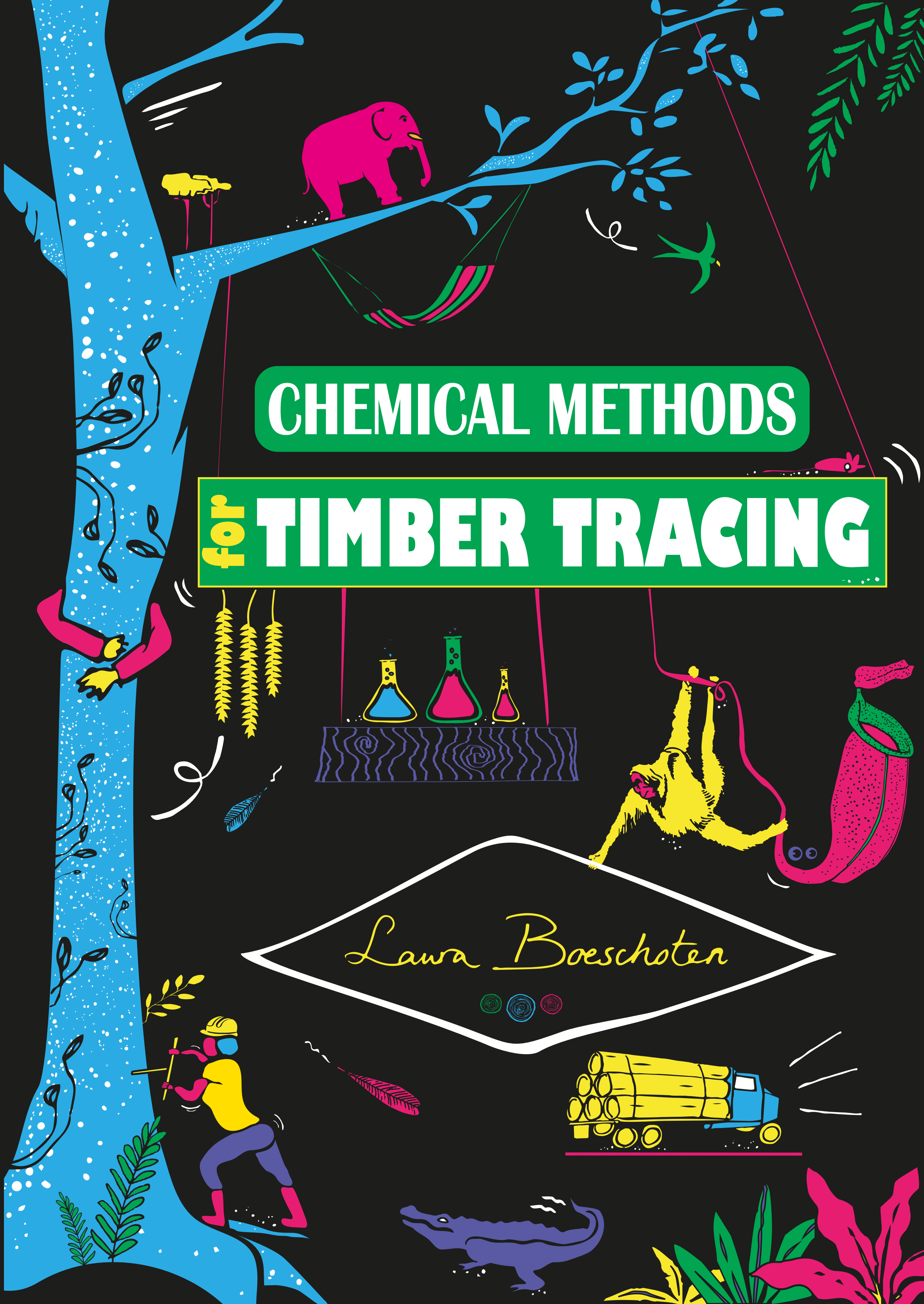


CHEMICAL METHODS for TIMBER TRACING

Laura Boeschoten

Propositions

1. Improving statistical tools is as important as expanding reference databases to increase tracing success.
   (this thesis)

2. Showing the potential of a tracing technique in science does not mean it can be applied in practice.
   (this thesis)

3. Counterintuitively, a pandemic can strengthen international scientific collaborations.

4. We can only achieve a truly diverse scientific community through positive discrimination of minority groups in science.

5. During the PhD trajectory more emphasis must be given to collaboration and less to independence, as scientific breakthroughs are hardly ever achieved alone.

6. If even researchers who study climate change impacts do not adopt a plant-based diet, it is unlikely that the rest of the world can be convinced to do so.

7. Upscaling nature conservation can only be achieved when both quantity and diversity of people that care about nature are increased.

Propositions belonging to the thesis, entitled

Chemical methods for timber tracing

Laura Emily Boeschoten
Wageningen, 22 September 2023

Chemical Methods for Timber Tracing

Laura Emily Boeschoten

Thesis committee

Promotor
Prof. Dr P.A. Zuidema
Personal chair at the Forest Ecology and Forest Management Group
Wageningen University & Research

Co-promotors:
Dr U.G.W. Sass-Klaassen
Associate professor, Forest Ecology and Forest Management
Wageningen University & Research
Lector at Hogeschool Van Hall Larenstein

Dr M. Vlam
Postdoc, Forest Ecology and Forest Management
Wageningen University & Research
Lecturer at Hogeschool Van Hall Larenstein

Other members:
Prof. Dr S.M. van Ruth, Wageningen University & Research and
University College Dublin, Ireland
Prof. Dr P. Hietz, University of Natural Resources and Life Sciences, Vienna, Austria
Dr V. Deklerck, Royal Botanic Gardens, Kew, UK
Dr M. Domínguez-Delmás, Universitat Autònoma de Barcelona, Spain

This research was conducted under the auspices of the C.T. de Wit Graduate School
of Production Ecology & Resource Conservation (PE&RC)

Chemical methods for timber tracing

Laura Emily Boeschoten

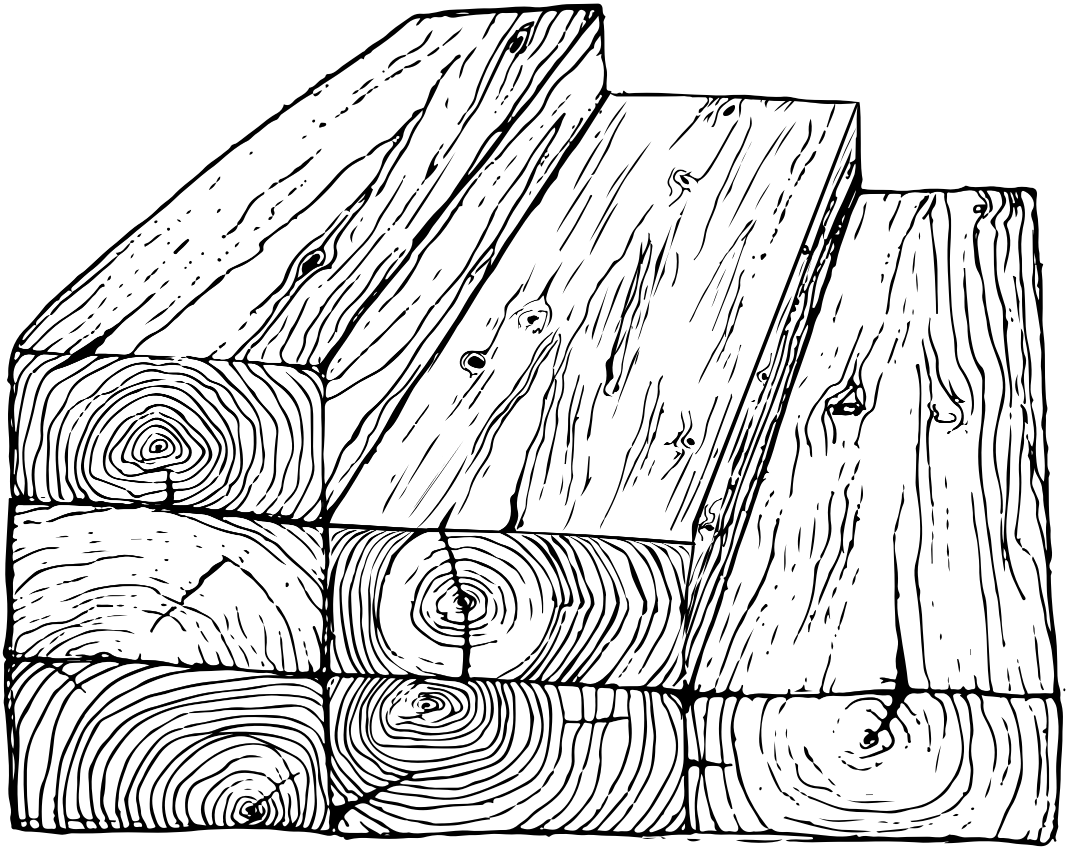
Thesis
submitted in fulfilment of the requirements for the degree of doctor at
Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 22 September 2023
at 4 p.m. in the Omnia Auditorium.

Laura Emily Boeschoten
Chemical methods for timber tracing,
184 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2023)
With references, with summary in English

ISBN 978-94-6447-768-9
DOI https://doi.org/10.18174/633639

Contents



Chapter 1

Introduction

1.1 The need for transparency in the timber supply chain

Illegal logging undermines sustainable timber production, therewith threatening ecosystems, economies, and human well-being. Estimations of the trade in illegally harvested timber range between 4% and 30% globally (Hoare & Uehara, 2022), making it the third largest transnational crime worldwide (May, 2017; Nellemann, 2012). In tropical countries the share of illegally traded timber is especially high, threatening the preservation of tropical forests and coinciding with violations of human rights (Hoare & Uehara, 2022; Nellemann, 2012; Kleinschmit et al., 2016). These challenges persist despite legislation being in place in most producer, processing, and consumer countries, which is aimed at combating this illegal timber trade (Hoare & Uehara, 2022; Gan et al., 2016).

Legislation remains the primary focus of most international efforts to stop illegal timber from being traded (Hoare & Uehara, 2022). Various large importer countries have implemented laws to prevent illegal timber from entering their markets, such as the US Lacey Act (since 2008), the Australian Illegal Logging Prohibition Act (2012), and the Japanese Clean Wood Act (2017). Although these legislations do not explicitly include origin verification requirements, they oblige traders to verify the legality of their timber. In the European Union (EU), new legislation that will be implemented in 2025 will take this one step further. The new anti-deforestation regulation (EUDR) will replace the EU Timber Regulation (EUTR, established in 2010) and it will require timber traders to provide the geolocation of their timber, thereby increasing the need for origin verification. This regulation goes beyond the previous legislation as it not only addresses timber legality but also aims to effectively combat deforestation.

In addition to these demand side initiatives, some major timber-producing and processing countries, including Indonesia (in 2016), South Korea (in 2018), Malaysia (in 2019), and Vietnam (in 2020), have also implemented policies to prohibit the import of illegal timber (Hoare and Uehara, 2022). Furthermore, China revisited its Forest Law in 2019, explicitly banning the purchase, processing, or transportation of timber from illegal sources. Although the implementation and regulation of the Chinese law is still unclear, it could be a huge step towards improving transparency and reducing the global illegal timber trade (Fuller et al., 2019). Enforcement of these legislations remains challenging however and fraud remains prevalent. Timber is traded through multiple countries by many different stakeholders, which makes it extremely hard to verify the documentation that accompanies timber after the wood has left the forest (Lowe et al., 2016).

The main types of timber fraud involve falsifying species and/or origin information (Gan et al., 2016). Species fraud includes trade of unauthorized species, for example species that are listed in one of the three CITES appendices (Convention on International Trade in Endangered Species of Wild Flora and Fauna). The CITES appendices include species for which trade is restricted because they are currently threatened with extinction or species that require controlled trade to avoid over-harvesting and future extinction threats. Species fraud is especially prevalent in genera where the distinction between species cannot be made easily based on wood characteristics, such as for Rosewood (*Dalbergia* spp.) and Ebony wood (*Dyospyros* spp.) (Jahanbanifard et al., 2019; Siriwat & Nijman, 2023). All Rosewood and Ebony wood from Madagascar has been listed on CITES appendix II, as there has been a rapid decline in the wild. However, it is hard to distinguish the Malagasy species from their close relatives based on wood anatomy and the illegal trade remains (Jahanbanifard et al., 2020; Siriwat & Nijman, 2023). For many traded timber species however, accurate species identification can be achieved based on available reference databases (Low et al., 2022).

Similar to species fraud, many examples of origin fraud exist as well, at varying geographical scales. An overview of the categories of fraud as reported in literature is presented in Figure 1.1, categorized based on the scale at which the fraud is taking place. The first scale at which fraud occurs (A) is within legal timber operations, when trees are harvested within a concession that should not have been cut. This includes exceeding cutting quotas for example, or cutting wider along road corridors. The second scale (B) is related to plantation establishment, for example by mixing illegal timber with plantation products. This can occur within a country, but also across continents. This has been the case for teak wood for example, which was harvested illegally from natural forests in Myanmar but sold as if it originated from a plantation in Brazil (Wortel & van Erp, 2020). The third scale (C) includes all cases where timber was logged illegally elsewhere within the country. It can then be laundered in many ways, such as by by mixing it with legal timber at the concession level or during further processing in mills, or by falsifying paperwork to make it seem legal. Lastly, the fourth scale (D) includes laundering practices across borders. The timber supply chain can be long and complex, crossing multiple borders where laundering practices can occur that obscure the real origin of the wood.

These various scales at which fraud occurs make the verification of timber origin extremely complicated, therewith hampering law enforcement. The only way to detect these types of false reported origin after the wood has entered the complex global trade network is through independent origin verification, as it cannot be checked by paperwork alone (Lowe et al., 2016; Hoare & Uehara, 2022). This is where forensic science can make a difference: forensic methods can verify the paperwork accompanying timber and improve transparency along the timber supply chain (Dormontt et al., 2015). This essential to detect illegal practices as well as to support initiatives aimed at increasing sustainable forest management, as non-certified products can be sold as if they were sustainably produced (Hoare & Uehara, 2022; Nurrochmat et al., 2016). Therefore I aimed to contribute to further development of these forensic methods that can verify trade claims of timber in this thesis.

