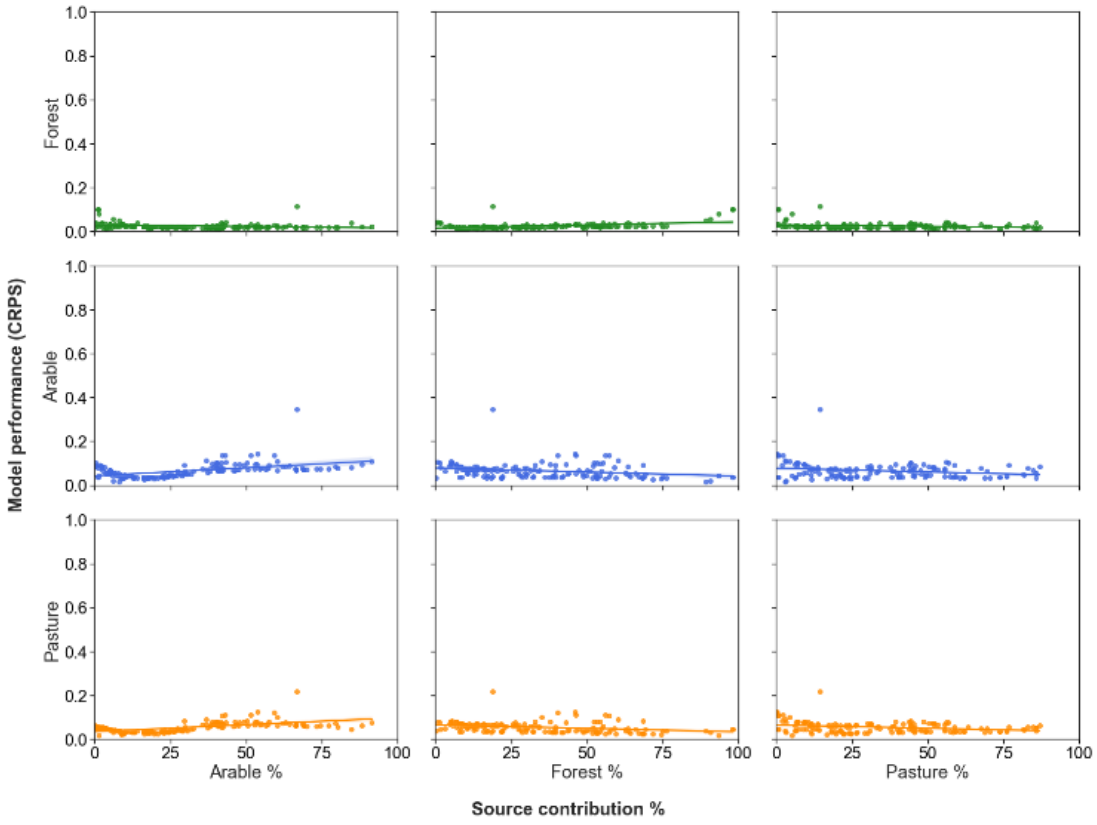


Fig. 8 The prediction bias analysis of $\delta^{15}N$ and $\delta^{13}C$ FA_{26} shows how the model's performance for each source is impacted by different source proportions, with the contribution of each source plotted against the mean model's performance (where higher CRPS values indicate lower performance). By comparing the linear regression slope to Figure 7, it is clear that that there is a decrease in predictive bias with less non-informative tracers

Identifying non-informative tracers by scaling and discrimination analysis (SDA)

The tracers' balance of source discrimination and source uncertainty is determined on a regular basis using boxplots and a Kruskal Wallis test (Fig. 2). However, tracers are not independent factors and work in the mixing model simultaneously. The current approach to tracer selection is to see if individual tracers can distinguish between sources. As an alternative, we investigated whether it is possible to distinguish various FA tracers based on mixing space.

The majority of $\delta^{13}C$ FA tracers (94%) had significantly different source distributions ($p<0.05$) (except $\delta^{13}C$ FA_{28} –FA_{30} in arable) (Fig. 9, A, Table 3, Left). This can lead to the assumption that each tracer has valuable information for the model. However, the difference between absolute source distributions of each tracer (distance of source distribution from 0) is caused by each FA tracer being depleted by approximately -1.1‰ $\delta^{13}C$ per two additional carbon atoms (similarly shown by Chikaraishi et al. 2004) (Fig. 3). Considering that MixSIAR uses relative (source-source) tracer vales rather than the absolute tracer value, tracers which demonstrate modification of all source distributions by direct isometric translation (e.g., every point/source of the mixing shape is moved in the same distance and in the same direction) can be considered mathematically non-unique in terms of mixing space.

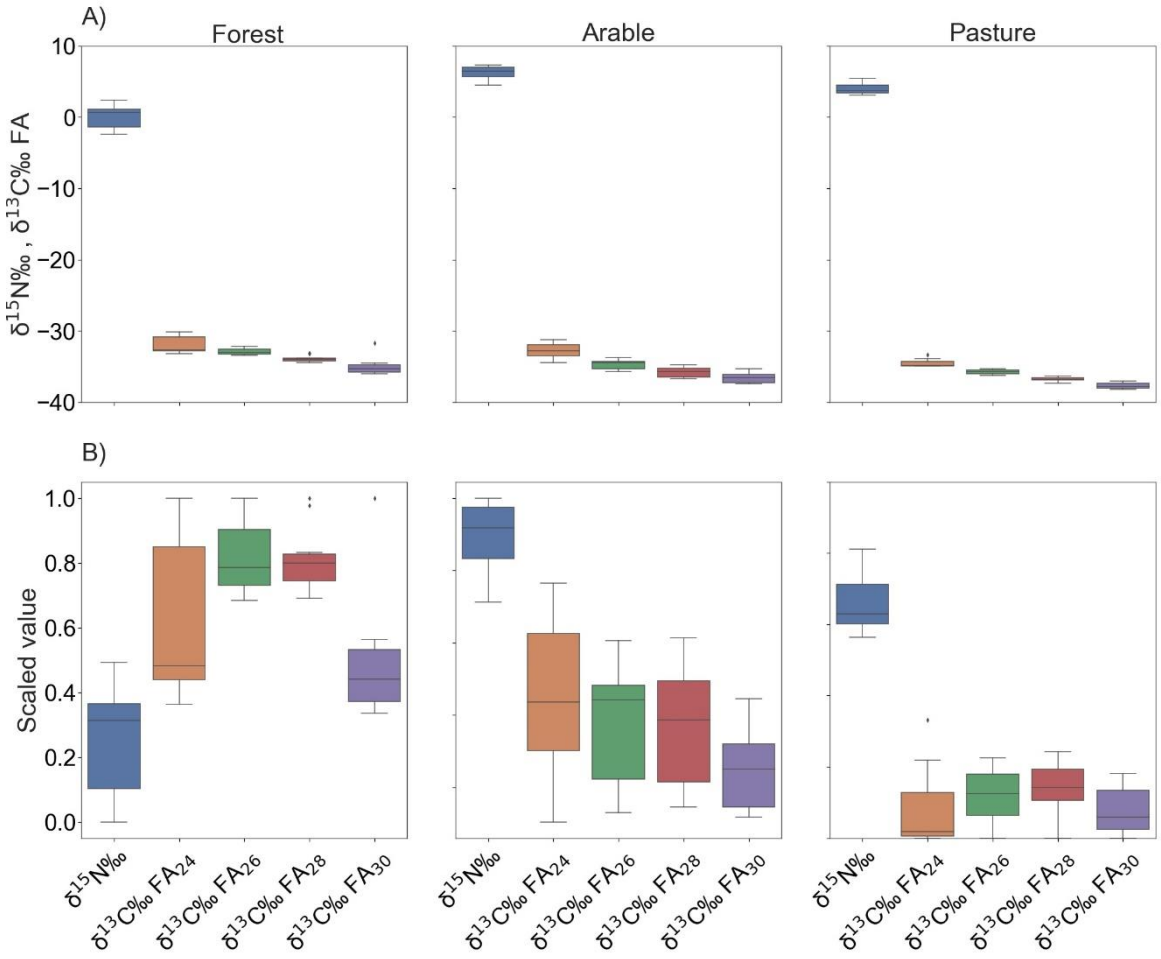


Fig. 9 A) $\delta^{15}N$ and $\delta^{13}C$ FA grouped by source, demonstrating the linear relationship of $\delta^{13}C$ FA and alkyl chain length B) $\delta^{15}N$ and $\delta^{13}C$ FA tracers scaled and grouped by source allows comparison of relative source-source distributions for each tracer. Tracer values are scaled using the max and minimum value of each tracer over all land uses

Table 3 Left: Kruskal Wallis (KW) test results of the comparison of each tracer's absolute source distribution, grey highlight indicates no significant difference between source distribution of tracer pairs, * indicates significant difference. Right: Kruskal Wallis test results of the comparison of each tracer's scaled source distribution

Non-scaled tracer values KW **Scaled tracer values KW**

Forest

	$\delta^{15}N$	FA_{24}	FA_{26}	FA_{28}	FA_{30}
$\delta^{15}N$	1.00				
FA_{24}	<0.00*	1.00			
FA_{26}	<0.00*	0.03*	1.00		
FA_{28}	<0.00*	<0.00*	0.00*	1.00	
FA_{30}	<0.00*	<0.00*	0.00*	0.00*	1.00

	$\delta^{15}N$	FA_{24}	FA_{26}	FA_{28}	FA_{30}
$\delta^{15}N$	1.00				
FA_{24}	<0.00*	1.00			
FA_{26}	<0.00*	0.11	1.00		
FA_{28}	<0.00*	0.11	0.90	1.00	
FA_{30}	<0.00*	0.20	<0.00*	<0.00*	1.00

Arable

	$\delta^{15}N$	FA_{24}	FA_{26}	FA_{28}	FA_{30}
$\delta^{15}N$	1.00				
FA_{24}	<0.00*	1.00			
FA_{26}	<0.00*	<0.00*	1.00		
FA_{28}	<0.00*	<0.00*	0.01*	1.00	
FA_{30}	<0.00*	<0.00*	<0.00*	0.06	1.00

	$\delta^{15}N$	FA_{24}	FA_{26}	FA_{28}	FA_{30}
$\delta^{15}N$	1.00				
FA_{24}	<0.00*	1.00			
FA_{26}	<0.00*	0.41	1.00		
FA_{28}	<0.00*	0.29	0.88	1.00	
FA_{30}	<0.00*	0.03*	0.07	0.13	1.00

Pasture

	$\delta^{15}N$	FA_{24}	FA_{26}	FA_{28}	FA_{30}
$\delta^{15}N$	1.00				
FA_{24}	<0.00*	1.00			
FA_{26}	<0.00*	<0.00*	1.00		
FA_{28}	<0.00*	<0.00*	<0.00*	1.00	
FA_{30}	<0.00*	<0.00*	<0.00*	<0.00*	1.00

	$\delta^{15}N$	FA_{24}	FA_{26}	FA_{28}	FA_{30}
$\delta^{15}N$	1.00				
FA_{24}	<0.00*	1.00			
FA_{26}	<0.00*	0.44	1.00		
FA_{28}	<0.00*	0.37	0.80	1.00	
FA_{30}	<0.00*	0.70	0.44	0.16	1.00

To provide an alternative and more robust line of evidence of non-unique mixing spaces of FA tracers, the tracer values were scaled between 0 and 1 across all sources. Scaling retains the relative location, shape, and distribution of the sources for each tracer, enabling comparison of the relative source locations of different tracers using Kruskal Wallis test and as such is a suitable tool for evaluating if tracers have significantly different mixing spaces.

The scaled value of $\delta^{15}N$ was shown to be significantly different ($p<0.05$) to all $\delta^{13}C$ FAs for all land uses (Fig. 9, B, Table 3, Right). In contrast, only 17% of FA tracers had significant differences between any of the scaled source values. Pasture had no significant differences between any scaled $\delta^{13}C$ FA. Forest and arable only had a significant difference between two and one pair of scaled tracers, respectively (Forest: $\delta^{13}C$ $FA_{26} - FA_{30:0}$ and $\delta^{13}C$ $FA_{28} - FA_{30}$. Arable: $\delta^{13}C$ $FA_{24} - FA_{30}$) (Fig. 9, B). The minimal but present uniqueness of mixing space for 20% of the FAs in the forest source can be assumed to be caused by a more biodiverse FA input, while the non-uniqueness of 10% of FAs in arable sources could be a result of the legacy tracer signal from crop rotation (Upadhayay et al. 2020).

$\delta^{13}C$ depletion during the FA elongation processes appears to be similar for all land uses, with any land-use-specific isotopic variation during FA elongation being negligible when compared to the intra-source variability. The linear relationship between $\delta^{13}C$ and FA alkyl length causes FA tracers to be direct isometry translations of each other and consequently, there is minimal significant differences between the relative source locations of each FA tracer (Fig. 9, B) and as such, the mixing space can be thought of as being non-unique for all $\delta^{13}C$ FA tracers. The similarities between scaled source values for all tracers are illustrated in Fig. 10. $\delta^{15}N$ is depicted to have to have non-translation transformations of the mixing shape compared to FAs. The similarities in the mixing shape for all FAs indicate that direct isometry translation is present between different FAs, making multiple FA tracers non-unique and non-informative.

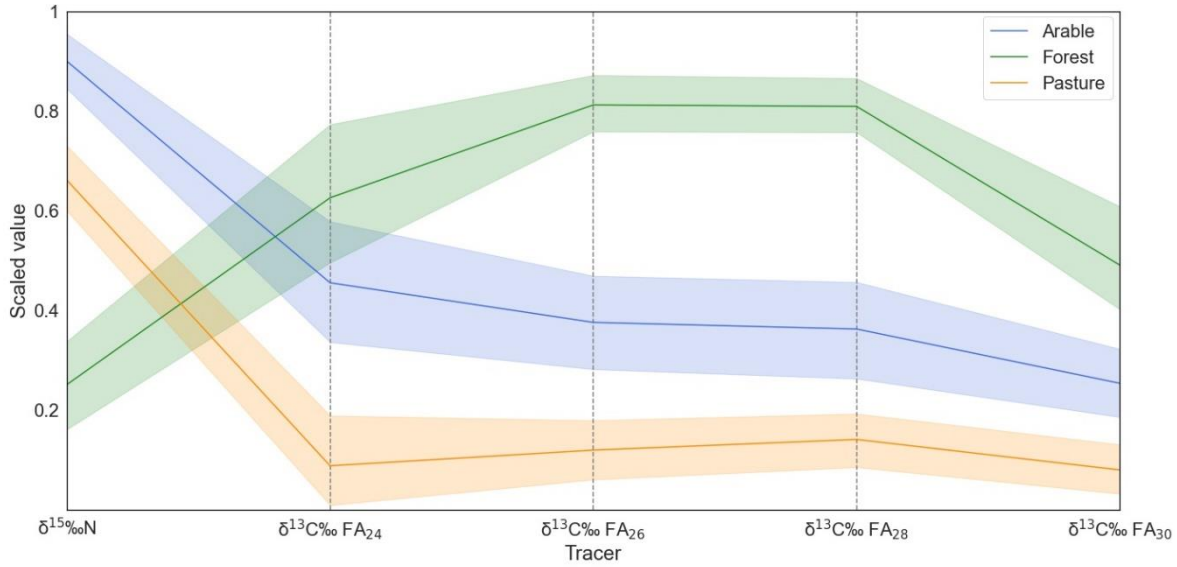


Fig. 10 Scaled $\delta^{15}N$ and $\delta^{13}C$ FA tracer values used to compare the relative source distributions for each tracer. Tracer values are scaled using the max and minimum value of each tracer over all land uses. The mixing space is shown to be similar for all FAs while $\delta^{15}N$ mixing space is shown to be modified by non-isometry translations

Considering MixSIAR uses the relative source-sediment-source positions for un-mixing, any modification of the mixing space by only direct isometry translation has a null effect on the mixing space. Therefore, any tracer with a mixing space that is a direct isometry translation of another tracer can be seen as almost identical and either one of the tracers is non-informative.

Non-informative tracers – Clone Tracer Analysis

To assess MixSIAR's performance when using tracers with identical mixing spaces, a non-informative clone tracer (an exact copy of a tracer used as an additional tracer) was used as a direct approach to test non-informative tracer behaviour. Three sequential additions of the clone tracer $\delta^{13}C$ FA_{26} were added to the $\delta^{15}N + \delta^{13}C$ FA_{26} tracer set (Fig. 11, A). Increasing the number of clone tracers decreased the model performance ($\delta^{15}N$ + 1x $\delta^{13}C$ FA_{26} CRPS: 0.034, 2x FA_{26}: 0.181, 3x FA_{26}: 0.199, 4x FA_{26}: 0.202). This effect can be attributed to the lack of any additional beneficial information when using tracers with non-unique mixing spaces, whilst the source uncertainty error induced by adding multiple clone tracers is propagated. In this study, these results disagree with the notion that MixSIAR handles non-informative tracers sufficiently (Smith et al. 2018).

Optimizing model performance strives to balance new beneficial source discrimination and the additional source uncertainty brought to the model by each additional tracer. Fig. 11, A, indicates that when using a clone tracer, the source uncertainty is propagated until the addition of a fourth tracer. The difference in CRPS between clone tracers and different FAs was used as a measure of information gain when using additional FA tracers. The mean CRPS of different FA tracer combinations with the same number of tracers displayed a similar trend to that when adding additional clone tracers ($\delta^{15}N$ + 1x $\delta^{13}C$ FA CRPS: 0.060, 2x FA: 0.081, 3x FA: 0.124, 4x FA: 0.138) (Fig. 11, B). Small non-translation modifications of FA mixing spaces resulted in the CRPS using additional different FA tracers being generally lower (mean 22%) than additional clone tracers. Therefore, from a mathematical perspective, different FAs are not completely non-informative. Although, from practical perspective additional FA tracers are essentially non-informative, as any beneficial information gained is outweighed by the error added from the propagation of source uncertainty.

Indeed, this approach is highly experimental, and it is unlikely that you will have truly identical tracers in the field. Nonetheless, this method demonstrates that non-informative tracers can add bias to a model, as additional FA tracers may bring limited additional information for unmixing. When using different FAs, our results demonstrate that the error gained by mixing spaces translation effects outweighs the information gained from non-translation modification. This, however, may not be the case for all catchments and tracers.

An intriguing area of investigation is how the balance of source discrimination and tracer mixing space similarities effect model performance. CSSI of FAs have a relatively narrow range of possible source values (ca. 10-40 ‰) compared to other tracers (e.g., geochemistry). When tracers with a higher degree of source discrimination, though identical mixing spaces, are modelled, the propagation of source uncertainty may be out weighted and potentially result in improved model performance.

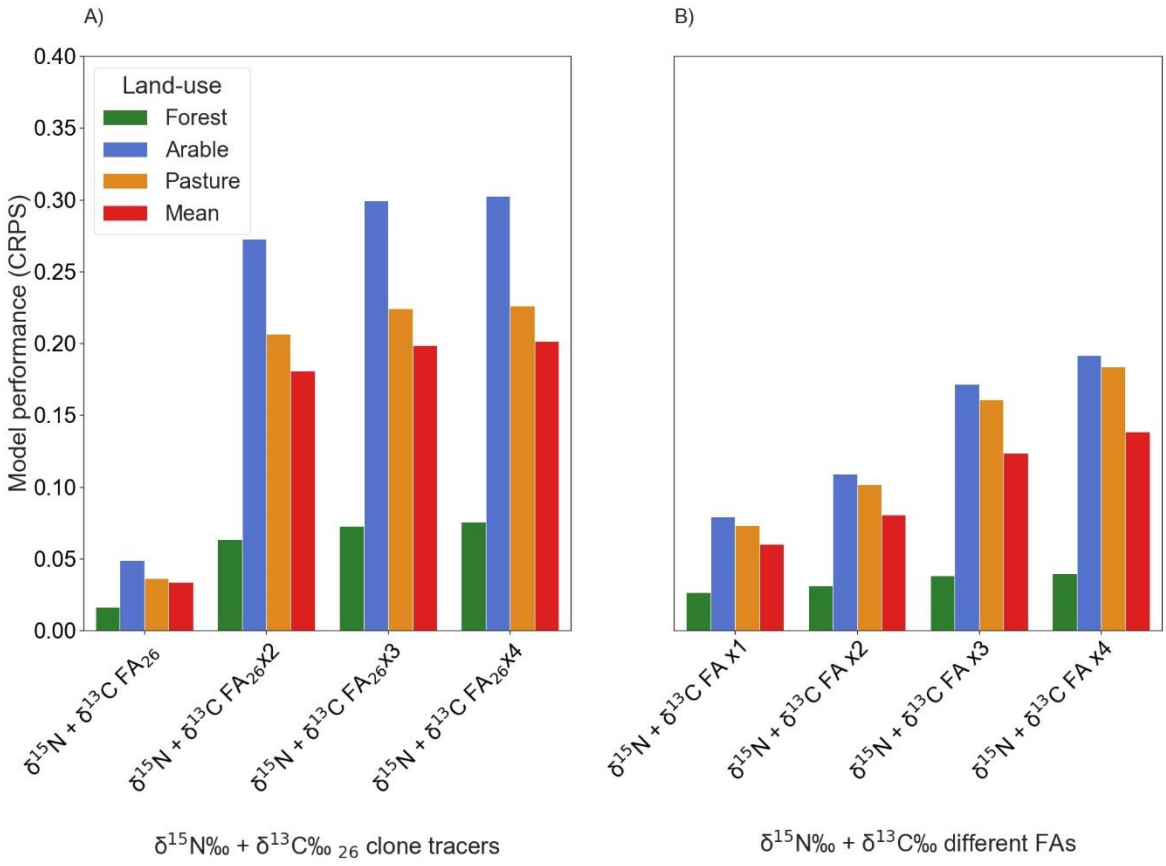


Fig. 11 Comparison between model performance using additional clone tracers ($\delta^{13}C$ FA_{26}) and additional different $\delta^{13}C$ FA tracers to the $\delta^{15}N$ + $\delta^{13}C$ FA_{26} tracer set. The mean CRPS of all tracer combinations with the same number of tracers is used for additional FA tracers to improve the representative of results (higher CRPS indicates lower performance)

3.4 Conclusion

Using mathematical mixtures, the addition of $\delta^{15}N$ to expand the CSSI FA mixing line, improved the model by 22% compared to using only $\delta^{13}C$ FAs. The evaluation of possible combinations of tracers indicated that $\delta^{15}N$ + $\delta^{13}C$ FA_{26} was the optimal tracer set and had a 62% improvement compared to $\delta^{15}N$ + all $\delta^{13}C$ Fas. LDA tracer selection is regularly used in the literature to select the optimal suite of tracers to increase model performance. However, in this case, the tracers selected by the LDA did not provide the optimal tracer selection. Additional $\delta^{13}C$ FA tracers had a negative influence on model performance, indicating that increasing the number of conservative tracers does not necessarily result in improved performance, as previously suggested when using a Bayesian framework. However, the reduction in number of tracers will increase the influence of any non-conservative tracers. As mathematical mixtures, by definition, do not contain non-conservative tracers, the potential influence of non-conservative tracers needs careful consideration when apportioning sediment sources.

Our results indicated there is a reduction of predictive bias when using a single FA tracer. Using a novel SDA test, additional Fas were shown to have non-unique mixing spaces. Considering MixSIAR uses the relative source-sediment-source positions for un-mixing, any tracer which exhibits a non-unique mixing space can be seen as non-informative. Using a clone tracer to evaluate MixSIAR's performance handling non-informative tracers resulted in strong evidence of MixSIAR's insufficient handling of tracers with non-unique mixing spaces. In particular, model performance decreased when using additional FA as well as clone tracers.

Land-use-specific sediment source apportionment using FA CSSIs requires a supplemental offset tracer that is not dependent on the C_3-C_4 discrimination pathway. Since a single FA CSSI had the best performance with an additional offset tracer, an alternative single tracer to FA CSSIs that uses the C3-C4 discrimination pathway for source discrimination such as bulk isotopes may be more accessible and have similar unmixing performance. However, the conservativeness and unmixing performance of these tracers need to be explored further; the latter can be evaluated confidently by using mathematical mixtures. Even though adding $\delta^{15}N$ as a tracer in this study outperformed the combination of several FA CSSI, $\delta^{15}N$ may be prone to isotopic fractionation during the degradation of molecules and thus may not meet the requirement of a conservative tracer under real world situations, where molecules are subject to transport and possible degradation. $\delta^{15}N$ may be useful in scenarios where the balance between beneficial information gained by improving source discrimination outweighs any effect of modification or fractionation of the tracer during sediment mobilization, transport and deposition processes. Here, we capitalized on the availability of $\delta^{15}N$ data (which is analysed simultaneously with bulk $\delta^{13}C$) to demonstrate the utility of additional tracers that have an alternative mixing space.

In fingerprinting applications, additional tracer selection steps should be considered, including: 1) checking the uniqueness of tracer mixing spaces by SDA, with the removal of tracers that show non-unique mixing spaces, and 2) where feasible, analysing all combinations and permutations of tracers using mathematical mixtures to further optimize tracer selection. Although computationally intensive, it can help identify the optimal tracer suite for modelling. Even though this method is applied to FA CSSI and $\delta^{15}N$ tracers in this study, this method is potentially appropriate for broader application to identify non-informative tracers. This includes multiple fingerprinting parameters (e.g. fallout radionuclides, spectra and geochemical tracers) in which the co-linearity of tracers is not uncommon. However, we suggest further exploration of mathematical mixtures to determine the effect of different error structures on model performance and the validity of organic matter or particle size corrections. We anticipate that the use of mathematical mixtures and tracer combinations as a decisive tracer selection step will enable a wider range of applications for sediment fingerprinting, improve our knowledge of the dynamics of soil and sediment in the environment, and enhance soil erosion mitigation techniques.

3.5 References

Agrawal S, Galy V, Sanyal P, Eglinton T (2014) C4 plant expansion in the Ganga Plain during the last glacial cycle: Insights from isotopic composition of vascular plant biomarkers. Org Geochem 67:58–71. https://doi.org/10.1016/j.orggeochem.2013.12.007

Alewell C, Birkholz A, Meusburger K et al (2016) Quantitative sediment source attribution with compound-specific isotope analysis in a C3 plant-dominated catchment (central Switzerland). Biogeosciences 13:1587–1596. https://doi.org/10.5194/bg-13-1587-2016

Amundson R, Austin AT, Schuur EAG et al (2003) Global patterns of the isotopic composition of soil and plant nitrogen. Global Biogeochem Cycles 17: https://doi.org/10.1029/2002GB001903

Bahadori M, Chen C, Lewis S et al (2019) A novel approach of combining isotopic and geochemical signatures to differentiate the sources of sediments and particulate nutrients from different land uses. Sci. Total Environ 655:129–140. https://doi.org/10.1016/j.scitotenv.2018.11.084

Bakker MM, Govers G, Rounsevell MDA (2004) The crop productivity-erosion relationship: An analysis based on experimental work. Catena 57:55–76. https://doi.org/10.1016/j.catena.2003.07.002

Barthod LRM, Liu K, Lobb DA et al (2015) Selecting Color-based Tracers and Classifying Sediment Sources in the Assessment of Sediment Dynamics Using Sediment Source Fingerprinting. J Environ Qual 44:1605–1616. https://doi.org/10.2134/jeq2015.01.0043

Batista PVG, Laceby JP, Evrard O (2022) How to evaluate sediment fingerprinting source apportionments. J Soils Sediments 22:1315–1328. https://doi.org/10.1007/s11368-022-03157-4

Batista PVG, Laceby JP, Silva MLN et al (2019) Using pedological knowledge to improve sediment source apportionment in tropical environments. J Soils Sediments 19:3274–3289. https://doi.org/10.1007/s11368-018-2199-5

Belmont P, Willenbring JK, Schottler SP et al (2014) Toward generalizable sediment fingerprinting with tracers that are conservative and nonconservative over sediment routing timescales. J Soils Sediments 14:1479–1492. https://doi.org/10.1007/s11368-014-0913-5

Ben Slimane A, Raclot D, Evrard O et al (2013) Fingerprinting sediment sources in the outlet reservoir of a hilly cultivated catchment in Tunisia. J Soils Sediments 13:801–815. https://doi.org/10.1007/s11368-012-0642-6

Blake WH, Boeckx P, Stock BC et al (2018) A deconvolutional Bayesian mixing model approach for river basin sediment source apportionment. Sci Rep 8:. https://doi.org/10.1038/s41598-018-30905-9

Blake WH, Ficken KJ, Taylor P et al (2012) Tracing crop-specific sediment sources in agricultural catchments. Geomorphology 139–140:322–329. https://doi.org/10.1016/j.geomorph.2011.10.036

Borselli L, Cassi P, Torri D (2008) Prolegomena to sediment and flow connectivity in the landscape: A GIS and field numerical assessment. Catena 75:268–277. https://doi.org/10.1016/j.catena.2008.07.006

Bouchez J, Gaillardet J, France-Lanord C et al (2011) Grain size control of river suspended sediment geochemistry: Clues from Amazon River depth profiles. Geochemistry, Geophysics, Geosystems 12:. https://doi.org/10.1029/2010GC003380

Bravo-Linares C, Schuller P, Castillo A (2018) First use of a compound-specific stable isotope (CSSI) technique to trace sediment transport in upland forest catchments of Chile. Sci. Total Environ 618:1114–1124. https://doi.org/10.1016/j.scitotenv.2017.09.163

Chikaraishi Y, Naraoka H, Poulson SR (2004) Hydrogen and carbon isotopic fractionations of lipid biosynthesis among terrestrial (C3, C4 and CAM) and aquatic plants. Phytochemistry 65:1369–1381. https://doi.org/10.1016/j.phytochem.2004.03.036

Collins AL, Burak E, Harris P et al (2019) Field scale temporal and spatial variability of δ13C, δ15N, TC and TN soil properties: Implications for sediment source tracing. Geoderma 333:108–122. https://doi.org/10.1016/j.geoderma.2018.07.019

Collins AL, Pulley S, Foster IDL et al (2017) Sediment source fingerprinting as an aid to catchment management: A review of the current state of knowledge and a methodological decision-tree for end-users. J Environ Manage 194:86–108. https://doi.org/10.1016/j.jenvman.2016.09.075

Collins AL, Walling DE (2004) Documenting catchment suspended sediment sources: Problems, approaches and prospects. Prog Phys Geogr 28:159–196. https://doi.org/10.1191/0309133304pp409ra

Collins AL, Walling DE, Leeks GJL (1997a) Source type ascription for fluvial suspended sediment based on a quantitative composite fingerprinting technique. Catena 29:27

Collins AL, Walling DE, Leeks GJL (1996) Composite fingerprinting of the spatial source of fluvial suspended sediment : a case study of the Exe and Severn river basins, United Kingdom. Géomorphologie relief processus environnement 2:41–53. https://doi.org/10.3406/morfo.1996.877

Collins AL, Walling DE, Leeks GJL (1997b) Fingerprinting the origin of fluvial suspended sediment in larger river basins: Combining assessment of spatial provenance and source type. Geografiska Annaler, Series A: Physical Geography 79:239–254. https://doi.org/10.1111/j.0435-3676.1997.00020.x

Collins AL, Walling DE, Sichingabula HM, Leeks GJL (2001) Suspended sediment source fingerprinting in a small tropical catchment and some management implications. Appl Geogr 21:387-412

Collister JW, Rieley G, Stern B et al (1994) Compound-specific ~3C analyses of leaf lipids from plants with differing carbon dioxide metabolisms. Org Geochem 21:619–627

Cooper RJ, Krueger T (2017) An extended Bayesian sediment fingerprinting mixing model for the full Bayes treatment of geochemical uncertainties. Hydrol Process 31:1900–1912. https://doi.org/10.1002/hyp.11154

Cooper RJ, Pedentchouk N, Hiscock KM et al (2015) Apportioning sources of organic matter in streambed sediments: An integrated molecular and compound-specific stable isotope approach. Sci. Total Environ 520:187–197. https://doi.org/10.1016/j.scitotenv.2015.03.058

Erdbrügger P, Fröhlich F (2020) The role of very long chain fatty acids in yeast physiology and human diseases. Biol Chem 402:25–38

Evrard O, Batista PVG, Company J et al (2022) Improving the design and implementation of sediment fingerprinting studies: summary and outcomes of the TRACING 2021 Scientific School. J Soils Sediments 22:1648–1661. https://doi.org/10.1007/s11368-022-03203-1

Evrard O, Laceby JP, Ficetola GF (2019) Environmental DNA provides information on sediment sources: A study in catchments affected by Fukushima radioactive fallout. Sci. Total Environ 665:873–881. https://doi.org/10.1016/j.scitotenv.2019.02.191

Evrard O, Poulenard J, Némery J et al (2013a) Tracing sediment sources in a tropical highland catchment of central Mexico by using conventional and alternative fingerprinting methods. Hydrol Process 27:911–922. https://doi.org/10.1002/hyp.9421

Evrard O, Poulenard J, Némery J et al (2013b) Tracing sediment sources in a tropical highland catchment of central Mexico by using conventional and alternative fingerprinting methods. Hydrol Process 27:911–922. https://doi.org/10.1002/hyp.9421

Fox JF, Papanicolaou AN (2007) The use of carbon and nitrogen isotopes to study watershed erosion processes. J Am Water Resour Assoc 43:1047–1064. https://doi.org/10.1111/j.1752-1688.2007.00087.x

García-Comendador J, Martínez-Carreras N, Fortesa J et al (2023) In-channel alterations of soil properties used as tracers in sediment fingerprinting studies. Catena 225:107036. https://doi.org/10.1016/j.catena.2023.107036

Gateuille D, Owens PN, Petticrew EL et al (2019) Determining contemporary and historical sediment sources in a large drainage basin impacted by cumulative effects: the regulated Nechako River, British Columbia, Canada. J Soils Sediments 19:3357–3373. https://doi.org/10.1007/s11368-019-02299-2

Gellis AC, Noe GB (2013) Sediment source analysis in the Linganore Creek watershed, Maryland, USA, using the sediment fingerprinting approach: 2008 to 2010. J Soils Sediments 13:1735–1753. https://doi.org/10.1007/s11368-013-0771-6

Gibbs MM (2008) Identifying source soils in contemporary estuarine sediments: A new compound-specific isotope method. Estuaries and Coasts 31:344–359. https://doi.org/10.1007/s12237-007-9012-9

Haddadchi A, Olley J, Laceby P (2014) Accuracy of mixing models in predicting sediment source contributions. Sci. Total Environ 497–498:139–152. https://doi.org/10.1016/j.scitotenv.2014.07.105

Haddadchi A, Ryder DS, Evrard O, Olley J (2013) Sediment fingerprinting in fluvial systems: review of tracers, sediment sources and mixing models. J. Sediment. Res. 28:560-578

Hancock GJ, Revill AT (2013) Erosion source discrimination in a rural Australian catchment using compound-specific isotope analysis (CSIA). Hydrol Process 27:923–932. https://doi.org/10.1002/hyp.9466

Hirave P, Glendell M, Birkholz A, Alewell C (2021) Compound-specific isotope analysis with nested sampling approach detects spatial and temporal variability in the sources of suspended sediments in a Scottish mesoscale catchment. Sci. Total Environ 755:. https://doi.org/10.1016/j.scitotenv.2020.142916

Issaka S, Ashraf MA (2017) Impact of soil erosion and degradation on water quality: a review. Geology, Ecology, and Landscapes 1:1–11. https://doi.org/10.1080/24749508.2017.1301053

Jacob J, Disnar JR, Boussafir M et al (2005) Pentacyclic triterpene methyl ethers in recent lacustrine sediments (Lagoa do Caçó, Brazil). Org Geochem 36:449–461. https://doi.org/10.1016/j.orggeochem.2004.09.005

Kemper JT, Rathburn SL, Friedman JM et al (2022) Fingerprinting historical tributary contributions to floodplain sediment using bulk geochemistry. Catena 214:. https://doi.org/10.1016/j.catena.2022.106231

Koiter AJ, Owens PN, Petticrew EL, Lobb DA (2013) The behavioural characteristics of sediment properties and their implications for sediment fingerprinting as an approach for identifying sediment sources in river basins. Earth Sci Rev 125:24–42

Laceby JP, Evrard O, Smith HG et al (2017) The challenges and opportunities of addressing particle size effects in sediment source fingerprinting: A review. Earth Sci Rev 169:85–103

Laceby JP, McMahon J, Evrard O, Olley J (2015) A comparison of geological and statistical approaches to element selection for sediment fingerprinting. J Soils Sediments 15:2117–2131. https://doi.org/10.1007/s11368-015-1111-9

Lake NF, Martínez-Carreras N, Shaw PJ, Collins AL (2022) High frequency un-mixing of soil samples using a submerged spectrophotometer in a laboratory setting—implications for sediment fingerprinting. J Soils Sediments 22:348–364. https://doi.org/10.1007/s11368-021-03107-6

Latorre B, Lizaga I, Gaspar L, Navas A (2021) A novel method for analysing consistency and unravelling multiple solutions in sediment fingerprinting. Sci. Total Environ 789:. https://doi.org/10.1016/j.scitotenv.2021.147804

Lavrieux M, Birkholz A, Meusburger K et al (2019) Plants or bacteria? 130 years of mixed imprints in Lake Baldegg sediments (Switzerland), as revealed by compound-specific isotope analysis (CSIA) and biomarker analysis. Biogeosciences 16:2131–2146. https://doi.org/10.5194/bg-16-2131-2019

le Gall M, Evrard O, Foucher A et al (2016) Quantifying sediment sources in a lowland agricultural catchment pond using 137Cs activities and radiogenic 87Sr/86Sr ratios. Sci. Total Environ 566–567:968–980. https://doi.org/10.1016/j.scitotenv.2016.05.093

Liu J, Han G (2021) Tracing Riverine Particulate Black Carbon Sources in Xijiang River Basin: Insight from Stable Isotopic Composition and Bayesian Mixing Model. Water Res 194:. https://doi.org/10.1016/j.watres.2021.116932

Lizaga I, Bodé S, Gaspar L et al (2021) Legacy of historic land cover changes on sediment provenance tracked with isotopic tracers in a Mediterranean agroforestry catchment. J Environ Manage 288:. https://doi.org/10.1016/j.jenvman.2021.112291

Lizaga I, Latorre B, Gaspar L, Navas A (2022) Combined use of geochemistry and compound-specific stable isotopes for sediment fingerprinting and tracing. Sci. Total Environ 832:. https://doi.org/10.1016/j.scitotenv.2022.154834

Lizaga I, Latorre B, Gaspar L, Navas A (2020) Consensus ranking as a method to identify non-conservative and dissenting tracers in fingerprinting studies. Sci. Total Environ 720:. https://doi.org/10.1016/j.scitotenv.2020.137537

Martínez-Carreras N, Krein A, Gallart F et al (2010) Assessment of different colour parameters for discriminating potential suspended sediment sources and provenance: A multi-scale study in Luxembourg. Geomorphology 118:118–129. https://doi.org/10.1016/j.geomorph.2009.12.013

Matheson JE, Winkler RL (1976) Scoring Rules for Continuous Probability Distributions. Source: Management Science 22:1087–1096

Motha JA, Wallbrink PJ, Hairsine PB, Grayson RB (2002) Tracer properties of eroded sediment and source material. Hydrol Process 16:1983–2000. https://doi.org/10.1002/hyp.397

Mukundan R, Radcliffe DE, Ritchie JC et al (2010) Sediment Fingerprinting to Determine the Source of Suspended Sediment in a Southern Piedmont Stream. J Environ Qual 39:1328–1337. https://doi.org/10.2134/jeq2009.0405

Nosrati K, Govers G, Ahmadi H et al (2011) An exploratory study on the use of enzyme activities as sediment tracers: biochemical fingerprints? J Sediment. Res. 26:136-151

Owens PN, Blake WH, Gaspar L et al (2016) Fingerprinting and tracing the sources of soils and sediments: Earth and ocean science, geoarchaeological, forensic, and human health applications. Earth Sci Rev 162:1–23

Papanicolaou T (2003) Soil fingerprinting in the Palouse Basin, USA, using stable carbon and nitrogen isotopes. International Journal of Sediment Research 18:278–284

Pedregosa F, Michel V, Grisel O et al (2011) Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12:2825–2830

Pedro HTC, Coimbra CFM, David M, Lauret P (2018) Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. Renew Energy 123:191–203. https://doi.org/10.1016/j.renene.2018.02.006

Pulley S, Collins AL (2018) Tracing catchment fine sediment sources using the new SIFT (SedIment Fingerprinting Tool) open source software. Sci. Total Environ 635:838–858. https://doi.org/10.1016/j.scitotenv.2018.04.126

Pulley S, Foster I, Antunes P (2015) The application of sediment fingerprinting to floodplain and lake sediment cores: assumptions and uncertainties evaluated through case studies in the Nene Basin, UK. J Soils Sediments 15:2132–2154. https://doi.org/10.1007/s11368-015-1136-0

Reiffarth DG, Petticrew EL, Owens PN, Lobb DA (2016) Sources of variability in fatty acid (FA) biomarkers in the application of compound-specific stable isotopes (CSSIs) to soil and sediment fingerprinting and tracing: A review. Sci. Total Environ 565:8–27

Reiffarth DG, Petticrew EL, Owens PN, Lobb DA (2019) Spatial differentiation of cultivated soils using compound-specific stable isotopes (CSSIs) in a temperate agricultural watershed in Manitoba, Canada. J Soils Sediments 19:3411–3426. https://doi.org/10.1007/s11368-019-02406-3

Smith HG, Karam DS, Lennard AT (2018) Evaluating tracer selection for catchment sediment fingerprinting. J Soils Sediments 18:3005–3019. https://doi.org/10.1007/s11368-018-1990-7

Stock BC, Jackson AL, Ward EJ et al (2018) Analyzing mixing systems using a new generation of Bayesian tracer mixing models. PeerJ 2018:. https://doi.org/10.7717/peerj.5096

Swales A, Gibbs MM (2020) Transition in the isotopic signatures of fatty-acid soil biomarkers under changing land use: Insights from a multi-decadal chronosequence. Sci. Total Environ 722:. https://doi.org/10.1016/j.scitotenv.2020.137850

Upadhayay HR, Bodé S, Griepentrog M et al (2017) Methodological perspectives on the application of compound-specific stable isotope fingerprinting for sediment source apportionment. J Soils Sediments 17:1537–1553

Upadhayay HR, Bodé S, Griepentrog M et al (2018a) Isotope mixing models require individual isotopic tracer content for correct quantification of sediment source contributions. Hydrol Process 32:981–989. https://doi.org/10.1002/hyp.11467

Upadhayay HR, Griepentrog M, Bodé S et al (2020) Catchment-wide variations and biogeochemical time lags in soil fatty acid carbon isotope composition for different land uses: Implications for sediment source classification. Org Geochem 146:. https://doi.org/10.1016/j.orggeochem.2020.104048

Upadhayay HR, Smith HG, Griepentrog M et al (2018b) Community managed forests dominate the catchment sediment cascade in the mid-hills of Nepal: A compound-specific stable isotope analysis. Sci. Total Environ 637–638:306–317. https://doi.org/10.1016/j.scitotenv.2018.04.394

Vale S, Swales A, Smith HG et al (2022) Impacts of tracer type, tracer selection, and source dominance on source apportionment with sediment fingerprinting. Sci. Total Environ 831:. https://doi.org/10.1016/j.scitotenv.2022.154832

Walling DE (2005) Tracing suspended sediment sources in catchments and river systems. Sci. Total Environ 344:159–184. https://doi.org/10.1016/j.scitotenv.2005.02.011

Walling DE (2013) The evolution of sediment source fingerprinting investigations in fluvial systems. J Soils Sediments 13:1658–1675

Xu Z, Belmont P, Brahney J, Gellis AC (2022) Sediment source fingerprinting as an aid to large-scale landscape conservation and restoration: A review for the Mississippi River Basin. J Environ Manage 324:116260. https://doi.org/10.1016/j.jenvman.2022.116260

Chapter 3

Using stable carbon isotopes of lignin-derived methoxy to improve historical apportionments of particulate organic matter and sediment sources incorporating multiple Suess corrections

Journal of Soils and Sediments

https://doi.org/10.1007/s11368-024-03765-2

Received: 13 November 2023 / Accepted: 24 February 2024

Using stable carbon isotopes of lignin-derived methoxy to improve historical apportionments of particulate organic matter and sediment sources incorporating multiple Suess corrections

****

Terry Cox[1], Patrick Laceby[2], Markus Greule[3], Frank Keppler[3,4], Christine Alewell[1]

Corresponding authors email: Terry.cox@unibas.ch

[1] Department of Environmental Sciences, University of Basel, Bernoullistrasse 30, 4056 Basel, Switzerland

[2] Environment and Protected Areas, Government of Alberta, 3535 Research Rd NW, Calgary, AB, Canada, T2L 1Y1

[3] Institute of Earth Sciences, Heidelberg University, Im Neuenheimer Feld 234-236, D-69120 Heidelberg, Germany

[4] Heidelberg Center for the Environment (HCE), Heidelberg University, Germany

****

Abstract

Purpose

Soil erosion models are essential to improving sediment management strategies. Sediment source fingerprinting is used to help validate erosion models. Fingerprinting sediment sources with organic isotopic tracers faces challenges from aquatic sources and co-linearity. To address these complexities, integrating another land-use-specific tracer is essential. Suess corrections incorporating multiple mean-residence-times are necessary to accurately model historical sediment apportionments. In previous studies, compound specific isotopic tracers indicated forest as the dominant source. We hypothesise there is an overestimation of forest contribution, attributed to the misclassification of particulate organic matter as forest.

Methods

In this study, we utilise stable carbon isotope ($\delta^{13}C$) values of fatty acid and the average chain length in combination with the $\delta^{13}C$ values of lignin derived methoxy groups as an additional tracer. We apply different Suess corrections to explore the effect of the changing atmospheric $\delta^{13}CO_2$ values on sediment apportionment. The performance of the unmixing model is evaluated with 300 mathematical mixtures. To determine shifts in sediment sources throughout the last 130 years, particulate organic matter contributions are determined, and removed to apportion sediment soil sources. We investigate the potential misclassification of forest contributions by merging particulate organic matter and forest sources to simulate tracers which are unable to discriminate.

Results

The inclusion of $\delta^{13}C$ values of lignin methoxy groups and the alkane average chain length as additional tracers successfully removed tracer co-linearity. Additionally, we used an updated concentration dependent point in polygon test to identify sediment with increased potential for incorrect source apportionments. Changes in the dominant sediment sources over time (Forest: pre-1990, Pasture: 1910-1940, Arable: post 1940) highlight the effect of policy induced land-use changes. Additionally, the inability to discriminate particulate organic matter and forest sources was revealed to cause a 37 % over estimation of forest contributions from 1944-1990.

Conclusion

Using $\delta^{13}C$ values of lignin methoxy groups as an additional tracer, we identified critical points in the 130-year sediment history of Lake Baldegg. Furthermore, we highlight the importance of incorporating multiple Suess effects. Through mathematical mixtures, we assessed the confidence that should accompany apportionment estimates. While merging forest and particulate organic matter sources did not result in forest as the dominant source over the last 130 years, but separating these sources resulted in more accurate apportionment. These insights offer valuable information to enhance the accuracy of sediment fingerprinting, which can then be used to assist soil erosion models employed for sediment mitigation policies.

Keywords

Sediment Tracing ·sediment apportionment · Swiss Lake Baldegg · MixSIAR · CSSI · Suess corrections ·

3.1 Introduction

Soil erosion and sedimentation are recognized globally as a critical problem (Pimentel 2006). On-site impacts of soil erosion include the degradation of soil structure (Zhang et al. 2007), depletion in soil carbon and nutrients (Bashagaluke et al. 2018), and a reduction in agricultural productivity (Bakker et al. 2007). However, off-site impacts (e.g., sedimentation), can be equally detrimental, leading to pollution and eutrophication of fresh and ocean waters (Zamparas and Zacharias 2014). The accurate identification and apportionment of sediment sources is necessary for developing targeted and effective sediment mitigation strategies to ensure the long-term sustainability of aquatic ecosystems and land resources (Collins and Walling 2004; Walling 2005; Owens et al. 2016).

The sediment fingerprinting method uses various characteristics of the soil and sediment to fingerprint possible sources, tracking back to potential land-uses, vegetation cover or geological bedrock. However, classic fingerprinting approaches using geochemical (Batista et al. 2019) or fallout radionuclide tracers (Evrard et al. 2013), are often not suitable to provide information on land-use-specific sources. Compound specific stable isotopes (CSSI) of carbon of long chain fatty acids (LCFA) and alkanes have been used to apportion the relative contribution of different land-uses (Gibbs 2008; Alewell et al. 2016; Upadhayay et al. 2018; Lavrieux et al. 2019).

The linear relationship between ^{13}C depletion and LCFA elongation during biosynthesis occurs similarly in all land-uses, often causing multiple ^{13}C LCFA tracers to have co-linearity, resulting in a problematic one-dimensional mixing line (Alewell et al. 2016; Lavrieux et al. 2019; Cox et al. 2023). In particular, Cox et al. (2023) demonstrated that one-dimensional mixing lines have the potential to cause inaccurate apportionment estimates due to misclassification of contributions of the central source(s) as contributions from sources at either endpoint. Additionally, Cox et al. (2023) demonstrated that additional co-linear tracers decrease the performance of the un-mixing model. Lavrieux et al. (2019) clearly illustrated the problematic mixing line when using $\delta^{13}C$ values of LCFA (C24-C28) as tracers at Lake Baldegg (Canton Luzern, central Switzerland). The $\delta^{13}C$ LCFA one-dimensional mixing line prevented meaningful sediment source estimates with the $\delta^{13}C$ LCFA of forest and grassland source fingerprints plotting at either endpoint of a mixing line, with orchards and arable sources plotting on a line in between these two endmember sources. Furthermore, sediment $\delta^{13}C$ $FA_{24,26}$ values were found to be extremely depleted in ^{13}C and outside the range of the source values potentially indicative of an unknown source.

Lavrieux et al. (2019) hypothesized that the missing source was of aquatic origins. The $\delta^{13}C$ FA_{28} values of source soils bracketed all sediment $\delta^{13}C$ FA_{28} values except 1956 and 1965. 1956 coincided with evidence of large turbidites in Lake Baldegg (Lotter et al. 1997), and 1965 had the highest TOC and TN levels over the 130-year record suggesting primary production was at its maximum (Lotter et al. 1997). Considering, the remaining sediments are within the $\delta^{13}C$ FA_{28} value source range, Lavrieux et al. (2019) presumed that the unknown source did not significantly contribute to the $\delta^{13}C$ FA_{28} sediment fingerprint for the majority of years. Nonetheless, for accurate land-use-specific source apportionment in Lake Baldegg, it is evident that the incorporation of an additional non-aquatic and land-use-specific tracer is necessary to expand the mixing line into a suitable mixing polygon. In this context, we propose the incorporation of $\delta^{13}C$ values of lignin-derived methoxy groups ($\delta^{13}C$ LMeO) as a promising novel tracer.

Lignin and its copper oxidation products have been employed to discriminate sediment sources (angiosperms vs gymnosperms, terrestrial vs marine, and land-uses) (Goñigoñi et al. 1998; Goñi et al. 2000; Kuzyk et al. 2008; Rezende et al. 2010). However, the application of lignin copper oxidation products has limitations. These include reduced source discrimination (Thevenot et al. 2010), modifications to the monomer composition during degradation (Dümig et al. 2009), and fractionation occurring during soil phase transitions (Hernes et al. 2007). Methoxy groups (MeO, molecular formular: CH_3O) comprise approximately 10-25% of lignin monomers and 2.5% of the total carbon content in the terrestrial biosphere (Galbally and Kirstine 2002). MeO predominantly originate from lignin (ether bound MeO groups) and pectin (ester bound MeO groups). While pectin, a polysaccharide, can be found in some algae species (Domozych et al. 2014), lignin is exclusively found in higher terrestrial plants (Boerjan et al. 2003). The selective removal of ester bound pectin MeO by alkaline hydrolysis allows for the analysis of the remaining terrestrial derived lignin MeO (LMeO) through the conversion of the residual LMeO groups to methyl iodide (MeI) using the Zeisel method (Zeisel 1885).

Two prerequisites are required for a successful fingerprinting tracer: the ability to discriminate between land-uses (e.g., forest, arable, pasture), and secondly, to have predictable or no modification of tracer signature from source to sink (e.g., conservative behavior) (Motha et al. 2002; Koiter et al. 2013; Belmont et al. 2014; García-Comendador et al. 2023). LMeO display not only large ^{13}C discrimination between C_3, C_4 and CAM plants (Keppler et al. 2004), but are also potentially appropriate for land-use-specific source discrimination between C_3 vegetation. The conservativeness of LMeO has yet to be explored fully for fingerprinting applications. Litter bag experiments in soils measuring bulk $\delta^{13}C$ MeO (pectin and lignin MeO) indicated small isotopic fractionation during degradation (Anhäuser et al. 2015), however, this was reasoned to be due to the preferential degradation of pectin. Analysis of the bulk MeO $\delta^{13}C$ continuum from wood to coal indicated significant fractionation during degradation, with Rayleigh coefficients of -15 ‰ being calculated. Considering we would not expect LMeO degradation on the same scale as the wood-to-coal continuum (remaining MeO fraction in coal ca. 10^{-4}) (Lloyd et al. 2021), we would expect isotopic fractionation to be minimal within LMeO concentrations in soil and sediments.

The utilization of CSSI tracers to apportion sediment sources has the implicit assumption that the sediment isotopic fingerprint originates from the input and mixing of the isotopic fingerprint of the possible terrestrial source soils (i.e., mineral horizons). However, this presupposes the absence of CSSI tracers from additional sources present in the sediment. While lake sediments are largely comprised of clastic materials (i.e., clay, silt and sand), they also contain an organic fraction in the form of mineral associated organic matter (MAOM), plant debris (terrestrial particulate organic matter- POM_{terr}) and particulate organic matter from aquatic sources (POM_{aq}). Since the primary goal of sediment fingerprinting is to provide information for soil erosion mitigation policies, the main aim of using CSSI tracers is the apportionment of the soil MAOM fraction and the subsequent conversion to soil proportions using concentration dependency. Thus, any contribution of the sediment CSSI signal from POM_{terr} or POM_{aq} will result in erroneous MAOM sediment source attribution.

POM_{aq} is known to be an important factor in lacustrine biochemical cycles and contributes to sediment organic matter (Xu et al. 2019; Wynants et al. 2021). However, to accurately apportion sediment sources, tracers are specifically used which are not substantially present in POM_{aq} (e.g., long chain fatty acids and lignin derived methoxy groups). However, non-MAOM sources of carbon (and CSSIs) such as POM_{terr}, can enter the watercourses through aeolian transport, leaching and surface wash-off during rain events. As suggested by Wiltshire et al. (2022), the inclusion of POM_{terr} in the sediment isotopic fingerprint might cause the over estimation of forest sediment input. While the physical separation of POM and MAOM has been achieved previously by density centrifugation (Cui et al. 2016), relatively large amounts of sample are required (5 g). Distinguishing between MAOM and POM_{terr} as separate endmembers using $\delta^{13}C$ LCFA is difficult due to the small isotopic difference (Wiltshire et al. 2022). Compared to LCFAs, LMeO have a much larger $\delta^{13}C$ discrimination between woody material and leaves (ca. -20 ‰) (Keppler et al. 2004), potentially resulting in a larger difference of $\delta^{13}C$ between MAOM and POM_{terr}. Although multiple studies utilize proxies, ratios and tracers to determine the POM_{aq} and POM_{terr} contribution to the sediment (Derrien et al. 2017), there are currently no studies attempting

to disentangle the POM_{terr} and MAOM contribution and subsequent land-use-specific apportionment of MAOM.

Currently in sediment fingerprinting research, there is extensive literature dedicated to tracer selection (Laceby et al. 2015; Smith et al. 2018; Batista et al. 2019; Vale et al. 2022; Cox et al. 2023). Despite the wealth of information surrounding this initial stage, the subsequent tracer selection steps often follows a reductionist approach, systematically eliminating tracers through a series of selection tests. Beyond whether to choose CSSI, geochemical or radionuclides tracers, there appears to be a lack in literature on additive tracer selection in which tracers are added for targeted discrimination. We propose a paradigm shift towards a more purpose-oriented tracer selection strategy, in which tracers are intentionally incorporated to serve a specific discrimination objective (Cox et al. 2023). Here, we use the average alkane chain length (ACL, 21-33 C odd alkanes). While the ACL may not be able to discriminate between specific land-uses (Wiltshire et al. 2022), ACL can be used as an additional tracer to enhance the discrimination of POM_{terr} and MAOM in the sediment. Moreover, the extraction and analysis of alkanes often coincide with CSSI analysis, providing readily available data for retrospective and future studies.

The use of CSSI $\delta^{13}C$ tracers for historical sediment apportionment requires Suess correction (i.e., correcting the isotopic values for the changing atmospheric $\delta^{13}CO_2$ composition), during the last 100 years due to anthropogenic fossil fuel burning (Verburg 2007; Gibbs et al. 2014). However, in sediment fingerprinting, Suess corrections are often either omitted under the assumption that variability induced by the Suess effect is negligible when compared to the source uncertainty (Brandt et al. 2018), or a single mean residence time (MRT) is used (Bravo-Linares et al. 2020). To accurately model the possible Suess effect on isotopic tracers, a range of MRTs should be assessed. The MRTs of FAs varies from decades to millennia (Lützow et al. 2006; Wiesenberg et al. 2010). While the MRT of LMeO has not been investigated, the MRT of lignin has been estimated to be comparatively shorter than FAs (5-26 years for pasture, 9-38 years for arable soils) (Heim and Schmidt 2007). As the exact MRT for each isotopic tracer is dependent on the soil ecosystem and soil properties (Schmidt et al. 2011), to accurately model the possible Suess effects, three MRTs (10 yr, 30 yr, and 100 yr) were considered when correcting for the Suess effect.

Point in polygon/mixing space tests (e.g., range/bracket tests) are a standard procedure in sediment fingerprinting (Collins et al. 2020). Polygons (mixing spaces) are drawn around the sources to identify sediment values which are located outside the mixing space. Sediment fingerprints outside the polygon then indicate that with the current sources (presuming no modification of the tracer), the mixture is highly implausible or even impossible. However, the sediment isotopic fingerprint is known to be impacted not only by the proportions of each source and the sources isotopic value, but also the concentration of the CSSI in each source (Upadhayay et al. 2018) and as such, this should be reflected in the possible mixing space.

In this study, first we test the accuracy and precision of $\delta^{13}C$ LMeO, ACL and $\delta^{13}C$ FA_{28} as a tracer set using mathematical mixtures. Second, the sediment contribution from the main land-uses was then

estimated for the last 130 years for a sediment core from Lake Baldegg. Third, we explore the impact of MRTs on Suess corrections, by comparing three different MRTs for both isotopic tracers. Fourth, we test the hypothesis that high forest contribution in sediment fingerprinting might be a result of the misclassification of POM_{terr}, if POM_{terr}, is not considered as a separate source. For this, we group forest and POM_{terr} into a single source to simulate a tracer which cannot discriminate between these sources and compare against the ungrouped results.

The objective of this study is to advance our understanding of CSSI sediment fingerprinting by addressing the challenges associated with these tracers, including the presence of one-dimensional mixing lines, aquatic and POM_{terr} sources of organic tracers, accounting for tracer MRTs when applying Suess, and the improbable high contribution of forest soils to sediment. This study aims to enhance the accuracy of land-uses specific unmixing models and ultimately facilitate better techniques for monitoring sediment erosion mitigation strategies. This will be accomplished by i) introducing an additional land-use-specific tracer, ii) incorporating multiple Suess effects, and iii) investigating the influence of POM_{terr} on apportionment estimates. This study aims to improve land-use specific sediment source apportionment using CSSI tracers to aid targeted sediment management interventions. This approach is particularly valuable for ensuring the maintenance of both soil health and water quality and can be applied to regions of similar temperate climate and agricultural practices.

3.2 Methods

Study site

Lake Baldegg is a glacial lake located in the central Swiss plateau situated 463 m above sea level (coordinates: 47◦1200 00 N, 8◦1504000 E). It extends from north to south with a maximum depth of 66 m and mean depth of 33 m. The Lake covers a surface area of 5.2 km^2 and holds a volume of 0.173 km^3. The water residence time of Lake Baldegg has been estimated to be 4.3 years, with the outflow located at its northern end. The catchment area surrounding the lake, excluding the lake surface area, spans 68.2 km^2 (Wehrli et al. 1997). On the year of source sampling (2016), 77% of the catchment was being used for agriculture and 12% was covered by forest and 5% urbanized (Lavrieux et al. 2019).

Over the last century, Lake Baldegg has shown an extensive history of eutrophication (Lotter et al. 1997). Since the beginning of the 20^{th} century, the discharge of untreated waste and agricultural runoff into the lake has increased due to population growth and the intensification of agriculture. The rise of nutrients led to an increase in primary production from 1885 to 1970. The geographical setting of Lake Baldegg prevents water column mixing and resulted in permanent anoxic conditions in the hypolimnion. The anoxic conditions prohibited bioturbation and resulted in seasonal material being preserved in laminated varved layers which have been used for core dating (Lotter et al. 1997; Wehrli et al. 1997). Additionally, the sediment core contained multiple turbidites. Most turbidites originate in the Baldegg catchment and are transported into the lake during heavy rainfall or flood events (Lotter et al. 1997).

Sampling

In this study, we reanalyzed source samples presented in Lavrieux et al. (2019) and POM_{terr} (leaves, pine needles, and the Oi organic horizon) from Hirave et al. (2020). Samples were stored dry at 4 ºC in a sample archive (Department of Environmental Sciences, University of Basel). Using a modified sediment connectivity index (Borselli et al. 2008), Lavrieux et al. (2019) selected source sample locations based on high connectivity and representing the main land-use types in the catchment (arable lands, permanent grasslands, temporary grasslands, forests, orchards). Importantly, only A horizons (mineral soil) were collected from each source sample location and sieved to 2 mm. In the case of forest samples, plant debris was removed from source samples.

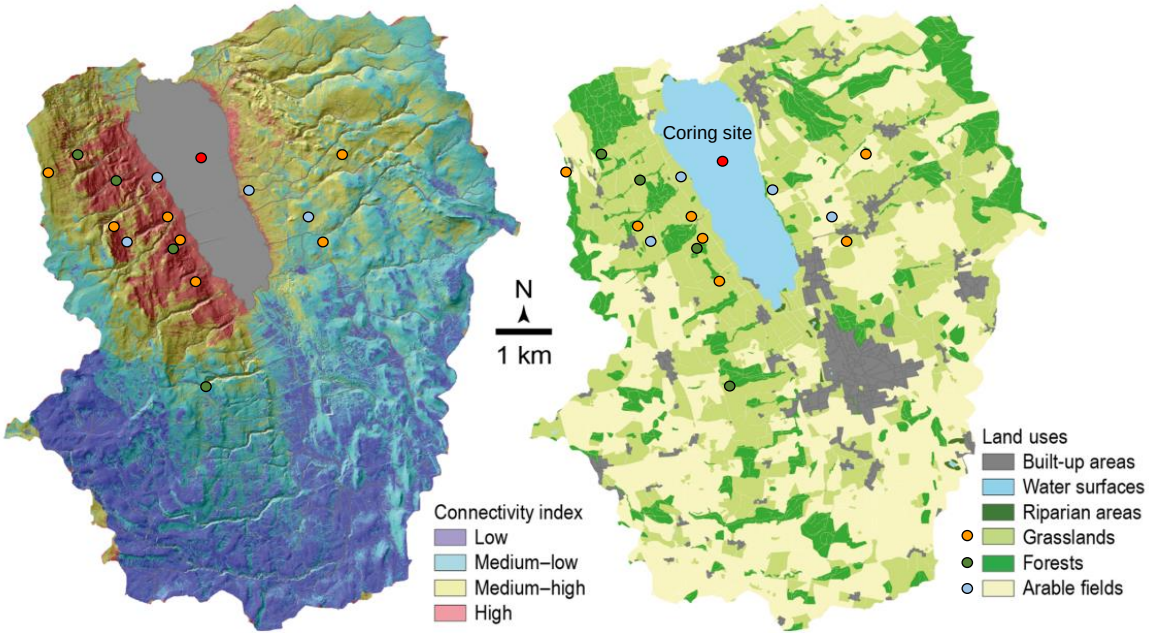


Fig. 1 Connectivity and land-use of the Lake Baldegg (Canton Luzern; Central Switzerland) catchment with sampling sites marked from Lavrieux et al. (2019) with source soil and coring sites marked with the 'O' symbol

Subsamples from the sediment core Ba-09-03 (Lavrieux et al. 2019) were reanalysed. As described in Lavrieux et al. (2019), the varved sediment allowed seasonal bi-annual dating back to 1885. Subsamples had a 3 cm thickness (Table S2) and the mean age is reported throughout. For further core sub sampling information see Lavrieux et al. (2019). Details on sampling the sediment core, as well as the full core age–depth model can be found in van Raden (2012) and Kind (2012).