1.2 Scientific methods to verify trade claims

Multiple scientific verification methods are already successfully applied to catch the first type of fraud: a misrepresentation of the species. This includes methods for species identification based on as wood anatomy (Lens et al., 2020), genetics (Bolson et al., 2015), Near-Infrared Spectroscopy (Snel et al., 2018), DART-TOFMS (Musah et al., 2015; Deklerck et al., 2019)

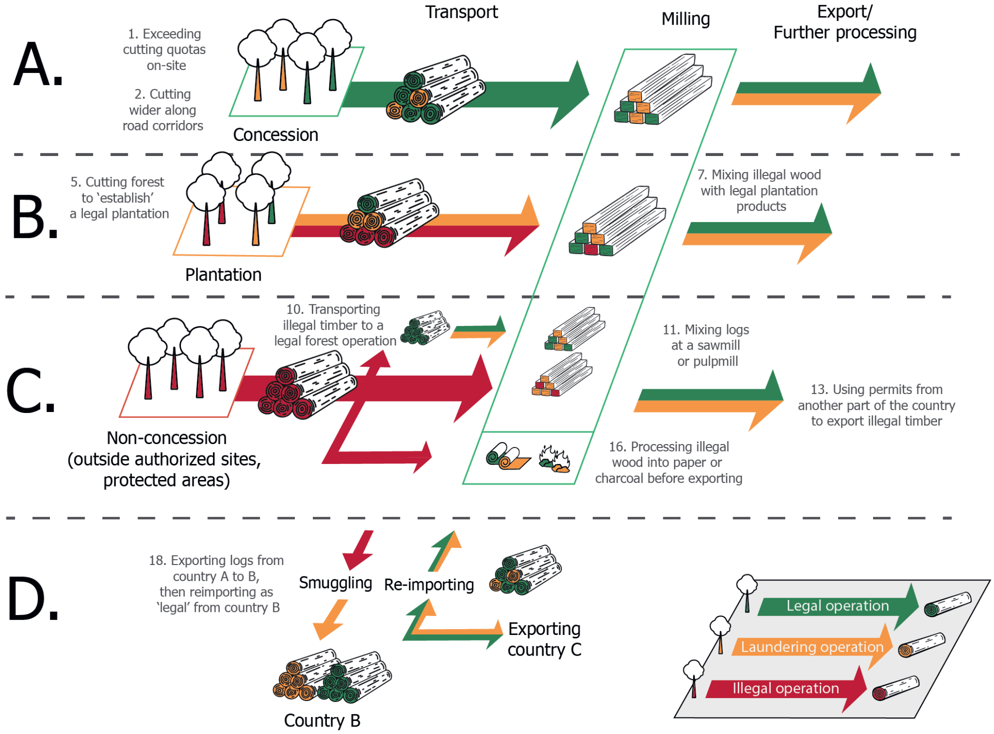


Figure 1.1: Examples of laundering practices throughout the timber supply chain. Color indicates legality of the timber: green is legal, red is illegal and orange is laundered timber that has become 'legal'. Sustainable and/or legal timber originates from well-managed concessions, is transported to a log yard and subsequently to a mill after which it is processed before or after export. The four rows correspond to the categories at which fraud occurs: A) represents examples of fraud occurring within the legal forest operation itself, B) represents examples of fraud related to plantation establishment, C) includes all types of fraud where timber was illegally harvested elsewhere in the country and mixed with legal logs as well as through further processing to obscure origin, and D) includes laundering practices that involve exporting timber to other countries and claiming origin elsewhere. Examples of laundering practices from literature are included as text, they correspond to the extended version of this figure in the General Discussion (Chapter 7). The figure is based on Nelleman et al. (2016) and (Lowe et al., 2016).

and even stable isotopes (Paredes-Villanueva et al., 2022). While challenges remain to identify species within complex genera with high similarities (Jahanbanifard et al., 2020; Gasson, 2011), these scientific methods offer the toolbox to achieve species identification with high accuracy (Low et al., 2022).

Origin identification of timber on the other hand still heavily relies on external documents, tags, or scans, leaving room for fraudulent practices. Even though various forensic methods are currently being developed to verify origin and some have shown good tracing results, they are still not widely implemented. Therefore, the focus of this thesis is on origin identification rather than species identification.

The available forensic methods are based on intrinsic wood properties, such as the genetic, structural or chemical variation within the wood (English et al., 2001; Deklerck, 2023; Dormontt et al., 2015; Lowe et al., 2016; Low et al., 2022). Tracing origin of timber always starts with a reference database, consisting of samples from known origin for which the intrinsic properties have been measured. To then trace the origin of a new sample, it is compared to the reference database and an origin claim can be accepted or rejected based on the similarity with the reference samples. The methods depend on natural variation within a tree species and a requirement is that more variation is present between sites than within sites. Given the diverse ways in which timber is laundered (Fig 1.1), a one-size-fits-all solution to catch all origin fraud occurring in the timber supply chain is unlikely. Consequently, a range of forensic methods for origin verification of timber may be necessary to effectively identify as many types of timber fraud as possible.

The most applied method for origin verification is based on genetic differences between trees (Low et al., 2022). Genetic methods operate on the principle that genetically related trees are also geographically closer. The method has high potential for verification of individual trees (DNA fingerprinting) as well as for origin verification based on population genetics (Dormontt et al., 2015). Furthermore, genetic reference databases are already available for timber species from all continents (Low et al., 2022) and specific populations were found to differ even at very small spatial scales (Vlam et al., 2018b; Ng et al., 2017).

Another method used for origin verification is dendrochronology, which involves analysing the unique pattern of tree rings. In this method, site-specific growth patterns are characterised and wood is then assigned to origin based on this chronology (Bridge, 2012). It has already been successfully applied to identify the origin of temperate species (Bridge, 2012; Bonde et al., 1995). However, the method cannot be applied for a lot of tropical timbers, as many tropical species do not produce regular annual rings (Brienen et al., 2016).

Additionally, chemical characteristics of wood also provide the basis for several origin verification methods. A method that is gaining a lot of momentum recently is based on stable isotope ratios in the wood (English et al., 2001; Watkinson et al., 2020; Gori et al., 2018). Geographical variation in the wood isotopic composition is caused by differences in climate, geology and/or deposition. As a result, the isotopic composition reflects the area where the tree was growing. However, results vary across locations and scales. While stable isotopes have shown promising differences in some locations in Gabon at distances of 240 km (Watkinson et al., 2022a), another study found insufficient variation for origin verification at distances of 14-200 km in the neighbouring country Cameroon (Vlam et al., 2018b). Furthermore, no geographical signal was found in Bolivia at larger spatial scales of hundreds of kilometers (Paredes-Villanueva et al., 2022), while in the United States researchers did find sufficient variation for origin verification at similar distances (Watkinson et al., 2020).

A second chemical method that has been proposed for origin verification more recently is DART-TOFMS (Deklerck, 2023; Paredes-Villanueva et al., 2018). The method has high potential for species identification (Deklerck et al., 2019) and may also provide a basis for origin verification (Deklerck et al., 2020; Finch et al., 2017). The mechanisms behind the geographical variation remain unknown however and will require more research before the method can be implemented widely (Deklerck, 2023).

A third method based on wood chemistry is Near Infrared Spectroscopy (NIRS). NIRS is based on variation in phytochemical properties, like DART-TOF-MS. These are characterised by wood absorption spectra when a sample is exposed to near infrared electromagnetic energy. NIR spectra have been applied to species identification (Bergo et al., 2016) and have now also been linked to timber origin (Sandak et al., 2011; Silva et al., 2018; Li et al., 2019). The number of studies is limited however, and the drivers of spatial variation in reflectance spectra within species need further research.

A fourth chemical method that is explored in detail in this thesis, is based on the multi-elemental composition of the wood. A large set of elements (such as Mg, Ca, K) is measured simultaneously by dissolving a small amount of wood and measuring it through mass spectrometry. Differences may occur in the multi-elemental composition of wood through differences in soil characteristics for example (Kabata-Pendias, 2011). The method has already been used in a variety of commodities, from capers (Pepi & Vaccaro, 2018) to cocoa (Acierno et al., 2020), with high potential at small spatial scales (Ma et al., 2016). However, it has not been previously applied to timber.

1.3 Remaining barriers in timber tracing

Despite this large toolbox of forensic methods that is currently under development, barriers in these scientific methods remain (Lowe et al., 2016; Low et al., 2022). First, the lack of georeferenced databases hampers direct implementation of the methods, even if the methods are shown to perform well: for many traded species, no georeferenced databases are available for forensic origin verification (Low et al., 2022; Gasson et al., 2021). This can be partly alleviated by increasing the sampling effort and the laboratory analyses, to expand these databases (Gasson et al., 2021). Furthermore, sharing and integrating databases amongst institutions would also optimise the use of existing data and move the tracing field further as a whole. However, it may take several decades before adequate reference databases are accessible for all main traded timber species (Low et al., 2022). Therefore, the use of existing data can also be optimized by developing statistical methods that can interpolate in between sampling points (Truszkowski et al., 2023), thereby reducing the required sample size for origin identification.

Second, small-scale tracing remains challenging, especially at distances of 100 km. This level of tracing is crucial for identifying a wide range of fraudulent practices (Fig 1.1), but no method has yet consistently demonstrated high spatial assignment accuracy across multiple species and locations. Both genetic methods and stable isotope ratios, which are currently the most advanced timber tracing methods, have shown promising results at small spatial scales for some specific locations (Vlam et al., 2018b; Gori et al., 2018), but have not been that successful for other areas or species (Ng et al., 2017; Paredes-Villanueva et al., 2022).

These challenges highlight the need for further research and development to improve the accuracy and applicability of origin verification methods for timber. Even though the potential is there to reduce illegal timber trade through forensic origin verification, its current application remains limited to a case-by-case basis, falling short of its full potential to effectively reduce illegal logging.

1.4 Main objective

In this thesis I aimed to address these barriers in scientific timber tracing and contribute towards its operationalization, with a specific focus on two chemical methods: multi-element analysis and stable isotope ratios. I tested the potential of these methods for timber tracing and assessed the mechanisms that drive geographical variation within species, in order to improve the capacity to verify timber origin and support the legal and sustainable timber trade. Furthermore, understanding the mechanisms by which these methods operate is crucial in determining research priorities and identifying the methods with the highest potential.

The thesis is composed of seven chapters: this Introduction (Chapter 1), five research chapters (Chapters 2-6) and a General Discussion (Chapter 7). In chapters 2 and 3 I investigate the potential of multi-element analysis, a tracing method that is novel for timber. In chapters 4 and 5, I apply stable isotope ratios for the origin assignment of timber and test their potential at a pan-tropical and a sub-country level. In chapter 6, I combine the chemical data with genetic data from the same trees to evaluate the tracing potential of three forensic methods individually as well as combined. In chapter 7 I then put these findings into perspective in the General Discussion. I compare the methods that are currently available for origin verification of timber and link them to the types of fraud that have been reported to occur along the timber supply chain. By doing so, I identify key challenges that still remain for the widespread implementation of forensic timber tracing.

1.5 Study species and sites

For this research I focus on three economically important tropical timber species: Azobé (*Lophira alata* Banks ex C.F. Gaertn, Ochnaceae) and Tali (*Erythrophleum ivorense* A. Chev. and *E. suaveolens* (Guill. & Perr.) Brenan, Fabaceae) from Central Africa and Red Meranti (*Shorea* spp., Dipterocarpaceae) from Borneo (Fig 1.2). They represent a single-species timber, Azobé, as well as two multi-species timbers; Tali, which is sourced from two botanical species, and Red Meranti, which is generally accepted to be sourced from 10 species (Klaassen, 2018). Both Azobé and Tali are highly durable timbers that are mainly used in road- and waterworks (bridges, decking), whereas Red Meranti is often used for window frames, paneling, flooring

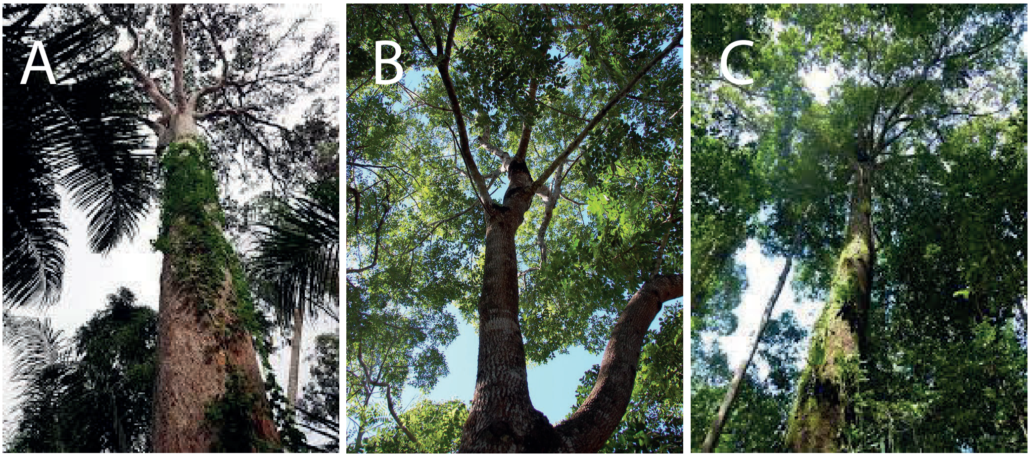


Figure 1.2: Pictures of the study species: A) Azobé (*Lophira alata*), B) Tali (*Erythrophleum ivorense*) and C) one of the Red Meranti species (*Shorea macrophylla*).

and plywood. None of the species is listed as endangered in the CITES appendices.

Sampling was conducted in the countries with the largest share of timber export in the regions: wood samples were collected at 22 sites spread across Cameroon (seven sites), Gabon (seven sites) and the Republic of the Congo (eight sites), between September 2019 and April 2022 (Fig 1.3A). These three countries together represent the most important sources of timber export in the Congo Basin, while at the same time illegal practices remain widespread throughout their timber supply chains (Hoare & Uehara, 2022). Additionally, samples were taken at nine sites across the provinces West, East and Central Kalimantan, Indonesia, between January 2020 and February 2022 (Fig 1.3B). Together with North Kalimantan and Sarawak and Sabah in Malaysia, these provinces represent the biggest share of Meranti export in the region. All sites were natural forest concessions, accessed in collaboration with the operating forestry companies. A subset of all sites and trees was used per chapter.

1.6 The Timtrace project

This thesis is part of the Timtrace project, which aims to develop and test methods to verify the claimed origin of tropical timber. Three main scientific methods are investigated for their potential for timber tracing: multi-element analysis, stable isotope ratios and genetic analysis. Through this research, Timtrace can help timber trading companies to support the paper trail showing they source timber legally and create transparency in their value chain. Additionally, it can help governmental authorities to verify trade claims and combat the illegal timber trade. And it enables environmental organizations to verify the origin of traded timber.

The Timtrace project includes two PhD studies: the work presented in this thesis as well as a second study by Barbara Rocha Venâncio Meyer-Sand, focusing on the potential of genetic trac-

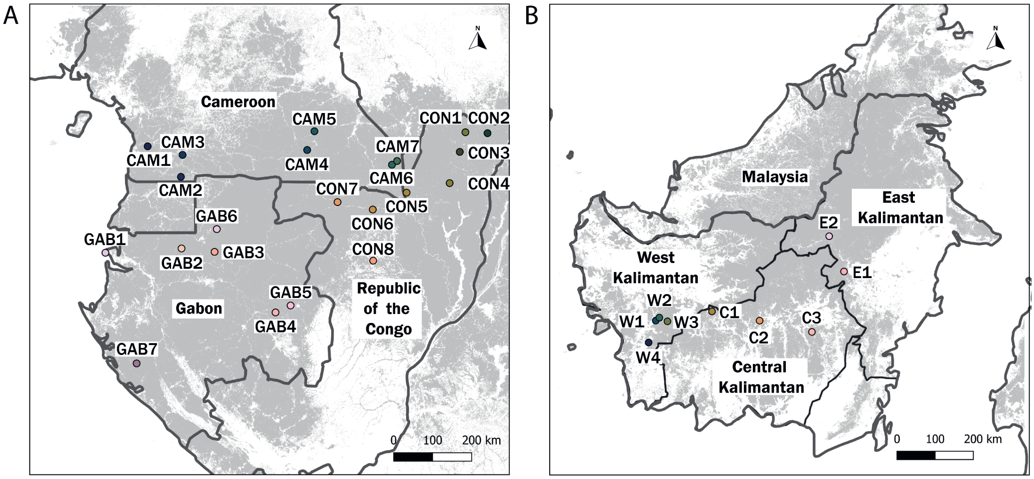


Figure 1.3: Maps of the study areas and sites across A) Central Africa and B) Borneo. Primary tropical forest extend from Global Forest Watch in light grey (Turubanova et al., 2018).

ing for the same tree species at the same locations. By developing relevant reference databases for both chemical and genetic methods, we are the first to directly compare and combine multiple tracing methods within one study (Chapter 6). This is an important step forward to improve the spatial scale at which timber tracing can be successfully applied.