Usually, sediment source attributions are limited to relative source contributions. To convert these relative contributions to absolute sediment accumulated from each source, the mean sediment accumulation rate (SAR) was taken as an average of sediment cores BA93-C and BA93-BA extracted from Lotter et al. (1997) who determined the mean SAR by varve counting and measuring the thickness of sediment deposit annually. The SAR was then used to convert source proportions to annually deposited sediment.

For improved source discrimination, temporary and permanent grassland were grouped into one pasture source. Although at the time of sampling, orchards represented only 1% of the catchment, aerial photographs demonstrate orchards to be a more significant land cover in the catchment since 1950. However, orchard source samples taken by Lavrieux et al. (2019) are from modern orchard plantations, which consist of highly dense fruit trees in rows. Historically, fruit trees have been grown in pastures or meadow fields at low density and have isotopic signatures that are similar to pastures or meadows. As such, the orchard source samples taken by Lavrieux et al. (2019) are not representative of the orchard land cover for the last 130 years and therefore have not been included as a sediment source.

Laboratory analysis

Sample preparation

Adapting the method of Greule and Keppler (2011), ester bound MeO groups (predominantly of pectin origin) were removed from the sediment and source soils by the conversion to methanol by alkaline hydrolysis. Briefly, 1 ml of 1 M NaOH was added to approximately 150 mg of soil/sediment in 1.5 ml vials, samples were capped and heated at 90 ºC for 4 hrs. Samples were then uncapped and dried at 60 ºC in a sand bath. To ensure hydroiodic acid (HI) is not used up by the neutralization of NaOH or the removal carbonates, 100 µl deionized water was added to each sample followed by acid fumigation for 24 hrs in 37 % fuming hydrochloric acid (HCl). Samples were then dried again at 60 ºC in a sand bath. Residual LMeO groups (ether bound) were converted to MeI by the addition of 500 µl HI 57 % (Sigma Aldrich, Stabilized) and heated for 1 hr at 130 ºC. Samples were left to equilibrate at room temperature for 1 hour before GC-FID and GC-IRMS analysis (Greule et al. 2009).

Lignin methoxy concentration

The analysis of MeI was conducted using static headspace analysis, adapting the procedure outlined in Greule et al. (2009). A headspace volume of 10-90 µl was manually injected (Hamilton, 100 µl, gas tight, side-port) into the Trace Ultra gas chromatograph (GC) with a flame ionization detector system (FID; Thermo Scientific, Walthalm, MA 02451, USA). Conditions of the GC-FID were set at 200 ºC inlet temperature, 1.8 ml/ min He flow rate, and an isothermal oven temperature at 65 ºC resulting in an approximate elution time of 4 minutes. The MeI analyte was quantified using an external calibration curve generated from compounds with known MeO concentrations (i.e., Vanillin, (Sigma Aldrich, 99 %) and HUGB3 (Greule et al. 2020)). The concentration of source samples was analysed with triplicate samples with single injection. Due to minimal sediment sample availability (ca. 200 mg), the concentration of sediment samples was analysed by triplicate injection of a single sample.

The calibration of MeO quantification ranged from 0.001 mg g^{-1} (dry soil weight) to 0.5 mg g^{-1} MeO, with a $r^2 > 0.97$ (Pearson's correlation) for all analysis. Deviation from a linear relationship was not found between concentration and counts using static headspace injection as reported by Lee et al. (2019) and as such, we found injection of head space a simple and effective methodology of quantification for soil content of LMeO.

Lignin methoxy isotopic values

The $\delta^{13}C$ isotopic composition of MeI was determined using a Trace Ultra GC instrument interfaced online through a GC-Isolink to a Conflo IV and Delta V Advantage isotope ratio mass spectrometer (Thermo Scientific, Walthalm, MA 02451, USA) via an oxidation reactor. A reduction stage was fitted to remove potential corrosive contamination (Feakins et al. 2013) (see Fig. S1 for instrument schematic). The inlet temperature was set to 200 ºC, the column flow rate (He) was 1.8 ml min^{-1}. The initial oven temperature was set to 30 ºC for 3.8 min, ramp at 30 ºC per minute until 100 ºC.

In this paper, all stable carbon isotope ratios are expressed in the conventional 'delta' δ notation, meaning the relative difference of the isotope ratio of a substance compared to the standard substance Vienna Peedee Belemnite (VPDB). Reference materials HUGB1 ($\delta^{13}C_{V\text{-}PDB}$ = -50.17 ± 0.08 ‰) and HUGB4 ($\delta^{13}C_{V\text{-}PDB}$ = -30.07 ± 0.10 ‰) were used for the normalization of sample isotopic ratios (Greule et al. 2019, 2020). HUGB3 ($\delta^{13}C_{V\text{-}PDB}$ = -29.30 ± 0.10 ‰) was treated identically to samples and used as a quality control sample throughout the run.

Triplicate injections of MeI were used to determine the instrumental uncertainty of isotopic analysis. We found the mean instrumental precision (1 SD) to be 0.16 ‰ over all sequences. Single injection of triplicate source samples had a mean SD of 0.3 ‰. Using a two-point calibration of HUGB1 and HUGB4, and HUGB3 as a quality control sample throughout. HUGB3 was shown to have a RSME of 0.4 ‰. For information on the FA and ACL method and analytical precision see Hirave et al. (2020) and Lavrieux et al. (2019).

Suess corrections

To account for a time lag between atmospheric changes and changes in the soil, different MRTs were compared (10 yr, 30 yr and 100 yr). Following an identical methodology to Lavrieux et al. (2019) isotopic values of both soils and sediments were corrected using the atmospheric CO_2 curve of Verburg (2007) for multiple MRTs. Eq. (1) was applied to correct isotopic values to the pre-industrial era (1840).

$$\delta^{13}C\ Soil_{(t+1)} = \left(1 - \frac{1}{R}\right)\delta^{13}C\ Soil_{(t)} + \frac{1}{R}(\delta^{13}C\ Soil_{(t0)} + (\delta^{13}C\ Atmosphere_{(t+1)} - \delta^{13}C\ Atmosphere_{(t0)})) \quad \text{Eq.1}$$

We assigned the variable 't' as the year of observation, with 't_0' corresponding to the year 1840. 'R' represents the MRT of the isotopic tracer in years (10 yr, 30 yr and 100 yr). We assumed that there were only marginal changes in the $\delta^{13}C$ values of atmospheric CO_2 prior to 1840. The calculated change in $\delta^{13}C$ values after 1840 are based on the atmospheric data from Verburg (2007). We also maintained the assumption of stability in the soil organic carbon pool size over time and no changes in the isotopic fractionation during CO_2 uptake due to increased atmospheric CO_2 concentrations.

Defining a mixing space and sediment classification

Point in polygon/ mixing space tests (e.g., range/bracket tests) are standard practice in sediment fingerprinting (Collins et al. 2020). They are designed to identify mixtures which are located outside of the possible mixing space of the sources, which suggests the mixture is not possible from the current sources and will result in erroneous unmixing. Although point in polygon tests are standard practices, they have yet to be adapted to concentration dependent tracers (e.g., CSSI). To define a concentration dependent mixing space, the open-source python script of Cox et al. (2023) was used to generate mathematical mixtures. For each Suess correction, 300 concentration dependent mathematical mixtures were created by weighting the mean source isotopic tracer values by their respective mean concentrations of source soils and multiplying by 300 randomly selected proportions with the condition that all proportions must sum to one.

A limitation of using the methodology of Cox et al. (2023) is that the source variation is not incorporated. We attempted to incorporate the source variation by expanding the mixing space to include all source samples. While MixSIAR uses the mean and SD, we suggest that even though highly improbable, sediment could be derived from an outlying source sample outside the SD of source value. Consequently, we have expanded to the mixing space beyond the SD to encompass all source values. As MixSIAR incorporates source variance, the mixing space is diffuse and lacks a precisely defined boundary. The determination whether a sediment is within the mixing space is probability based, and as such is non-binary. The position of the sediment fingerprint may be around the diffuse boundary of the possible mixing space (mixing horizon). Mixing horizons can be defined as the boundary in which the unmixing results are sensical. Beyond this horizon (in a three-source model of sources X, Y and Z), the model's calculation is dominated by a reduction in contribution X, leading to imputation of Y and Z source contributions to satisfy the condition that the sum of contributions should equal 1. Importantly, this is the case even when the sediment moves away from the mixing space and the X, Y and Z sources. This must be taken into consideration when interpreting unmixing results with the comparison of bi-plots.

Using bi-plots, we classified sediment values into three classes: A) inside the mixing space, B) located around the mixing space horizon in the expanded mixing space, and C) outside the expanded mixing space. Sediment are classified on their lowest class in all bi-plots (A→B→C). Class C sediments are excluded from interpretation due to a high probability of erroneous mixing results. The coherence of the apportionment of sediments in Class B are tested using a range test. The range test consists of determining if the sediment apportionment is within the range of the Class A sediments on either side (the years before and after). This method of evaluating coherence apportionment relies on the assumption that the transition of sediment sources occurs gradually, and the sediment being evaluated exhibits comparable source contributions to the bracketing samples which may not hold true for all cases. Sediment apportionments which show non-coherence are interpreted with caution and potentially excluded from interpretation.

Unmixing

Source apportionment was estimated using the open-source R package, MixSIAR (Stock et al. 2018). Using $\delta^{13}C$ of FA_{28}, LMeO and the ratio ACL (alkanes 21-31 C) as a tracer set, MixSIAR run with concentration dependency (mean concentrations of each source) and uninformative priors. As demonstrated by Lavrieux et al. (2019) only $\delta^{13}C$ FA_{28} source values bracketed sediment values (except 1965 and 1956), as such $\delta^{13}C$ $FA_{24,26}$ were excluded from the tracer set. Additionally, $\delta^{13}C$ FA source soils in the Baldegg catchment displayed co-linearities. The use of multiple tracers with co-linearities has been demonstrated to decrease model performance (Cox et al. 2023). Given that ACL represents a ratio, the ACL concentration was set to value of 1 for all sources to remove its concentration dependency. All MixSIAR runs used the same model parameters: chains = 3, chain length = 3,000,000, thin = 500, burn = 2,700,000 with a "very long" run time. The convergence of the mixing model was assessed by using the Gelman-Rubin diagnostic, with model output being rejected if variables scored >1.05. Due to minimal amount of sediment core samples available (ca. 200 mg), a single mixture of each sample is unmixed using the "process only" error structure, in which the variation in the mixtures is assumed to be fully dependent on the weighted source variation.

There are some often overlooked limitations when running the "process only" error structure in MixSIAR. Bayesian mixing models incorporate probability distributions for both sources and mixtures. In this study only a single sediment replicate was available to represent the mixture, as such we cannot empirically determine a probability distribution for each layer in a sediment core. In such cases, MixSIAR users rely on the "prosses only" error structure, where the model derives the probability distribution for the mixture from the variance of the sources (Smith et al. 2018). The intricate nature of complex mixing systems inevitably yields disparities in both source and mixture variances, and as such, the "process only" error structure has limitations in incorporating the natural variation in the sediment samples. Nevertheless, the "process only" error structure is frequently applied due to lack of sediment replicates or the need for higher resolution apportionment estimates, such as apportioning sediment sources for each layer in a sediment core.

Evaluation of unmixing performance

To evaluate the accuracy and precision of the model, mathematical mixtures created to determine the mixing space were unmixed in identical manner to sediment samples. Results of the estimated proportions of mathematical mixtures were then compared to the known mixture proportions. As advocated by Batista et al. (2022), the evaluation of MixSIAR's probabilistic output should be probabilistic rather than deterministic. As such, we use the continuously ranked probability score (CRPS) (Matheson and Winkler 1976). Model comparisons between tracer sets were evaluated using the continuously ranked probability skill score (CRP skill score) (Pedro et al. 2018; Cox et al. 2023).

Removal of the POM_{terr} fingerprint from sediment

Source apportionment estimates were then converted into g cm^{-2} y^{-1} deposited by mineral sources using Eq. (2). First, source proportions were corrected to reflect contribution from only mineral sources (i.e., the removal of the POM_{terr} fraction) by converting mean source estimates to the percentage of the total MAOM fraction (MAOM fraction = Forest % + arable % + pasture %). To determine the sediment accumulated per year by each source, the MAOM apportionments were then multiplied by the mean SAR extracted from Lotter et al. (1997). The uncertainty of the sediment accumulated per source was calculated by the standard deviation (SD) of the MixSIAR outputs being treated identically to the apportionment estimates using Eq. (3). Furthermore, it's essential to note that the exclusion of POM_{terr} was carried out using mean values, potentially introducing an unquantified source of additional uncertainty.

$$Sediment\ accumulated\ per\ year\ by\ source = (s\%/(F\% + A\% + P\%)) * SAR \quad \text{Eq.2}$$

$$SD\ of\ sediment\ accumulated\ per\ year\ by\ source = (sSD\ /(FSD + ASD + PSD)) * SAR \quad \text{Eq.3}$$

In this context, 's%' is the model estimates for either forest, arable or pasture (the source undergoing transformation). *F%, A%, P%* denotes the model estimates for all MAOM sources. SAR is extrapolated from Lotter et al. (1997). Since SAR were only available from 1885 to 1990, source apportionment estimates were only converted into g cm^{-2} y^{-1} between these years.

3. Results and Discussion

Isotopic fingerprints of soils, POM_{terr} and sediments

Mean soil $\delta^{13}C$ values of lignin methoxy groups ranged from -39.2 to -47.8 ‰. Arable MeO had the highest ^{13}C content ($\delta^{13}C$: -41.7 ‰, SD: 1.9 ‰) followed by forest ($\delta^{13}C$: -43.2 ‰, SD: 0.54 ‰) then pasture ($\delta^{13}C$: -44.1 ‰, SD: 2.0 ‰) (Table 1). Similar to Lavrieux et al. (2019) discussing $\delta^{13}C$ FA enrichment in arable soils, we suggest the enriched ^{13}C MeO arable signature can be reasoned to be a result of rotational crop system in Switzerland, with a legacy maize (C_4) signal contributing to the arable fingerprint. The observed similarities between forest and pasture can be explained by the dominance of C_3 species in both sources. This suggests that the $\delta^{13}C$ variance between C_3 plants in LMeO may be less than in FA. In general, $\delta^{13}C$ MeO values of the sediment fall within the source soil isotopic range (min-max ± SD) (Fig. 2, Table1). However, sediment from the years 1971 and 1958 are enriched in ^{13}C MeO compared to the source soils and outside the mixing space (Fig. 2). In these years, large turbidites are recorded in the sediment core representing a flood or earthquake event. Such events may trigger subsurface erosion of older deep soils potentially containing isotopically enriched LMeO. The $\delta^{13}C$ LMeO of POM_{terr} had a mean value of -52.5 ‰ (SD: 5.9 ‰), with maple leaves being ^{13}C depleted in comparison to the rest of the POM_{terr}. POM_{terr} $\delta^{13}C$ LMeO values are within range of $\delta^{13}C$ LMeO values

of leaves previously reported by Keppler et al. (2004). The $\delta^{13}C$ LMeO values of soils is novel and not available in the current literature. Our results report the $\delta^{13}C$ LMeO values of soil was between wood (-25 to -28 ‰) and tree leaves (mean ~-60 ‰, ranging from -40 to -77‰ including C_3 and C_4 plants) (Keppler et al. 2004; Greule et al. 2009, 2020). Considering that soil organic carbon is a mixture of both woody (roots and above ground woody material) and leaf material, our findings of soil $\delta^{13}C$ LMeO values of -39.2 to -47.8 ‰ are within a reasonable range.

Table 4 Summary of $\delta^{13}C$ values and concentrations of LMeO and FA_{28} for sources and sediment (see Table S1 for the full ungrouped dataset and Fig. S2 for PCA).

	$\delta^{13}C$ MeO (‰)		MeO [mg g^{-1}]		$\delta^{13}C$ FA_{28} (‰)		FA_{28} [mg g^{-1}]		ACL		SAR
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	g cm^{-2}
Source											
Forest	-43.2 (n=4)	0.54	1.53	0.36	-34.4(n=4)	0.63	0.12	0.03	28.7 (n=2)	0.1	-
Arable	-41.7 (n=4)	1.97	1.19	0.50	-35.5(n=4)	0.28	0.07	0.04	29.5 (n=2)	0.1	-
Pasture	-44.1 (n=7)	2.04	1.80	0.46	-36.5(n=7)	0.33	0.10	0.03	29.7 (n=4)	0.3	-
POM_{terr}	-52.5 (n=4)	5.83	28.15	9.53	-33.7(n=4)	0.69	0.02	0.03	27.1 (n=3)	0.6	-
Sediment											
2010	-45.8	0.37	0.50	0.01	-34.8	0.1	0.03	-	29.1	-	-
2000	-45.7	0.13	0.42	0.09	-34.7	0.3	0.04	-	29.0	-	-
1990	-47.7	0.82	0.44	0.00	-34.8	0.4	0.03	-	28.9	-	0.10
1983	-47.8	0.24	0.40	0.01	-34.6	0.3	0.04	-	29.2	-	0.12
1977	-45.6	0.19	0.48	0.03	-34.4	0.4	0.02	-	29.4	-	0.11
1971	-38.5	0.61	0.33	0.02	-36.1	0.4	0.04	-	28.7	-	0.15
1965	-43.9	0.34	0.61	0.07	-37.8	0.1	0.07	-	29.0	-	0.12
1958	-35.8	0.22	0.15	<0.01	-34.1	0.5	0.02	-	29.0	-	0.08
1951	-39.1	0.28	0.33	0.03	-34.3	0.2	0.02	-	29.2	-	0.09
1945	-45.7	1.31	0.24	0.01	-34.0	0.1	0.01	-	29.1	-	0.09
1939	-41.8	0.02	0.26	<0.01	-35.9	0.1	0.02	-	29.3	-	0.10
1933	-41.5	0.86	0.25	<0.01	-35.6	0.4	0.01	-	29.2	-	0.10
1927	-41.3	0.6	0.26	0.02	-35.9	0.1	0.02	-	29.2	-	0.11
1921	-40.4	2.87	0.33	0.01	-35.1	0.2	0.01	-	29.1	-	0.10
1915	-43.3	0.19	0.31	0.01	-36.3	0.4	0.02	-	29.1	-	0.10
1909	-44.0	0.13	0.32	0.01	-36.0	0.1	0.01	-	29.0	-	0.10
1903	-44.1	0.08	0.29	<0.01	-35.8	0.5	0.01	-	28.8	-	0.10
1899	-40.9	0.31	0.25	0.01	-33.6	0.5	0.01	-	28.8	-	0.15
1891	-45.2	0.6	0.22	0.00	-33.6	0.3	0.01	-	28.9	-	0.12
1885	-41.9	0.16	0.11	0.01	-34.7	0.3	0.01	-	28.8	-	0.20
1880	-47.5	0.17	0.44	0.03	-34.1	0.3	0.01	-	29.0	-	-

The $\delta^{13}C$ FA_{28} values of source soils bracket all sediment values apart from 1965 and 1958 (min-max ± SD). The inclusion of POM_{terr} expanded the range to bracket 1958, with only 1965 to not be within range. 1965 has been suggested to be the peak of eutrophication (Lotter et al. 1997), and as such may contain non-terrestrial derived FA_{28} as production of LCFAs of aquatic origin have been hypothesized to be produced in hyper-eutrophic lakes (Van Bree et al. 2018; Lavrieux et al. 2019). The large discrimination of POM_{terr} and MAOM using ACL indicates that ACL hold significant potential as a proficient tracer to aid with the estimation of POM_{terr} contribution to the sediment. Although only a fraction of the source samples was analysed for alkanes by Lavrieux et al. (2019) (Table 1), ACL from source soils presented here agree with ACL reported in literature (Cooper et al. 2015; Chen et al. 2016; Wiltshire et al. 2023).

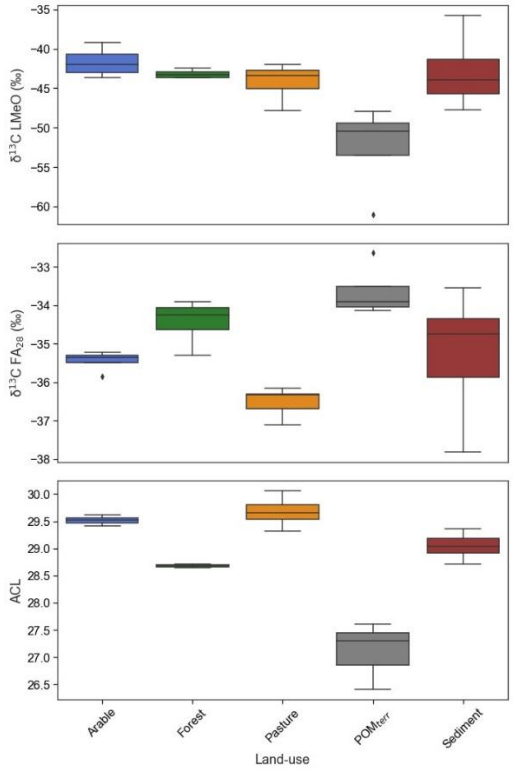


Fig. *2* Isotopic values of $\delta^{13}C$ LMeO (this study) and $\delta^{13}C$ FA_{28} and ACL (from Hirave et al. (2021) and Lavrieux et al. (2019)) of source soils, POM_{terr}, and sediment. The boxes represent 25, 50 and 75% quantiles with whiskers showing a 1.5 interquartile range

Defining a mixing space and Suess corrected sediment classification

Fingerprinting with $\delta^{13}C$ FA tracers often present a one-dimensional mixing line leading to potentially inaccurate sediment apportionment estimates. In our findings, incorporating $\delta^{13}C$ LMeO values as an additional tracer expands the FA mixing line presented in Lavrieux et al. (2019), creating a more suitable mixing space for adequate unmixing (Fig. 3). This expansion of the mixing space allows for the use of mixing models with significantly reduced misclassification of the central source as contribution from source(s) from either end point (Cox et al. 2023).

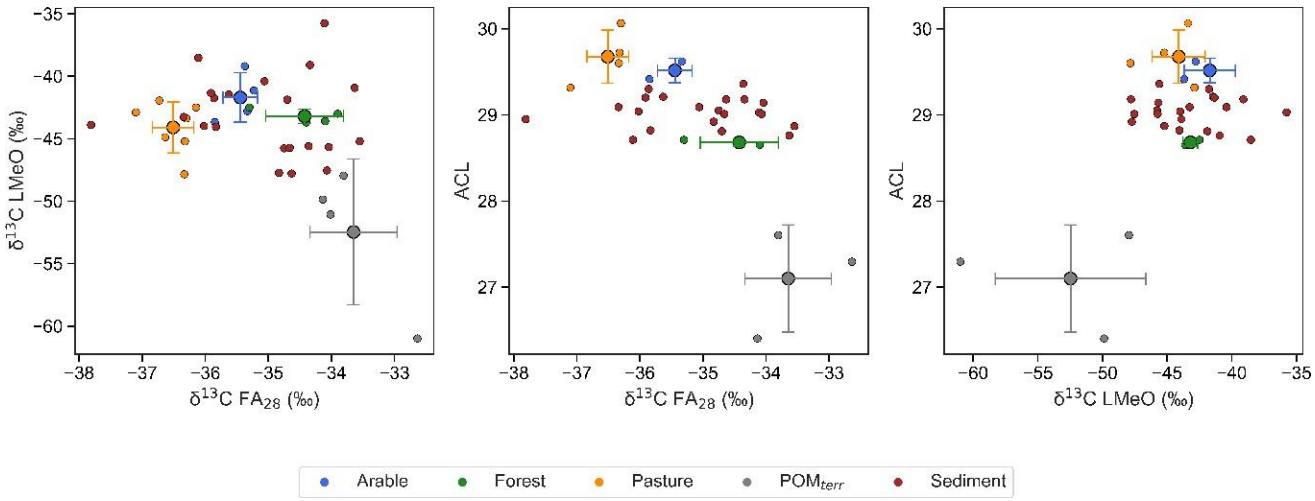


Fig. 3 Bi-plots of soil, POM_{terr} and sediment fingerprints illustrating the expansion of the mixing space using ACL and δ^{13}C LMeO tracers of non-Suess corrected values. Error bars are shown on the mean source values and calcauted as 1 SD of the source variation

Concentration dependent mathematical mixtures were used to define the mixing space and identify sediment samples outside the mixing space. The fingerprints of mathematical mixtures illustrated on Fig. 4 demonstrate how the mathematical mixtures are highly weighted by the concentration dependency of δ^{13}C LMeO values. Interestingly, the mixing space (green polygon-defined by the mathematical mixtures) is not limited to the convex hull of source means, suggesting traditional tests of conservatism should be questioned when concentration dependency is involved. Considering only the mean source values were used to generate the mixtures, we would assume the incorporation of source variation in the generation of mathematical mixtures would increase the mixing space (yellow polygon) and provide a more accurate mixing space of possible sediment fingerprints.

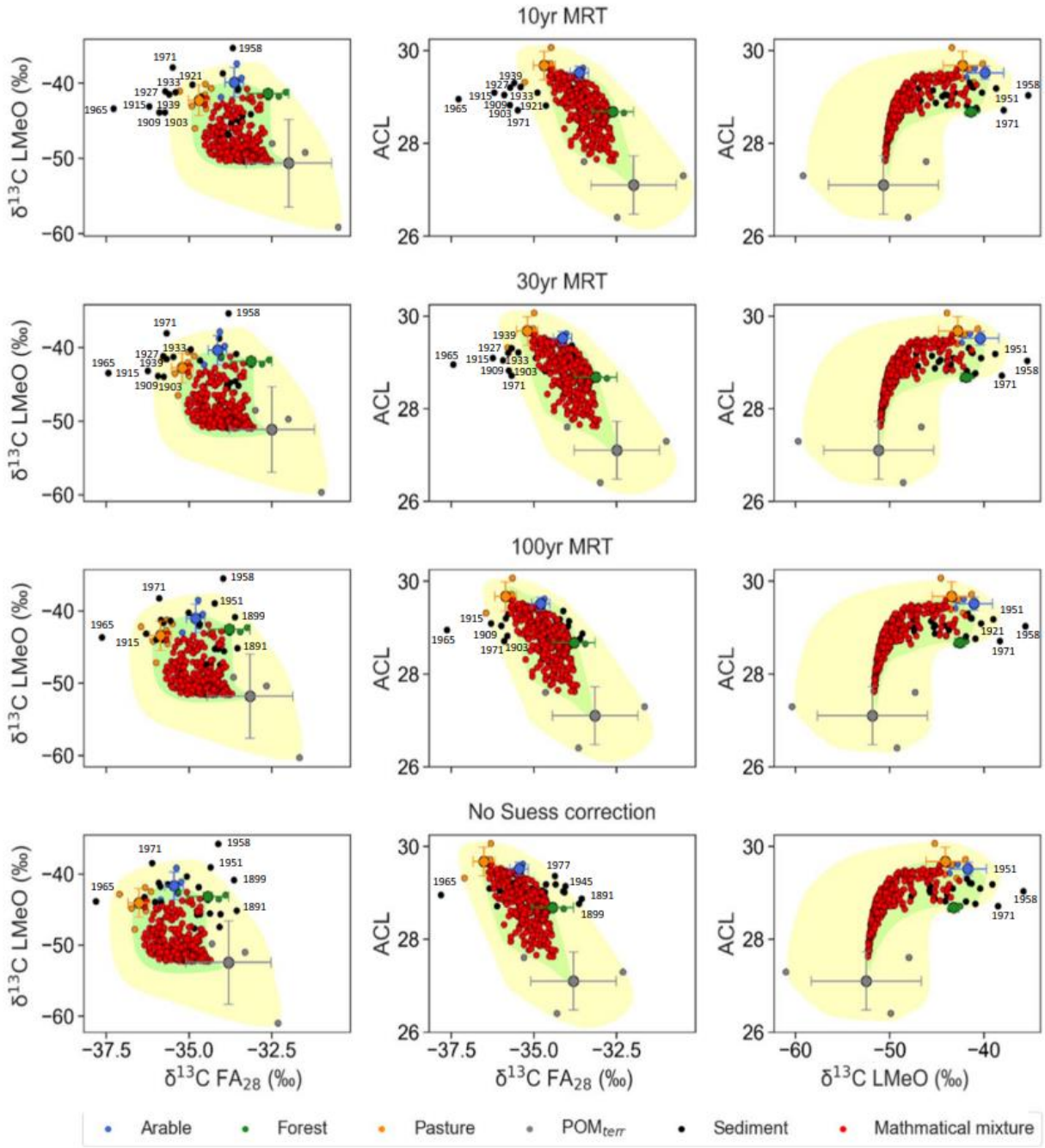


Fig. 4 Bi-plots of sources before and after Suess correction using three different MRTs (10, 30, and 100 yr). The mean mixing space is defined as the area around the location of the mathematical mixtures (green polygon). The green mixing space is then expanded to incorperate all source samples (yellow polygon). Full labeled bi-plots are available in supplementary (Fig. S3, Fig. S4, Fig. S5 and Fig. S6)

Isotopic fingerprints of soils, POM_{terr} and sediment were corrected for the Suess effect using thee different MRTs (Table S2 and S3). Suess corrected sediment samples are classsifed into the following groups: sediments are within the mathmatical mixture mixing space (Class A, green), located around the mixing horizon (Class B, orange), or outside the expanded mixing space (Class C, red) (Fig. 5). Without any Suess correction,14 % (1945,1951 and 1977) are in Class B (around the mixing horizon)

and 24% of sediment values are in Class C (outside the mixing space, Fig. 5). Results of the 10yr MRT show 14 % of sediments in Class B and a higher number of sediment values in Class C (38 %). The 30yr MRT corrected sediments had 19 % of samples in Class B and 24 % in Class C. The 100yr MRT corrections have a higher number of sediment values within the mixing space with 29 % in Class B and only 14 % being in Class C, and. All Suess corrections depict the 1971,1965 and 1958 to be outside the mixing space (Class C). 1970 and 1958 and coincide with high SAR (Lotter et al. 1997). 1965 was suggested to be the peak of eutrophication and as such, may contain significant concentrations of non-terrestrial material (Lotter et al. 1997).