1.7 Thesis outline

Chapter 2

In *Chapter 2* we tested how physical and chemical soil properties (clay content, soil organic matter, pH) as well as tree-physiological processes (heartwood formation) drive variation in wood elemental composition. An important step in evaluating the potential of multi-element analysis for timber tracing is to understand the drivers of wood elemental composition. However, little is known about the underlying mechanisms that drive the uptake and allocation of many elements in wood. Therefore, we evaluated the relationship between soil chemical and physical properties and wood elemental composition. We studied the reactive soil element pools and the elemental composition in sapwood and heartwood for 40 Azobé trees at two of our study sites in Cameroon. We hypothesised that wood elemental composition was linked to soil chemical and physical properties, as they drive the soil reactive pool (Rieuwerts, 2007; Antoniadis et al., 2017; Bukata & Kyser, 2008). From there, trees take up the elements after which they are taken up in the sapwood cells. We hypothesised that further translocation can occur at heartwood formation, through selective recycling of elements from the inner sapwood, which is transformed to heartwood, to the outer sapwood (Smith et al., 2014). As a result, for essential elements we expected higher concentrations of essential elements in the sapwood compared to the heartwood (Meerts, 2002; Hevia et al., 2017). For the non-essential elements an opposite pattern was expected: concentrations will be higher closer to the centre of the tree because the wood cation binding capacity decreases from the pith to the outer wood (Herbauts et al., 2002; Momoshima & Bondietti, 1990; Amais et al., 2021). With this study we aimed to understand what drives within-species variation in elemental composition. This can help to identify geographical regions where multi-element analysis may have the highest potential for timber tracing.

Chapter 3

In *Chapter 3* we put multi-element analysis for timber tracing to the test by analyzing its potential for tracing across all countries and species sampled in the project: Azobé and Tali from Central Africa and Red Meranti from Borneo. We tested the potential for origin assignment as well as origin verification. Wood samples from 991 trees were measured using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). Element concentrations were analysed to chemically group similar sites (clustering) and assess accuracy of tracing samples to their origin (Random Forest assignment models). We hypothesised that trees occurring in the same locations would be chemically more similar, compared to trees further away. Furthermore, we expected large-scale patterns to be shared among tree species as their elemental composition depends on the availability of elements in the soil. Even if species then show distinct chemical patterns (Amais et al., 2021), the variation between sites would still be shared between them. The objective of the study was to test the performance of multi-element analysis in practice. Therefore we also simulated a real-life tracing case by assigning blind samples to their most likely origin.

Chapter 4

In *Chapter 4* I continued with stable isotope ratios. We tested the potential of a pantropical δ^{18}O isoscape for tracing timber at a coarse spatial scale (i.e. country level). As mentioned, fraud occurs on many levels and cross-country fraud is prevalent across the tropics (Nellemann, 2012). With this study we aimed to reduce the dependency on reference databases (barrier 1) by using spatial interpolation methods, as a prediction map can be used to identify the most likely isotopic composition in between sampling points. We based the prediction model for δ^{18}O on physiological insights into the main determinants of δ^{18}O in wood of tropical trees (Barbour, 2007), being the isotopic signature of rainfall (root uptake), total rainfall ("amount effect") and drought (evapotranspiration). Based on these relevant covariables and by combining them with a database of isotopic measurements in tropical tree species developed through multiple research projects, including 27 species in 20 countries, we developed a prediction map of δ^{18}O (isoscape) across the tropics. We developed the isoscape using a novel statistical technique, Quantile Regression Forests, which is based on machine learning and explicitly includes prediction uncertainty. As a second step, we used the δ^{18}O isoscape to evaluate the spatial scale at which δ^{18}O-isotopic tracing is expected to differentiate wood samples within species. We hypothesised it would be most useful at a country scale (100's of km), as the drivers of isotopic variation also occur at that scale.

Chapter 5

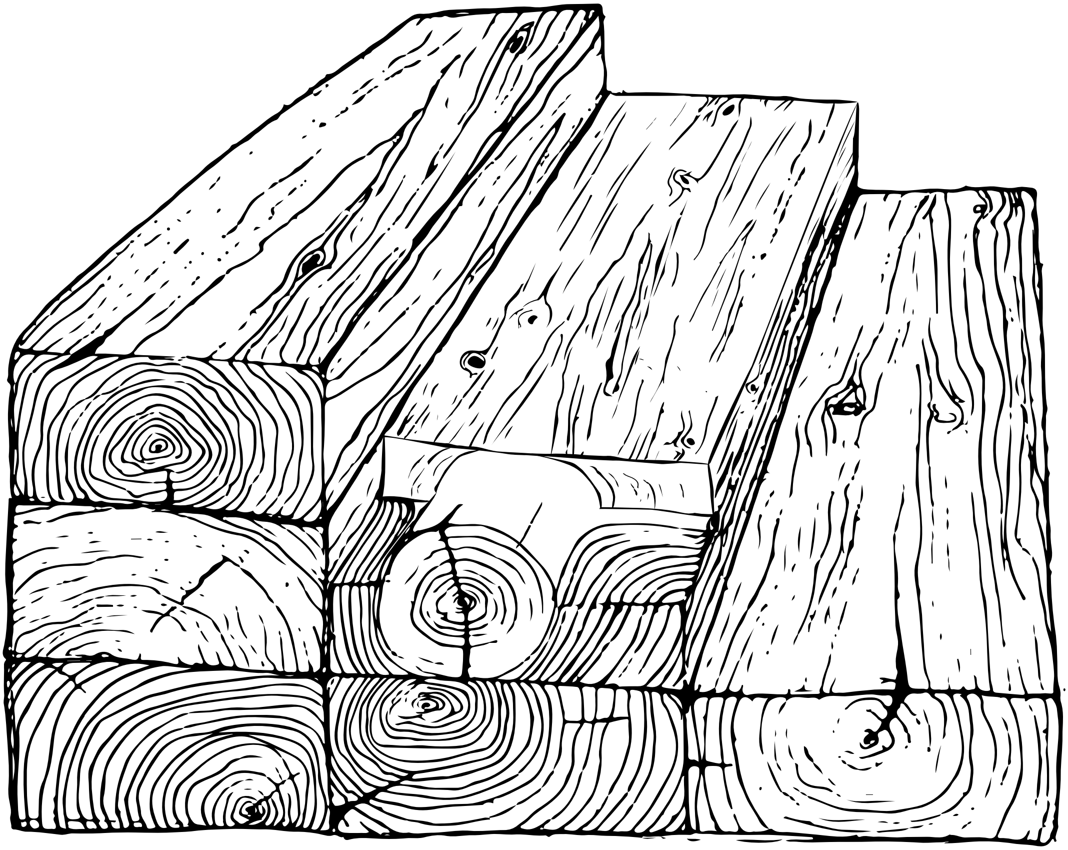
In *Chapter 5* we then investigated the potential of multiple stable isotope ratios (δ^{18}O, δ^{2}H, δ^{13}C and δ^{34}S) for timber tracing at a smaller scale: within and between countries in Central Africa. We tested tracing potential in two timber species (Azobé and Tali) at 17 of our study sites. In Central Africa the urgency for accurate timber tracing is especially high due to high timber illegality (Hoare & Uehara, 2022), but the variation in the drivers of stable isotope ratios is quite limited (climate, geology and sulphur deposition). This might hamper the use of stable isotopes for tracing purposes. Previous studies were not conclusive on the potential for small-scale tracing in the region, as Vlam et al. (2018b) found no potential at 15-200 km, whereas Watkinson et al. (2022a) did find isotopic differences at 240 km. However, large-scale studies (more than 2-3 sites) that test local and regional variation were lacking. Therefore we aimed to better quantify local and regional variation in the region and assess whether stable isotopes can accurately trace timber origin at a sub-country scale.

Chapter 6

In *Chapter 6*, we then integrated three promising timber tracing methods (multi-element analysis, stable isotopes, genetic methods) to test their individual as well as their shared potential for origin assignment. No tracing method has yet shown consistent and reliable results at small spatial scales (100 km). Therefore we aimed to compare and combine three tracing methods for tracing Azobé at 13 of our study sites in Central Africa. For a subset of trees at all sites we (1) detected genome-wide chloroplast Single Nucleotide Polymorphisms (cpSNPs), measured (2) three stable isotope ratios (δ^{18}O, δ^{2}H and δ^{34}S) and (3) analysed the multi-elemental composition. We hypothesised that combining tracing methods would improve assignment accuracy, due to the different mechanisms that drive geographic variation within species: each method has its own strengths and limitations and the variation at a landscape scale might be complementary across methods. Therefore, we expected tracing accuracy to improve when combining multiple methods. With this research we aimed to improve the understanding of how these tracing methods performed individually when applied to the same trees, and we aimed to test their potential for combined use.

Chapter 7

In the final *Chapter 7* I synthesise the findings of this thesis and relate them to the main objective: developing effective tools for origin verification of tropical timber. Furthermore, I provide a more in-depth discussion on the strengths and limitations of all methods currently applied and/or under development for timber tracing. I link the available methods to the challenges occurring in the timber trade. This provides a perspective on the implications of our findings and what scientific development is still needed to apply these forensic methods in practice. I continue with a reflection on the next steps that can be taken to develop the novel multi-element analysis into a successful method for timber tracing, as it was first described in this thesis. Lastly, I discuss how these methods can be integrated into a system of checks along the supply chain to effectively realise a reduction in illegal timber trade.



Chapter 2

Clay and soil organic matter drive wood multi-elemental composition of a tropical tree species: implications for timber tracing

Laura E. Boeschoten, Ute Sass-Klaassen, Mart Vlam, Rob N.J. Comans, Gerwin F. Koopmans, Barbara Rocha Venâncio Meyer-Sand, Steve N. Tassiamba, Martin T. Tchamba, Herman T. Zanguim, Pascaline T. Zemtsa, Pieter A. Zuidema

Published in Science of the Total Environment, 849 (2022) 157877, https://doi.org/10.1016/j.scitotenv.2022.157877

Abstract

Forensic methods to independently trace timber origin are essential to combat illegal timber trade. Tracing product origin by analysing the multi-element composition has been successfully applied in several commodities, but its potential for timber is not yet known. To evaluate this potential the drivers of wood multi-elemental composition need to be studied. Here we report on the first study relating wood multi-elemental composition of forest trees to soil chemical and physical properties.

We studied the reactive soil element pools and the multi-elemental composition in sapwood and heartwood for 37 Azobé (*Lophira alata*) trees at two forest sites in Cameroon. A total of 46 elements were measured using ICP-MS. We also measured three potential drivers of soil and wood elemental composition: clay content, soil organic matter and pH. We tested associations between soil and wood using multiple regressions and multivariate analyses (Mantel test, db-RDA). Finally, we performed a Random Forest analysis of heartwood elemental composition to check site assignment accuracy.

We found elemental compositions of soil, sapwood and heartwood to be significantly associated. Soil clay content and organic matter positively influenced individual element concentrations (for 13 and 9 elements out of 46 respectively) as well as the multi-elemental composition in wood. However, associations between wood and topsoil elemental concentrations were only significant for one element. We found close associations between element concentrations and composition in sapwood and heartwood. Lastly, the Random Forest assignment success was 97.3%.

Our findings indicate that wood elemental composition is associated with that in the topsoil and its variation is related to soil clay and organic matter content. These associations suggests that the multi-elemental composition of wood can yield chemical fingerprints obtained from sites that differ in soil properties. This finding in addition to the high assignment accuracy shows potential of multi-element analysis for tracing wood origin.

2.1 Keywords

Element uptake, Heartwood formation, ICP-MS, Soil chemistry, Timber forensics, Wood chemistry

2.2 Introduction

Illegal timber trade is widespread, causing harm to the environment and the economies of exporting countries as well as fostering social inequality. To stop this, increased transparency in the timber supply chain is needed, in particular for tropical timbers (Hoare, 2015; Nellemann, 2012). Major forms of fraud in the sector are the false reporting of origin and mixing legally and illegally sourced timber (Kleinschmit et al., 2016; Brancalion et al., 2018). At present, these types of fraud are hard to detect, because methods to independently verify timber origin are limited and are fully depended on external documents and tags (Bisschop, 2012; Lowe et al., 2016). Therefore, a variety of tracing methods is under development to trace timber origin based on intrinsic wood properties, such as techniques using DNA (Vlam et al., 2018b), stable isotopes (Gori et al., 2018) and metabolic profiles (Deklerck et al., 2020). They reduce the dependency on external documents and tags which are currently used to track origin (Dormontt et al., 2015; Gasson et al., 2021). All of these methods rely on natural variation in wood characteristics within a species range. This variation determines the discrimination power as well as the spatial accuracy of the method (Vlam et al., 2018b).

Methods based on the chemical analysis of the multi-elemental composition are already successfully applied for tracing the geographical origin of a wide variety of commodities (Drivelos & Georgiou, 2012), but not yet for wood. Wine origin for example was distinguished on a scale of 100 km (Gonzálvez et al., 2009) and green tea even at an even finer scale of 10 km (Ma et al., 2016). The method relies on the quantification of a large set of elements after digestion of the organic material, measured simultaneously by analytical techniques like Inductively Coupled Plasma Mass Spectrometry (ICP-MS). This chemical profile then provides a specific 'fingerprint' per area to enable verification of the geographical origin. However, additional challenges arise when applying the method to wood. As wood is the product of a continuous growing process of trees over many years, inter-annual fluctuations in physiological processes within the tree trunk can lead to variation in its elemental composition (Hietz et al., 2015). Additionally, further element allocation processes in the trunk and the proportion of different cell types can affect the elemental composition of the heartwood, the traded part of the tree (Hietz et al., 2015). These processes may reduce differences in elemental composition between trees and thus obscure possible individual or site-specific elemental fingerprints in wood. So far, the drivers of spatial variation in wood elemental composition remain poorly understood: earlier studies showed a lack of longitudinal or latitudinal patterns in wood chemical composition (Kagawa et al., 2007), while others showed associations with bedrock type (Durand et al., 1999).

To assess the potential of wood multi-elemental analysis for tracing timber, it is important to understand what drives wood elemental composition. Yet, little is known about the underlying mechanisms that drive the uptake and allocation of many elements found in wood, especially so for many non-essential elements such as metals and rare earth elements (Watanabe et al., 2007; Brun et al., 2010). Such knowledge is essential, for instance because non-essential elements are not actively attracted by roots. Therefore their concentration and composition in wood may reflect soil conditions and thus provide a site-specific chemical fingerprint. Their composition is likely to be affected by physical-chemical soil properties that affect the binding of elements in soil, such as pH, clay content and soil organic matter (Hevia et al., 2017; Tyler & Olsson, 2001) These properties may then provide a first indication for the potential of using elemental composition for tracing wood. This is aided by the increasing availability and quality of global soil maps (Poggio et al., 2021).