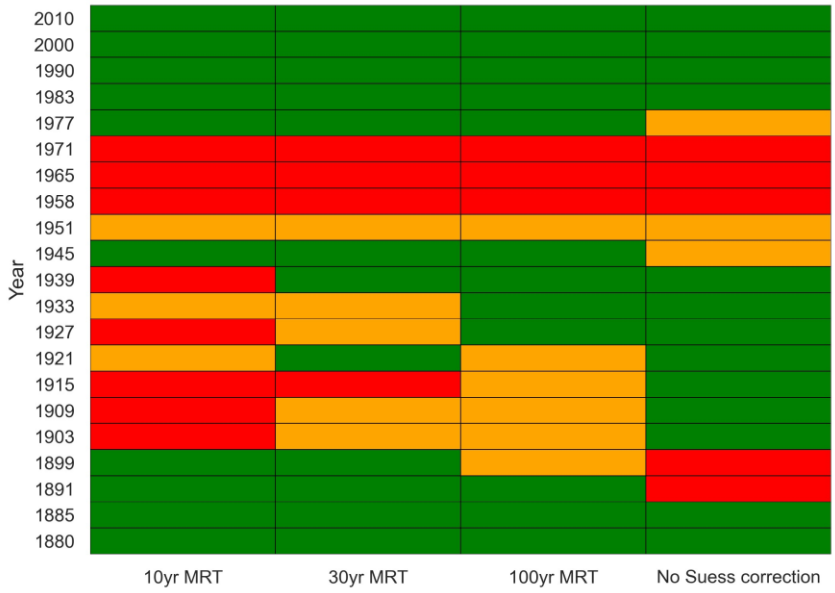


Fig. 5 The identification of Suess corrected sediment samples which are within the mixing space (Class A, green), located around the mixing horizon (Class B, orange), and outside the expanded mixing space (Class C,red)

Particle size is another possible area of uncertainty and reasoning for Class C sediments, the particle size distribution was not assessed by Lavrieux et al. (2019). The depleted isotope tracers of sediment years associated with turbidites were speculated to arise from the sediment originating from subsurface erosion. However, it may also be a result of particle size selectivity during these events (Laceby et al. 2017). As there was insufficient material available to assess for particle size or subsurface contribution analysis, future research will have to assess these potential areas of uncertainty.

While multiple MRT were applied to accurately model the effect of different Suess corrections, here we used the same MRT for both FA and LMeO and for each land-use type, however in reality this is unlikely. To improve the accuracy of modelling the Suess effect on historical records, multiple iterations of the different MRT for all tracers and sources separately could be applied. However, we find our approach as an appropriate starting point for using multiple MRT Suess corrections for CSSI fingerprinting. While we have modelled the Suess effect by applying a wide range of mean residency times, sediment which

is outside the mixing space might also arise from an overcorrection using a MRT of 10 yrs. Additionally, it's important to acknowledge the relatively low number of samples associated sediment apportionment using CSSI tracers may impede capturing the full source variance or even missing sources.

Assessment of model and tracer selection performance

As suggested by Cox et al. (2023), concentration dependent mathematical mixtures (n=300) (Fig. 4) were used to test performance of the model and tracer suites. Considering mathematical mixtures are generated from only source values, Suess corrections do not significantly affect the mixing space (illustrated in Fig. 4) and as such we find the evaluation of the non-Suess corrected mathematical mixtures applicable to all MRTs. Figure 6 (A) shows the estimated proportions vs the known proportions of the mathematical mixtures using the $\delta^{13}C$ FA_{28} and $\delta^{13}C$ LMeO tracer set, with the 1:1 line indicating a perfect fit. Results show the over estimation of arable contribution (until ca. 50 % contribution), consequently causing an underestimation in forest and pasture contribution estimates. POM_{terr} has a relatively accurate and highly certain estimate at lower contributions, while the performance of estimates slightly decreases with increasing POM_{terr} contribution. Extremely low known proportions of POM_{terr} of two mixtures appears to have highly inaccurate estimates, both mixtures contain almost 0 % contribution from POM. Using CRPS, the performance of the model was evaluated and forest, pasture, and POM_{terr} had similar performances, with a median CRPS of 0.06, 0.05 and 0.05, respectively. Arable was demonstrated to have the lowest performance (CRPS 0.11), demonstrating the effect of a lower source discrimination. The higher CRPS score of arable can be attributed to more centralised position and reduced source discrimination between the other source (PCA supplied in Fig. S2), causing a higher likelihood for misclassification. The reduced source discrimination may be result of the legacy isotopic signal of forest present after the deforestation and conversion to arable fields.

The inclusion of ACL improves the model performance (Fig. 6 (B)) (Forest CRPS: 0.04, arable: 0.06, pasture: 0.03, and POM_{terr}:0.03). Using the CRP skill score, the inclusion of ACL improved the average model performance by 39 % (Forest CRP skill score: 38 %, arable: 43 %, pasture: 37 %, POM_{terr}: 39 %). Additionally, the higher sensitivity of the arable estimates to changing source contributions is illustrated in Fig. 6 (B). However, arable estimates still displayed an overestimation and underestimation of lower and higher known contributions, respectively. Apart from two extremely low POM_{terr} contribution mixtures, POM_{terr} estimates are very accurate. Results of unmixing the mathematical mixtures had no non-sensical results caused by mixtures being outside of the 'traditional' convex hull mixing space, demonstrating the need to reassess the current approach to the frequently used point in polygon tests for concentration dependent tracers.

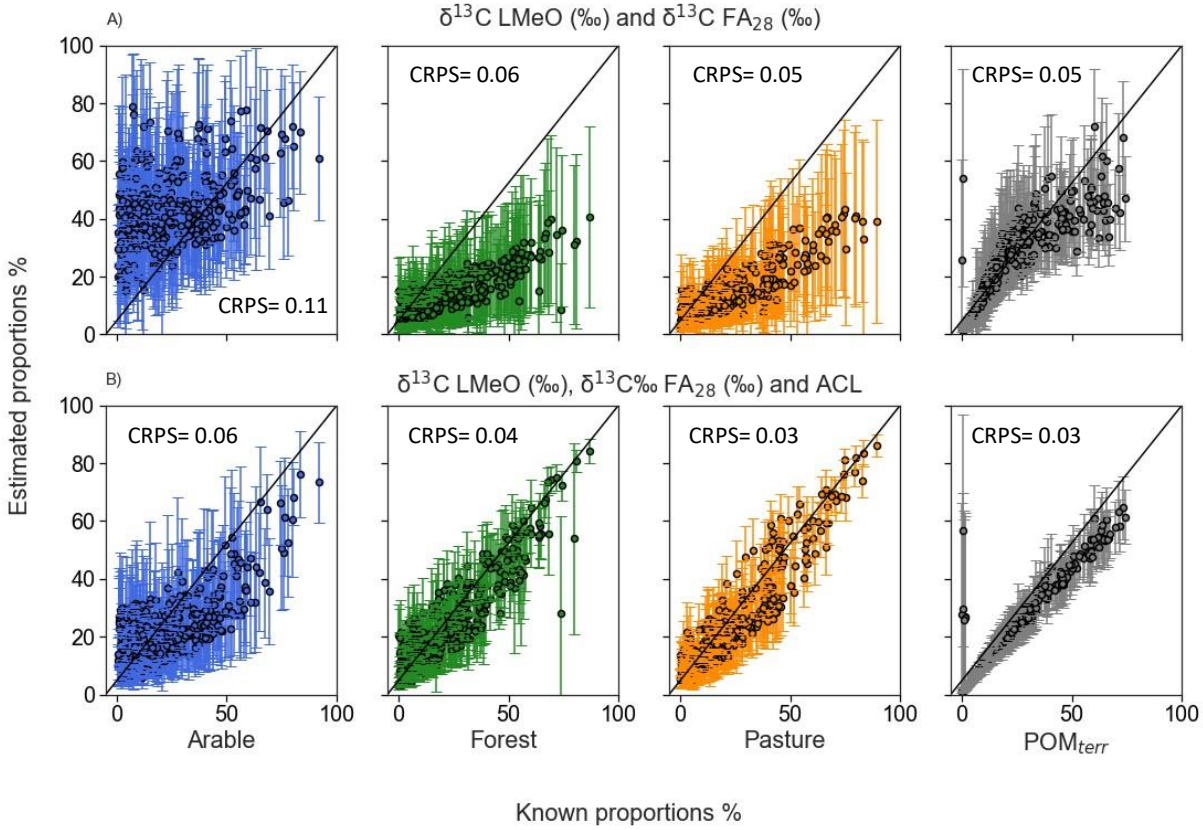


Fig. 6 Estimated proportions vs known mixture proportion for different tracer sets, with no Suess correction being applied. The solid line indicating perfect fit (Estimated proportion= known proportion). A) δ^{13}C LMeO and δ^{13}C $FA_{28,}$ B) δ^{13}C LMeO, δ^{13}C FA_{28} and ACL. The model performance is denoted by the CRPS value

One of the options when using the MixSiAR model is the flexible Bayesian framework incorporating adaptable error structures. The “process error” structure is often used in the sediment fingerprinting literature, where the variation in the mixtures is only dependent on variation in the sources, while not incorporating the variation in the mixture (e.g., target sediment) (Smith et al. 2018). While this error structure may not be optimal for accounting for variation in sediment core samples, the use of mathematical mixtures allows for the evaluation of the errors associated with the apportionment and how much confidence should be applied to the model results. Although both mathematical mixtures and sediments are unmixed identically, the omission of sediment variance limits the capability of the MixSIAR model to propagate the inherent variance within the sediment samples to the posterior distributions in the model outputs. For catchment sediment systems, the residual error only structure is preferable since it better represents erosion and sediment transport processes (Smith et al. 2018). Here, the “process only” error structure was selected due to limited sediment sample availability, which is a frequent issue for sediment cores due to their limited sample mass. Considering the potential implications of not incorporating the variation in the sediment mixtures, future sediment core fingerprinting approaches should incorporate multiple sediment mixtures either by i) sampling multiple

cores, ii) increasing the number of samples per core sample or iii) grouping core sections by their age and/or depth. Future work testing different error structures using mathematical mixtures that incorporate both sediment and source variance could significantly aid in improving the fingerprinting method.

Relative proportion of POM_{terr} to sediment

Here, we used the $\delta^{13}C$ LMeO, $\delta^{13}C$ FA_{28} and ACL tracer set to determine the relative contribution of POM_{terr} to the sediment record. To determine the relative contribution of POM_{terr} to the sediment, mineral source soils and POM_{terr} were unmixed simultaneously using MixSIAR. Three different MRT scenarios are used to model the possible Suess effect on the apportionment estimates. Class B sediment are checked for coherence with the assumption that they should fit within the range of years bracketing (Table S4). Furthermore, sediment values which are Class C (Fig. 5) are not considered in Fig. 7(A) in our interpretation of historical POM_{terr} trends.

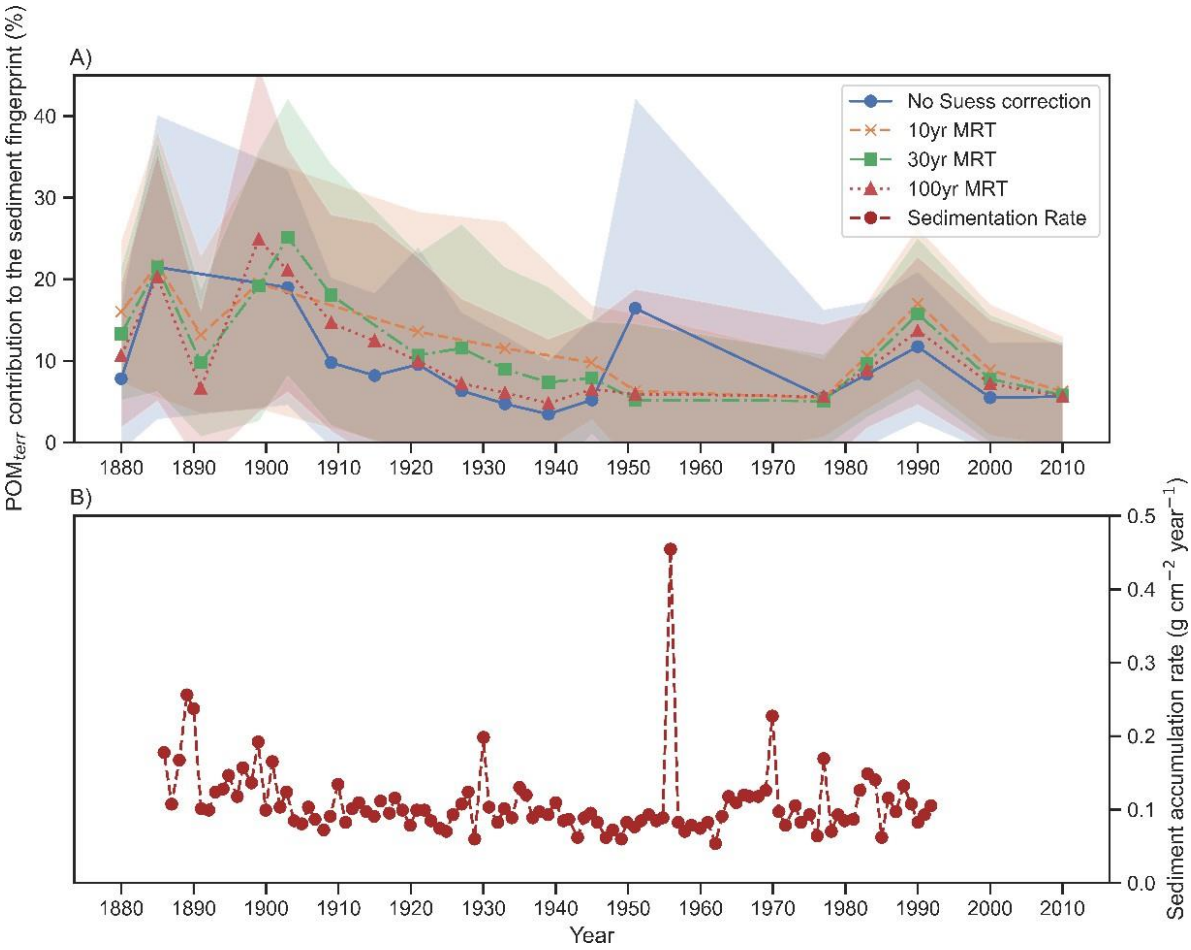


Fig. 7 A) Estimated proportion of POM_{terr} to the sediment core before and after three different Suess corrections (10, 30, and 100 yr). Sediment samples in Class C are removed from the figure. SD of POM_{terr} contributions are shown as error band B) SAR extracted from Lotter et al. (1997)

All MRTs show a similar trend of relatively high contributions of POM_{terr} from 1885 to 1909 (Fig. 7(A)). The comparison to the high SAR (Fig. 7(B)) indicates this may be a result of flood events eroding the forest floor of debris. There is a gradual decline in POM_{terr} from 1909 to 1939 in all models with the exception of 10yr MRT model where the Suess corrected POM_{terr} contribution appears to further decline until 1951. For the other three models, the year 1939 demonstrates an increased contribution of POM_{terr}. Around this time the Wahlen plan (a Swiss self-sufficiency plan) entered into force, deforesting large areas in Switzerland and around Lake Baldegg to increase agricultural production (Federal Statistical Office 1949). This may have caused an increase in sediment input from forest soils, but also eroding and washing off formerly forest humus layers and POM_{terr}. The interpretation of POM_{terr} contributions from the 1951 to 1971 is hindered by the exclusion of samples lying outside the mixing space. These exclusions are suggested to arise from subsurface erosion events during periods of elevated SAR, as illustrated in Fig. 7(B). Additionally, in the case of 1965, the detected signal may be influenced by the peak of eutrophication, potentially incorporating contributions from algae.

For all models apart from 10 yr MRT, the shift in the estimated historical POM_{terr} contribution matches with aerial photographs and pollen records (Van Der Knaap et al. 2000). These records illustrate the increase in trees and shrubs around the lake shore and stream banks from 1940 onwards. The misalignment of the 10yr MRT with the catchment's historical trends suggests errors when applying a 10-year MRT correction to the isotopic tracers. 1977 appears to be another notable time in which POM_{terr} drastically increases for all Suess corrected models, this increase in POM_{terr} coincides with an increase in pollen from trees in the sediment record (Van Der Knaap et al. 2000). Despite the relatively low percentage of POM_{terr} introduced, the depleted ^{13}C LMeO values and to a lesser extent, the depleted ^{13}C FA values from POM_{terr} can potentially have a substantial effect on the isotopic record in the sediment. Consequently, this could result in the apportionment of organic carbon rather than soil if the POM_{terr} source is not grouped separately. While the exact cause for increased POM_{terr} is somewhat speculative, the use of mathematical mixtures allows for the evaluation of the model and can help determine how much confidence we should have in the model output. Results of the model evaluation in Fig. 6 illustrates the increase in uncertainty with increasing POM_{terr} contribution, which corroborates with Fig. 7(A), adding evidence to suggest mathematical mixtures are a valuable and representative tool of sediment unmixing. Additionally, Fig. 6 illustrates an underestimation of predicted POM_{terr} contribution (ca. <5 %), as such we would assume our estimates of POM_{terr} in the sediment are also slightly underestimated.

Apart from the non-Suess corrected sediment of 1951, all sediments that are classified as Class B (Fig. 5, Table S4) displayed non-outlying apportionments (i.e., the estimated sediment apportionment is within the range of the Class A POM_{terr} estimates either side of the class B sediment) suggesting coherent results and confidence can be applied the apportionment estimates. Interestingly, the non-Suess corrected sediment of 1951 suggests a high contribution of POM, although bi-plots illustrate 1951 being located on the opposite side from the POM_{terr} source. The high contribution of sediment from 1951 is erroneous and a result of the sediment being out of the mixing space. As such we hypothesize that as the sediments move away from the arable, pasture and forest sources (in the direction away

from the mixing space) in the bi-plots, the model overcompensates and solves the model by increasing the estimated POM_{terr} contribution. While this method of determining non-coherent results has limitations, we suggest that the identification of sediment with outlying results using the concentration dependent mixing space is evidence for the need to re-evaluate concentration dependent point in polygon tests.

Improved sediment source apportionment by including POM_{terr} discrimination

When using $\delta^{13}C$ LMeO, $\delta^{13}C$ FA, and ACL as a tracer set, all source apportionment estimates of mathematical mixtures, apart from arable, are shown to be highly reliable (Fig. 6) and can be interpreted with confidence. Arable estimates can be seen as less reliable, with potential under and over estimation for high and low contributions, respectively. Considering arable contributions are minimal until 1940, taking over/under estimation into account, would not drastically change pre-1940 arable contributions. However, mathematical mixtures indicate that arable apportionment might be underestimated after 1940, as arable land increases its dominance as a sediment source. Apportionment of Class B sediments (Table S5) show no outlying results, except for 1951 POM_{terr} (Table S4), suggesting no additional erroneous results.

SAR were not available in the years post 1990 and pre-1885, and therefore cannot be converted to sediment deposited by each land-use. Additionally, it is important to recognize that the removal of POM_{terr} was accomplished using mean values, potentially introducing an additional uncertainty that was not incorporated. Nonetheless, the unmixing of the mineral soil sources demonstrates similar trends for all Suess corrections (Fig. 8, see Table S5 for values of sediment delivered by land-use). For all Suess corrections, the end of the 19th century is dominated by high forest input. While these years are not identified as turbidites by Lotter et al. (1997), the high SAR suggest high flow events likely resulting in large amounts of POM_{terr} (Fig. 7) and forest soil potentially being transported into the lake (Alewell et al. 2016). Although all models show an increase in pasture until around 1945, land-use statistics display no significant change in the relative proportion of productive land cover (arable and pasture) in the surrounding catchment in the years 1912 and 1923 (Federal Statistical Office 1912, 1922). The high pasture contribution can then be reasoned to be a result of two possible scenarios: pasture fields being converted temporally or permanently to arable fields, with the legacy pasture signal being dominant. However, if this was true, the legacy effect of pasture would decrease proportionally to the increase in arable contribution, but this is not the case. Another rationale is the intensification of the harvesting pasture for silage, hay and barn feeding during World War one, as farming resources were diverted to support the war effort.

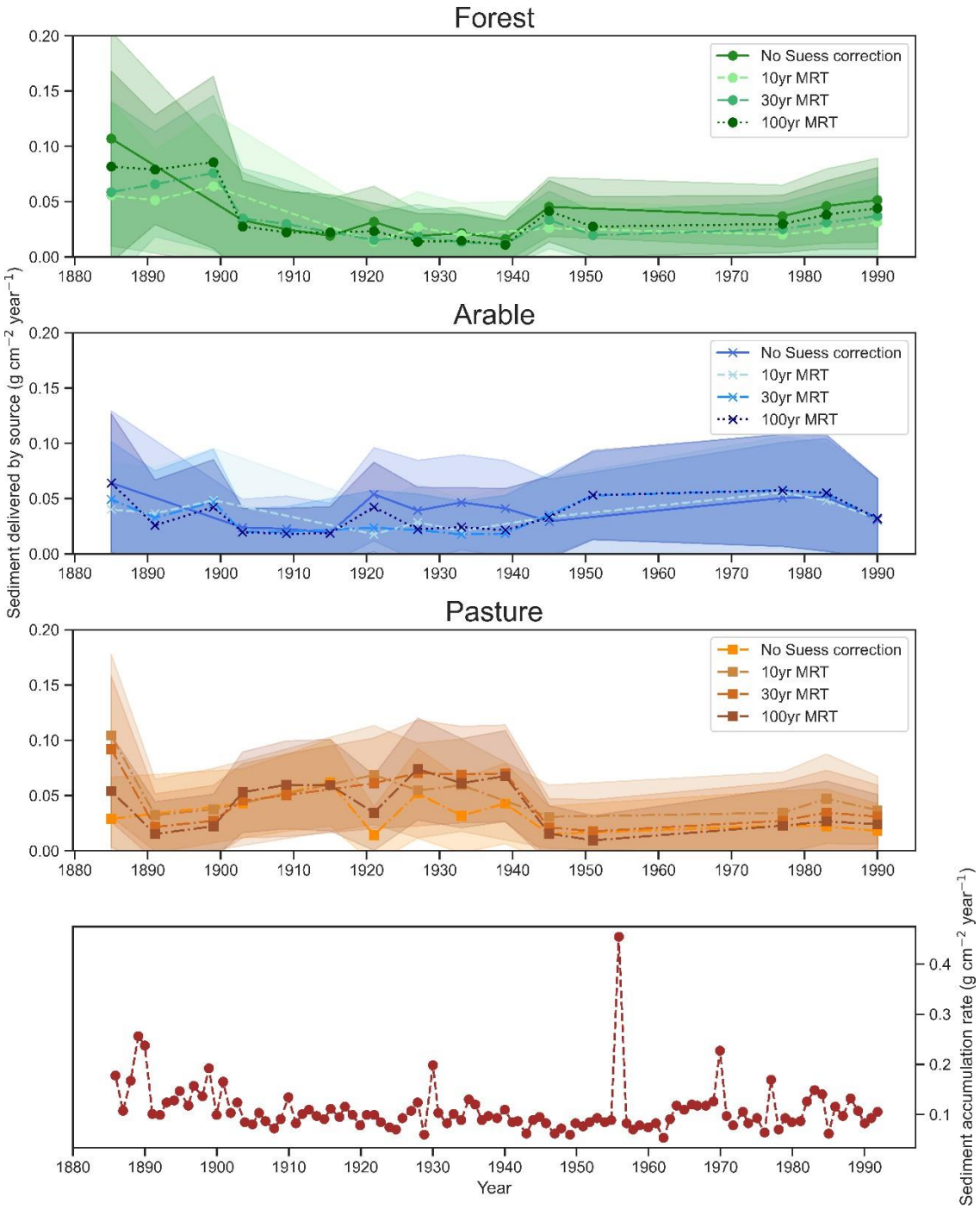


Fig. 8 Estimated sediment delivered by each land-use source after POM_{terr} correction, before and after using three different Suess corrections (10, 30, and 100 yr), and the SAR extracted from Lotter et al. (1997)

Sediment from 1920 shows a notable change of apportionment using different MRTs, demonstrating the sensitivity of $\delta^{13}C$ tracers to the Suess corrections. The sensitivity of apportionment result to the varying MRT Suess corrections can be suggested to be dependent on source discrimination. All models except the 10 yr MRT depict a peak in forest contributions around 1939. This can be credited to the

enforcement of the Wahlen plan, leading to a 3.6-fold increase in open land in the Baldegg catchment (Federal Statistical Office 1949). Contributions of arable sources appears to be low and stable until 1939, at which point arable becomes the dominant source. The timing of the arable dominance coincides with the introduction of maize in the catchment in the 1940s (Federal Statistical Office 1949; Van Der Knaap et al. 2000). The sediment record then potentially indicates maize cultivation as the main sediment source after the 1940s. Another explanation of increased arable contribution might be due to the increasing use of agricultural machinery as well as the increased connectivity of sources to the watercourse (e.g. land consolidation, removal of linear landscape features, improvements in drainage systems). Connectivity models of the Baldegg catchment indicated linear landscape features, especially roads, are important regulators of sediment transport (Batista et al. 2022). Additionally, it is important to acknowledge that while sediment samples potentially associated with subsurface erosion and outside the mixing space were removed from the historic interpretation, subsurface/ channel bank erosion is ubiquitous in sediment transport. As such our estimations of sediment delivery by each source may be overestimated due to the exclusion of channel bank erosion as a potential source.

Sediment source apportionment of sediment without POM_{terr} discrimination

The difference in the historical sediment source trend when merging forest and POM_{terr} is illustrated in Fig. 9. The difference between the three and four source model was calculated by subtracting the mean apportionment of the three-source model (Table S6) from the four-source model (Table S5, see Table S7 for differences and Fig. S7 for sediment deposition). The 10 yr MRT Suess corrected sediment are illustrated to be highly affected by the merging of forest and POM_{terr} and depicts an inconsistent outcome to other MRT corrections. Again, we suggest this is a result of the 10 yr MRT being a less accurate representation of the MRT of the tracers. The 30 yr MRT, 100 yr MRT and non-Suess-corrected sediment contributions demonstrate similar forest sedimentation rates when in comparison with the three-source model, with relatively minor difference between the three and four source model until 1945 (mean difference: -5 %). From 1945 until 1990, the three-source model then estimates a higher sediment contribution of forest (mean: 0.0133 g cm^{-2} y^{-1} difference, 37 %), and lower contribution of arable (mean: -0.014 g cm^{-2} y^{-1}, -29 %) and pasture (mean: -0.002 g cm^{-2} y^{-1}, -12 %). The lower contribution of arable can be reasoned to be a result of a reduced source discrimination between forest and arable (Fig. S2).

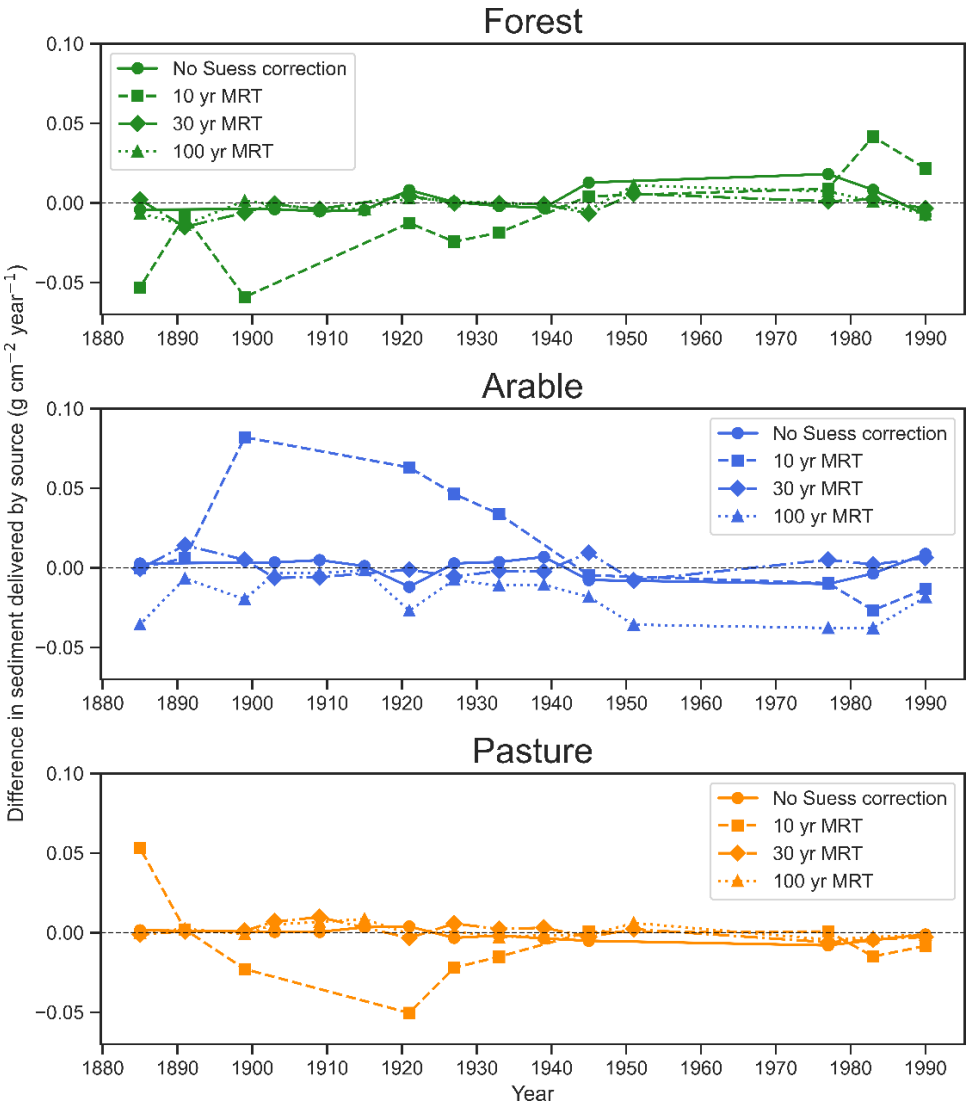


Fig. 9 The difference in the mean estimated sediment delivered using the three or four source model before and after three different Suess corrections (10yr, 30yr, and 100 yr MRT) (three-source model – four-source model = difference)

Sediment fingerprinting using CSSI tracers is a semi-empirical method to determine the relative contribution of different land-uses. Interestingly, results regularly report forest as a major source of sediment (Alewell et al. 2016; Chen et al. 2016; Upadhayay et al. 2020; Wiltshire et al. 2022). While this may be true in specific cases, in general, this appears to contradict our current understanding of the mechanistic process involved in soil erosion as well as soil erosion modeling (Borrelli et al. 2021; Wiltshire et al. 2022). In general, forests should be less susceptible to soil erosion due to the tree canopy cover with additional understory or ground vegetation and a humus layer cover (Blanco-Canqui and Lal 2010).

As stated previously, aerial images since 1940 show an increase in trees and shrubs along the stream bank. Without the separation of POM_{terr} and forest, apportionments suggest a higher sediment contribution from forests. A more coherent explanation for the increase in forest contribution since 1945 can be contributed to the fact that POM_{terr} from newly planted trees around the lake shore could have

the potential to contribute to the sediment isotopic fingerprint. Without the source separation of forest and POM_{terr}, the POM_{terr} contribution is then mistakenly identified as forest input.

In the Baldegg catchment, our initial hypothesis suggests that the misclassification of POM_{terr} as forest contribution would result in forest being the dominant source, this was not supported. However, we demonstrated the separation of POM_{terr} improves the model performance by reducing the overestimation of forest (37 %). While our hypothesis was not supported in the Baldegg catchment, it is possible that for catchments, that are dominated with forests and/or increased connectivity of POM_{terr} sources to the watercourse, there could be an increased potential for erroneous source apportionment results. In any case, this hypothesis does require more testing. Although the dominant source of sediment did not significantly change (Forest: pre-1990, Pasture: 1910-1940, Arable: post 1940), we recommend the inclusion of the POM_{terr} as a separate source of CSSI tracers as being fundamental to helping ensure accurate and representative sediment apportionments.

3.4 Conclusion

The use of lignin $\delta^{13}C$ LMeO values as an additional land-use-specific tracer is highly effective at improving source discrimination and subsequent expansion of the mixing space. While the results of the historic apportionment are highly credible and fit well with land-use history, the conservativeness of $\delta^{13}C$ LMeO during transport and deposition requires further investigation. Furthermore, the use of two different types of CSSI tracers increases the representativeness of the soil and reduces the potential bias associated with specific tracers. The preparation of samples for $\delta^{13}C$ FA analysis is a time and organic solvent consuming procedure. In contrast the preparation of samples for $\delta^{13}C$ LMeO analysis is a relatively rapid but also a solvent free method, allowing for higher sample throughput and the subsequent increase in source heterogeneity testing, all while being more environmentally friendly. Furthermore, the minimal sample mass required for MeO analysis enables the replication of sediment analysis, facilitating the empirical determination of mixture variance. Consequently, this method can be used in the future to enhance the conclusiveness of the mixing model by directly accounting for mixture variance, rather than being estimated by source variance.

A range of MRTs are used to model the possible Suess effect on CSSI fingerprints and should be applied as a standard protocol for historical use of CSSI tracers. Mathematical mixtures are an effective tool to define a concentration dependent mixing space, and identify sediment values which may give erroneous results. Additionally, they allow for the evaluation of the degree of confidence which should be put in the model output. Although the exact explanation for historical sediment trends remains speculative, the overall story of sediment contribution becomes significantly more credible when using POM_{terr} corrected apportionments. Although our hypothesis was not supported, as the merging of POM_{terr} and forest did not result in a change of forest being the dominant source, our results did demonstrate the possibility of an overestimation of forest contributions in the literature. Therefore, we suggest the discrimination of POM_{terr} and forest is important to enhance the accuracy of sediment fingerprinting applications when using CSSIs or other organic tracers in order to enable future policies to be based on the most reliable data available.

By determining the past historic trends of sediment sources, we gain insight into the causes of soil erosion. Additionally, land-use specific apportionments allow the discrimination between natural and anthropogenic events leading to soil erosion. While this study was focused on Lake Baldegg, we find it applicable to other temperate regions with similar intensive agricultural practices. However, limitations may be present when using CSSI tracers for regions dominated by subsurface erosion. Nonetheless, this information is vital for effectively evaluating soil erosion models and consequently, their use in designing or evaluating sediment mitigation strategies. Improving the accuracy of land-use specific apportionment estimates is therefore of the utmost importance to prevent the negative consequences of soil erosion and sedimentation.

3.5 References

Alewel C, Birkholz A, Meusburger K, Schindler Wildhaber Y, Mabit L (2016) Quantitative sediment source attribution with compound-specific isotope analysis in a C3 plant-dominated catchment (central Switzerland). Biogeosciences 13(5):1587–1596. https://doi.org/10.5194/bg-13-1587-2016

Anhäuser T, Greule M, Zech M, Kalbitz K, McRoberts C, Keppler F (2015) Stable hydrogen and carbon isotope ratios of methoxyl groups during plant litter degradation. Isotopes Environ Health Stud 51(1):143–154. https://doi.org/10.1080/10256016.2015.1013540

Bakker MM, Govers G, Jones RA, Rounsevell MDA (2007) The effect of soil erosion on Europe's crop yields. Ecosystems 10(7):1209–1219. https://doi.org/10.1007/s10021-007-9090-3

Bashagaluke JB, Logah V, Opoku A, Sarkodie-Addo J, Quansah C (2018) Soil nutrient loss through erosion: Impact of different cropping systems and soil amendments in Ghana. PLoS ONE 13(12). https://doi.org/10.1371/journal.pone.0208250

Batista PVG, Fiener P, Scheper S, Alewel C (2022) A conceptual-model-based sediment connectivity assessment for patchy agricultural catchments. Hydrol Earth Syst Sci 26(14):3753–3770. https://doi.org/10.5194/hess-26-3753-2022

Batista PVG, Laceby JP, Evrard O (2022) How to evaluate sediment fingerprinting source apportionments. J Soils Sediments 22(4):1315–1328. https://doi.org/10.1007/s11368-022-03157-4

Batista PVG. Laceby JP, Silva MLN, Tassinari D, Bispo DFA, Curi N, Davies J, Quinton JN (2019) Using pedological knowledge to improve sediment source apportionment in tropical environments. J Soils Sediments 19(9):3274–3289. https://doi.org/10.1007/s11368-018-2199-5

Belmont P, Willenbring JK, Schottler SP, Marquard J, Kumarasamy K, Hemmis JM (2014) Toward generalizable sediment fingerprinting with tracers that are conservative and nonconservative over sediment routing timescales. J Soils Sediments 14(8):1479–1492. https://doi.org/10.1007/s11368-014-0913-5

Blanco-Canqui H, Lal R (2010) Principles of soil conservation and management. Principles of Soil Conservation and Management. Springer, Netherlands. https://doi.org/10.1007/978-1-4020-8709-7

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Borrelli P, Alewell C, Alvarez P, Anache JAA, Baartman J, Ballabio C, Bezak N, Biddoccu M, Cerdà A, Chalise D, Chen S, Chen W, De Girolamo AM, Gessesse GD, Deumlich D, Diodato N, Efthimiou N, Erpul G, Fiener P, Panagos P (2021) Soil erosion modelling: A global review and statistical analysis. Sci Total Environ 780. https://doi.org/10.1016/j.scitotenv.2021.146494

Borselli L, Cassi P, Torri D (2008) Prolegomena to sediment and flow connectivity in the landscape: A GIS and field numerical assessment. Catena 75(3):268–277. https://doi.org/10.1016/j.catena.2008.07.006

Brand C, Benmansour M, Walz L, Nguyen LT, Cadisch G, Rasche F (2018) Integrating compound-specific δ13C isotopes and fallout radionuclides to retrace land use type-specific net erosion rates in a small tropical catchment exposed to intense land use change. Geoderma 310:53–64. https://doi.org/10.1016/j.geoderma.2017.09.008

Bravo-Linares C, Schuller P, Castillo A, Salinas-Curinao A, Ovando-Fuentealba L, Muñoz-Arcos E, Swales A, Gibbs M, Dercon G (2020) Combining isotopic techniques to assess historical sediment delivery in a forest catchment in central Chile. J Soil Sci Plant Nutr 20(1):83–94. https://doi.org/10.1007/s42729-019-00103-1

Chen F, Fang N, Shi Z (2016) Using biomarkers as fingerprint properties to identify sediment sources in a small catchment. Sci Total Environ 557–558:123-133. https://doi.org/10.1016/j.scitotenv.2016.03.028

Collins AL, Blackwell M, Boeckx P, Chivers CA, Emelko M, Evrard O, Foster I, Gellis A, Gholami H, Granger S, Harris P, Horowitz AJ, Laceby JP, Martinez-Carreras N, Minella J, Mol L, Nosrati K, Pulley S, Silins U, Zhang Y (2020) Sediment source fingerprinting: benchmarking recent outputs, remaining challenges and emerging themes. J Soils Sediments 20(12):4160–4193. https://doi.org/10.1007/s11368-020-02755-4

Collins AL, Walling DE (2004) Documenting catchment suspended sediment sources: Problems, approaches and prospects. Prog Phys Geogr 28(2):159–196. https://doi.org/10.1191/0309133304pp409ra

Cooper RJ, Pedentchouk N, Hiscock KM, Disdle P, Krueger T, Rawlins BG (2015) Apportioning sources of organic matter in streambed sediments: An integrated molecular and compound-specific stable isotope approach. Sci Total Environ 520:187–197. https://doi.org/10.1016/j.scitotenv.2015.03.058

Cox T, Laceby JP, Roth T, Alewell C (2023) Less is more? A novel method for identifying and evaluating non-informative tracers in sediment source mixing models. J Soils Sediments 23:3241-3261. https://doi.org/10.1007/s11368-023-03573-0

Cui X, Bianchi TS, Hutchings JA, Savage C, Curtis JH (2016) Partitioning of organic carbon among density fractions in surface sediments of Fiordland, New Zealand. J Geophys Res Biogeosci 121(3):1016–1031. https://doi.org/10.1002/2015JG003225

Derrien M, Yang L, Hur J (2017) Lipid biomarkers and spectroscopic indices for identifying organic matter sources in aquatic environments: A review. Water Res 112:58–71. https://doi.org/10.1016/j.watres.2017.01.023

Domozych DS, Sørensen I, Popper ZA, Ochs J, Andreas A, Fangel JU, Pielach A, Sacks C, Brechka H, Ruisi-Besares P, Willats WGT, Rose JKC (2014) Pectin metabolism and assembly in the cell wall of the charophyte green alga Penium margaritaceum. Plant Physiol 165(1):105–118. https://doi.org/10.1104/pp.114.236257

Dümig A, Knicker H, Schad P, Rumpel C, Dignac MF, Kögel-Knabner I (2009) Changes in soil organic matter composition are associated with forest encroachment into grassland with long-term fire history. Eur J Soil Sci 60(4):578–589. https://doi.org/10.1111/j.1365-2389.2009.01140.x

Evrard O, Poulenard J, Némery J, Ayrault S, Gratiot N, Duvert C, Prat C, Lefèvre I, Bonté P, Esteves M (2013) Tracing sediment sources in a tropical highland catchment of central Mexico by using conventional and alternative fingerprinting methods. Hydrol Process 27(6):911–922. https://doi.org/10.1002/hyp.9421

Feakins SJ, Rincon M, Pinedo P (2013) Analytical challenges in the quantitative determination of 2H/1H ratios of methyl iodide. RCM 27(3):430–436. https://doi.org/10.1002/rcm.6465

Federal Statistical Office (1912) Statistique de la superficie arrêtée le 1er juillet 1912, Bern, Switzerland

Federal Statistical Office (1924) IIe Statistique de la superficie de la Suisse 1923/24, Bern, Switzerland

Federal Statistical Office (1949) Der Schweizerische Ackerbau in der Kriegszeit, Eidgenössische Anbauerhebungen 1939–1947, Bern, Switzerland

Galbally IE, Kirstine W (2002) The production of methanol by flowering plants and the global cycle of methanol. J Atmos Chem 43:195-229. https://doi.org/10.1023/A:1020684815474

García-Comendador J, Martínez-Carreras N, Fortesa J, Company J, Borràs A, Palacio E, Estrany J, (2023) In-channel alterations of soil properties used as tracers in sediment fingerprinting studies. Catena 225. https://doi.org/10.1016/j.catena.2023.107036

Gibbs MM (2008) Identifying source soils in contemporary estuarine sediments: A new compound-specific isotope method. Estuaries Coast 31(2):344–359. https://doi.org/10.1007/s12237-007-9012-9

Gibbs M, Swales A, Olsen G (2014) Suess effect on biomarkers used to determine sediment provenance from land-use changes. Proceedings – International Symposium on Managing Soils for Food Security and Climate Change Adaption and Mitigation 371–375.

Goñiigoñii MA, Ruttenberg KC, Eglinton TI (1998) A reassessment of the sources and importance of land-derived organic matter in surface sediments from the Gulf of Mexico. Geochim Cosmochim Acta 62(18):3055-3075

Goñii MA, Goñii G, Thomas KA (2000) Sources and transformations of organic matter in surface soils and sediments from a tidal estuary (North inlet, south Carolina, USA). Estuaries 23:548–564. http://organic.geol.sc.edu/ninlet/ninlet.htm

Greule M, Keppler F (2011) Stable isotope determination of ester and ether methyl moieties in plant methoxyl groups. Isot Environ Health Stud 47(4):470–482. https://doi.org/10.1080/10256016.2011.616270

Greule M, Moossen H, Geilmann H, Brand WA, Keppler F (2019) Methyl sulfates as methoxy isotopic reference materials for *δ* 13 C and *δ* 2 H measurements. RCM 33(4):343–350. https://doi.org/10.1002/rcm.8355

Greule M, Moossen H, Lloyd MK, Geilmann H, Brand WA, Eiler JM, Qi H, Keppler F (2020) Three wood isotopic reference materials for *δ*2H and *δ*13C measurements of plant methoxy groups. Chem Geol 533. https://doi.org/10.1016/j.chemgeo.2019.119428

Greule M, Mosandl A, Hamilton JTG, Keppler F (2009) A simple rapid method to precisely determine 13C/12C ratios of plant methoxyl groups. RCM 23(11): 1710–1714. https://doi.org/10.1002/rcm.4057

Heim A, Schmidt MWI (2007) Lignin turnover in arable soil and grassland analysed with two different labelling approaches. Eur J Soil Sci 58(3):599–608. https://doi.org/10.1111/j.1365-2389.2006.00848.x

Hernes PJ, Robinson AC, Aufdenkampe AK (2007) Fractionation of lignin during leaching and sorption and implications for organic matter “freshness.” Geophys Res Lett 34(17). https://doi.org/10.1029/2007GL031017

Hirave P, Glendell M, Birkholz A, Alewell C (2021) Compound-specific isotope analysis with nested sampling approach detects spatial and temporal variability in the sources of suspended sediments in a Scottish mesoscale catchment. Sci Total Environ 755. https://doi.org/10.1016/j.scitotenv.2020.142916

Hirave P, Wiesenberg GLB, Birkholz A, Alewell C (2020) Understanding the effects of early degradation on isotopic tracers: Implications for sediment source attribution using compound-specific isotope analysis (CSIA). Biogeosci 17(8):2169–2180. https://doi.org/10.5194/bg-17-2169-2020

Keppler F, Kalin RM, Harper DB, Mcroberts WC, Hamilton JTG (2004) Carbon isotope anomaly in the major plant C1 pool and its global biogeochemical implications Carbon isotope anomaly in the major plant C 1 pool and its global biogeochemical implications. European Geosciences Union 1(2). www.biogeosciences.net/bg/1/123/

Kind J (2012) Ferromagnetic resonance spectroscopy and Holocene Earth’s magnetic field variations in sediment from Swiss lakes. PhD thesis, ETH Zurich, Switzerland

Koiter AJ, Owens PN, Petticrew EL, Lobb DA (2013) The behavioural characteristics of sediment properties and their implications for sediment fingerprinting as an approach for identifying sediment sources in river basins. Earth-Sci Rev 125:24–42. https://doi.org/10.1016/j.earscirev.2013.05.009

Kuzyk ZZA, Goñi M A, Stern GA, Macdonald RW (2008) Sources, pathways and sinks of particulate organic matter in Hudson Bay: Evidence from lignin distributions. Mar Chem 112(3–4):215–229. https://doi.org/10.1016/j.marchem.2008.08.001

Laceby JP, Evrard O, Smith HG, Blake WH, Olley JM, Minella JPG, Owens PN (2017) The challenges and opportunities of addressing particle size effects in sediment source fingerprinting: A review. Earth-Sci Rev 169:85–103. Elsevier B.V. https://doi.org/10.1016/j.earscirev.2017.04.009

Laceby JP, McMahon J, Evrard O, Olley J (2015) A comparison of geological and statistical approaches to element selection for sediment fingerprinting. J Soils Sediments 15(10):2117–2131. https://doi.org/10.1007/s11368-015-1111-9

Lavrieux M, Birkholz A, Meusburger K, Wiesenberg GLB, Gilli A, Stamm C, Alewell C (2019) Plants or bacteria? 130 years of mixed imprints in Lake Baldegg sediments (Switzerland), as revealed by compound-specific isotope analysis (CSIA) and biomarker analysis. Biogeosci 16(10):2131–2146. https://doi.org/10.5194/bg-16-2131-2019

Lee H, Feng X, Mastalerz M, Feakins SJ (2019) Characterizing lignin: Combining lignin phenol, methoxy quantification, and dual stable carbon and hydrogen isotopic techniques. Org Geochem 136. https://doi.org/10.1016/j.orggeochem.2019.07.003

Lloyd MK, Trembath-Reichert E, Dawson KS, Feakins SJ, Mastalerz M, Orphan VJ, Sessions A L, Eiler JM (2021) Methoxyl stable isotopic constraints on the origins and limits of coal-bed methane. AAAS 374(6569):894–897. https://www.science.org

Lotter AF, Sturm M, Teranes JL, Wehrli B (1997) Varve formation since 1885 and high-resolution varve analyses in hypertrophic Baldeggersee (Switzerland). Aquat Sci 59(4):304–325. https://doi.org/10.1007/BF02522361

Lützow MV, Kögel-Knabner I, Ekschmitt K, Matzner E, Guggenberger G, Marschner B, Flessa H (2006) Stabilization of organic matter in temperate soils: Mechanisms and their relevance under different soil conditions - A review. Eur J Soil Sci 57(4):426–445. https://doi.org/10.1111/j.1365-2389.2006.00809.x

Matheson JE, Winkler RL (1976) Scoring rules for continuous probability distributions. Manag Sci 22(10):1087–1096. https://about.jstor.org/terms

Motha JA, Wallbrink PJ, Hairsine PB, Grayson RB (2002) Tracer properties of eroded sediment and source material. Hydrol Process 16(10):1983–2000. https://doi.org/10.1002/hyp.397

Owens PN, Blake WH, Gaspar L, Gateuille D, Koiter AJ, Lobb DA, Petticrew EL, Reiffarth DG, Smith HG, Woodward JC (2016) Fingerprinting and tracing the sources of soils and sediments: Earth and

ocean science, geoarchaeological, forensic, and human health applications. Earth-Sci Rev 162:1–23. https://doi.org/10.1016/j.earscirev.2016.08.012

Pedro HTC, Coimbra CFM, David M, Lauret P (2018) Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts. Renew Energy 123:191–203. https://doi.org/10.1016/j.renene.2018.02.006

Pimentel D (2006) Soil erosion: A food and environmental threat. Environ Dev Sustain 8(1):119–137. https://doi.org/10.1007/s10668-005-1262-8

Rezende CE, Pfeiffer WC, Martinelli LA, Tsamakis E, Hedges JI, Keil RG (2010) Lignin phenols used to infer organic matter sources to Sepetiba Bay - RJ, Brasil. Estuar Coast. Shelf Sci 87(3):479–486. https://doi.org/10.1016/j.ecss.2010.02.008

Schmidt MWI, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, Kleber M, Kögel-Knabner I, Lehmann J, Manning DAC, Nannipieri P, Rasse DP, Weiner S, Trumbore SE (2011) Persistence of soil organic matter as an ecosystem property. Nature 478 (7367):49–56. https://doi.org/10.1038/nature10386

Smith HG, Karam DS, Lennard AT (2018) Evaluating tracer selection for catchment sediment fingerprinting. J Soils Sediments 18(9):3005–3019. https://doi.org/10.1007/s11368-018-1990-7

Stock BC, Jackson AL, Ward EJ, Parnell AC, Phillips DL, Semmens BX (2018) Analyzing mixing systems using a new generation of Bayesian tracer mixing models. PeerJ 2018(6). https://doi.org/10.7717/peerj.5096

Thevenot M, Dignac MF, Rumpel C (2010) Fate of lignins in soils: A review. Soil Biol Biochem 42(8):1200–1211. https://doi.org/10.1016/j.soilbio.2010.03.017

Upadhayay HR, Bodé S, Griepentrog M, Bajracharya RM, Blake W, Cornelis W, Boeckx P (2018) Isotope mixing models require individual isotopic tracer content for correct quantification of sediment source contributions. Hydrol Process 32(7):981–989. https://doi.org/10.1002/hyp.11467

Upadhayay HR, Lamichhane S, Bajracharya RM, Cornelis W, Collins AL, Boeckx P (2020) Sensitivity of source apportionment predicted by a Bayesian tracer mixing model to the inclusion of a sediment connectivity index as an informative prior: Illustration using the Kharka catchment (Nepal). Sci Total Environ 713. https://doi.org/10.1016/j.scitotenv.2020.136703

Upadhayay HR, Smith HG, Griepentrog M, Bodé S, Bajracharya RM, Blake W, Cornelis W, Boeckx P (2018) Community managed forests dominate the catchment sediment cascade in the mid-hills of Nepal: A compound-specific stable isotope analysis. Sci Total Environ 637–638:306–317. https://doi.org/10.1016/j.scitotenv.2018.04.394

Vale S, Swales A, Smith HG, Olsen G, Woodward B (2022) Impacts of tracer type, tracer selection, and source dominance on source apportionment with sediment fingerprinting. Sci Total Environ 831. https://doi.org/10.1016/j.scitotenv.2022.154832

Van Bree LGJ, Peterse F, Van der Meer MTJ, Middelburg JJ, Negash AMD, De Crop W, Cocquyt C, Wieringa JJ, Verschuren D, Sinninghe Damsté JS (2018) Seasonal variability in the abundance and stable carbon-isotopic composition of lipid biomarkers in suspended particulate matter from a stratified equatorial lake (Lake Chala, Kenya/Tanzania): Implications for the sedimentary record. Quat Sci Rev 192:208–224. https://doi.org/10.1016/j.quascirev.2018.05.023

Van Der Knaap WO, Van Leeuwen JFN, Fankhauser A, Ammann B (2000) Palynostratigraphy of the last centuries in Switzerland based on 23 lake and mire deposits: chronostratigraphic pollen markers, regional patterns, and local histories. Rev Palaeobot Palynol 108.

van Raden U (2012): High resolution Swiss lake records of climate change, PhD Thesis, ETH Zurich (Switzerland),

Verburg P (2007) The need to correct for the Suess effect in the application of δ13C in sediment of autotrophic Lake Tanganyika, as a productivity proxy in the Anthropocene. J Paleolimnol 37(4):591–602. https://doi.org/10.1007/s10933-006-9056-z

Walling DE (2005) Tracing suspended sediment sources in catchments and river systems. Sci Total Environ 344(1-3 SPEC. ISS.):159–184. https://doi.org/10.1016/j.scitotenv.2005.02.011

Wehrli B, Lotter AF, Schaller T, Sturm M (1997) High-resolution varve studies in Baldeggersee (Switzerland): Project overview and limnological background data. Aquatic Sci 59(4):285–294. https://doi.org/10.1007/BF02522359

Wiesenberg GLB, Gocke M, Kuzyakov, Y (2010) Fast incorporation of root-derived lipids and fatty acids into soil - Evidence from a short term multiple 14CO2 pulse labelling experiment. Org Geochem 41(9):1049–1055. https://doi.org/10.1016/j.orggeochem.2009.12.007

Wiltshire C, Glendell M, Waine TW, Grabowski RC, Meersmans J (2022) Assessing the source and delivery processes of organic carbon within a mixed land use catchment using a combined n-alkane and carbon loss modelling approach. J Soils Sediments 22(5):6129–1642. https://doi.org/10.1007/s11368-022-03197-w

Wiltshire C, Waine TW, Grabowski RC, Meersmans J, Thornton B, Addy S, Glendell M (2023) Assessing n-alkane and neutral lipid biomarkers as tracers for land-use specific sediment sources. Geoderma 433. https://doi.org/10.1016/j.geoderma.2023.116445

Wynants M, Patrick A, Munishi L, Mtei K, Bodé S, Taylor A, Millward G, Roberts N, Gilvear D, Ndakidemi P, Boeckx P, Blake W (2021) Soil erosion and sediment transport in Tanzania: Part II sedimentological evidence of phased land degradation. Earth Surf Process Landf 46:3112–3126

Xu J, Lyu H, Xu X, Li Y, Li Z, Lei S, Bi S, Mu M, Du C, Zeng S (2019) Dual stable isotope tracing the source and composition of POM during algae blooms in a large and shallow eutrophic lake: All contributions from algae? Ecol Indic 102:599–607. https://doi.org/10.1016/j.ecolind.2019.03.014

Zamparas M, Zacharias I (2014) Restoration of eutrophic freshwater by managing internal nutrient loads. A review. Sci Total Environ 496:551–562. https://doi.org/10.1016/j.scitotenv.2014.07.076

Zeisel S (1885) Über ein Verfahren zum quantitativen Naohweise von Methoxyl. Monatsh Chem 6:989-997. https://doi.org/10.1007/BF01554683

Zhang GS, Chan KY, Oates A, Heenan DP, Huang GB (2007) Relationship between soil structure and runoff/soil loss after 24 years of conservation tillage. Soil Tillage Res 92(1–2):122–128. https://doi.org/10.1016/j.still.2006.01.006

Chapter 4

Isotopic analysis (δ13C and δ2H) of lignin methoxy groups in forest soils to identify and quantify lignin sources

Science of the Total Environment

https://doi.org/10.1016/j.scitotenv.2024.175025

Received: 23 April 2024/ Accepted: 23 July 2024

Isotopic Analysis ($\delta^{13}C$ and $\delta^{2}H$) of Lignin Methoxy Groups in Forest Soils to Identify and Quantify Lignin Sources

****

Terry Cox[1], Anna Wieland[2], Markus Greule[2], Frank Keppler[3], Annika Einbock[1], Christine Alewell[1]

Corresponding authors email: Terry.cox@unibas.ch

[1] Department of Environmental Sciences, University of Basel, Bernoullistrasse 30, 4056 Basel, Switzerland

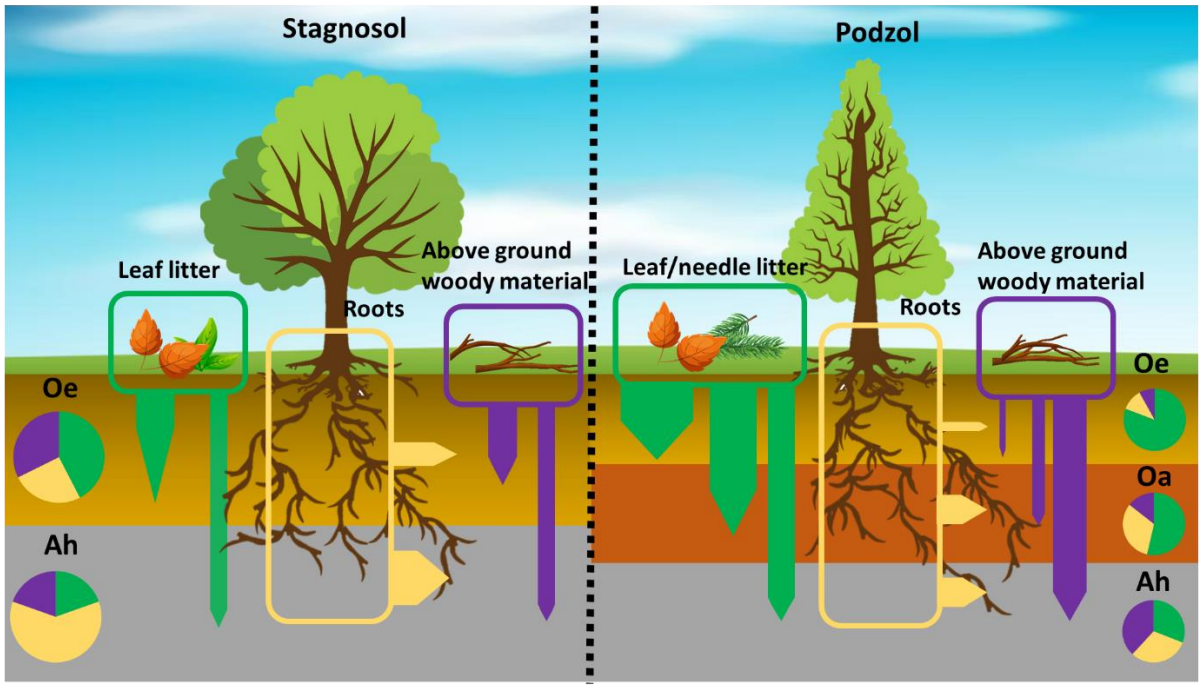
[2] Institute of Earth Sciences, Heidelberg University, Im Neuenheimer Feld 234-236, D-69120 Heidelberg, Germany

[3] Heidelberg Center for the Environment (HCE), Heidelberg University, Germany

****

Abstract

The relative apportion of above and below ground carbon sources is known to be an important factor in soil organic matter formation. Although lignin is the most abundant aromatic plant material in the terrestrial biosphere, our understanding of lignin source contributions to soil organic matter (SOM) is limited due to the complex molecular structure and analysis of lignin. In this study, we novelly apply the dual isotopic analysis ($\delta^{13}C$ and $\delta^{2}H$ values) of lignin methoxy groups (LMeO) with the Bayesian mixing model, MixSIAR, to apportion lignin sources in two contrasting soil types, a podzol and a stagnosol. Results of the isotopic analysis of LMeO demonstrate the ability of $\delta^{2}H$ LMeO values to discriminate between above and below ground lignin sources, while $\delta^{13}C$ LMeO values discriminated between photosynthesising and non-photosynthesising tissues. In the stagnosol subsurface horizons, a decreasing proportion of the leaf litter lignin was observed with increasing organic matter degradation, cumulating in the Ah horizon being dominated by lignin from roots. The podzol sites indicated a similar reduction in leaf litter lignin with an increase in organic matter degradation and depth. However, the Ah horizon was shown to accumulate lignin from the above ground woody material. Furthermore, given the significant abundance of LMeO groups in the terrestrial biosphere and the extremely depleted $\delta^{13}C$ LMeO values in leaf litter, we employed a mass balance approach to determine the extent in which the $\delta^{13}C$ bulk enrichment generally associated with isotopic fractionation during organic matter decomposition can be attributed to the shift in lignin sources. Analysis reveals that 14 % and 11 % of bulk $\delta^{13}C$ enrichment can be attributed to the transition in LMeO sources from leaf litter to roots in the stagnosol and podzol, respectively. Thus, models relying on $\delta^{13}C$ enrichment with depth as an indicator of carbon turnover may be partially overestimating rates.



Keywords Source apportionment, Stable isotopes, MixSIAR, Serine, Carbon soil dynamics

****

4.1 Introduction

Vascular plant-derived organic compounds constitute 75 % of the carbon content in soils (Hedges, 1992). Lignin, the second most prevalent plant-derived compound class after cellulose (Boerjan et al., 2003), plays a crucial role in the formation of soil organic matter and the decomposition of plant litter (Guo et al., 2021; Heim and Frey, 2004; Moore et al., 1999). Despite being the most abundant aromatic plant material in terrestrial ecosystems and representing around 30 % of the organic carbon in the global terrestrial biosphere (Boerjan et al., 2003), our understanding of lignin sources and mixing dynamics from plants to soil organic matter (SOM) is constrained by the complex molecular structures and analysis of lignin (Koegel-Knabner, 2002; Hatfield and Fukushima, 2005). However, understanding the relative contribution of above and below-ground sources of lignin into SOM can provide further insights in carbon and nutrient cycling. The source contributions of SOM are a complex process of input, transformations, transport and outputs (Prescott and Vesterdal, 2021). There seems to be no clear consensus on the distribution and dynamics of lignin in soils. Most studies indicate a decrease in lignin concentration and an increase in degradation with depth (Guggenberger and Zech, 1994; Rumpel et al., 2002; Peinemann et al., 2005; Heim and Schmidt, 2007), while others have shown an accumulation in deeper soils (Wedin et al., 1995; Sanger et al., 1997; Feng and Simpson, 2007; Mason et al., 2009). As lignin is transferred from plant biomass to soils, the composition of its monomers (monolignols) are partially preserved. However, the use of monolignols for source apportionment has limitations: minimal source discrimination (Thevenot et al., 2010), and the modification of the monomer composition during degradation (Dümig et al., 2009), irreversible sorption (Hernes et al., 2013) and soil sorption/leaching phase transitions (Hernes et al., 2007). Furthermore, monolignols from non-lignin sources (e.g.,

tannins) have been observed (Hedges and Parker, 1976; Goni et al., 2000). The isotopic composition of specific compounds is a valuable tool for understanding carbon dynamics in soils (Glaser, 2005). However, the ^{13}C values of lignin also has limited discrimination between plant tissues (Wedin et al., 1995; Schweizer et al., 1999; Fernandez et al., 2003; Dignac et al., 2005). Moreover, during degradation, the isotopic composition of lignin in both leaf litter and roots exhibit isotopic fractionation (Wedin et al., 1995). Methoxy groups (MeO, molecular formular: -OCH3) make up approximately 2.5 % of the total carbon in the terrestrial biosphere (Galbally and Kirstine, 2002), and 15–20 % of lignin monomers. While MeO predominantly originate from lignin (as ether bound MeO groups), the polysaccharide pectin is also known to be a major source of MeO in the environment (ester bound MeO groups). Considering we focus on tracing only lignin sources, alkaline hydrolysis can be used to selectively remove the ester bound MeO, enabling the analysis of the residual lignin derived methoxy groups (LMeO) by the Zeisel reaction (Zeisel, 1885; Greule et al., 2009; Cox et al., 2024). A large $\delta^{13}C$ LMeO discrimination between leaves and woody material has been observed (Keppler et al., 2004), suggesting the possibility of accurately attributing lignin sources within organic and mineral soil horizons. Accurate apportionment of sources requires that the tracer(s) (also known as: properties or characteristics) remain unaltered (or predictable) throughout degradation or transport (Motha et al., 2002; Koiter et al., 2013; Belmont et al., 2014). During a 27-month litter bag experiment, European beech and Sycamore maple bulk methoxy groups (lignin and pectin derived MeO) showed no obvious isotopic fractionation. However, litter bags from Norway spruce, Scots pine and Mountain ash exhibited small isotopic fractionation (Anhauser et al., 2015). The relatively small correlation between $\delta^{13}C$ MeO values and MeO content was attributed to the preferential degradation of pectin. The difference between species may be a result of variation of lignin and pectin abundance in species. Analysis of $\delta^{13}C$ MeO values during the degradation continuum from wood to coal revealed notable fractionation (Lloyd et al., 2021). However, it's important to note that we wouldn't expect LMeO degradation to occur on the same scale as the wood-to-coal continuum where the remaining MeO fraction in coal is only ~ 10− 4 and occurring over time scales of million years (Lloyd et al., 2021). Unlike bulk wood, cellulose, or lignin, MeO has the benefit of nonexchangeable hydrogen (Lee et al., 2019), resulting in the δ^2 H MeO values being stable during metabolism and having multiple applications (Keppler et al., 2007; Keppler and Hamilton, 2008; Lee et al., 2019; Wieland et al., 2024). Experimental evidence for the stability of δ^2 H MeO values has been demonstrated during litter bag degradation experiments (Anhauser et al., 2015) and microbial inoculations of garden biomass (Lu et al., 2022), in which the δ^2 H MeO values remained unchanged and were shown not to correlate with MeO degradation. While the conservativeness of $\delta^{13}C$ LMeO values are not as well established as δ^2 H LMeO values, in a two-source mixing model (e.g., using above and below ground sources as end-members), the deviation of the mixtures from the two-source mixing line allows isotopic fractionation and the mixing of sources to be disentangled (Lutz and Van Breukelen, 2014a, 2014b). To gain further understanding in lignin sources and mixing dynamics of SOM, we novelly use the dual isotopic analysis ($\delta^{13}C$ and δ^2 H values) of LMeO and apply the Bayesian framework mixing model, MixSIAR, to apportion lignin sources in organic and soil horizons. Additionally, considering the relatively high percentage of LMeO in terrestrial biosphere which has been observed to reach up to 2 % of the OC in soils (Kosaka and Honda, 1956), and extremely $\delta^{13}C$ depleted values of

leaf litter, we explore the possibility that the well-known, but to some extent, unexplained ^{13}C enrichment with organic matter decomposition can be partially explained by the change in lignin sources from leaf litter to the roots.