To address this knowledge gap and improve our understanding of how physical and chemical soil properties (clay content, soil organic matter, pH) as well as tree physiological processes (heartwood formation) drive variation in wood elemental composition, we applied a five step approach: 1) we evaluated associations of individual element concentrations in sapwood with their concentration in soil and these three soil properties, 2) we tested the associations of sapwood elemental composition with soil elemental composition and these soil properties, 3) we evaluated changes in wood elemental composition from sapwood to heartwood, 4) we tested the assignment success of heartwood based on the multi-elemental composition and 5) we visualised the spatial variation in relevant soil properties within the distributional range of our study species. Steps 1-3 yield an understanding of the associations between soil and wood elemental composition, and thus the mechanism by which chemical tracing may operate. Steps 4-5 aimed to provide a first proof of principle of elemental timber tracing. We conducted the study on Azobé, *Lophira alata* (*L. alata*), a major African timber species.

The first two steps aim to investigate the relationships between soil and wood elemental composition in two ways. For step 1, we analysed the concentrations of each element by itself in soil and wood, hereafter called the individual element concentrations. For step 2, we assessed all the occurring elements simultaneously, hereafter referred to as elemental composition in soil and wood. This step addresses element stoichiometry: it shows whether elemental composition in wood and soil co-varies across trees (Scharnweber et al., 2016; Xing et al., 2015).

We tested the following hypotheses, illustrated in Figure 1. Soil organic matter and clay minerals are expected to provide a buffer function in the soil as they retain cations at their reactive surfaces, preventing such cations from leaching (Figure 2.1a) (Rieuwerts, 2007; Soong et al., 2020; Antoniadis et al., 2017). Both soil constituents are thus hypothesized to have a positive effect on the geochemically reactive soil cations pools, from which trees can potentially take up the associated cations. The binding of cations to these reactive soil particles also depends on soil pH (Bukata & Kyser, 2008). However, since the variation in pH of very acidic tropical soils is usually small, no strong effect of soil pH on the wood element concentrations is expected here (Rieuwerts, 2007; DeWalle et al., 1991; Yavitt et al., 2009). From the reactive soil pools, the elements are then taken up by the roots and transported through the sapwood so we hypothesize that sapwood elemental composition primarily reflects the reactive soil pools (Vitousek, 1984). This association with soil elemental composition was also observed in other commodities (Wang et al., 2020; Baroni et al., 2015; Pilgrim et al., 2010) as well as in wood (Lira-Martins et al., 2019; Heineman et al., 2016; Buras et al., 2020; DeWalle et al., 1991). Yet, trees are not just passive monitors of the reactive soil pools: soil properties and environmental factors can shift the association between soil and plant elemental compositions (Martin et al., 2006; Wyttenbach et al., 1995; Hevia et al., 2017; Russell et al., 2017; Bukata & Kyser, 2008).

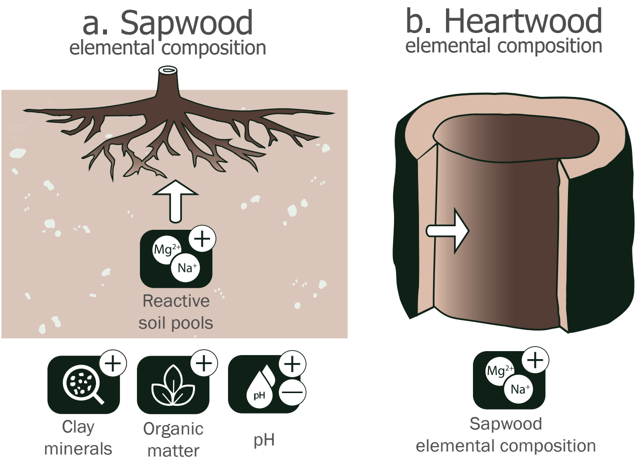


Figure 2.1: Schematic overview of processes driving sapwood (a) and heartwood (b) elemental composition of tropical trees. Hypothesised direction of the effects is indicated (positive + or no strong effect expected +/-). a) Sapwood elemental composition is expected to be a function of the reactive soil pools, which are driven by clay content, organic matter content and pH. For elements that occur predominantly as anions, such as P and As, soil organic matter and clay play a smaller role in buffering reactive element pools than for cationic elements. From the reactive soil pools, elements (both cations and anions) are taken up by the roots and transported and incorporated into the sapwood. b) Elemental composition in the heartwood is affected by translocation of elements from sapwood to heartwood during heartwood formation.

We therefore test for their effect on wood elemental composition.

After uptake, within-tree differences in wood elemental composition are expected due to the transition from sapwood to heartwood (Figure 2.1b) (Hillis, 1999; Colin-Belgrand et al., 1996; Amais et al., 2021). During this process, physiologically relevant elements can be relocated through selective recycling from the inner sapwood, which is transformed to heartwood, to the outer sapwood (Smith et al., 2014). This relocation results in higher concentrations of essential elements in sapwood compared to heartwood (Meerts, 2002; Hevia et al., 2017). For the non-essential elements an opposite pattern is expected: concentrations will be higher closer to the centre of the tree because the wood cation binding capacity decreases from the pith to the outer wood (Herbauts et al., 2002; Momoshima & Bondietti, 1990; Amais et al., 2021).

2.3 Materials and Methods

Study species and sample collection

The study was conducted on the commercial Central African timber species Azobé (*Lophira alata* Banks ex C.F. Gaertn, Ochnaceae). It occurs in evergreen and moist deciduous forests from Guinea to the Democratic Republic of Congo (GBIF.org, 2021), mostly on Ferrasols and

Table 2.1: Site topsoil properties (mean ± sd); basement rock (from Gazel (1956)), pH, organic matter content (OM; g/kg) and clay content (g/kg).

	Latitude	Longitude	Basement	pH-$CaCl_2$	OM (g/kg)	Clay (g/kg)
S1	2.796	11.071	Gneiss	3.75 ± 0.26	91.64 ± 20.31	418.05 ± 110.97
S2	2.981	10.265	Granite	3.63 ± 0.24	97.39 ± 23.48	245.59 ± 92.35

Acrisols. Both soil types are characterised by low nutrient reserves and acidity. The species produces mainly superficially spreading roots and forms associations with arbuscular mycorrhyzae (Doumenge & Sene, 2012). Mineral crystals are found in its parenchyma cells, which are potentially important storage locations for trace elements (Doumenge & Sene, 2012).

Sampling was conducted in two logging concessions in southern Cameroon in September and October 2019. The concessions are approximately 100 km apart and experience similar climatic conditions (locations in Table 2.1). Average annual rainfall is 2250 mm and annual average temperature is 25 °C.

We sampled heartwood and sapwood from 20 Azobé trees at each concession (so 40 in total), as well as the corresponding topsoil. Trees within each concession were between 100 m and 5 km apart. They were of at least 30 cm diameter breast height (DBH) and were either standing or recently felled. Sapwood and heartwood samples were collected as increment cores (Haglöf Increment borer 350 mm x 5,15 mm; n = 18), wood chunks (n = 12) or wood powder samples gained with an electrical drill (n = 10). All samples were taken at least 20 cm into the tree. Additionally, topsoil was sampled within 1 m from each tree trunk. Because *L. alata* roots are expected to grow mainly superficially, topsoil was assumed to be the most important for element acquisition by roots. Three soil subsamples of 20 cm deep were dug, which were then mixed and manually cleaned of large roots and leaves. The resulting soil samples were air-dried in the field.

Wood processing and chemical analysis

For the chemical analysis a subsample of heartwood and sapwood was cut from every increment core and wood chunk with the Chopper2000 (König, 2022), spanning at least 3-5 cm of wood to ensure the samples include wood formed during multiple years (estimated at 5 years or more). *L. alata* shows a clear colour difference between sapwood and heartwood, allowing for easy distinction. The resulting samples were analysed using an adapted protocol for large element screening (SOP-A-1120). For this purpose, 1.0 g of wood was digested using a microwave technique (CEM Mars 6), using 3 mL of 70% HNO_3 and 5 mL of ultrapure Milli-Q H_2O. The remaining solutions contained no traces of organic constituents.

Element concentrations in the digests were determined by ICP-MS (NexION 350D, PerkinElmer). Two multi-element solutions (1.0 mg/L) and a Hg solution (10 mg/L) were used to prepare calibration standards, including 60 elements: Ag, Al, As, Ba, Be, Bi, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Ga, Gd, Ge, Hf, Ho, K, La, Li, Lu, Mg, Mn, Mo, Na, Nb, Nd, Ni, P, Pb, Pr, Rb, Re, Sb, Sc, Se, Si, Sm, Sn, Sr, Ta, Tb, Te, Th, Ti, Tl, Tm, U, V, W, Y, Yb, Zn, Zr and Hg. Spiked and non-spiked certified reference material BCR 482 Lichen were used to guarantee the accuracy of the measurements. Rh was used as internal standard. The detection limit was determined per element per measurement cycle, calculated as three times the intensity of that element in a blank standard. Only elements measured at a concentration above their detection limit in more than half of the wood samples were included in further data analyses. Two elements (Fe, essential element and Al, 'other element') were measured but results were considered to be unreliable due to the high occurrence of inferences in the ICP-MS measurements. Consequently, we have no quantitative information on the occurrence of Fe- and Al-(hydr)oxides in our soil samples, which are responsible for the binding of oxyanions in tropical soils (Mendez et al., 2022). For three trees the internal standard was not sufficiently high in the sapwood or heartwood measurement, therefore the final dataset contained 37 trees.

Topsoil processing and analysis

Soil samples were dried for at least 24 hours at 40 °C upon arrival in the laboratory, after which they were sieved to grain size <2 mm. In every soil sample, pH was measured in a settling suspension after shaking the samples in a 0.01 M $CaCl_2$ solution at a solution-to-soil ratio of 10 L/kg for two hours at 20 °C (Houba et al., 2000). Soil organic matter and clay mineral contents were determined according to standard analytical procedures (Houba et al., 1995). Soil organic matter content was measured by loss-on-ignition in a muffle furnace at 550 °C. Clay mineral content was determined using the sieve and pipette method. To measure the reactive elemental pool in topsoil, we used the 0.43 M HNO_3 extraction method, following Groenenberg et al. (2017). The soil samples were shaken for four hours in a 0.43 M HNO_3 solution at 20 °C using a solution-to-soil ratio of 10 L/kg. After centrifuging and filtration, element concentrations were measured by ICP-MS as described above for the wood samples.

Statistical analysis

Associations between individual element concentrations in soil and sapwood

All statistical analyses were performed in R version 4.0.2 (R Core Team, 2021). For step 1, individual concentrations of elements in topsoil and sapwood were correlated using Spearman rank correlations. Additionally, the effect of soil properties on the sapwood concentration was tested per element by modelling sapwood element concentration as a function of clay content, soil organic matter, pH and the soil concentration of the respective element in multiple linear regressions. Interactions between clay content, soil organic matter, pH and soil element concentration were also included in the full model. Site was not included as factor in these models, because we had no interest in distinguishing sites. All variables except for pH were log-transformed prior to analysis. A small number (10-12 mg/kg for all elemental concentrations and 10-12 g/kg for soil organic matter and clay content) was added in order to shift all values to above zero as some element concentrations that fell below their detection limit were recorded as zero. Then all variables were centred and scaled to zero mean and one standard deviation. Each full model was dredged to generate a list of candidate models ranked by AICc. Full model averaging was performed using candidate models within the top two ΔAICc units (Burnham & Anderson, 2002). All interactions were tested but only 'clay content x organic matter' occurred in the candidate models. Model dredging and averaging were conducted with the MuMIn package (Barton, 2020).

Multivariate analysis of elemental composition

To test the associations between topsoil and sapwood elemental composition (step 2), we calculated distance matrices for all sapwood and topsoil samples based on scaled element concentrations. Distance matrices were calculated using Chord distances:

$$CD_{i,h} = \sqrt{\sum_{j=1}^{p}((\frac{a_{i,j}}{\sum_{j=1}^{p} a_{i,j}^2}) - (\frac{a_{h,j}}{\sum_{j=1}^{p} a_{h,j}^2}))^2} \tag{2.1}$$

where $\mathrm{CD_{i,h}}$ is the Chord distance between sample i and h. The data matrix has p element columns and $\mathrm{CD_{i,h}}$ includes the differences in all p elements: $\mathrm{a_{i,j}}$ (the concentration of element j in sample i) minus $\mathrm{a_{h,j}}$ (the concentration of element j in sample h), both standardized to the summed concentration of all measured elements within the sample.

In this equation, differences in total concentration among samples are removed and this thus provides a measure of relative element concentrations within a sample (McCune & Grace, 2002). The resulting distance matrices (for topsoil and sapwood) were correlated by Mantel tests (9999 iterations) based on Spearman correlations.

To test whether any soil properties (clay content, soil organic matter, pH) significantly explained the relative elemental composition in the sapwood, we conducted a distance based Redundance Analysis (db-RDA). This calculates the variance explained by the three soil properties. Mantel tests and db-RDA were conducted using the vegan package (mantel and capscale) (Oksanen et al., 2020).

Correlating element concentrations and elemental composition in sapwood and heartwood

To address step 3, the associations of elemental concentrations in sapwood vs. heartwood were tested using Spearman rank correlations. Additionally, Mantel tests were used to correlate elemental compositions of sapwood and heartwood. As in the analysis of topsoil and sapwood correlations, this was calculated based on Chord distances. Lastly, we tested whether elemental concentrations were higher in sapwood or heartwood, by calculating radial discrepancy (RD (%) = (sapwood - heartwood)/sapwood * 100, Heineman et al. (2016)).

Potential for timber tracing: site assignment and modelling of soil properties within the species range

In step 4 we aimed to verify whether elemental composition can be used to assign wood to its respective origin in these two study sites using a Random Forest classification model. In this analysis, we assigned heartwood samples to their most likely origin based on the 46 measured elements. Furthermore, the elements that were most important to distinguish between the two sites were identified based on variable importance in the Random Forest analysis. This was done using the randomForest package (Liaw & Wiener, 2002)

Modelling soil properties within the species distribution range of L. alata

As 5th and final step, clay and soil organic matter – two soil variables that turned out to be important in determining wood elemental composition – were plotted within the species range of *L. alata*. This was done in order to show the variation in these two soil properties and thus potential for tracing within the range of *L. alata*. The species distribution range was modelled using Maxent (Phillips et al., 2006), based on GBIF occurrence data (GBIF.org, 2021) and 19 bioclimatic variables from worldclim.org (Fick & Hijmans, 2017). Clay content and soil organic carbon (SOC) were obtained from Soilgrids within this range (Poggio et al., 2021). We divided SOC by 0.58 to obtain soil organic matter values, to compare it to the soil organic matter measured in this study (Pribyl, 2010). Species distribution modelling was conducted with the dismo package (Hijmans et al., 2020), maps were plotted with the raster package (Hijmans, 2021a).

2.4 Results

Soil and wood properties

Soils at the study sites were acid, which is typical for the region (pH between 3.06 and 4.38; Table 2.1). Clay content showed the largest variation of the three soil properties (125.38 - 543.49 g/kg). It was higher on average in site 1 (S1) but still overlapped across the two sites. The average in soil organic matter content was similar at both sites, ranging from 45.70 - 139.90 g/kg.

Single element concentrations varied by four orders of magnitude, from 0.12 µg/kg for Dy to almost 1 g/kg for K. Out of the 60 measured elements, 46 were detected in over half of the heartwood and sapwood samples of *L. alata*. All of these elements were also detected in the corresponding soil samples. We categorized them in three categories: 1) 10 essential elements (K, Mg, Ca, Mn, Ni, Cu, Zn, Mo, Si and P); 2) 16 rare earth elements (Sc, Y, La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb and Lu) and 3) 20 'other elements' (Li, Na, Rb, Cs, Be, Sr, Ba, Ti, Cr, Co, Zr, Cd, Hg, Ga, Sn, Pb, Ge, As, Th and U). Twelve elements were omitted from the analyses because they fell below the detection limit in more than 50% of the samples (all 'other elements'): Ag, Bi, Hf, Nb, Re, Sb, Se, Ta, Te, Tl, V, W.

Soil to sapwood regression models

Step 1 of our five- step approach yielded multiple regressions of sapwood element concentration as a function of soil properties. The regressions contained a significant explanatory variable in only 16 out of 46 elements (Figure 2.2). Among the regression variables included in the models, clay content was significant in most of the models (13 of 46 elements), followed by soil organic matter (9 of 46) and pH (2 of 46). The effect of clay on sapwood element concentration was mostly positive (11 of 13 elements). For soil organic matter, the effect was positive for 6 of the 9 elements where it was significant (alkali metals and lanthanides). The observed minor effect of pH is likely due to its very limited variability in this acidic range where the pH-dependent solubility of most major and trace elements is relatively high. An effect of the reactive element concentration in the topsoil on elemental concentration in sapwood was only found in one multiple regression model (Zr, 'other element' class).

In addition to the regression models, we also checked for associations between soil and wood elemental concentrations using Spearman correlations. These showed somewhat contrasting patterns to the regression models: 14 out of 46 correlations between element concentrations in topsoil and sapwood were significant (Figure S2.1). Of those correlations, 13 were positive and one negative.

In the regression models, the effect of soil properties on sapwood elemental concentrations was different for the three predefined elements groups. Within the category of essential elements we did not find a strong pattern in the effect of the soil properties. Clay content and organic matter had a strong effect on the sapwood concentration of only two essential elements, Mo and Mn (Figure 2.2). In contrast, we found more consistent significant effects of soil properties on the sapwood concentrations of the rare earth elements. For half of the rare earth elements, clay content or the combination of clay content and organic matter had a strong positive effect on the sapwood element concentrations. Among the drivers of sapwood concentrations, clay content turned out to be most important as it was included in all significant models for rare earth elements. Within the category of 'other elements', model results were quite variable. We found no consistent effect of soil properties on sapwood elemental composition there, neither when grouped based on atomic mass or element type in the periodic system.

Elemental composition in topsoil and sapwood

Step 2 involved conducting multivariate analyses in order to assess drivers of wood elemental composition. We found significant correlations of Chord distances of topsoil and sapwood elemental concentrations, i.e. when element concentrations were standardized to the total element concentration per sample (Figure 2.3A; multivariate scaling shown in Figure S2.2A and B).

Our db-RDA analysis of the soil properties showed that clay content and soil organic matter together explained 22.34% of the variation in sapwood elemental composition, while pH was not significant (Figure 2.4). Most variation (19.6%) was explained by the first axis. Consistent with results for the element concentrations, clay content was also the most important driving variable explaining variation in sapwood elemental composition: it defined most of the first RDA axis.

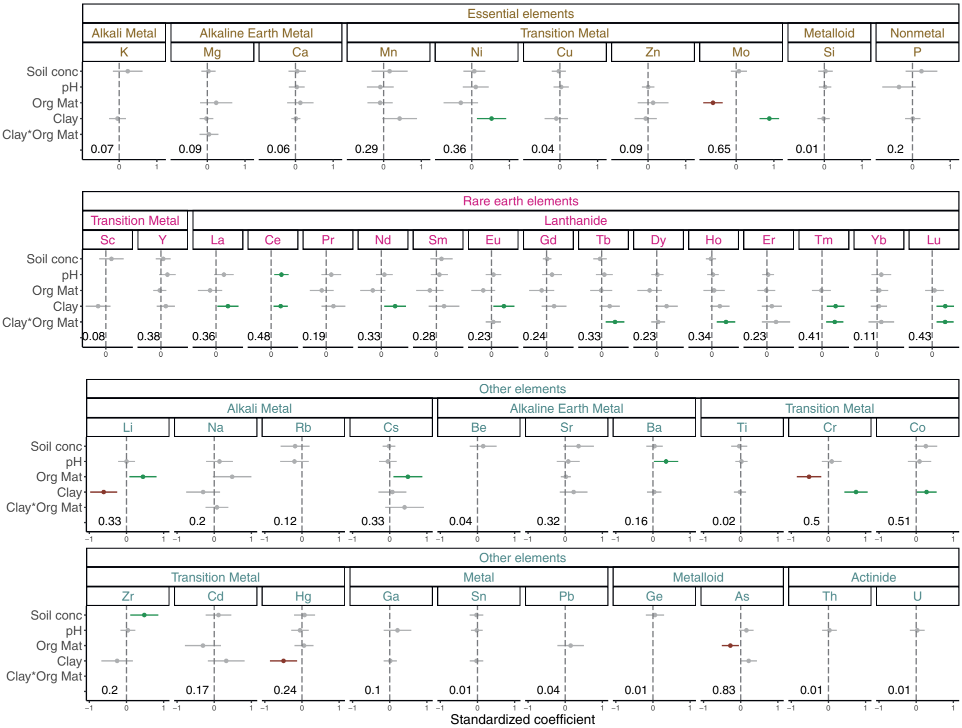


Figure 2.2: Multiple regression models of sapwood element concentrations as a function of soil element concentration, soil properties and their interactions (step 1). Elements are grouped and colour-coded by category, and within category ordered by type and weight in the periodic system. Explanatory variables are concentration of the respective element in topsoil (soil conc, mg/kg), pH, organic matter content (g/kg), clay content (g/kg) and interaction of clay and organic matter. Coefficient plots include slope estimates (mean and 95 %CI), averaged across top-ranked models. The CI of the red and green circles does not overlap with zero and therewith denotes a strong and large effect, green for a positive and red for a negative effect. The grey circles were included in the top models but their CI overlaps with zero, therefore they do not indicate a strong effect on the response variable. Numbers in the panels represent averaged adjusted R^2 of the best models.

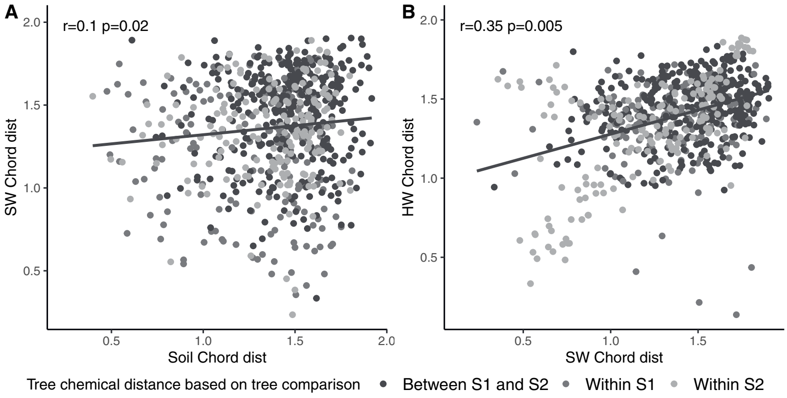


Figure 2.3: Associations between chemical composition in soil, sapwood (SW) and heartwood (HW), based on Mantel tests (results shown per panel). Soil and sapwood (A; step 2) and sapwood and heartwood (B; step 3) distances based on Chord distance. Grey shades of symbols indicate whether trees were from different sites (dark grey), both from S1 (medium dark grey) or both from site 2 (S2) (light grey). Grey lines show significant relations.

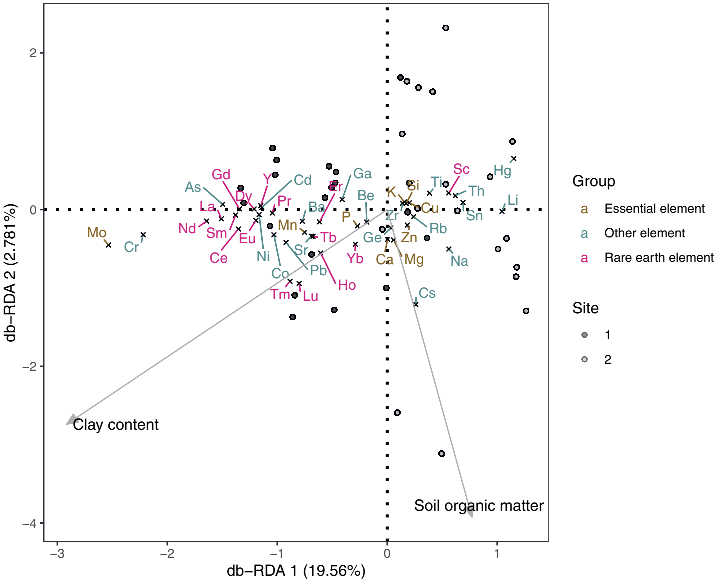


Figure 2.4: Distance based Redundancy Analysis (db-RDA) of sapwood elemental composition (step 2). Relative position of trees (dots) is based on Chord distances of square root transformed element concentrations, coloured by site. Weight and direction of the element concentrations in the sapwood is indicated as crosses, element categories colour-coded as in Fig 2.2. Vectors indicate the weight and the direction of the two significant soil properties. Percentages in axis labels denote the percentage of total variance explained.

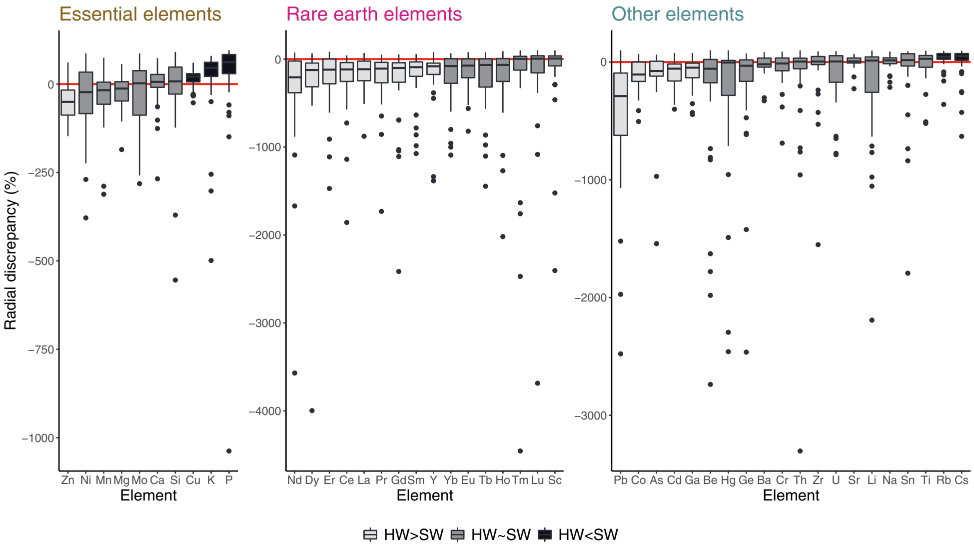


Figure 2.5: Radial discrepancy plot of element concentrations in heartwood (HW) compared to sapwood (SW; step 3). Grey shades indicate whether 3 of the 4 quantiles fell above, around or below zero. Values below zero indicate higher concentrations in the heartwood (lightest grey), whereas values above zero indicate higher concentrations in the sapwood (darkest grey). Values around zero indicate similar sapwood and heartwood concentrations (medium dark grey).

Elemental composition in sapwood and heartwood

In step 3 we found significant associations of elemental composition in sapwood and heartwood (Figure 2.3B); multivariate scaling shown in Figure S2.2B and C). This indicates that elemental composition in the sapwood was largely conserved after heartwood formation. This association was also reflected by the large share of positive correlations (35 out of 46) between sapwood and heartwood concentrations of individual elements (Figure S2.3).

Comparisons of concentration differences between sapwood and heartwood, expressed as the radial discrepancy, revealed that concentrations were higher in sapwood for three of the 10 major essential elements (Cu, K and P, Figure 2.5). Two elements from the ‘other elements’ category (Rb and Cs), known to be similar to K in terms of plant uptake and chemical behaviour (Zhu & Smolders, 2000), were also more abundant in the sapwood. In the rare earth element category, over half of the elements was more abundant in the heartwood (9/16 elements) and the rest did not differ between the two wood compartments (7/16).

Tracing potential: origin assignment and spatial variation within the distributional range of *L. alata*

Step 4 then provided a first evaluation of the assignment success of wood elemental tracing at our two study sites. The Random Forest classification was very successful: 36 of the 37

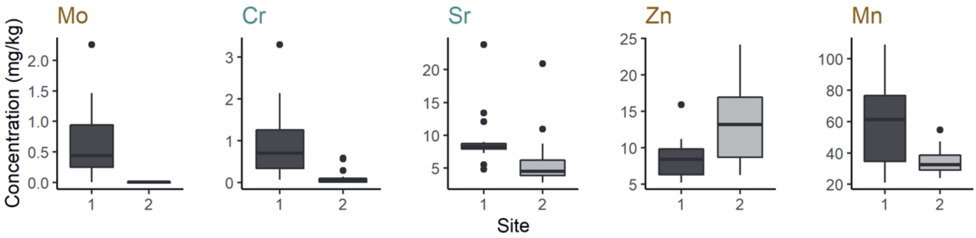


Figure 2.6: Element concentrations (mg/kg) of the five most important elements of the Random Forest analysis, in decreasing order of importance (Random Forest assignment accuracy of 97.3%; step 4). Elements are colour-coded per category as in Fig 2.2.

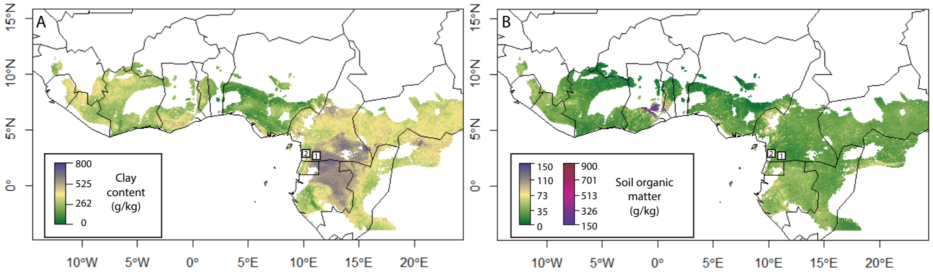


Figure 2.7: Variation in clay content (A; g/kg) and soil organic matter (B; g/kg) within the predicted distributional range of *L. alata* (step 5). Predicted range was based on Maxent modelling using observation data from GBIF and 19 bioclimatic variables. Note the two colour scales for soil organic matter, to better show variation across most of the species range (values above 150 g/kg are only found in one region in Ghana). Numbers indicate the research sites.

trees were correctly assigned (success rate of 97.3%). The five most important elements in the Random Forest analysis were in the essential element category (Mo, Zn, Mn) and ‘other element’ category (Cr, Sr; concentrations shown in Figure 2.6). The highest ranking rare earth element in terms of importance was La, at place number nine.

In step 5, we further evaluated the potential of chemical tracing for our study species beyond the two study sites, by mapping the spatial variation of the two most important drivers of wood chemical composition (clay content and soil organic matter). We found soil clay content to differ five-fold within the distributional range of *L. alata*, with large differences at small scales such as in southern Cameroon (Figure 2.7A). Regional differences in soil organic carbon were also found although not as pronounced as for clay content (Figure 2.7B), except for one region in Ghana.