4.2 Methods

Sites and sampling strategy

Two sites were selected for the investigation of lignin mixing dynamics. These research sites are situated in the Black Forest in Southern Germany (Site 1: 47.658100 N, 7.784390 W; Site 2: 47.652130 N,7.776240 W). Both sampling sites were previously studied to understand the early-stage isotopic degradation of very long chain fatty acids (Hirave et al., 2020). To account for spatial variations, three spatial replicates (A, B, C) were excavated using a spade (30 × 30 cm) at each sampling site (June 2022) with an approximate distance of 10–20 m between them. To further improve the representativeness, each spatial replicate was a composite mixture of three subsamples taken at 2 m apart. The soil type at Site 1 is classified as a stagnosol (IUSS Working Group WRB, 2022). At Site 1A and 1C, *Fagus sylvatica* was the predominant vegetation, while Site 1B was stocked with a combination of *Abies alba* (Silver fir) and *Fagus sylvatica* (European beech). Site 2 soil was characterized as a podzol (IUSS Working Group WRB, 2022) and was predominantly stocked with *Pinus sylvestri*s (Scots pine) and *Calluna vulgaris* (heather). Site 2B had an additional presence of *Fagus sylvatica* (European beech). leaves, while Site 2C featured an additional cover of *Quercus petraea* (Sessile oak) litter. Using the WRB classification system, four distinct classes were identified and collected: Oi (comprising of only leaves and needles, i.e., photosynthesizing tissues), Oe (partially decomposed layer), Oa (characterized by a very dark layer of well-decomposed humus, only present in the podzol), and the Ah horizon (mineral soil). In addition, non-photosynthesising tissues such as above-ground woody material (AGW), i.e., twigs and branches with a diameter < 1 mm) and large roots (LR, diameter 1–5 mm) were collected through all horizons. Soil aggregates were removed from the AGW and root samples using an ultrasonication bath. All samples then underwent homogenization using a ball mill.

Sample preparation and analysis

Carbon content and isotope analysis

Bulk stable carbon isotopes and carbon content were measured by mass spectrometry coupled to an SL elemental analyser (Integra2, Sercon, UK). Throughout this study, all stable carbon isotope ratios are expressed in the conventional 'delta' (δ) notation, representing the relative difference in the isotope ratio of a substance compared to the standard substance Vienna Peedee Belemnite (VPDB).

Lignin methoxy groups extraction

Prior to conversion of LMeO to methyl iodide (MeI) by the Zeisel reaction (Zeisel, 1885), ester bound MeO groups were removed using alkaline hydrolysis (Greule and Keppler, 2011; Cox et al., 2024). To ensure similar amounts of analyte in each sample, 1 ml of 1 M NaOH was added to 2–50 mg of sample (~5 mg for Oi, Oe and Oa horizons, and ~10 mg for LR, FR and AGW, and ~30 mg for the Ah horizon) in 1.5 ml vials, vials were capped and heated at 90 ◦C for 4 h. Afterwards, samples were uncapped and

dried at 60 ◦C in a sand bath. To ensure hydroiodic acid (HI) was not used up by the neutralization of NaOH or the removal carbonates within the mineral soil samples (Harris et al., 2001), 100 µl deionized water was added to each sample followed by acid fumigation for 24 h in 37 % fuming hydrochloric acid (HCl). The remaining ether bound LMeO groups were converted to MeI by the addition of 500 µl of HI (57 %, Sigma Aldrich, Stabilized) and heated for 1 h at 130 ◦C. After allowing the samples to equilibrate at room temperature for 1 h, samples underwent concentration and isotopic analysis.

Concentration of LMeO

After the conversion of LMeO to MeI, the concentration of MeI was determined utilizing static headspace injection (Greule et al., 2009). A headspace volume ranging from 10 to 90 µl was manually injected (Hamilton, 100 µl, gas-tight, side-port) into the Trace Ultra gas chromatograph (GC) equipped with a flame ionization detector (FID; Thermo Scientific, Waltham, MA 02451, USA) system. GC-FID conditions were set as: 200 ◦C inlet temperature, a helium flow rate of 1.8 ml/min, and an isothermal oven temperature of 65 ◦C, resulting in an approximate elution time of 4 mins. The MeI analyte was quantified using an external calibration curve generated from compounds with known MeO contents, vanillin (Sigma Aldrich, 99 %) and birch wood reference material, HUGB3 (Greule et al., 2020). Additional vanillin and HUGB3 samples were treated identical to samples and used as quality control samples. The concentration of the samples was analysed by the single injection of triplicate samples. The calibration ranged from 0.001 to 0.5 mg MeO, with a r 2 > 0.97 (Pearson's correlation) for all analysis. Using the HUGB3 and the vanillin quality control samples, we found the mean precision (1 SD) to be 0.18 % and 0.41 %, respectively (root mean square error (RMSE); Vanillin 0.52 %: HUGB3 0.18 %).

Dual Isotopic composition of LMeO

The $\delta^{13}C$ isotopic composition of MeI generated by the Zeisel reaction (Zeisel, 1885) was determined utilizing a Trace Ultra GC instrument, connected online via a GC-Isolink to a Conflo IV and Delta V Advantage isotope ratio mass spectrometer (Thermo Scientific, Waltham, MA 02451, USA). To ensure that non-desirable compounds (e.g., HI) are not transported into the instrument components upstream of the combustion stage, a reduction stage was installed between the Nafion water trap and the He backflush valve as described in Cox et al. (2024). The inlet temperature was set at 200 ◦C, and the helium (He) column flow rate was maintained at 1.8 ml/min. The initial oven temperature was set to 30 ◦C for 3.8 min, with a ramp at 30 ◦C per minute until reaching 100 ◦C.

Reference materials HUGB1 ($\delta^{13}C$ VPDB = − 50.17 ± 0.08 ‰) and HUGB4 ($\delta^{13}C$ V-PDB = − 30.07 ± 0.10 ‰) were employed for the normalization of sample isotopic ratios (Greule et al., 2019, Greule et al., 2020). HUGB3 ($\delta^{13}C$ V-PDB = − 29.30 ± 0.10 ‰) underwent identical treatment to the samples and served as a quality control sample throughout the sequence. Using a two-point calibration of HUGB1 and HUGB4, the quality control samples of HUGB3 were observed to have a RMSE of 0.5 ‰.

The δ^{2} H isotopic composition of MeI was determined using an HP 6890 N gas chromatograph (Agilent, Santa Clara, USA) equipped with an auto sampler A200S (CTC Analytics, Zwingen, Switzerland), coupled to a DeltaPLUSXL isotope ratio mass spectrometer (Thermo Fisher Scientific, Bremen,

Germany) via a thermo conversion reactor (ceramic tube (Al2O3), length 320 mm, 0.5 mm i.d., reactor temperature 1450 ◦C) and a GC Combustion III Interface (ThermoQuest Finnigan, Bremen, Germany). The GC was fitted with a Zebron ZB-5MS capillary column (Phenomenex, Torrance, USA) (30 m × 0.25 mm i.d., df 1 μm). GC conditions were set as: split injection (4:1), initial oven temperature at 30 ◦C for 3.8 min, ramp at 30 ◦C/min to 100 ◦C. Helium was used as carrier gas at a flow of 0.6 ml/min constant flow.

All hydrogen isotope ratios are reported in comparison to the Vienna Standard Mean Ocean Water (VSMOW). Reference methyl sulfate salt material HUGB2 (δ^2 H VSMOW = − 102.0 ± 1.3 ‰) and HUGB3 (δ^2 H VSMOW = − 272.9 ± 1.5 ‰) were used for the normalization of sample isotopic ratios (Greule et al., 2019; Greule et al., 2020). Additional samples of HUGB3 and Tineo wood reference material, HUGB5 (δ^2 H VSMOW = − 191.7 ± 0.8 ‰) were treated identically to samples and used as a quality control sample throughout the sequence (RMSE: HUGB3 3.2 ‰; HUGB5 5.3 ‰).

Unmixing and model evaluation

Contribution of the Oi, AGW and root sources to the organic and mineral horizons (Oe, Oa and A) were modelled using the open-source R package, MixSIAR (Stock et al., 2018). While simple linear equation can be used for source apportionments, MixSIAR incorporates uncertainty and source variance into the model estimates. MixSIAR was run with concentration dependency and uninformative priors. Using the molecular weights of hydrogen (1.008 Da), carbon (12.011 Da), and methoxy groups (31.034 Da), the concentrations of the δ^{13}C and δ^2 H LMeO tracers were calculated as mg of carbon/hydrogen derived from LMeO groups per gram of organic carbon (OC). All MixSIAR runs used the same model parameters: chains = 3, chain length = 3000, thin = 100, burn = 200,000 with a ‘long’ run time. The stagnosol and podzol sites were unmixed independently, using the spatial replicates (A, B, and C) to generate source distributions for each site. The variance of each horizon was incorporated into the model by using the “Residual only” error structure and using horizon spatial replicates as mixture replicates. The performance of the model was evaluated using 150 concentration dependent mathematical mixtures with known source proportions. Mathematical mixtures were generated using the open-source python script of Cox et al. (2023) using Eq. (1). Where V is the mean isotopic value of the tracer t, C refers to the mean concentration for all (∀) tracers in a set (∈) of tracers T in source S. SO refers to the number of sources and P refers to the known proportions of the mathematical mixtures.

$$V_t = \frac{\sum_s^{so}(C_{s,t} \times P_{s,t} \times V_{s,t})}{\sum_s^{so}(C_{s,t} \times P_{s,t})}\ ,\ \forall\, t \in T \qquad \text{Eq.1}$$

However, the use of “Residual only” error structure, which incorporates both source and mixture variance, requires a distribution of mathematical mixtures rather than a single value. Similar to Vale et al. (2022), replicate tracer values (xi) for each mixture were generated according to Eq. (2) by randomly drawing samples from a normal distribution using the tracer value of the single mixture (Vt) generated by Eq. (1) as the mean. The variance of the mixture distribution (σ2) was set as the mean SD of horizons

from the same soil sites. Estimated proportions were then compared to the proportions used to generate the mathematical mixtures.

$$V_{t_i} \sim N(Vt, \sigma 2), i = 1,2, \dots ,10, \forall t \in T \quad \text{Eq.2}$$

$\delta^{13}C$ Non-MeO mass balance and contribution of $\delta^{13}C$ LMeO to the $\delta^{13}C$ bulk enrichment

To determine the $\delta^{13}C$ non-LMeO values, the mass balance equation Eq. (3) was applied. Where fLMeO is the fraction of LMeO in the total organic carbon. The molecular weights (MW) are used to convert the concentration of LMeO to the concentration of carbon derived from LMeO.

$$\delta^{13}C\ Non\ MeO = \frac{\delta^{13}C\ bulk - \delta^{13}C\ MeO * \left(f_{LMeO} * \left(\frac{MW\ C}{MW\ MeO}\right)\right)}{1 - \left(f_{LMeO} * \left(\frac{MW\ C}{MW\ MeO}\right)\right)} \quad \text{Eq.3}$$

The enrichment of $\delta^{13}C$ bulk from Oi to the Ah horizon attributed to LMeO was calculated by two methods; i) Eq. (4), calculating the difference between the Oi and Ah horizon for both the $\delta^{13}C$ bulk and $\delta^{13}C$ Non-MeO fraction, with the difference being the enrichment caused by the LMeO fraction.

$$\Delta\delta^{13}C\ Bulk\ \because LMeO\ = \left(1 - \frac{\delta^{13}C\ Non{-}LMeO\ Oi - \delta^{13}C\ Non{-}LMeO\ Ah}{\delta^{13}C\ Bulk\ Oi - \delta^{13}C\ Bulk\ Ah}\right) * 100 \quad \text{Eq.4}$$

ii) Plotting a categorical line plot of the $\delta^{13}C$ Bulk and $\delta^{13}C$ non-LMeO values in relation to each horizon. As a slope of 0 would indicate no enrichment has taken place, the slope angle can be used as a semi-quantitative proxy for degradation. The $\delta^{13}C$ enrichment induced by LMeO is then seen as the disparity between these two slopes, and is reported as a percentage of the $\delta^{13}C$ Bulk, Eq. (5). See Fig. S1 for further information. The uncertainty reported throughout represents the SD of the three spatial replicate sites (A, B, C) for each site.

$$\Delta\ \delta^{13}C\ Bulk\ \because LMeO\ = (1 - (Slope^{Horizons}_{^{13}C\ Bulk} / Slope^{Horizons}_{^{13}C\ Non\ MeO})) * 100 \quad \text{Eq.5}$$

4.3 Results

Measured $\delta^{13}C$ and $\delta^{2}H$ LMeO values of lignin sources and mixtures

The Oi horizon contained the lowest $\delta^{13}C$ LMeO values in both sites (stagnosol; $\delta^{13}C$: − 58.8 ± 0.9 ‰: podzol; $\delta^{13}C$: − 53.9 ± 2.7 ‰) (Fig. 1, Table S1.1, S1.2, S2). The most ^{13}C enriched LMeO values were found in LR of the stagnosol ($\delta^{13}C$: − 32.4 ± 1.5 ‰) and the AGW of the podzol ($\delta^{13}C$: − 34.0 ± 2.8 ‰). The Oi horizon $\delta^{13}C$ LMeO values were within the range of tree leaves previously reported (Keppler et al., 2004). Our results of the $\delta^{13}C$ LMeO values of woody tissue were slightly more depleted compared to LMeO values previously reported (Greule et al., 2009; Greule et al., 2020). The $\delta^{13}C$ LMeO values

of FR and LR were found not to be different in the podzol sites, while differences were found in the stagnosol (Mann-Whitney-U test, p value: stagnosol 0.04; podzol 0.33). The results of the concentration analysis are available in the supplementary material (Fig. S2, Table S1.1, S1.2, S2).

$\delta^{13}C$ LMeO values of organic horizons (stagnosol: Oe; podzol: Oe, Oa) ranged from − 45.1 to − 50.8 ‰. While the $\delta^{13}C$ LMeO of soil organic horizons has not been assessed before, soil organic horizons are a mixture of both leaves and non-photosynthesising tissues (roots and above ground woody material), as such our results of $\delta^{13}C$ LMeO values are within a credible range (Fig. 1, Fig. 2). Mineral soil horizons (Stagnosol Ah; $\delta^{13}C$: − 41.7 ± 1.7 ‰: Podzol Ah; $\delta^{13}C$: − 38.8 ± 1.3 ‰) were similar to those reported by Cox et al. (2024) for soils in Switzerland ($\delta^{13}C$ LMeO: − 39.2 to − 47.8 ‰). Additionally, the organic and mineral horizons demonstrated ^{13}C LMeO enrichment with increasing organic matter decomposition.

Stagnosol non-photosynthesising tissues (LR, FR, and AGW) displayed significant differences in $\delta^{13}C$ LMeO values, whereas no differences were observed in podzol (Kruskal Wallis, p value: stagnosol 0.03; podzol 0.8). Differences were observed between non-photosynthesising tissues (LR, FR, AGW) and photosynthesising tissues (Oi) in both sites (Mann-Whitney-U test, p value: stagnosol

Similar to the $\delta^{13}C$ LMeO values, $\delta^{2}H$ values of the Oi horizon was found to be most isotopically depleted in 2 H in the podzol ($\delta^{2}H$: − 208.4 ± 1.3 ‰). In the stagnosol, both the Oi horizon ($\delta^{2}H$: − 195.4 ± 9.5 ‰) and AGW ($\delta^{2}H$: − 196.0 ± 3.4 ‰) were observed to be the most 2 H depleted. The most isotopically enriched 2 H LMeO values were found in the FR of the stagnosol ($\delta^{2}H$: − 159.4 ± 4.6 ‰), and the LR of the podzol ($\delta^{2}H$: − 153.8 ± 18.9 ‰). We found roots to be more enriched in 2 H compared to the Oi horizon, aligning with differences in $\delta^{2}H$ values previously reported for non-exchangeable bulk hydrogen (Debond et al., 2012; Ruppenthal et al., 2015; Guidi et al., 2023). No significant difference was found between the $\delta^{2}H$ LMeO values of LR and FR in both sites (Mann-Whitney-U test, p value: stagnosol 0.27; podzol 0.51).

Comparable to the $\delta^{13}C$ LMeO values, the $\delta^{2}H$ LMeO values of organic and mineral horizons are also located between the lignin sources (Fig. 1, Fig. 2). However, in comparison to the stagnosol Ah horizon at Sites A and B, the Site C Ah horizon was found to exceptionally 2 H enriched ($\delta^{2}H$: − 144.2 ± 6.4 ‰) and outside the sources mixing space (Fig. 2, S2). Both sites demonstrated an enrichment of 2 H LMeO with increasing organic matter decomposition. In contrast to $\delta^{13}C$ LMeO values, AGW exhibited $\delta^{2}H$ LMeO values more similar to the Oi horizon than root material. Differences were not observed between the $\delta^{2}H$ LMeO values of non-photosynthesising and photosynthesising tissues in the stagnosol but were present in the podzol (Mann-Whitney-U test, p value: stagnosol 0.07; podzol <0.01) Fig.1).

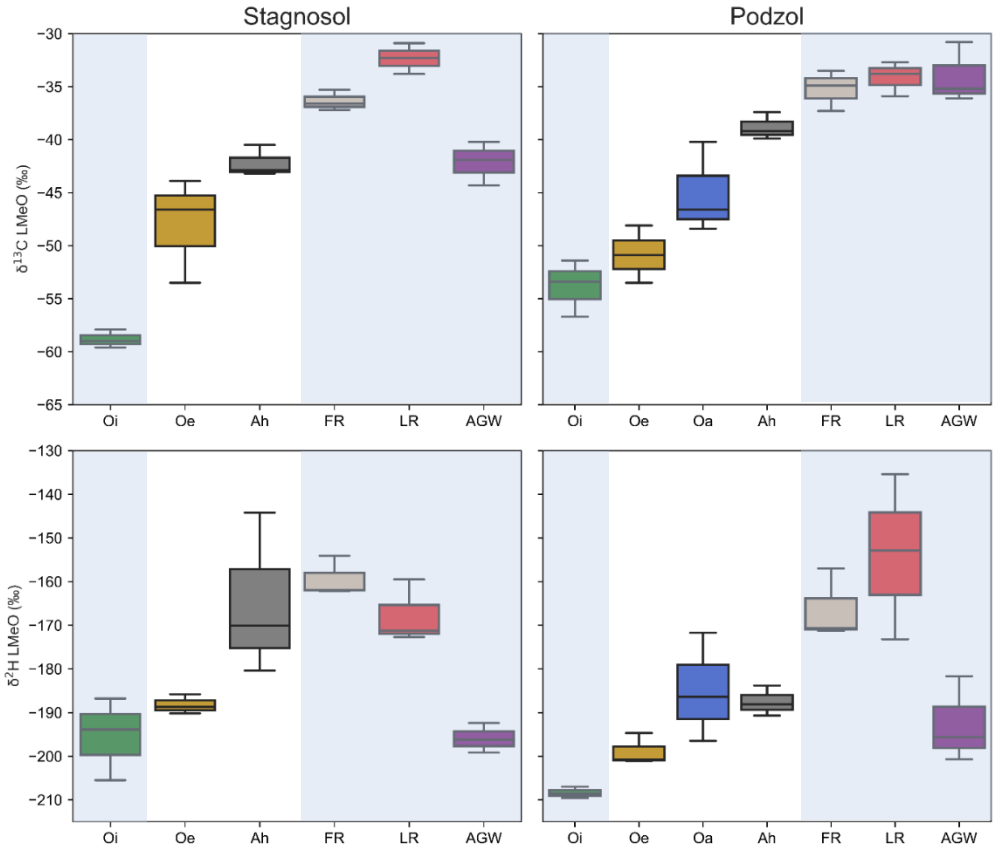


Fig. 1 $\delta^{13}C$ and $\delta^{2}H$ LMeO values of lignin sources and horizons. The boxes represent 25, 50 and 75% quantiles with whiskers showing a 1.5 interquartile range. Sources are highlighted in blue

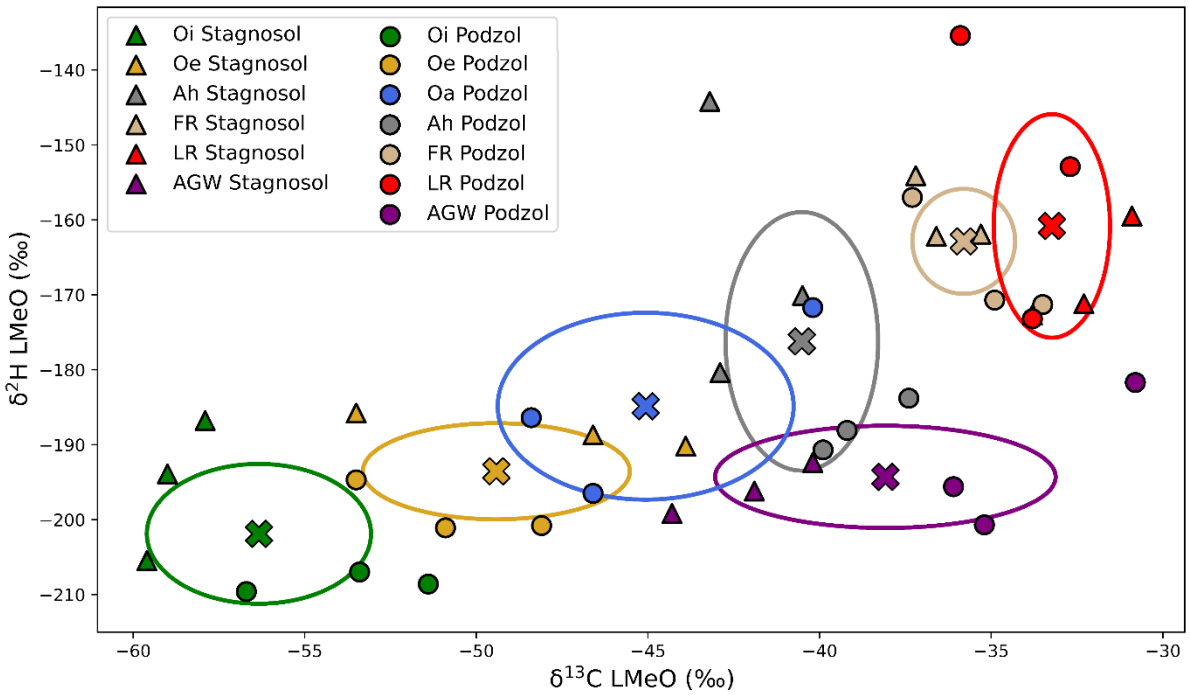


Fig. 2 Bi-plot of the $\delta^{13}C$ and $\delta^{2}H$ LMeO values of lignin sources and horizons. Mean values of both sites are denoted by the 'X' symbol, with eclipses illustrating the standard deviation of the spatial replicates of each source or horizon

Lignin source apportionment and model evaluation

The model evaluation using mathematical mixtures indicated excellent model performance with model estimates compared to known proportions having a Pearson correlation coefficient of r 2 > 0.97. In the stagnosol, mathematical mixtures containing high proportions of AGW were shown to be underestimated (~10 %) (Fig. S3). Mathematical mixtures of the podzol demonstrate an overestimation of the lower contributions of Oi, and a small but consistent underestimation of root contribution by ~ 5 % (Fig. S4).

Source contributions in the stagnosol Oe horizon were found to dominated by both AGW (median values: 32 %) and Oi material (42 %), while contribution from root material was estimated at 25 % (Fig. 3). Contribution from Oi (19 %) and AGW (19 %) sources decreased in the Ah horizon, with roots being the dominate source (55 %) (see Table S3 for source apportionment summary).

In the Oe horizon of the podzol, the Oi horizon lignin was the dominate lignin source (80 %), with lower input from roots (11 %) and AGW (8 %). In the Oa horizon, there was a decrease in Oi contribution (52 %) accompanied by an increase in root contribution (31 %), while AGW had a smaller contribution (14 %). The Ah horizon of the podzol showed an increase in AGW lignin contribution (37 %), with similar contributions from Oi and roots (Oi: 30 %; Roots: 30 %). (See Table S4 for source apportionment summary).

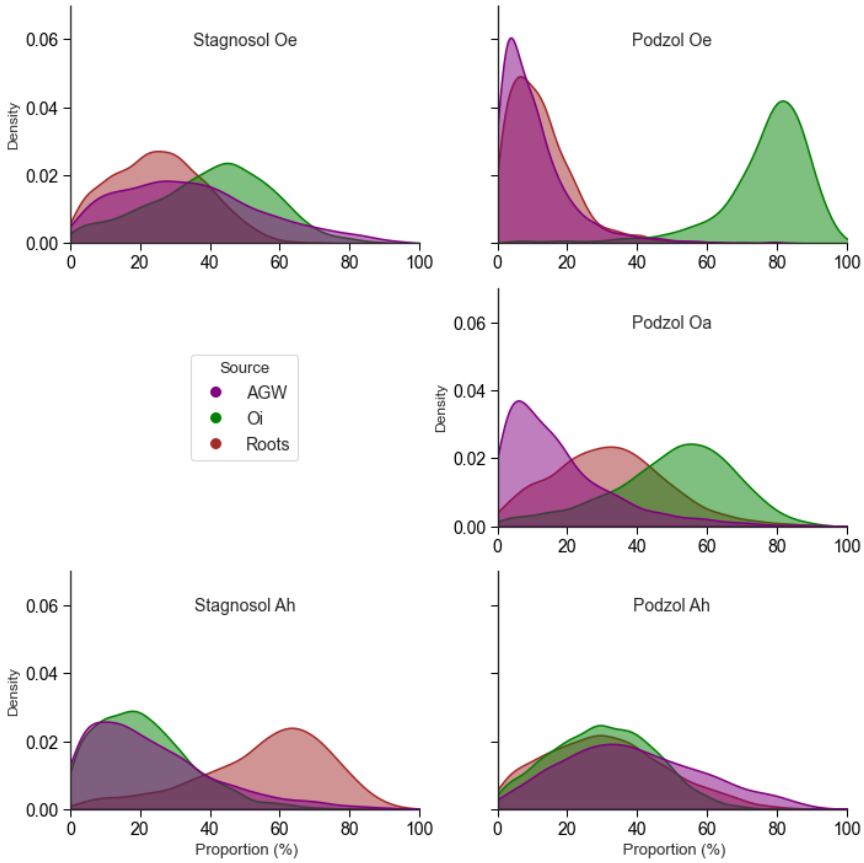


Fig. 3 Density plots of the contributions of lignin source (Oi, AGW and Root) to the Oe, Oa and Ah horizon estimated using MixSIAR

Calculated $\delta^{13}C$ values of the non-LMeO fraction

Using the mass balance equation Eq. (3), we calculated the $\delta^{13}C$ values of the non-LMeO fraction. For the stagnosol and podzol Oi horizon (Table S1 and S2), the $\delta^{13}C$ non-LMeO values were determined to be − 29.8 ± 0.3 ‰ and − 29.2 ± 0.2 ‰, respectively. In the Oe horizon, the $\delta^{13}C$ non-LMeO values were − 29.0 ± 0.6 ‰ for the stagnosol and − 29.3 ± 0.2 ‰ for the podzol. The podzol Oa horizon exhibited $\delta^{13}C$ values of the non-LMeO fraction at − 29.0 ± 0.3 ‰. The stagnosol Ah horizon displayed $\delta^{13}C$ non-LMeO values of 27.7 ± 0.7 ‰, while the podzol Ah horizon $\delta^{13}C$ non-LMeO values of − 27.8 ± 0.1 ‰. The $\delta^{13}C$ non-LMeO values for stagnosol AGW were − 28.7 ± 0.7 ‰, and − 27.3 ± 2.0 ‰ for the podzol AGW. The roots δ13C non-LMeO values were − 28.7 ± 0.7 ‰ and − 29.4 ± 1.3 ‰ in the stagnosol and podzol, respectively.

Calculated bulk $\delta^{13}C$ enrichment caused by the LMeO fraction

The difference in $\delta^{13}C$ enrichment between the Oi to Ah horizon for both the $\delta^{13}C$ bulk and $\delta^{13}C$ non-LMeO fraction values was used as the basis for estimating the enrichment induced by LMeO. Applying Eq. (4), the percentage of $\delta^{13}C$ bulk enrichment caused by LMeO was calculated at 14 ± 4 % for the stagnosol and 11 ± 4 % for the podzol. Using Eq. (5), an identical proportion of the $\delta^{13}C$ bulk enrichment was attributed to the LMeO fraction in the stagnosol (14 ± 4 %). However, the LMeO-induced ^{13}C bulk enrichment calculated using Eq. (5) was only 4 ± 3 % in the podzol.

4. Discussion

$\delta^{13}C$ LMeO and δ^{2} H end member discrimination

The dual isotope ($\delta^{13}C$ and δ^{2} H values) analysis of LMeO revealed the ability of $\delta^{13}C$ LMeO values to discriminate between woody and photosynthesising tissues, while δ^{2} H LMeO values demonstrated the capacity to differentiate between above and below ground tissues (Fig. 4). The dual isotope approach then allowed for the discrimination between Oi, AGW and root lignin sources. In the following discussion, we find it acceptable to use the term Oi interchangeably with leaves. While in this study, we did not measure fresh leaves, we would not expect differences in $\delta^{13}C$ LMeO values between fresh and undecomposed leaf litter. Additionally, our results of Oi are similar to fresh leaves $\delta^{13}C$ LMeO values previously reported (Keppler et al., 2004).