2.5 Discussion

We found that wood element concentrations and composition of a tropical tree species were associated with soil clay content and organic matter but the individual element concentrations were not directly linked with element concentrations in the topsoil. We also found that relative elemental composition was retained across topsoil, sapwood and heartwood. Furthermore, we observed shifts in individual element concentrations from sapwood to heartwood. Finally, we could correctly assign heartwood to one of the two study sites with 97.3% accuracy .

Elemental composition of soil is reflected in wood

We hypothesized that wood elemental composition would be driven by soil composition for elements taken up by the roots. In our study, the relative elemental composition of the topsoil was indeed retained in the sapwood as well as the heartwood (step 2, Figure 2.3A). This is in line with findings on other commodities such as bananas and honey, for which the elemental composition of the commodities was also linked to soil element composition including multiple elements (Wang et al., 2020; Baroni et al., 2015).

In contrast, individual topsoil element concentrations, as extracted by 0.43M HNO_3 (Groenenberg et al., 2017), were not a good single predictor for the sapwood concentrations in our regression and correlation models (step 1, Figure 2.2). This is consistent with results from previous studies on nutrients in soil and wood (Bukata & Kyser, 2008; Johnson et al., 2001; Russell et al., 2017), although contrasting findings have also been reported (Heineman et al., 2016; Lira-Martins et al., 2019). A possible explanation is that the studies reporting significant associations between soil and wood concentrations included a wide range of soil element concentrations or studied multiple tree species (Heineman et al., 2016; Lira-Martins et al., 2019). In our study, concentration differences were relatively small and between-site distance was <100 km. Therefore, research of these relations across larger soil gradients is needed to provide further insights into the influence of soil elemental concentrations on wood chemistry.

Clay content and soil organic matter drive wood element concentrations and elemental composition

Individual wood element concentrations were found to be associated more with chemical and physical soil properties than with the individual element concentrations in the soil (step 1, Figure 2.2). This is consistent with other studies indicating the importance of soil properties for wood elemental composition (Soong et al., 2020; Hevia et al., 2017). Among the 18 regression models of wood element concentration with significant explanatory variables, clay content and soil organic matter were most commonly included as explanatory variable. That effect was mostly positive for both soil properties, in line with our hypothesised framework (Figure 2.1). The effect of clay content and soil organic matter was also found in the relative concentrations of these elements across soil and wood: variation in the relative elemental composition was explained by clay content and soil organic matter (Step 2, Figure 2.4).

This positive effect of clay content and soil organic matter on wood element concentrations contrasts that of similar research on annual crops, where it is often found that clay content and soil organic matter decrease cation mobility and thus uptake by plants (Antoniadis et al., 2017). This discrepancy is likely explained by the longevity of trees. On time scales of days to weeks, relevant to annual crops, a high number of cation exchange sites in the soil results in decelerated cation mobility and thus cation availability (Römkens et al., 2009; Antoniadis et al., 2017; Kabata-Pendias, 2011). For trees however, those fixed cations can also be taken up from less directly available pools as they gradually become available through weathering or by active displacement of cations from soil exchange sites over the period of multiple years or decades (Russell et al., 2017). Thus, at longer time scales, a larger number of cation binding sites may prevent short-term leaching of cations and then a higher concentration of clay minerals and soil organic matter may act as a long-term reservoir of elements (Rieuwerts, 2007; Soong et al., 2020). Even though the applied HNO_3-extraction method has been developed to represent these longer-term available element pools (Groenenberg et al., 2017), we did not find a direct relationship between soil and wood concentrations (Figure 2.2). This may be caused by the limited gradients in availability-controlling soil properties (i.e., pH, soil organic matter, clay minerals) in this study.

Even though we found an effect of some soil properties on wood concentrations, there were also 28 of 46 individual element models where we found no effect of any of our tested soil properties (step 1, Figure 2.2). Testing these relations across larger gradients of clay content, organic matter and pH is needed to verify whether this absence of soil effects on wood element concentrations is rigorous. However, finding such a gradient in a single tree species may be challenging, as soil conditions are a major driver of species occurrence. Further studies should also include analyses of clay mineralogy as this can potentially affect the availability of trace elements in tropical soils through weathering of clay minerals (Antoniadis et al., 2017; Rieuwerts, 2007). Furthermore, we recommend addressing the role of Fe- and Al-(hydr)oxides in future studies as they are especially important for elements forming oxyanions like phosphate and arsenic (Mendez et al., 2022). Together, these recommendations will generate a more complete understanding of the links between soil and wood chemistry for tropical trees.

Heartwood and sapwood: elemental composition is retained, but individual elements differ

In addition to soil chemistry, elemental composition in the heartwood is also determined by stem physiological processes such as the transition from sapwood to heartwood. We found that individual concentrations of three essential (Cu, K, P) as well as two non-essential elements that are similar in uptake and behaviour to K (Rb, Cs) were consistently higher in sapwood compared to heartwood (step 3, Figure 2.5). These higher sapwood concentrations are commonly found in tree species, and likely result from the active translocation of nutrients during heartwood formation (Hietz et al., 2015; Andrews et al., 1999; Meerts, 2002).

Furthermore, many rare earth elements followed the hypothesised pattern of higher concentrations in the heartwood. This is unlikely caused by active translocation, but may be explained by the increasing amounts of pectates in a gradient from bark to pith, which bind cations (Amais et al., 2021). In spite of the substantial shifts in some individual element concentrations from sapwood to heartwood, the relative elemental composition of the sapwood was retained in the heartwood. This indicates that the stoichiometric relations among elements are largely retained in the radial direction of the tree trunk, even after the transition from sapwood to heartwood. Yet, it is unclear whether this also holds for wood along the height of a tree trunk (Scharnweber et al., 2016). This needs to be verified to ascertain that wood originating from different heights along the trunk has similar elemental composition.

Implications for timber tracing based on multi-element analysis

Our findings help to evaluate the potential of multi-element analysis for tracing of *L. alata*. The results of steps 1-3 suggest potential for elemental tracing of timber in areas where soil chemical composition varies spatially. Even though we did not find strong associations for individual element concentrations, we did find that the reactive elemental composition in the topsoil was reflected in sapwood and heartwood. We also found wood chemical composition to be associated with chemical and physical soil properties. The soil property in which our two sites differed most was clay content, although it was not significantly different (Table 2.1). Thus, our results suggest that the elemental fingerprint of wood in our study species reflects that in the topsoil and is associated with general soil properties.

Our proof of principle to use elemental composition to assign samples to sites (step 4) showed a high assignment accuracy: 97.3%. This value is comparable to that obtained using elemental tracing in other commodities (Gonzálvez et al., 2009; Ma et al., 2016). We found three essential elements to be important in distinguishing our two study sites (Mo, Zn, Mn; Figure 2.6). In contrast with findings of elemental tracing for other commodities, rare earth elements on the other hand were not included in the top-five elements in the Random Forest analysis (Drivelos & Georgiou, 2012). Clearly, these findings are based on only one species, sampled at just two sites. Thus, the accuracy obtained for this first evaluation should be interpreted cautiously and cannot be extrapolated to larger areas or other species. Nevertheless, this high accuracy value – in combination with the soil-wood associations of elemental composition – does suggest that elemental tracing of tropical timber has a high potential to reach required levels accuracy for forensic tracing, and doing so at relatively short distances (100 km). To further evaluate the potential and prepare this technique for practical forensic application, we recommend testing soil-wood chemical associations and assignment tests in larger areas and for other tree species. Based on our finding that wood elemental composition is most strongly related to soil clay content and soil organic matter, regions with high variation in these soil properties, e.g. southern Cameroon, are particularly suitable for such next steps (step 5, Figure 2.7).

If the associations between soil and wood elemental composition found here are similar for other traded timbers in the region, this would allow pooling reference data from various species for tracing studies. Pooling samples would generate an advantage of this method over other tracing methods that require species-specific databases, e.g. genetic tracers or metabolites (Vlam et al., 2018b; Deklerck et al., 2020), as it would greatly reduce the effort to collect reference data.

2.6 Acknowledgments

This study was supported by the Dutch Research Council (NWO-TTW-OTP-16427). We thank Cameroon United Forest and their field teams for facilitating the fieldwork. We thank Greet van Bemmel for assistance in the laboratory work.

Supporting information

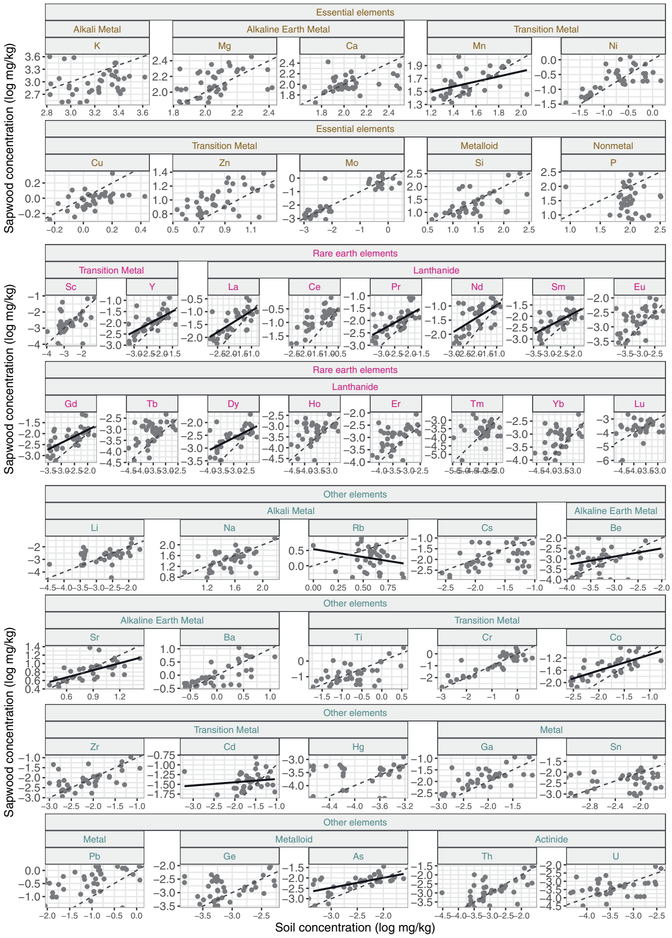


Figure S2.1: Spearman rank correlations of element concentrations in soil and sapwood (log mg/kg) of the two study sites. The elements are ordered by category (indicated in pink, blue and brown), their type in the periodic system and then by weight within that type/category. The solid black line shows the correlation coefficient, only significant correlations are drawn. The grey dotted lines depict a 1:1 relationship between soil and sapwood.

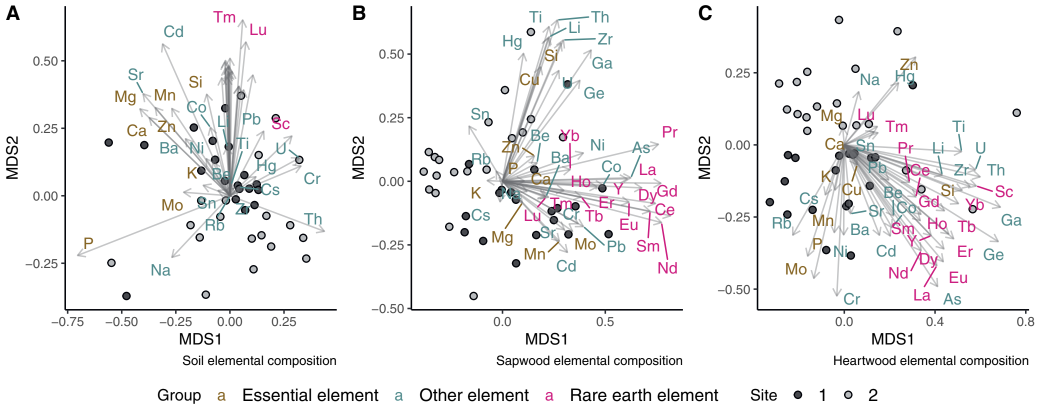


Figure S2.2: MDS of elemental composition in soil, sapwood and heartwood based on Chord distances. The weight and direction of the element concentrations is indicated as arrows, with the element categories indicated in pink, blue and brown. Dots represent the sampled trees originating from the two sites in dark and light grey.

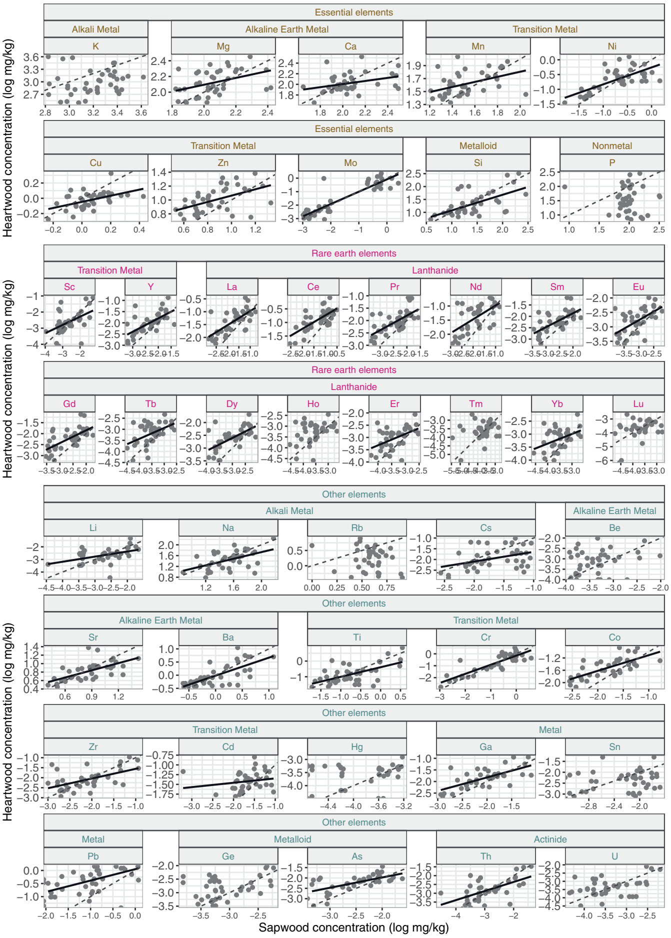
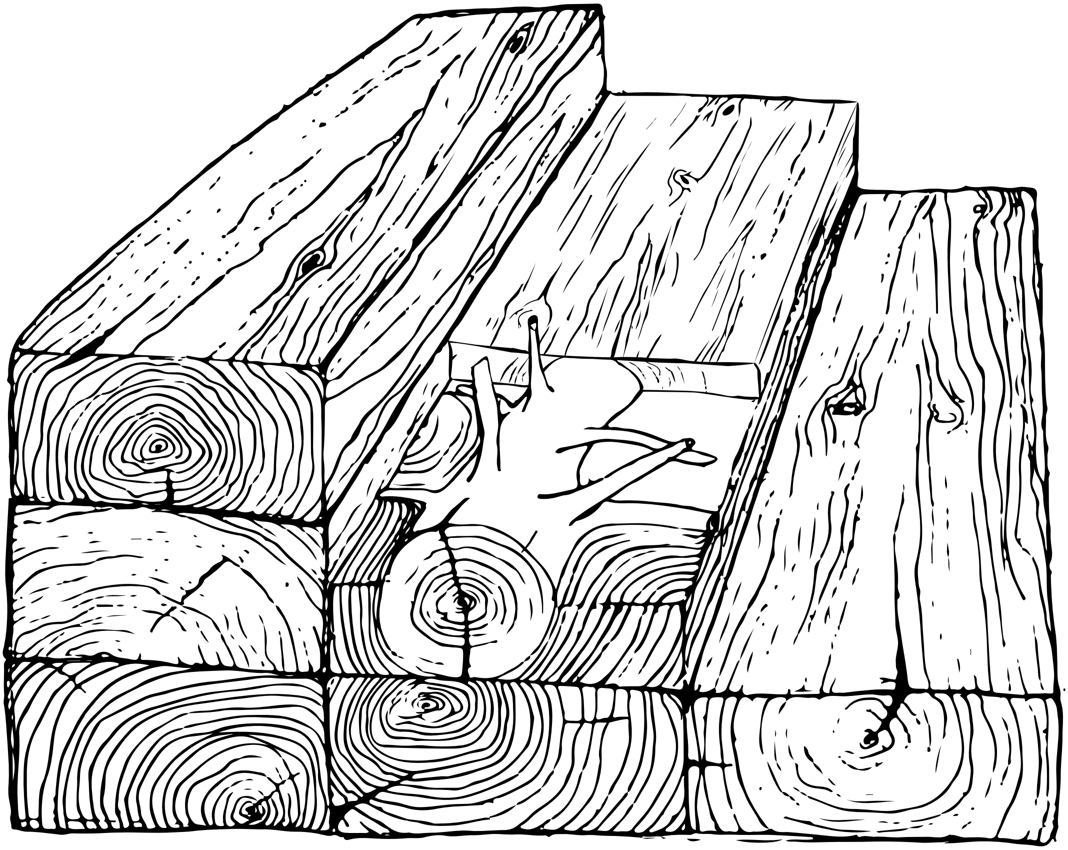


Figure S2.3: Spearman rank correlations of element concentrations in sapwood and heartwood (log mg/kg) of the two study sites. The elements are ordered by category (indicated in pink, blue and brown), their type in the periodic system and then by weight within that type/category. The solid black line shows the correlation coefficient, only significant correlations are drawn. The grey dotted lines depict a 1:1 relationship between sapwood and heartwood.



Chapter 3

A new method for the timber tracing toolbox: applying multi-element analysis to determine wood origin

Laura E. Boeschoten, Mart Vlam, Ute Sass-Klaassen, Barbara Rocha Venâncio Meyer-Sand, Ulfa Adzkia, Gaël U.D. Bouka, Jannici C.U. Ciliane-Madikou, Nestor L. Engone Obiang, Mesly Guieshon-Engongoro, Joël J. Loumeto, Dieu-merci M.F. Mbika, Cynel G. Moundounga, Rita M.D. Ndangani, Dyana Ndiade Bourobou, Mohamad M. Rahman, Iskandar Z. Siregar, Steve N. Tassiamba, Martin T. Tchamba, Bijoux B.L. Toumba-Paka, Herman T. Zanguim, Pascaline T. Zemtsa, Pieter A. Zuidema

Published in Environmental Research Letters, 18 (2023) 054001, https://doi.org/10.1088/1748-9326/acc81b

Abstract

To effectively reduce illegal timber trade, law enforcers need forensic methods to independently verify claims of wood origin. Multi-element analysis of traded plant material has the potential to be used to trace the origin of commodities, but for timber it has not been tested at relevant large scales. Here we put this method to the test, by evaluating its tracing accuracy for three economically important tropical timbers: Azobé and Tali in Central Africa (22 sites) and Red Meranti on Borneo (9 sites). Wood samples from 991 trees were measured using Inductively Coupled Plasma Mass Spectrometry (ICP-MS) and element concentrations were analysed to chemically group similar sites (clustering) and assess accuracy of tracing samples to their origin (Random Forest Models). For all three timbers, we found distinct spatial differences in chemical composition. In Central Africa, tracing accuracy was 86-98% for regional clusters of chemically similar sites, with accuracy depending on the tracing question. These clusters were 50-800 km apart and tracing accuracy was highest when combining the two timbers. Tracing accuracy of Red Meranti on Borneo was 88% at the site level. This high accuracy at a small scale may be related to the short distances at which differences in soil type occur on Borneo. A blind sample analysis of 46 African timber samples correctly identified the origin of 70-72% of the samples, but failed to exclude 70% of the samples obtained from different species or outside the study area. Overall, these results illustrate a high potential for multi-element analysis to be developed into a timber tracing tool which can identify origin for multiple species and can do so at a within-country scale. To reach this potential, reference databases need to cover wider geographic areas and represent more timbers.

Keywords

Illegal logging, Wood chemistry, Random Forest, Blind sample analysis, Tropical timber, ICP-MS, Timber forensics

3.1 Introduction

Illegal timber trade ranks among the most profitable wildlife crimes and the false declaration of origin is one of the main types of timber fraud (Hoare, 2015; Dormontt et al., 2015). This harms people, ecosystems and local economies and impairs initiatives to increase sustainable forest management and timber trade (Hoare & Uehara, 2022). The lack of methods to independently verify timber origin limits effective law enforcement and full supply chain transparency, as timber origin is currently verified by external documents and tags. Therefore, novel methods to trace timber origin based on intrinsic wood properties, rather than external documents, can play a crucial role in stopping illegal timber trade (Lowe et al., 2016).

The multi-element analysis of wood is a promising new method for origin identification (Boeschoten et al., 2022). A large number of elements (such as Mg, Ca, La) are measured simultaneously using mass spectrometry and based on this elemental composition, an origin-specific fingerprint is defined. By comparing the chemical fingerprint to an established geo-referenced dataset, the most likely origin of a sample can be determined. This can help to answer two relevant tracing questions: origin verification and assignment (Deklerck, 2023). The verification question addresses whether a wood sample came from a certain (claimed) origin. On the other hand, the assignment question addresses which location a wood sample most likely originated from in case sample origin is entirely unknown. Both questions are relevant in tracing but require a different statistical approach. Multi-element analysis was already found to be useful for assignment and verification of a variety of commodities, such as green asparagus, bananas and tea (Gonzalvez et al., 2009; Richter et al., 2019; Ma et al., 2016), and has been successful at both regional and continental scales (Joebstl et al., 2010; Baroni et al., 2015). For timber, further research across different countries and species is essential to understand how the method operates (Boeschoten et al., 2022).

Multi-element analysis may provide advantages in origin determination compared to other timber tracing methods under investigation, such as stable isotopes, genetic or DART-TOFMS analyses (Lee et al., 2015; Deklerck et al., 2020; Vlam et al., 2018b). As wood chemical composition is thought to be influenced by soil chemistry (Boeschoten et al., 2022), the chemical profile of multiple species might overlap. Even if species differ in elemental composition (Amais et al., 2021), the variation between sites could still be reflected in multiple species. In that case, their reference data could be combined, which would reduce the reference data collection effort drastically (Gasson et al., 2021). This is contrary to other methods that need a unique fingerprint for each species.

Challenges are expected in the application of elemental analysis for timber tracing as the chemical composition of timber is more complex than most other commodities. First, wood formation is a process of many years and the traded part of the tree, the heartwood, is thus influenced by environmental and ontogenetic changes over the tree‘s lifetime (Scharnweber et al., 2016; Hevia et al., 2017; Hietz et al., 2015). Many other commodities are only affected by the environmental conditions of one year, which may lead to a more unique fingerprint per region. Second, the method is only useful for tracing timber when variation is present with defined differences between origins (Hevia et al., 2017). The applicability of the method thus depends on natural variation in elemental composition within a species distribution range, but limited research is available on this for trace elements in wood. Wood elemental composition has been associated with soil physical and chemical properties (Boeschoten et al., 2022), so the method may be especially successful in regions where soil chemistry exhibits strong spatial variation.

This is the first study to put multi-element analysis for timber tracing to practice. We did so by determining its applicability for three economically important tropical timber species: two from Central Africa, traded as Azobé (*Lophira alata*) and Tali (*Erythrophleum ivorense* and *E. suaveolens*), and Red Meranti from Borneo (*Shorea* spp.). We first analysed whether the two timbers from Central Africa shared their variation in elemental composition across the study region. Then we developed classification models for all timber species based on the elemental composition, addressing both origin verification and assignment accuracy. We also investigated at what scale we found tracing potential for these three timber species. In order to understand what determines elemental variation across a landscape, we visualised how key elements varied across the study area and assessed whether wood elemental compositions correlated with soil variables that affect element uptake. Lastly, we simulated a real-life tracing case by applying the classification models to a set of blind samples, to determine their most likely origin.

3.2 Material and methods

Study design

The study was conducted on three valuable internationally traded tropical timber species: two from Central Africa (Azobé; *Lophira alata* Banks ex C.F. Gaertn, Ochnaceae and Tali; *Erythrophleum ivorense* A. Chev. and *E. suaveolens* (Guill. & Perr.) Brenan, Fabaceae), and one from South-East Asia (Red Meranti; *Shorea* spp., Dipterocarpaceae). They represent a single-species timber, Azobé, and two multi-species timbers. Tali timber is sourced from two botanical species and Red Meranti is generally accepted to be sourced from 10 species (Klaassen, 2018). The botanical species that make up Tali and Red Meranti are hard to distinguish in the field, therefore they were only identified to species level if leaf, flower and/or fruit material was available. Otherwise samples were identified at genus level. As many individual trees did not have branches below three meters and we had no equipment available for sampling higher branches, fresh leaves could not be collected at every tree. Furthermore, as the aim of this study was to test chemical differences at the timber level, we did not pursue further species-level analysis. All three timbers grow in evergreen and moist deciduous forests.

Samples in Central Africa were taken from 22 study sites across the main timber exporting countries: Cameroon (seven sites), Gabon (seven sites) and the Republic of the Congo (eight sites), between September 2019 and April 2022 (see Fig 3.2A). Samples were taken from both species at all sites except for three locations: at two sites only Azobé was sampled and at one site only Tali was sampled due to low occurrence of the other species. Samples on Borneo were taken from nine study sites across the provinces West, East and Central Kalimantan, Indonesia, between January 2020 and February 2022 (see Fig 3.2B). All sites were natural forest concessions, accessed in collaboration with the operating forestry companies.

Sample collection

At each site, we sampled heartwood from 20-30 trees per timber species. Sampled trees within one site were located between 100 m and 5 km apart. Trees were either standing or recently felled and were of at least 30 cm diameter at breast height (DBH). A heartwood sample was collected from each tree as an increment core (Haglöf Increment borer 350 mm x 5,15 mm; n = 238), with a FAMAG plug cutter of 15 mm diameter (n = 562), as a wood chunk (n = 23) or as a wood powder sample obtained with an electrical drill (n = 170). All samples were taken 10-20 cm into the tree. Additionally, GPS-coordinates and DBH were recorded. This resulted in a geolocated database of 179 Red Meranti, 420 Azobé and 394 Tali samples. For every tree, we also obtained soil variables at their growth locations from www.soilgrids.org (Poggio et al., 2021): pH, CEC, clay content and soil organic matter content.

Chemical analysis

Chemical analysis was performed following Boeschoten et al. (2022). In short, a 1.0 g subsample of heartwood was cut from 3-5 cm of wood and dissolved in 70% HNO_3 by heating in a microwave (CEM Mars 6). In the resulting solutions, element concentrations were determined by ICP-MS (NexION 350D, PerkinElmer). Lichen and/or Rye grass were used as certified reference material to guarantee measurement accuracy. The lowest detection limit per element was calculated as three times the intensity of that element in a blank standard. If elements were found in quantities below the detection limit in more than 200 samples (the equivalent of a quarter of the trees), they were excluded. This resulted in a multi-elemental composition of 41 elements at concentrations varying between 0.001 µg/kg (terbium) and 6 g/kg (potassium) in the wood samples.

Statistical analysis

Species differences across sites

All statistical analyses were performed in R version 4.1.0 (R Core Team, 2021). Differences in the multi-elemental composition of Azobé and Tali were tested by distance based Redundancy Analysis (db-RDA) using the vegan package (Oksanen et al., 2020), based on Chord distances. In addition to this multivariate analysis, we investigated whether individual elemental concentrations differed between the two species by constructing mixed effect models for all elements using site as random factor. These were performed using the lme4 package (Bates et al., 2015).

Random Forest classification of sites and clusters

We performed multiple Random Forest (RF) analyses for site and cluster classification using the randomForest package (Liaw & Wiener, 2002), see Table 3.1 for an overview. The two potential tracing questions were addressed: assignment (*where does this sample most likely originate from?*) and verification (*does this sample come from location X?*). First we tested origin assignment for the three timber species separately (three site-level RF models). Additional to the absolute element concentrations we included elemental ratios, based on the elements that were most important in the RF assignment models. All 1:1 ratios for the top 10 elements of each RF model were tested as additional variables in the RF assignment models, as well as ratios that were previously described to vary geographically (Hevia et al., 2017). Only ratios that improved assignment success were included. Assignment success per model was then calculated as: 100% minus the RF out-of-bag error rate. We also fitted a fourth RF model to assign site of origin combining both Azobé and Tali data.

We continued with a cluster analysis to test similarities between sites. Sites were aggregated into clusters based on divisive clustering using the cluster package (Maechler et al., 2021). Site distances were calculated using Mahanalobis distances, based on standardised mean elements and elemental ratios, using the HDMD package (McFerrin, 2013).

A second set of RF assignment analyses was then performed to assign samples to clusters (cluster-level models). Similar to the site assignment, we developed three species specific RF assignment models (one for each species) and a fourth where Azobé and Tali were combined. As clustering resulted in an imbalanced sample design, we applied SMOTE resampling using the UBL package (Branco et al., 2016) and we set aside 30% of the samples in each iteration before resampling to ensure full independence of the test samples. Assignment success of the final models was evaluated using 50-fold cross-validation, calculated as: 100% minus the percentage of incorrect assigned trees.

Lastly, a third set of RF classification analyses was performed to develop verification models rather than assignment models. Eight RF verification models were developed, one model per cluster for both Azobé and Tali. These models each classified two groups: one group included all reference samples from that cluster and the other group included all other reference samples. Before fitting the models, SMOTE resampling was applied to each reference dataset and 30% was set aside per iteration to ensure full independence of the test samples.

Soil drivers of chemical composition

To test which soil variables were linked to wood chemical variation, we focused on the top five elements or ratio's of the cluster level RF models. Soil data from Soilgrids was centred and scaled (Poggio et al., 2021). For each of the five top variables, a full model was dredged to generate a list of candidate models ranked by AICc using the MuMIn package (Barton, 2020). Full model averaging was performed using candidate models within the top two ΔAICc units (Burnham & Anderson, 2002).

Blind sample collection and classification

Parallel to the main sampling campaign, a second set of samples was arranged by a third party. They collected additional samples of which 46 were anonymized and labelled as Azobé and Tali. No blind samples were taken of Meranti. Each sample fitted one of three categories: 1) sampled within the same reference sites during the main sampling campaign, 2) sampled within 20 km of the reference sites but not from the same campaign, 3) sampled outside the reference sites and/or from species not covered in the reference samples. Multi-element composition was measured in the same way as the reference samples.