The presence of the enriched ^{13}C in bulk woody material compared to leaf material has been suggested to arise from external parameters such as stress, seasonal variations, temperature, and light, affecting the degree of isotopic fractionation during CO2 uptake with different plant tissue growth occurring at different stages (Cernusak et al., 2009). However, as MeO have a significant contribution to bulk $\delta^{13}C$ values, a proportion of the enrichment in bulk $\delta^{13}C$ values can be attributed to the enrichment in $\delta^{13}C$ MeO values (Keppler et al., 2004). Additionally, the large isotopic difference between leaf litter and woody material $\delta^{13}C$ LMeO values (~30 ‰) implies the involvement of more complex processes involving isotopic fractionation.

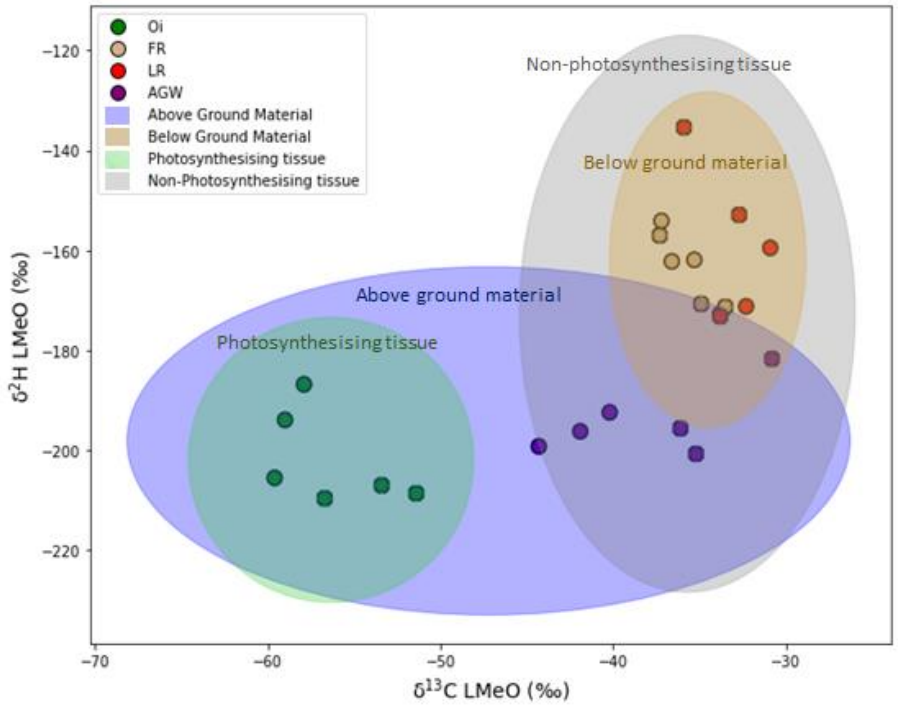


Fig. 4 The dual isotopic analysis demonstrates tissue type clustering. δ^{13}C LMeO values distinguish between photosynthesising and non-photosynthesising tissues, while δ^{2}H LMeO values differentiate between above and below ground tissues. This approach enables discrimination between Oi, AGW, and root tissues.

As lignin serves as a structural component of plants, it is not considered to be transported through the plant's vascular system (xylem and phloem) (Boerjan et al., 2003). However, the precursors of MeO groups (e.g., serine) have been reported to be both transported and synthesized across various plant tissues (Ros et al., 2014). As described in Greule et al. (2021) and Lloyd et al. (2023), MeO groups originate from the methylene group of serine. Serine has been shown to be a product of photorespiration (glycolate pathway), the Calvin cycle, Glycolysis (glycerate pathway), and the formate pathway (Kisluik, 1955; Ros et al., 2014).

The CH_2-unit (methylene group) originates from the position C-3 of serine and provides the carbon atom and two hydrogen atoms for the methoxy group, the third H comes from NADH (Greule et al., 2021; Lloyd et al., 2023). The CH_2 unit of serine is found to be slightly depleted in ^{2}H and corresponds to the plant source water (up to − 50 mUr, Augusti et al., 2006; Zhang et al., 2002). However, when certain flavoproteins transfer the third hydrogen from NADH to from the CH_3 unit strong hydrogen isotope fraction occurs (− 580 to − 790 mUr (Martin et al., 2004; Billault et al., 2001), leading to a total ^{2}H fractionation of about − 200 mUr between lignin methoxy groups and precipitation (Greule et al., 2021). Carbon depletion in plant tissues relative to atmospheric CO_2 is due to fractionation processes during diffusion into leaves and carbon fixation via RuBisCo (Francey and Farquhar, 1982). Leaves of C^{3} plants

are more ^{13}C depleted than stems, roots and seeds (Cernusak et al., 2009) and lignin is found to be more ^{13}C depleted than whole organic material or cellulose (Wilson and Grinsted, 1977; Cernusak et al., 2009). However, lignin methoxy groups showed an exceptionally depleted ^{13}C signature (Wieland et al., 2022, 2024; Keppler et al., 2004) were the stronger fractionation in lignin methoxy groups could occur in the last step of the C1 metabolism. Here the CH_3 unit is transferred from N_5-CH_3-THF to homocysteine to from methionine. This L-methionine synthase might include a large ^{13}C depletion, resulting in strongly depleted L-methionine.

We see strong differences between photosynthesising and nonphotosynthesising tissues in $\delta^{13}C$ LMeO values similar as reported by Cernusak et al. (2009) and large differences between above and below plant material in $\delta^{2}H$ LMeO values. There are four possible scenarios for the observed LMeO isotopic values, which involve precursor production or transport: i) all serine is formed in the chloroplast and are transported from the leaves through the phloem to lignifying cells, ii) serine formation significantly occurs in non-photosynthesising plant tissues through different pathways, iii) methionine is formed in the chlorophyll and transported to different plant compounds or iv) a mixture of these formation pathways. The third hydrogen (from NADH) is product of either photosynthesis or the pentose phosphate cycle, with two possible scenarios here: a) the NADH used in different plant tissues comes for the same product or b) photosynthesis cells use NADH from the photosynthesis while non-photosynthesising tissues use NADH from the pentose phosphate cycle (Schnell et al., 2012; Shin, 2004).

Combining the different scenarios for the methoxy formation we have then following possibilities: i–a; would mean that $\delta^{13}C$ LMeO and $\delta^{2}H$ LMeO values should be similar in different plant tissues, but in scenario i-b, $\delta^{2}H$ LMeO values are NADH-H source pathway dependent. NADH formed by photosynthesis has been observed to be extremely (> 100 ‰) depleted in ^{2}H compared to the pentose phosphate pathway (Luo et al., 1991; Schmidt et al., 2003). As such, the third hydrogen (from NADH) may result in overprinting and concealing any $\delta^{2}H$ differences in serine formation pathways.

ii-a Differences in $\delta^{13}C$ LMeO values, however as above, differences in $\delta^{2}H$ values may be overprinted by the third NADH-H from the same source. While ii-b would lead to differences in both $\delta^{2}H$ and $\delta^{13}C$ LMeO values.

iii-a; Would imply uniform $\delta^{2}H$ and $\delta^{13}C$ LMeO values across all plant tissues, which contradicts observed data.

iv) The isotopic signature of LMeO, influenced by both transported serine and serine formed through various pathways, would be dictated by the dominant process, whether it is transportation or formation. Therefore, it becomes challenging to precisely identify the specific process contributing more significantly to the isotopic signature.

Using the results of the dual-isotope approach of this study, we suggest scenario ii to be the most probable pathway for explaining the patterns of $\delta^{13}C$ LMeO values, as large differences between wood and leaf $\delta^{13}C$ LMeO values are observed. However, the variability between leaf and AGW $\delta^{2}H$ LMeO values are much smaller - indicating a similar NADH-H source (a), whereas roots are strongly enriched in $\delta^{2}H$ indicating a different NADH-H source (b). The discrimination between above and below ground

tissues suggests LMeO precursors formed by photosynthesising tissues might not be transported to the roots. As such, it is likely that the isotopic variation of LMeO observed in our data arises from several serine formation pathways and different NADH sources (photosynthesis or the pentose phosphate cycle) of above and below wood material. Presently, there is limited understanding of the isotopic fractionation associated with each serine formation pathways. For more discussion regarding the potential isotope fractionation of LMeO methyl groups derived from serine and involved in C1 carbon metabolism we refer to the Supplementary Information.

The ^{13}C depleted LMeO values observed in leaves in comparison to the AGW could also come from a partially open photorespiration cycle, in which isotopically enriched serine is taken out of the 3PGA cycle and transported to the xylem and used in the woody tissue (Lloyd et al., 2023). Additionally, another plausible explanation may be the reassimilation of isotopically light CO_2 produced by photorespiration, resulting in a fraction of isotopically depleted CO_2 being re-assimilated back into the leaves and into subsequently converted to serine. Kinetic fractionation effects would consequently lead to the rapid metabolism of the isotopically depleted serine near the assimilation site, resulting in isotopically depleted ^{13}C values in leaves. Furthermore, as non-photosynthesising tissues cannot re-assimilate respired CO2, the ^{13}C LMeO of heterotopic would be enriched in ^{13}C compared to the phototropic tissue, as our findings demonstrate. While the proportion of reassimilated CO_2 from tree leaves is currently unknown, rice and wheat plants have shown to significantly re-assimilate respired CO_2 (24–38 %) (Busch et al., 2013).

Organic horizons isotopes

In general, the $\delta^{13}C$ and $\delta^{2}H$ LMeO values observed in the organic and mineral horizons were within the range defined by the LMeO sources. However, the mineral Ah horizon of the Stagnosol site C demonstrated exceptionally depleted $\delta^{2}H$ LMeO values and was located outside the possible mixing space of the sources. As all other Ah horizons were within the source range, we propose the outlying site C Ah horizon is due to the contribution of lignin from a missing source in the Site C Ah horizon. As this Ah horizon demonstrates an isotopic fingerprint more comparable to the root material than leaf material, we suggest the outlier is from roots of a plant/ tree species not included in the Oi and root sampling. Therefore, as our sampling appears to of missed a potential source, the site C Ah horizon is left out the following source apportionment.

While $\delta^{2}H$ MeO values have been demonstrated to be stable during MeO degradation (Anhauser et al., 2015; Lu et al., 2022), the stability of $\delta^{13}C$ LMeO during degradation processes of organic matter is less investigated. Fig. 2 illustrates the Stagnosol organic and mineral horizons plotting on the linear mixing line between Oi and root sources, suggesting minimal contribution from AGW. In a two-source system, if ^{13}C LMeO fractionation was to occur with compound degradation, there would be a deviation from the expected mixing line. This deviation would be characterized by an enrichment of ^{13}C compared to the prediction outlined by the mixing line (Fig. S5) (Lutz and Van Breukelen, 2014a, 2014b). This is not observed in the Stagnosol, suggesting the isotopic stability of $\delta^{13}C$ LMeO values and its applicability for tracing lignin sources.

In the case of the podzol Ah horizon, a deviation from the two-source mixing line is observed. Two reasons for this are proposed: i) Contribution from AGW material in the Ah horizon, ii) ^{13}C fractionation occurring during LMeO degradation. The LMeO fraction remaining in the podzol is significantly higher than that remaining in the stagnosol (Fig. S2, Table S1 and S2). If isotopic fractionation occurred with LMeO degradation, one would expect to observe this also in the stagnosol, as this is not observed, we propose that the deviation from the mixing line is the result of contribution from AGW. Furthermore, the formation of podzol soils is reliant on precipitation being considerably greater than evapotranspiration, leading to the leaching of organic compounds. Consequently, we propose that lignin from the slower degrading above ground woody material is leached through the organic horizons and accumulating in the Ah horizon (Fig. 3) and stabilized by organic mineral associations (Schmidt et al., 2000).

Apportionment of lignin sources

The relative above and below ground contributions are reported to be an important factor in SOM formation (Freschet et al., 2013; Griepentrog et al., 2015; Berhongaray et al., 2019). Our observations at the stagnosol site reveal a relatively small contribution of root and AGW tissues in the upper Oe horizon, with both AGW and root lignin contributions being similar (Fig. 3). As AGW degradation occurs at a slow pace (Freschet et al., 2012), the accumulation of AGW lignin can be attributed to the minimal leaching and translocation of organic compounds in the stagnosol. This discrepancy between lignin sources in the Oe and Ah horizons highlights the reduced leaching processes occurring in the stagnosol, ensuing in a reduced input of lignin from above-ground sources in lower horizons. Our estimates of lignin contribution from roots (56 %) align well with global estimates of OM input, which suggest that root litter accounts for 48 % of the plant litter OM input in the forest soils (Freschet et al., 2013) with a higher proportion of root litter (45 %) being transferred to the SOM than leaf litter (8 %) (Jackson et al., 2017). In the stagnosol our results of roots being the dominate source aligns with that previously reported using lignin monomers (Wang et al., 2018).

In the podzol profile, we observed a dominance of lignin derived from the Oi layer in the Oe horizon, with a gradual decrease of Oi contribution towards the Ah horizon. This trend aligns with expectations, as the distance from the Oi horizon increases with depth through the soil horizons. Notably, lignin from AGW was found to be minimal until the Ah horizon. Our findings in the Ah horizon illustrate the gradual leaching of the slowly degrading AGW lignin through the Oe and Oa horizons, ultimately accumulating in the Ah horizon due to mineral associations. While the mechanism for lignin stabilisation by physicochemical processes has not been elucidated (Rumpel et al., 2004; Spielvogel et al., 2008), the stability of lignin has been revealed to be related to the fine mineral and clay fraction (Thevenot et al., 2010). While meta-analysis has indicated a higher proportion of root litter OM being transferred to the SOM than leaf litter OM (Freschet et al., 2013), in the case of the podzol Ah horizon this is not withstanding. However, analysis of deeper mineral horizons would likely show a further increase of root lignin proportion. Our result show that while plant roots are an important source of carbon to SOM (Berhongaray et al., 2019; Meena et al., 2019), the relative contribution of lignin sources is dependent on the specific characteristics of each site.

Driving force behind the enrichment of bulk ^{13}C in organic horizon

The mass balance equation, Eq. (3), was applied to determine the $\delta^{13}C$ values of the non LMeO fraction of organic and mineral horizons. Eq. (4) and Eq. (5) were applied to estimate the amount of ^{13}C enrichment explained by LMeO groups. Using Eq. (4) our results indicate that the changing LMeO sources explain 14 % and 11 % of the enrichment of ^{13}C Bulk from the Oi to Ah horizons in podzol and stagnosol, respectively. Estimates using Eq. (5) showed identical results for the stagnosol, however using Eq. (5), only 4 % of the podzol bulk ^{13}C enrichment could be explained by LMeO. The variance between these two values can be assumed to be a result of the non-linear ^{13}C enrichment found in the podzol. While various factors contribute to the enrichment of bulk ^{13}C with soil depth and organic matter degradation, our study reveals that a significant portion of the enrichment, up to 14 %, can be attributed to the transition of lignin sources from Oi to woody tissues. This discovery holds considerable importance, particularly in terms of carbon sequestration and understanding the movement of carbon with organic and soil horizons. Our findings suggest that models relying on ^{13}C enrichment with depth as an indicator of carbon turnover might be exaggerating turnover rates. Importantly, Keppler et al. (2004) and Cox et al. (2024) demonstrated that needle litter $\delta^{13}C$ LMeO values are more similar to woody material than leaf litter. This similarity between woody material may diminish the effect of shifting lignin sources in coniferous forests, in which case only turnover rates in deciduous forests may be significantly overestimated.

Conclusion

The appointment of lignin sources in SOM regularly faces challenges from the complex structures of lignin. Our methodology represents a significant advancement over previous techniques by offering a relatively rapid and solvent free dual isotopic ($\delta^{13}C$ and $\delta^{2}H$ values) analysis of LMeO enabling the discrimination of different lignin sources, and ultimately the apportionment of lignin in SOM. Here, we were able to show that the relative contribution of lignin sources is not only dependent on the specific characteristics of each site but also dependent on the different horizons within the soil profile. This not only enhances our understanding of carbon dynamics and sequestration in soils but also helps elucidate the intricate mechanisms these processes. Our results contribute to the understanding and modelling of carbon turnover and sequestration in soils. While the mechanism for the large LMeO isotopic discrimination between different plant tissues is currently not well understood, the dual isotopic analysis of LMeO in specific trees may hold the potential for understanding serine forming pathways, and potentially how these pathways respond to external factors (e.g., temperature, drought).

Our findings suggest that the substantial $\delta^{13}C$ enrichment observed with depth partly arise from the changes in the composition of lignin sources, particularly from leaf litter to root tissues, which is usually attributed to organic matter decomposition. As such, the use of bulk $\delta^{13}C$ values as a tool to determine carbon turnover may result in an overestimation. Interestingly, this effect may not be as dominant in coniferous forests in which the needle litter may have a $\delta^{13}C$ LMeO values more similar to woody tissues.

References

Anhauser, T., Greule, M., Zech, M., Kalbitz, K., McRoberts, C., Keppler, F., 2015. Stable hydrogen and carbon isotope ratios of methoxyl groups during plant litter degradation. Isotopes Environ. Health Stud. 51 (1), 143–154. https://doi.org/ 10.1080/10256016.2015.1013540.

Augusti, A., Betson, T.R., Schleucher, J., 2006. Hydrogen exchange during cellulose synthesis distinguishes climatic and biochemical isotope fractionations in tree rings. New Phytol. 172 (3), 490–499. https://doi.org/10.1111/j.1469-8137.2006.01843.x.

Belmont, P., Willenbring, J.K., Schottler, S.P., Marquard, J., Kumarasamy, K., Hemmis, J. H., 2014. Toward generalizable sediment fingerprinting with tracers that are conservative and nonconservative over sediment routing timescales. J. Soil. Sediment. 14 (8), 1479–1492. https://doi.org/10.1007/s11368-014-0913-5.

Berhongaray, G., Cotrufo, F.M., Janssens, I.A., Ceulemans, R., 2019. Below-ground carbon inputs contribute more than above-ground inputs to soil carbon accrual in a bioenergy poplar plantation. Plant and Soil 434, 363–378. https://doi.org/10.1007/ s11104-018-3850-z.

Billault, I., Guiet, S., Mabon, F., Robins, R., 2001. Natural deuterium distribution in longchain fatty acids is nonstatistical: a site-specificstudy by quantitative 2 HNMR spectroscopy. Chembiochem 2, 425–431.

Boerjan, W., Ralph, J., Baucher, M., 2003. Lignin biosynthesis. Annu. Rev. Plant Biol. 54, 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938.

Busch, F.A., Sage, T.L., Cousins, A.B., Sage, R.F., 2013. C3 plants enhance rates of photosynthesis by reassimilating photorespired and respired CO2. Plant Cell Environ. 36, 200–212. https://doi.org/10.1111/j.1365-3040.2012.02567.x.

Cernusak, L.A., Tcherkez, G., Keitel, C., Cornwell, W.K., Santiago, L.S., Knohl, A., Barbour, M.M., Williams, D.G., Reich, P.B., Ellsworth, D.S., Dawson, T.E., Griffiths, H.G., Farquhar, G.D., Wright, I.J., 2009. Why are non-photosynthetic tissues generally 13C enriched compared with leaves in C3 plants? Review and synthesis of current hypotheses. Funct. Plant Biol. 36 (3), 199–213. https://doi.org/ 10.1071/FP08216.

Cox, T., Laceby, J.P., Roth, T., Alewell, C., 2023. Less is more? A novel method for identifying and evaluating non-informative tracers in sediment source mixing models. J. Soil. Sediment. 23, 3241–3261. https://doi.org/10.1007/s11368-023- 03573-0.

Cox, T., Laceby, J.P., Greule, M., Keppler, F., Alewell, C., 2024. Using stable carbon isotopes of lignin-derived methoxy to improve historical apportionments of particulate organic matter and sediment sources incorporating multiple Suess corrections. J. Soil. Sediment. https://doi.org/10.1007/s11368-024-03765-2.

Debond N., Fogel M.L., Morrill P.L., Benner R., Bowden R., Ziegler S. (2012) Variable dD values among major biochemicals in plants: implications for environmental studies. Geochim. Cosmochim. Acta111:117–127. doi:https://doi.org/10.1016/j.gca.201 2.10.043.

Dignac, M.F., Bahri, H., Rumpel, C., Rasse, D.P., Bardoux, G., Balesdent, J., Girardin, C., Chenu, C., Mariotti, A., 2005. Carbon-13 natural abundance as a tool to study the dynamics of lignin monomers in soil: an appraisal at the Closeaux experimental field (France). Geoderma 128 (1–2), 3–17. https://doi.org/10.1016/j. geoderma.2004.12.022.

Dümig, A., Knicker, H., Schad, P., Rumpel, C., Dignac, M.-F., Kogel-Knabner, ¨ I., 2009. Changes in soil organic matter composition are associated with forest encroachment into grassland with long-term fire history. Eur. J. Soil Sci. 60, 578–589. https://doi. org/10.1111/j.1365-2389.2009.01140.x.

Feng, X., Simpson, M.J., 2007. The distribution and degradation of biomarkers in Alberta grassland soil profiles. Org. Geochem. 38, 1558–1570. https://doi.org/10.1016/j. orggeochem.2007.05.001.

Fernandez, I., Mahieu, N., Cadisch, G., 2003. Carbon isotopic fractionation during decomposition of plant materials of different quality. Global Biogeochem. Cycles 17 (3). https://doi.org/10.1029/2001gb001834. Francey, R.J., Farquhar, G.D., 1982. An explanation of 13C/12C variation in tree rings. Nature 297 (28–3), 1.

Freschet, G.T., Weedon, J.T., Aerts, R., van Hal, J.R., Cornelissen, J.H.C., 2012. Interspecific differences in wood decay rates: insights from a new short-term method to study long-term wood decomposition. J.Ecol. 100, 161–170. https://doi.org/ 10.1111/j.1365-2745.2011.01896.x.

Freschet, G.T., Cornwell, W.K., Wardle, D.A., Elumeeva, T.G., Liu, W., Jackson, B.G., Onipchenko, V.G., Soudzilovskaia, N.A., Tao, J., Cornelissen, J.H.C., 2013. Linking litter decomposition of above- and below-ground organs to plant-soil feedbacks worldwide. J. Ecol. 101 (4), 943–952. https://doi.org/10.1111/1365-2745.12092.

Galbally, I.E., Kirstine, W., 2002. The production of methanol by flowering plants and the global cycle of methanol. J. Atmos. Chem. 43, 195–229.

Glaser, B., 2005. Compound-specific stable-isotope (δ13C) analysis in soil science. J. Plant Nutr. Soil Sci. 168, 633–648. https://doi.org/10.1002/jpln.200521794.

Goni, M.A., Yunker, M.B., Macdonald, R.W., Eglinton, T., 2000. Distribution and sources of organic biomarkers in arctic sediments from the Mackenzie River and Beaufort shelf. Mar. Chem. 71, 23–51. https://doi.org/10.1016/S0304-4203(00)00037-2.

Greule, M., Keppler, F., 2011. Stable isotope determination of ester and ether methyl moieties in plant methoxyl groups. Isotopes Environ. Health Stud. 47, 470–482. https://doi.org/10.1080/10256016.2011.616270.

Greule, M., Mosandl, A., Hamilton, J.T.G., Keppler, F., 2009. A simple rapid method to precisely determine 13C/12C ratios of plant methoxyl groups. Rapid Commun. Mass Sp. 23, 1710–1714. https://doi.org/10.1002/rcm.4057.

Greule, M., Moossen, H., Geilmann, H., Brand, W.A., Keppler, F., 2019. Methyl sulfates as methoxy isotopic reference materials for δ13C and δ2 H measurements. Rapid Commun. Mass Sp. 33, 343–350. https://doi.org/10.1002/rcm.8355.

Greule, M., Moossen, H., Lloyd, M.K., Geilmann, H., Brand, W.A., Eiler, J.M., Qi, H., Keppler, F., 2020. Three wood isotopic reference materials for δ2 H and δ13C measurements of plant methoxy groups. Chem. Geol. 533 https://doi.org/10.1016/j. chemgeo.2019.119428.

Greule, M., Wieland, A., Keppler, F., 2021. Measurements and applications of δ2 H values of wood lignin methoxy groups for paleoclimatic studies. Quat. Sci. Rev. 268 https:// doi.org/10.1016/j.quascirev.2021.107107.

Griepentrog, M., Eglinton, T.I., Hagedorn, F., Schmidt, M.W.I., Wiesenberg, G.L.B., 2015. Interactive effects of elevated CO2 and nitrogen deposition on fatty acid molecular and isotope composition of above- and belowground tree biomass and forest soil fractions. Glob. Chang. Biol. 21, 473–486. https://doi.org/10.1111/gcb.12666.

Guggenberger, G., Zech, W., 1994. The science of the Total environment dissolved organic carbon in forest floor leachates: simple degradation products or humic substances? Sci. Total Environ. 152, 24–47. https://doi.org/10.1016/0048-9697(94) 90549-5.

Guidi, C., Lehmann, M.M., Meusburger, K., Saurer, M., Vitali, V., Peter, M., Brunner, I., Hagedorn, F., 2023. Tracing sources and turnover of soil organic matter in a longterm irrigated dry forest using a novel hydrogen isotope approach. Soil Biol. Biochem. 184, 109113 https://doi.org/10.1016/j.soilbio.2023.109113.

Guo, L., Deng, M., Yang, S., Liu, W., Wang, X., Wang, J., Liu, L., 2021. The coordination between leaf and fine root litter decomposition and the difference in their controlling factors. Glob. Ecol. Biogeogr. 30, 2286–2296.

https://doi.org/10.1111/geb.13384. Harris, D., Horw´ ath, W.R., Van Kessel, C., 2001. Acid fumigation of soils to remove carbonates prior to total organic carbon or carbon-13 isotopic analysis. Soil Sci. Soc. Am. J. 65, 1853–1856. https://doi.org/10.2136/sssaj2001.1853.

Hatfield, R., Fukushima, R.S., 2005. Can lignin be accurately measured? Crop. Sci. 45 (3), 832–839. https://doi.org/10.2135/cropsci2004.0238. Hedges, J.I., 1992. Global biogeochemical cycles: progress and problems. Mar. Chem. 39, 67–93. https://doi.org/10.1016/0304-4203(92)90096-S.

Hedges, J.I., Parker, P.L., 1976. Land-delved organic matter in surface sediments from the Gulf of Mexico. Pergamon. Pres. 40, 1019–1029. https://doi.org/10.1016/0016- 7037(76)90044-2.

Heim, A., Frey, B., 2004. Early stage litter decomposition rates for Swiss forests. Biogeochemistry 70, 299–313. https://doi.org/10.1007/s10533-003-0844-5.

Heim, A., Schmidt, M.W.I., 2007. Lignin turnover in arable soil and grassland analysed with two different labelling approaches. Eur. J. Soil Sci. 58, 599–608. https://doi. org/10.1111/j.1365-2389.2006.00848.x.

Hernes P.J., Robinson A.C., Aufdenkampe A.K. (2007) Fractionation of lignin during leaching and sorption and implications for organic matter "freshness." Geophys. Res. Lett. 34(17):17401. doi:https://doi.org/10.1029/2007GL031017.

Hernes, P.J., Kaiser, K., Dyda, R.Y., Cerli, C., 2013. Molecular trickery in soil organic matter: hidden lignin. Environ. Sci. Technol. 47, 9077–9085. https://doi.org/ 10.1021/es401019n.

Hirave P., Wiesenberg G.L.B., Birkholz A., Alewell C. (2020) Understanding the effects of early degradation on isotopic tracers: implications for sediment source attribution using compound-specific isotope analysis (CSIA). Biogeoscience17:2169–2180. doi: https://doi.org/10.5194/bg-17-2169-2020.

IUSS Working Group WRB, 2022. World Reference Base for soil resources. International soil classification system for naming soils and creating legends for soil maps. In: International Union of Soil Sciences (IUSS), 4th edition. Vienna, Austria.

Jackson, R.B., Lajtha, K., Crow, S.E., Hugelius, G., Kramer, M.G., Pineiro, G., 2017. The ecology of soil carbon: pools, vulnerabilities, and biotic and abiotic controls. Annu. Rev. Ecol. Evol. Syst. 48, 419–464. https://doi.org/10.1146/annurev-ecolsys112414-054234.

Keppler, F., Hamilton, J., 2008. Tracing the geographical origin of early potato tubers using stable hydrogen isotope ratios of methoxyl groups. Isotopes Environ. Health Stud. 44 (4), 337–347. https://doi.org/10.1080/10256010802507383.

Keppler, F., Kalin, R.M., Harper, D.B., McRoberts, W.C., Hamilton, J.T.C., 2004. Carbon isotope anomaly in the major plant C1 pool and its global biogeochemical implications. Biogeoscience 1, 123–131. https://doi.org/10.5194/bg-1-123-2004.

Keppler, F., Harper, D.B., Kalin, R.M., Meier-Augenstein, W., Farmer, N., Davis, S., Schmidt, H.-L., Brown, D.M., Hamilton, J.T.G., 2007. Stable hydrogen isotope ratios of lignin methoxyl groups as a paleoclimate proxy and constraint of the geographical origin of wood. New Phytol. 176, 600–609. https://doi.org/10.1111/j.1469- 8137.2007.02213.x.

Kisluik, 1955. A study of the mechanism of serine biosynthesis. J. Biol. Chem. 214 (1), 47–57.

Koegel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. Soil Biol. Biochem. 34 (2), 139–162. https://doi.org/10.1016/S0038-0717(01)00158-4.

Koiter, A.J., Owens, P.N., Petticrew, E.L., Lobb, D.A., 2013. The behavioural characteristics of sediment properties and their implications for sediment fingerprinting as an approach for identifying

sediment sources in river basins. Earth Sci. Rev. 125, 24–42. https://doi.org/10.1016/j.earscirev.2013.05.009.

Kosaka, J., Honda, C., 1956. Methoxyl content of humus. Soil Sci. Plant Nutr. 2, 59–62. https://doi.org/10.1080/00380768.1956.10431858.

Lee, H., Galy, V., Feng, X., Ponton, C., Galy, A., France-Lanord, C., Feakins, S.J., 2019. Sustained wood burial in the Bengal fan over the last 19 my. Proc. Natl. Acad. Sci. U. S. A. 116 (45), 22518–22525. https://doi.org/10.1073/pnas.1913714116.

Lloyd, M.K., Trembath-Reichert, E., Dawson, K.S., Feakins, S.J., Masalerz, M., Orphan, V. J., Sessions, A.L., Eiler, J.M., 2021. Methoxyl stable isotopic constraints on the origins and limits of coal-bed methane. Science 374, 894–897. https://doi.org/ 10.1126/science.abg0241.

Lloyd, M.K., Stein, R.A., Ibarra, D.E., Barclay, R.S., Wing, S.L., Stahle, D.W., Dawson, T. E., Stolper, D.A., 2023. Isotopic clumping in wood as a proxy for photorespiration in trees. PNAS 120 (46), e2306736120. https://doi.org/10.1073/pnas.2306736120.

Lu, Q., Jia, L., Awasthi, M.K., Jing, G., Wang, Y., He, L., Zhao, N., Chen, Z., Zhang, Z., Shi, X., 2022. Variations in lignin monomer contents and stable hydrogen isotope ratios in methoxy groups during the biodegradation of garden biomass. Sci. Rep. 12, 8734. https://doi.org/10.1038/s41598-022-12689-1.

Luo, Y.-H., Steinberg, L., Suda, S., Kumazawa, S., Mitsui, A., 1991. Extremely low D/H ratios of Photoproduced hydrogen by Cyanobacteria. Plant Cell Physiol. 32 (6), 897–900. https://doi.org/10.1093/oxfordjournals.pcp.a078158.

Lutz, S.R., Van Breukelen, B.M., 2014a. Combined source apportionment and degradation quantification of organic pollutants with CSIA: 1. Model derivation. Environ. Sci. Technol. 48, 6220–6228. https://doi.org/10.1021/es405400w.

Lutz, S.R., Van Breukelen, B.M., 2014b. Combined source apportionment and degradation quantification of organic pollutants with CSIA: 2. Model validation and application. Environ. Sci. Technol. 48, 6220–6228. https://doi.org/10.1021/ es405400w.

Martin, G.J., Lavoine-Hanneguelle, S., Mabon, F., Martin, M.L., 2004. The fellowship of natural abundance 2 H-isotopomers of monoterpenes. Phytochemistry 65, 2815–2831. https://doi.org/10.1016/j.phytochem.2004.07.015.

Mason, S.L., Filley, T.R., Abbott, G.D., 2009. The effect of afforestation on the soil organic carbon (SOC) of a peaty gley soil using on-line thermally assisted hydrolysis and methylation (THM) in the presence of 13C-labelled tetramethylammonium hydroxide (TMAH). J. Anal. Appl. Pyrolysis 85, 417–425. https://doi.org/10.1016/j. jaap.2008.11.005.

Meena, A., Bidalia, A., Hanief, M., Dinakaran, J., Rao, K.S., 2019. Assessment of aboveand belowground carbon pools in a semi-arid forest ecosystem of Delhi. India. Ecol. Process. 8, 8. https://doi.org/10.1186/s13717-019-0163-y.

Moore, T.R., et al., 1999. Litter decomposition rates in Canadian forests. Glob. Chang. Biol. 5, 75–82.

Motha, J.A., Wallbrink, P.J., Hairsine, P.B., Grayson, R.B., 2002. Tracer properties of eroded sediment and source material. Hydrol. Process. 16, 1983–2000. https://doi. org/10.1002/hyp.397.

Peinemann, N., Guggenberger, G., Zech, W., 2005. Soil organic matter and its lignin component in surface horizons of salt-affected soils of the Argentinian Pampa. Catena 60, 113–128. https://doi.org/10.1016/j.catena.2004.11.008.

Prescott, C.E., Vesterdal, L., 2021. Decomposition and transformations along the continuum from litter to soil organic matter in forest soils. For. Ecol. Manage. 498, 119522 https://doi.org/10.1016/j.foreco.2021.119522.

Ros, R., Munoz-Bertomeu, ˜ J., Krueger, S., 2014. Serine in plants: biosynthesis, metabolism, and functions. Trends Plant Sci. 19 (9), 564–569. https://doi.org/ 10.1016/j.tplants.2014.06.003.

Rumpel, C., Kogel-Knabner, ¨ I., Bruhn, F., 2002. Vertical distribution, age, and chemical composition of organic carbon in two forest soils of different pedogenesis. Org. Geochem. 33, 1131–1142. https://doi.org/10.1016/S0146-6380(02)00088-8.

Rumpel, C., Eusterhues, K., Kogel-Knabner, ¨ I., 2004. Location and chemical composition of stabilized organic carbon in topsoil and subsoil horizons of two acid forest soils. Soil Biol. Biochem. 36, 177–190. https://doi.org/10.1016/j.soilbio.2003.09.005.

Ruppenthal, M., Oelmann, Y., del Valle, H.F., Wilcke, W., 2015. Stable isotope ratios of nonexchangeable hydrogen in organic matter of soils and plants along a 2100-km climosequence in Argentina: new insights into soil organic matter sources and transformations? Geochim. Cosmochim. Acta 152, 54–71. https://doi.org/10.1016/ j.gca.2014.12.024.

Sanger, L.J., Anderson, J.M., Little, D., Bolger, T., 1997. Phenolic and carbohydrate signatures of organic matter in soils developed under grass and forest plantations following changes in land use. Eur. J. Soil Sci. 48, 311–317. https://doi.org/ 10.1111/j.1365-2389.1997.tb00551.x.

Schmidt, H.-L., Werner, R.A., Eisenreich, W., 2003. Systematics of 2 H patterns in natural compounds and its importance for the elucidation of biosynthetic pathways. Phytochem. Rev. 2, 61–85.

Schmidt, M.W.I., Knicker, H., Ko`e, Gel-Knabner I., 2000. Organic matter accumulating in Aeh and Bh horizons of a Podzol chemical characterization in primary organomineral associations. Org. Geochem. 31, 727–734. https://doi.org/10.1016/S0146- 6380(00)00045-0.

Schnell, M.J., White, D., Drummond, J., Fuqua, C., 2012. The Physiology and Biochemistry of the Prokaryotes, 4th ed. Oxford University Press.

Schweizer, M., Fear, J., Cadisch, G., 1999. Isotopic (13C) fractionation during plant residue decomposition and its implications for soil organic matter studies. Rapid Commun. Mass Sp. 13, 1284–1290. https://doi.org/10.1002/(SICI)1097-0231 (19990715)13:133.0.CO;2-0.

Shin, M., 2004. How is ferredoxin-NADP reductase involved in the NADP photoreduction of chloroplasts? Photosynth. Res. 80, 307–331. https://doi.org/10.1023/B: PRES.0000030456.96329.f9.

Spielvogel, S., Prietzel, J., Kogel-Knabner, ¨ I., 2008. Soil organic matter stabilization in acidic forest soils is preferential and soil type-specific. Eur. J. Soil Sci. 59, 674–692. https://doi.org/10.1111/j.1365-2389.2008.01030.x.

Stock, B.C., Jackson, A.L., Ward, E.J., Parnell, A.C., Phillips, D.L., Semmens, B.X., 2018. Analyzing mixing systems using a new generation of Bayesian tracer mixing models. PeerJ 6, e5096. https://doi.org/10.7717/peerj.5096.

Thevenot, M., Dignac, M.F., Rumpel, C., 2010. Fate of lignins in soils: a review. Soil Biol. Biochem. 42, 1200–1211. https://doi.org/10.1016/j.soilbio.2010.03.017.

Vale, S., Swales, A., Smith, H.G., Olsen, G., Woodward, B., 2022. Impacts of tracer type, tracer selection, and source dominance on source apportionment with sediment fingerprinting. Sci. Total Environ. 831, 154832 https://doi.org/10.1016/j. scitotenv.2022.154832.

Wang, X., Tian, Q., Li, Q., Liao, C., He, M., Liu, F., 2018. Lignin characteristics in soil profiles in different plant communities in a subtropical mixed forest. J. Plant Ecol. 11, 560–568. https://doi.org/10.1093/jpe/rtx028.

Wedin, D.A., Tieszen, L.L., Dewey, B., Pastor, J., 1995. Carbon isotope dynamics during grass decomposition and soil organic matter formation. Ecology 76 (5), 1383–1392.

Wieland, A., Greule, M., Roemer, P., Esper, J., Keppler, F., 2022. Climate signals in stable carbon and hydrogen isotopes of lignin methoxy groups from southern German beech trees. Clim. Past 18 (8), 1849–1866. https://doi.org/10.5194/cp-18-1849- 2022.

Wieland, A., Romer, ¨ P., Torbenson, M., Greule, M., Urban, O., Caslavsky, J., Pernicova, N., Trnka, M., Büntgen, U., Esper, J., Keppler, F., 2024. Tree-ring stable isotopes in cellulose and lignin methoxy groups reveal different age-related behaviour. Quat. Intern. 693, 38–48. https://doi.org/10.1016/j.quaint.2024.02.004.

Wilson, A.T., Grinsted, M.J., 1977. 12C/13C in cellulose and lignin as paleothermometers. Nature 265, 133–135. Zeisel, S., 1885. Über ein Verfahren zum quantitativen Naohweise yon Methoxyl. Monatsh. Chem. 6, 989–997. https://doi.org/10.1007/BF01554683.

Zhang B.L., Billault I., Lo X.B., Mabon F., Remaud G., Martin M.L. (2002) Hydrogen isotopic profile in the characterization of sugars. Influence of the metabolic pathway. J Agric. Food Chem. 50(6):1574–1580. doi:https://doi.org/10.1021/jf010776z.

Isotopic analysis (δ13C and δ2H) of lignin methoxy groups in forest soils to identify and quantify lignin sources

Supplementary

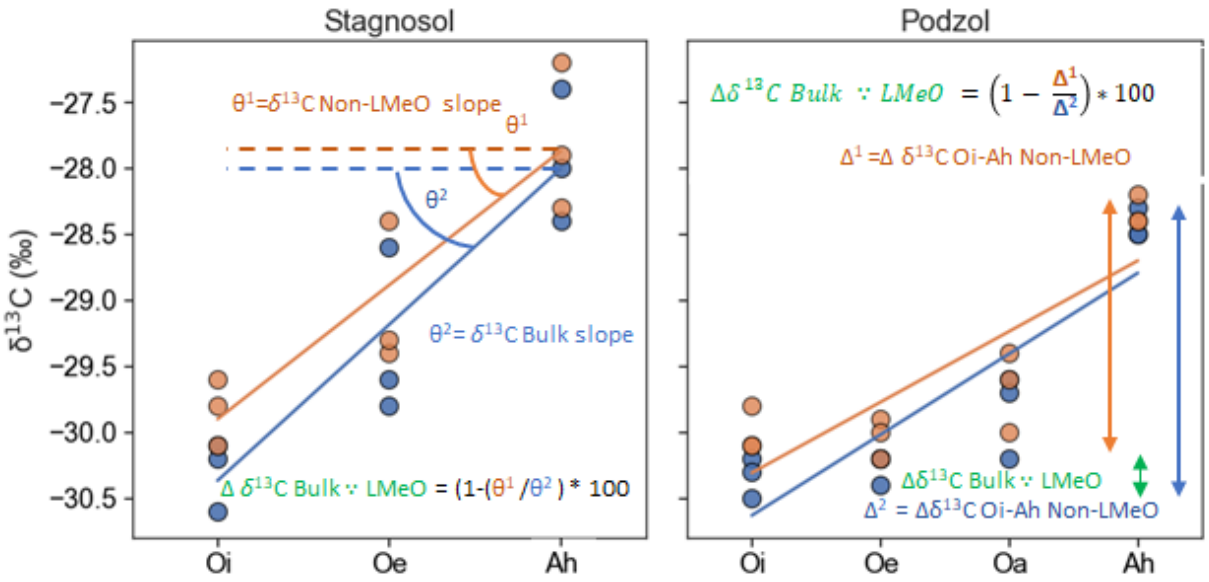


Fig. S1 Demonstration of Equations 4 and 5 is provided through the examples of the stagnosol and podzol sites. In the stagnosol, Equation 5 is exemplified where the slope of 0 (dashed line) indicates no isotopic enrichment during organic matter decomposition. The slope observed in the categorical line plot serves as a semi-quantitative measure for isotopic enrichment. The disparity between the slopes of δ^{13}C bulk and δ^{13}C non-LMeO represents the enrichment attributed to LMeO. The podzol site demonstrates the application of Equation 4. Instead of relying on the slope, the distinction between the Oi and Ah horizons is employed to determine isotopic enrichment. The variation between δ^{13}C bulk and δ^{13}C non-LMeO represents the enrichment induced by LMeO. In both scenarios, the enrichment is expressed as a percentage of δ^{13}C Bulk enrichment.

LMeO concentration

The highest concentration of LMeO was observed in the AGW at both sites (stagnosol: 61.4 ± 18.2 mg g OC^{-1}; podzol: 76.4 ± 7.4 mg g OC^{-1}), followed by the Oe horizon in the stagnosol (42.7 ± 6.7 mg g OC^{-1}) and large roots in the podzol (50.4 ± 6.1 mg g OC^{-1}) (Fig S1). Fine roots exhibited the low LMeO concentration at both sites (stagnosol: 29.8 ± 10.0 mg g OC^{-1}; podzol: 36.0 ± 4.6 mg g OC^{-1}). Large roots displayed a higher LMeO content compared to fine roots in both sampling sites.

The stagnosol showed a decrease in LMeO concentration from the Oe (42.7 ± 6.7 mg g OC^{-1}) to Ah (29.6 ± 0.4 mg g OC^{-1}). The LMeO concentration in the podzol was found to be similar in the Oe and Oa horizons (Oe: 30.8 ± 2.3 mg g OC^{-1}; Oa: 27.8 ± 4.6 mg g OC^{-1}), with a decrease LMeO content in the Ah horizon to 24.6 ± 3.7 mg g OC^{-1}.

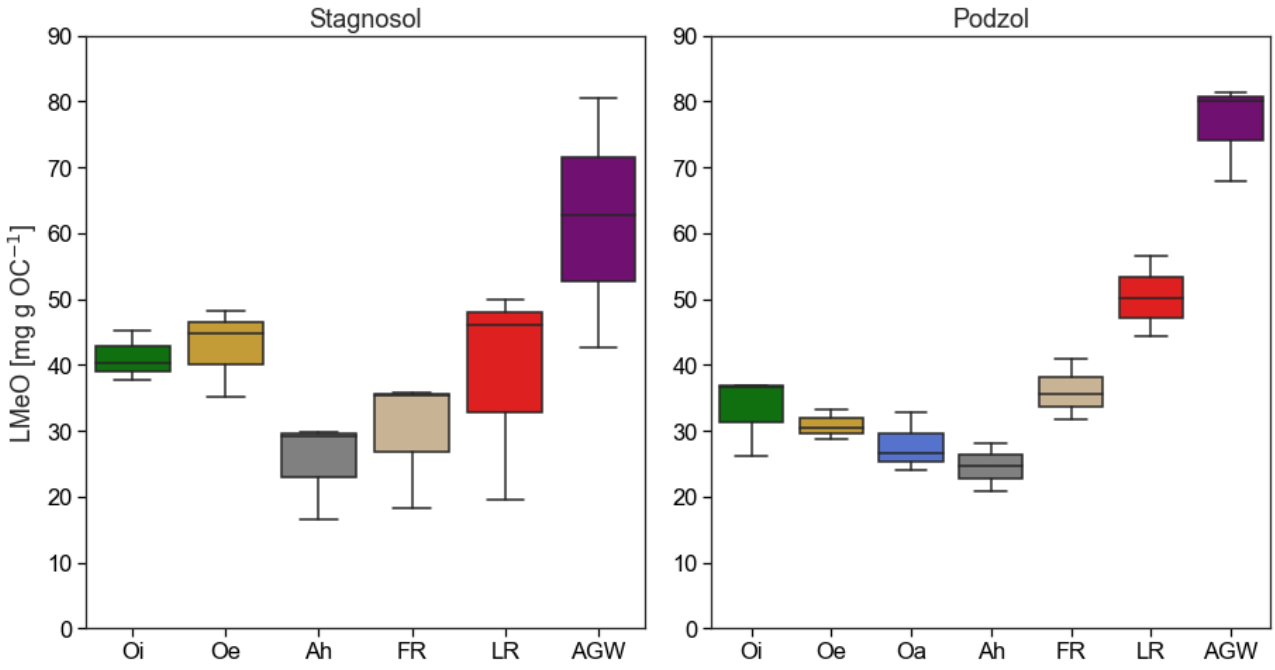


Fig. S2 Concentrations of LMeO in lignin sources and horizons. The boxes represent 25, 50 and 75% quantiles with whiskers showing a 1.5 interquartile range

Table S1 Summary of isotopic values and concentrations of LMeO in lignin sources and horizons of site 1, stagnosol (See Table S3 for full data set).

	OC %		$\delta^{13}C$ Bulk (‰)		LMeO [mg g OC^{-1}]		$\delta^{13}C$ LMeO (‰)		$\delta^{2}H$ LMeO (‰)		$\delta^{13}C$ Non-LMeO (‰)	
Stagnosol	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Oi	45.9	1.6	-30.3	0.3	41.1	3.7	-58.8	0.9	-195.4	9.5	-29.8	0.3
Oe	35.6	6.2	-29.4	0.6	42.7	6.7	-48.0	4.9	-188.3	2.3	-29.0	0.6
Ah	6.1	0.4	-27.9	0.7	29.6	7.6	-41.7	1.7	-175.3	7.2	-27.7	0.7
FR	48.9	1.8	-29.0	0.8	29.8	10.0	-36.3	1.0	-159.4	4.6	-28.9	0.8
LR	47.2	2.4	-28.6	0.8	38.5	16.5	-32.4	1.5	-167.8	7.2	-28.5	0.8
AGW	45.4	1.5	-29.0	0.8	61.4	18.2	-42.2	2.1	-196.0	3.4	-28.7	0.7
Roots	48.0	2.1	-28.8	0.8	34.1	13.1	-34.4	2.4	-163.6	7.1	-28.7	0.7

Table S2 Summary of isotopic values and concentrations of LMeO in lignin sources and horizons of site 2, podzol (See Table S3 for full data set).

	OC %		δ^{13}C Bulk (‰)		LMeO [mg g OC^{-1}]		δ^{13}C LMeO (‰)		δ^{2}H LMeO (‰)		δ^{13}C Non-LMeO (‰)	
Podzol	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Oi	46.9	0.5	-30.3	0.1	33.2	6.1	-53.9	2.7	-208.4	1.3	-29.2	0.2
Oe	42.4	2.7	-30.3	0.1	30.8	2.3	-50.8	2.7	-198.9	3.6	-29.3	0.2
Oa	37.5	3.3	-29.8	0.3	27.8	4.6	-45.1	4.3	-184.8	12.5	-29.0	0.3
Ah	12.4	0.6	-28.5	0.2	24.6	3.7	-38.8	1.3	-187.5	3.5	-27.8	0.1
FR	48.9	0.1	-29.9	0.3	36.0	4.6	-35.2	1.9	-166.3	8.1	-29.0	0.3
LR	49.0	0.5	-31.2	1.7	50.4	6.1	-34.1	1.6	-153.8	18.9	-29.9	1.7
AGW	46.8	0.8	-29.1	2.0	76.4	7.4	-34.0	2.8	-192.6	9.8	-27.3	2.0
Roots	49.0	0.4	-30.5	1.3	43.2	9.2	-34.7	1.7	-160.1	14.7	-29.4	1.3

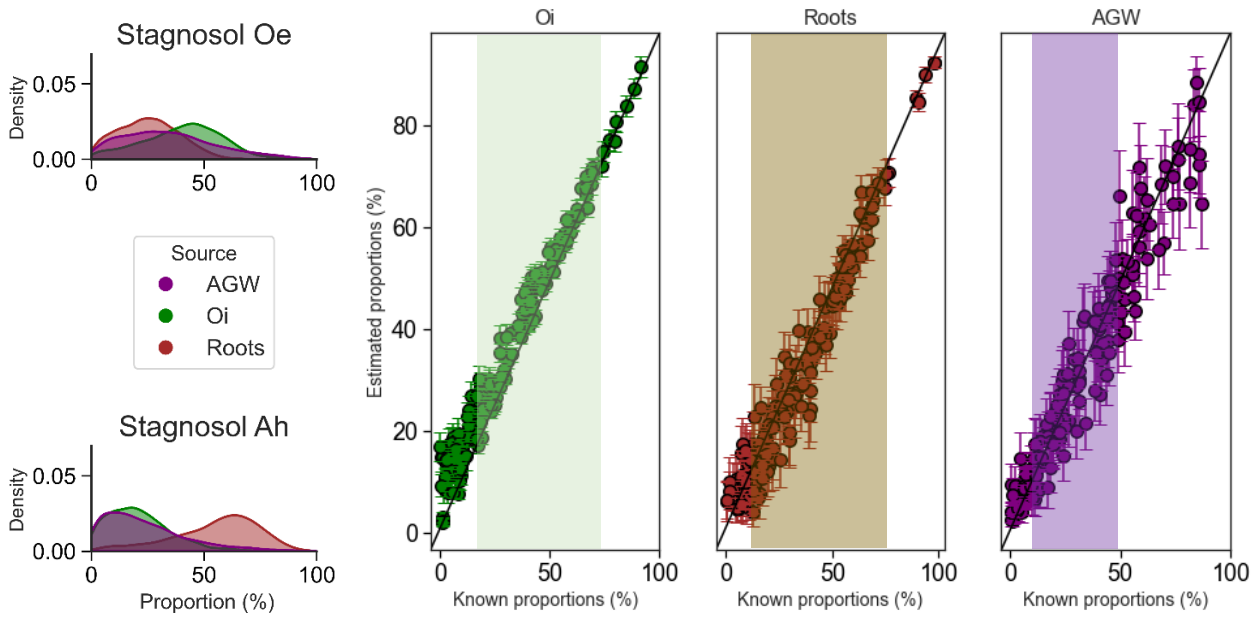


Fig. S3 150 concentration-dependent mathematical mixtures were generated to test the performance of the unmixing of the Stagnosol horizons using MixSIAR. The 25% and 75% IQR of 'real' mixtures are highlighted in the colour boxes. In general, results indicate excellent model performance (Pearson's r^2: Oi 0.99, AGW 0.97, and Roots,0.98), however an underestimation of AGW contribution (ca. 10%) in the higher AGW proportions is demonstrated.

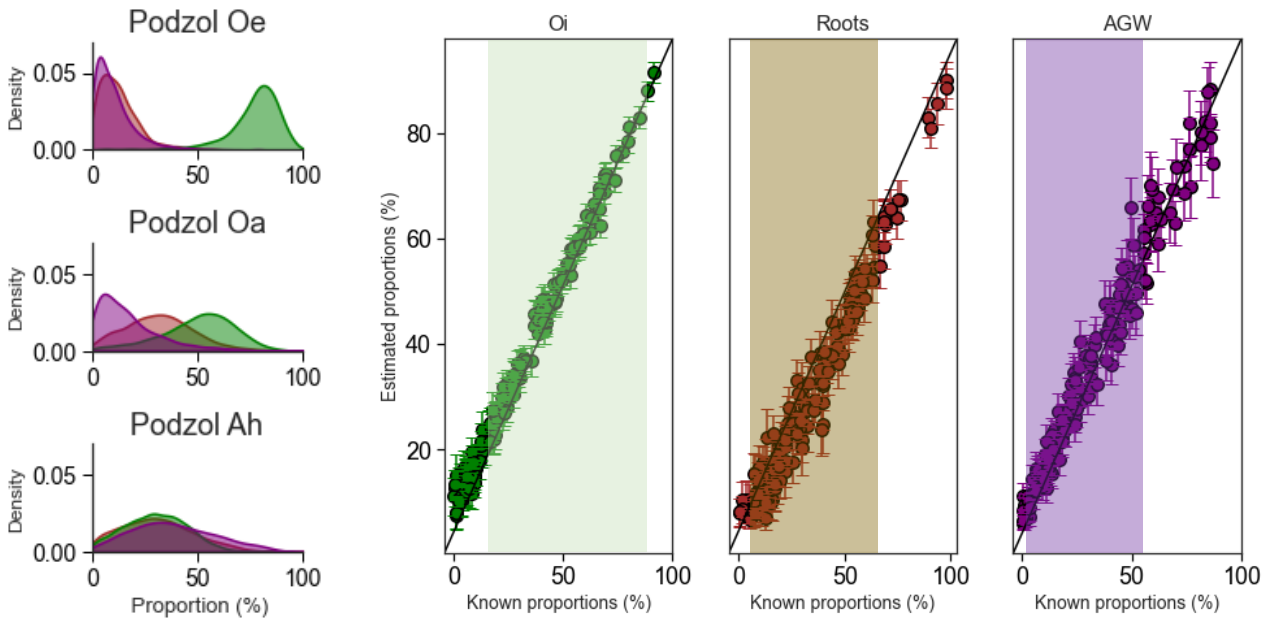


Fig. S4 Results of mathematical mixtures indicate excellent model performance (Pearson's R^2: Oi > 0.99, AGW 0.98, and Roots 0.98), however a consistent underestimation of root lignin contribution (ca. 5%). The 25% and 75% IQR of 'real' mixtures are highlighted in the colour boxes

Table S3 Summary of source apportionment of the stagnosol

	Source contributions (%)					
Stagnosol		**Mean**	**median**	**SD**	**lower IQR**	**upper IQR**
Oe	**Oi**	40	42	17	29	52
	Roots	25	25	13	15	35
	AGW	34	32	21	18	47
Ah	**Oi**	21	19	14	10	29
	Roots	55	59	19	44	69
	AGW	24	19	18	9	33

Table S4 Summary of source apportionment of the Podzol

	Source contributions (%)					
Podzol		**Mean**	**Median**	**SD**	**Lower IQR**	**Upper IQR**
Oe	**Oi**	76	80	14	72	85
	Roots	13	11	10	6	17
	AGW	11	8	11	4	14
Oa	**Oi**	50	52	18	40	76
	Roots	32	31	17	20	61
	AGW	18	14	16	6	51
Ah	**Oi**	30	30	15	19	41
	Roots	31	30	17	18	42
	AGW	39	37	20	24	52

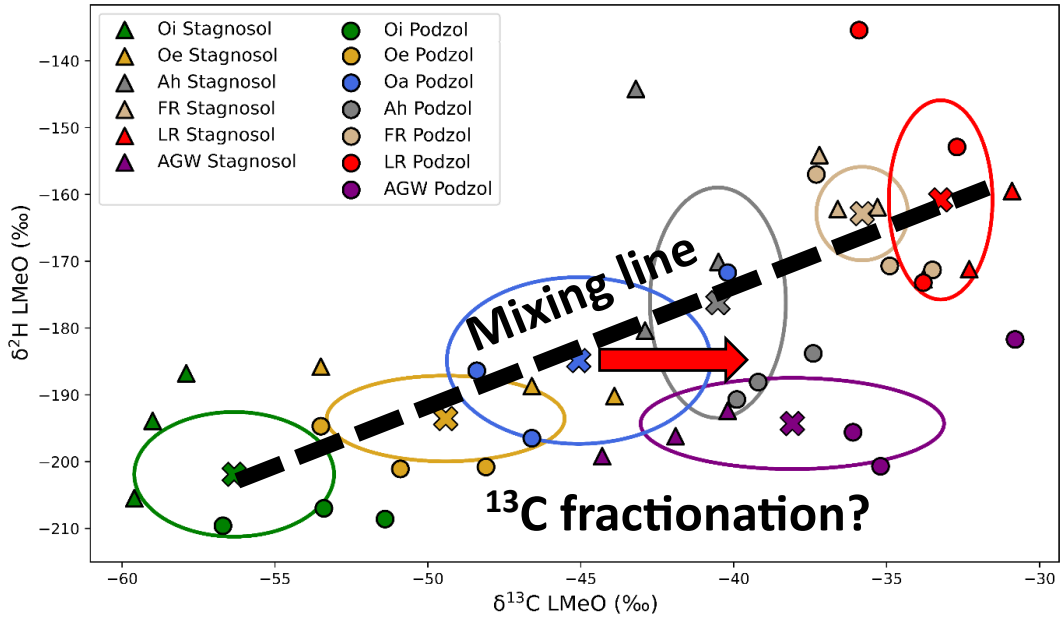


Fig. S5 Bi-plot of $\delta^{13}C$ and $\delta^{2}H$ LMeO values of lignin sources and horizons, with mean values of both sites being denoted by the 'X' symbol. Eclipse illustrates the SD of each source or horizon. In a two-source system (Oi and roots), mixtures will be located on the mixing line. Deviation from this mixing line in a two-source system illustrates ^{13}C fractionation. In a three-source system the deviation is a result of ^{13}C fraction or contributing from the third source, AGW

Chapter 5

Final remarks and outlook

5.1 Final remarks

Given the increasing problem of human-induced accelerated soil erosion and the significant influence of land-use types, land-use-specific sediment source apportionment may play a crucial role in implementing effective erosion mitigation strategies in the future. Hence, it is crucial for sediment fingerprinting to be as accurate and representative of the environmental system as possible to yield correct information for land managers and researchers. In this thesis, I have developed tools that can be utilized for both evaluating and enhancing sediment fingerprinting using CSSI tracers and overcome some of the problems related to using CSSI tracers.

I emphasized the necessity of an additional tracer to extend the one-dimensional mixing line (P1), preventing the misclassification of sources located between source endmembers. I also developed a method to identify tracers that may negatively impact model output. By employing concentration-dependent mixtures and a 'brute force' method, I illustrated that maximizing the number of conservative tracers in mixing models does not always improve model output (P2). Furthermore, the use of concentration-dependent mixtures enables the assessment of confidence associated with the model output (P3). Without this step, model outputs remain unevaluated, potentially leading to erroneous interpretations of the unmixing results.

As demonstrated in Chapter 2, an additional land-use-specific tracer is required to expand the one-dimensional $\delta^{13}C$ FA mixing line. In Chapter 3, I utilized the $\delta^{13}C$ LMeO values and the alkane average chain length (P1). Using this novel tracer, I was able to discriminate between POM_{terr} and MOAM fractions, allowing for the removal of the POM_{terr} fraction and enabling the apportionment of the MOAM fraction only (P4). This has significant implications regarding the use of CSSI tracers for sediment source appointments, suggesting that the reoccurring high estimates of forest soils to the sediment is overestimated due to the misclassification of POM_{terr} input as forest soil input.

Additionally, when dealing with historical sediment and isotopic tracers, Suess corrections are required. In Chapter 3, for the first time, I provided a more representative picture of sediment dynamics by incorporating multiple tracer turnover times (P5). I also used mathematical mixtures to assess model performance (P3) and develop an updated concentration-dependent mixing space-range test. Results of this suggest that the conventional point-in-polygon test is not sufficient for concentration-dependent tracers.

The conservativeness of tracers in sediment source apportionment is a fundamental and necessary tracer property. As such, the dual isotopes ($\delta^{13}C$ and $\delta^{2}H$) of LMeO were used to assess the conservativeness of $\delta^{13}C$ LMeO values during organic matter degradation from the litter layer to SOM. By using the isotopic stability of $\delta^{2}H$ MeO values during degradation, I disentangled isotopic fractionation from source mixing and demonstrated the stability of $\delta^{13}C$ LMeO values. Furthermore, the dual isotope analysis enabled the discrimination of the litter layer, above-ground woody material, and root lignin, allowing the apportionment of lignin sources in organic and mineral horizons. I then applied this method to two soil types (stagnosol and podzol), revealing that lignin mixing and source dynamics are soil type specific.

Additionally, I explored the possibility that $\delta^{13}C$ bulk enrichment during organic matter degradation in forest organic and soil layers can be partially explained by the transition of lignin sources from leaf litter to woody material. Results indicated that 14% and 11% of the bulk $\delta^{13}C$ enrichment in the podzol and stagnosol, respectively, can be explained by the changing of lignin sources. Our findings suggest that the use of $\delta^{13}C$ enrichment in soils as a proxy for carbon turnover may overestimate turnover.

5.2 Outlook

Chapter 2 introduces a tool for assessing model performance using concentration-dependent mixtures, which was developed and tested. In Chapter 4, this tool was enhanced to incorporate the variability of the mixtures. However, mean source values are used to generate the mixtures, neglecting source variance. Enhancing the model evaluation method by generating concentration-dependent mixtures that incorporate source variance would significantly improve its effectiveness and representativeness of 'real' sediment data. Furthermore, mathematical mixtures may be used to investigate how the balance of source discrimination and tracer mixing space similarities effect model performance. As $\delta^{13}C$ FA tracers have a relatively narrow range of possible source values compared to other tracers (e.g., geochemistry). When tracers with a higher degree of source discrimination, despite having identical mixing spaces, are modelled, the propagation of source uncertainty may be outweighed, potentially resulting in improved model performance.

In Chapter 3, I integrated tracer turnover times into the Suess corrections by employing a range of turnover times (10, 30, and 100 years) for all isotopic tracers. However, in reality, tracers are unlikely to possess identical turnover times due to variations in both the tracer type and land-use characteristics. A potential solution for addressing the variance in turnover times specific to tracers and land use may involve employing a 'brute force' method, wherein all plausible combinations of land-use and tracer turnover times are passed through the model.

In this thesis, I utilized $\delta^{13}C$ LMeO values as an additional tracer for sediment source apportionment. The application of $\delta^{2}H$ LMeO for land-use specific sediment source apportionment remains unexplored and could potentially make a significant contribution. Furthermore, this thesis marks the novel application of the LMeO isotopic values in the field of soil science. Given the high abundance of LMeO in the terrestrial biosphere, I eagerly await further use and research of LMeO isotopes resulting in a deeper understanding of the environment.

Acknowledgements

I would like to express my gratitude to my supervisor, Christine Alewell for their guidance and support throughout this research. I also thank my research group for their encouragement and assistance. Special thanks to Universität Basel for providing resources. The thesis text was intermittently rephrased with the assistance of ChatGPT version 3.5.