Using the RF classification models of the reference dataset, both tracing questions were addressed for all blind samples (origin verification and assignment, explained above). Origin verification was tested based on the eight RF cluster verification models (V1-V8, Table 3.1). Then, before origin assignment, it was tested whether the blind samples should be assigned at all or whether they should be excluded based on their chemical composition using multiple outlier statistics (univariate, PCA, Chord distances and isotree distances). Subsequently, all blind samples were assigned to their most likely cluster (model A4 for both Azobé and Tali) and site (A5 for Azobé and A8 for Tali) of origin (Table 3.1). After verification and assignment, we performed two robustness tests, to take into account uncertainty in the classifications because of (1) a small measurement error or (2) uncertainty in the RF models themselves. If the result of one of the robustness tests was that a blind sample was not classified to a single origin in more than 80% of the cases, the sample was recorded as ‘not assigned’. A detailed description of the blind samples analysis is attached in Supplementary Note S3.1.

3.3 Results

Species and site differences in Central Africa

The two timbers from Central Africa, Azobé and Tali, had a significantly different chemical fingerprint (Fig 3.1). This was also reflected in individual mixed effect models: 35 of the 41 element concentrations were significantly different (Table S3.1). Across our study area, 28 elements were more abundant in Azobé, whereas seven were more abundant in Tali. However, the second db-RDA axis also reflected a significant common origin pattern. Both timbers shared high or low concentrations of certain elements (such as As, Co or Y) at the same sites. This reflects a common spatial pattern across the study area.

Origin verification and assignment models

In Central Africa, the cluster analysis identified eight chemical clusters, including 1-7 sites each (Fig 3.2A).The distances between sites from the same cluster were 20-340 km, whereas some sites that were assigned to different chemical clusters were also as close as 50 km. Furthermore, all chemical clusters of sites were geographically well defined, i.e., there was no spatial mixing of clusters. On Borneo there were five chemical clusters of 1-4 sites each (Fig 3.2B). However, in contrast to Central Africa, on Borneo the sites that were chemically similar were not always

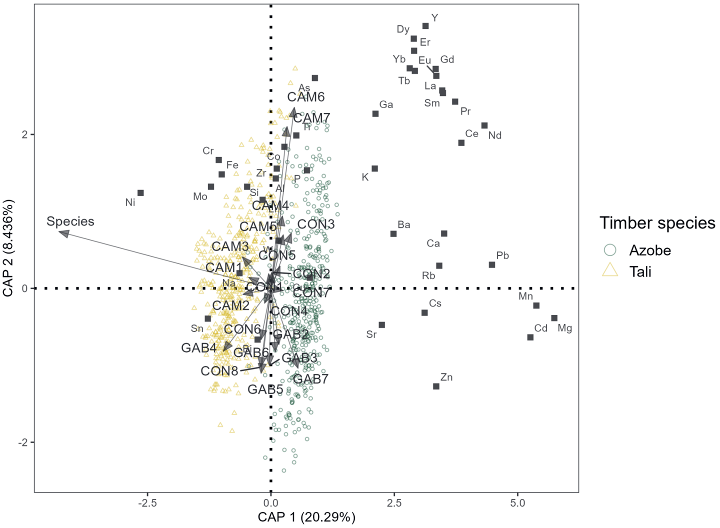


Figure 3.1: Distance based Redundancy Analysis (db-RDA) of heartwood elemental composition of the African timber species Azobé and Tali. Green and yellow dots indicate relative position of the wood samples, based on Chord distances of square root transformed element concentrations, coloured by timber species. Grey squares indicate weight and direction of the element concentrations in the wood. Vectors indicate the weight and the direction of the significant explanatory variables: timber species and sites of origin. Percentages in axis labels denote the percentage of total variance explained by this component. The Indonesian Red Meranti was not included as it does not co-occur with the other two species.

geographically close: one cluster consisted of one site in East, one in Central and two sites in West Kalimantan. Additionally, from three sites in West Kalimantan that were located within 40 km, one was chemically distinct from the other two.

The eight Random Forest verification models for clusters in Central Africa were successful in verifying origin (Table 3.1). They yielded a verification accuracy between 91.6% (green cluster, Republic of the Congo, model V7) and 97.6% (darkblue cluster, west Gabon, model V4, Table 3.1), with 96.0% accuracy on average. Because of the lack of geographical correspondence in the clusters on Borneo, no cluster-level verification was performed. Accuracy was slightly lower in the RF assignment analysis (Table 3.1). The individual model for Red Meranti performed best at assigning the site (model A7, 87.6% correct assignment) and cluster of origin (model A3, 92.6%). As expected, the cluster assignments were better than the site assignments, because the grouping into clusters with high similarity increases the chance of correct classification. Successful site pairwise assignment in the RF models went up with increasing distance in Central Africa but not on Borneo (Fig S3.1).

We tested two more assignment models for the samples from Central Africa, in which we combined Azobé and Tali in one model to check their combined accuracy. The model that

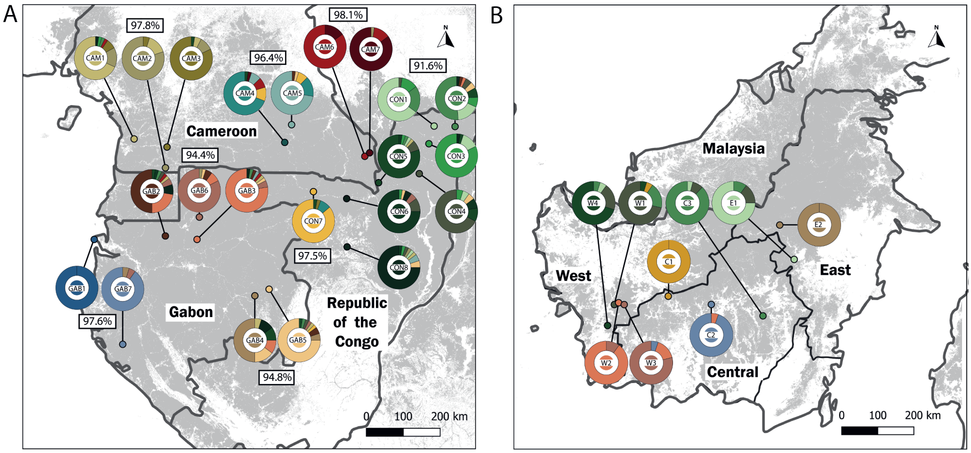


Figure 3.2: Confusion charts of the Random Forest assignment models for (A) the 22 study sites in Central Africa, model A8 in Table 3.1, and (B) the nine study sites on Borneo, model A7. Colours in the inner circle indicate the colour of the unique site. Colours in the outer circle indicate to which site the trees of that location were allocated, clusters from the cluster analysis are grouped and indicated in the different colour shades. No cluster-level verification was performed on Borneo because of the lack of geographical correspondence in the clusters. Primary tropical forest extent from Global Forest Watch in light grey (Turubanova et al., 2018).

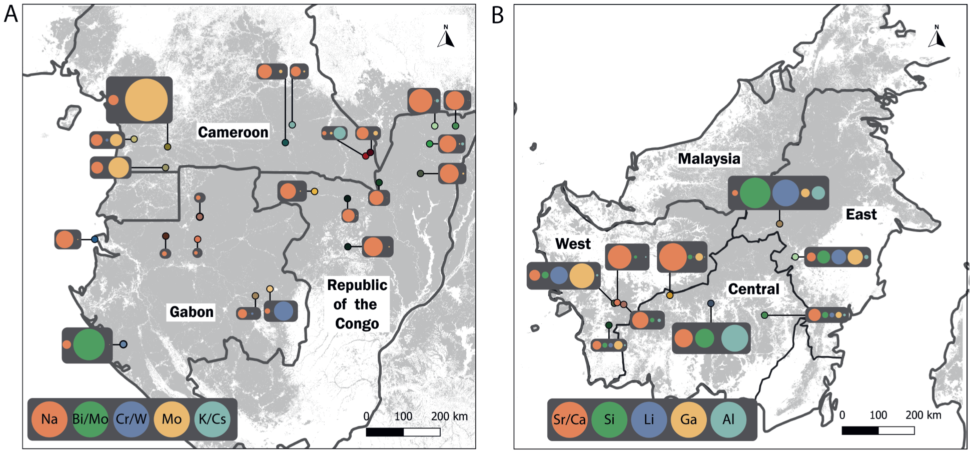


Figure 3.3: Site mean relative wood elemental concentrations of the five most important variables in the Random Forest models A4 (A, Central Africa) and A3 (B, Borneo) respectively. Site mean concentrations are scaled within each element, therefore the size of the circles indicates relative abundance per element at each site. Primary tropical forest extent from Global Forest Watch in light grey (Turubanova et al., 2018).

assigned the eight clusters yielded a correct assignment rate of 88.0% (model A4), which was comparable to the individual models (model A1 and A2). Interestingly, the variable timber species was not included in the most important distinguishing variables of the combined models (model A4 and A8). This again indicates that for certain elements, which are apparently distinguishing, the differences between sites were larger than the species differences.

Table 3.1: Success rates of the Random Forest cluster verification (model V1-V8), cluster assignment (model A1-A4) and site assignment models (model A5-A8). The verification models included both Azobé and Tali reference data and are presented per country (Rep o/t Congo as Republic of the Congo). The assignment models included reference data of each species individually and a combined model with both Azobé and Tali. Clusters are depicted in Fig 3.2. The number of clusters, sites and trees indicate their respective totals as included in the models.

1. Verification models

Country	Cluster-level success			# sites	# trees
Cameroon	(V1)	Khaki	97.8%	3	120
	(V2)	Lightblue	96.4%	2	79
	(V3)	Red	98.1%	2	79
Gabon	(V4)	Darkblue	97.6%	2	40
	(V5)	Orange	94.4%	3	120
	(V6)	Lightbrown	94.8%	2	60
Rep o/t Congo	(V7)	Green	91.6%	7	280
	(V8)	Yellow	97.5%	1	36
Average			96.0%		

2. Assignment models

	Cluster-level success		# clusters	Site-level success		# sites	# trees
Azobé	(A1)	85.7%	8	(A5)	80.1%	21	418
Tali	(A2)	87.2%	7	(A6)	68.8%	20	394
Red Meranti	(A3)	92.6%	5	(A7)	87.6%	9	179
Azobé and Tali	(A4)	88.0%	8	(A8)	74.4%	22	812

Soil chemistry and wood elemental concentrations across the study area

In the RF assignment models, both essential and non-essential elements as well as some ratios were most important for site classifications (Fig 3.3). There was no overlap in the set of key elements in Central Africa and Borneo. All multiple regression models associated wood element concentrations with soil variables but model fit was low overall (av.adj. R^2 between 0.11 and 0.47, Fig 3.4). The low model fit likely results from the variation of wood chemical composition within a site, which is not covered at the same resolution in the gridded soil data. As the soil data originates from interpolated maps, small-scale variation is not well represented in the explanatory variables. Therefore the models mostly describe the large-scale patterns across the study areas.

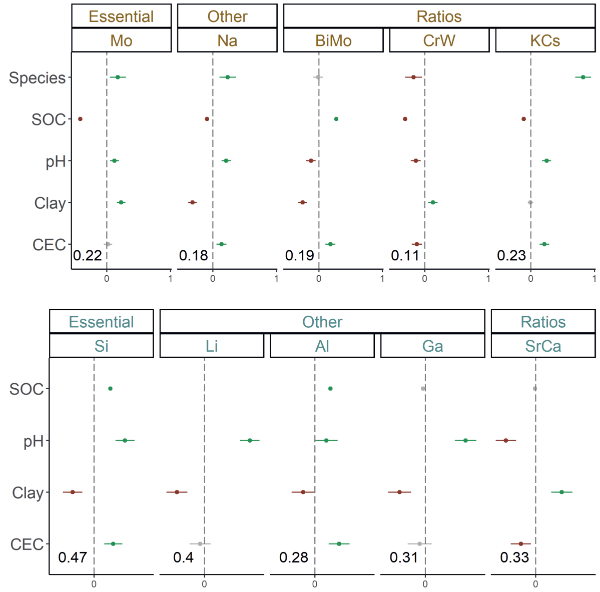


Figure 3.4: Multiple regression models of the five most important elements in the Random Forest assignment models in (A) Central Africa, model A8, and (B) Borneo, model A7. Element concentration was tested as a function of: soil organic carbon content (g/kg), pH, clay content (g/kg) and cation exchange capacity (mmol(c)/kg; all from Poggio et al., 2021), plus timber species (in Central Africa only). Coefficient plots include slope estimates (mean and 95% CI), averaged across top-ranked models. The CI of the red (negative) and green (positive) circles does not overlap with zero and therefore denotes a strong effect. A positive species effect indicates a higher value in Tali compared to Azobé. The grey circles were included in the top models but their CI overlaps with zero, therefore they do not indicate a strong effect on the response variable. Numbers in the panels represent av. adj. R^2 of the best models.

Blind sample verification and assignment

Regarding the first question of verification, 70% of the blind samples was included in the correct cluster. Only five samples were included in the incorrect cluster (false positives). From these five, four samples were from a different species (Fig 3.5A and Table S3.2) and the fifth was a Tali sample that was included in the correct cluster but also in a neighbouring cluster, both in Cameroon. However, the incorrect exclusion error of the verification analysis was higher (false negatives): nine of the 36 blind samples originating within the reference clusters were excluded from their own cluster. The cluster in south-west Cameroon stood out as four out of the five samples (80%) were incorrectly excluded from that cluster, whereas in the other clusters a maximum of 30% (3 out of 10) of the blind samples was incorrectly excluded, with an average of 14% of incorrect exclusions per cluster.

As for the second question, blind sample assignment, the first step of exclusion went well for the blind samples that originated from within the sites with reference samples (category 1; correct exclusion of 90% for Azobé and 77% for Tali; Table S3.2). However, it did not exclude samples from outside the study area or from different species well, seven out of 10 were not excluded (category 3). Furthermore, about half of the samples (6/13) that were collected close to the reference sites (<20 km; category 2) were excluded based on their chemical composition even though they should have been included. After exclusion, assignment to cluster level was of similar success as the RF self-assignment: 72% of the blind samples was correctly assigned to its cluster of origin (Fig 3.5B and Table 3.1). Similar to the verification analysis, most incorrect assignments were from the cluster in south-west Cameroon: none of the five samples was correctly assigned. Furthermore, assignment success at the site level was considerably lower compared to that at cluster level: only 30% of the blind samples was assigned to the correct site of origin.

3.4 Discussion

What drives variation in timber multi-elemental composition?

This study provides insights into the drivers of wood multi-elemental composition for multiple species and continents. As expected, co-occurring timber species differed in their chemical composition (Fig 3.1 and S3.1, Amais et al., 2021). Nevertheless, they also shared a spatial pattern for many essential and non-essential elements (Fig 3.1). This overlapping spatial variation was reflected in the combined RF assignment models: their assignment success was similar to the individual models (Table 3.1) and timber species was not amongst the most important variables.

In tracing studies on other commodities, elemental composition was mostly associated with soil type and soil chemistry (Baroni et al., 2015; Wang et al., 2020), but studies investigating this are limited. In wood, the important variables for assignment in the Random Forest models were associated with soil physical and chemical properties (Fig 3.4). The direction of the effects mostly reflected the buffer capacity of the soils: a positive effect of pH and cation exchange capacity on the wood elemental concentrations reflects that chemical concentrations in wood are higher on soils with a high nutrient retention capacity (Fig 3.4). These results underline that soil chemistry is an important driver of wood elemental composition (Boeschoten et al., 2022).

What is the performance of multi-elemental analysis for timber tracing?

The Random Forest classification models confirmed that there is potential for multi-elemental tracing of timber (Fig 3.2). In Central Africa, distinct chemical patterns were found in two timber species (Azobé and Tali) within the three most important timber exporting countries.

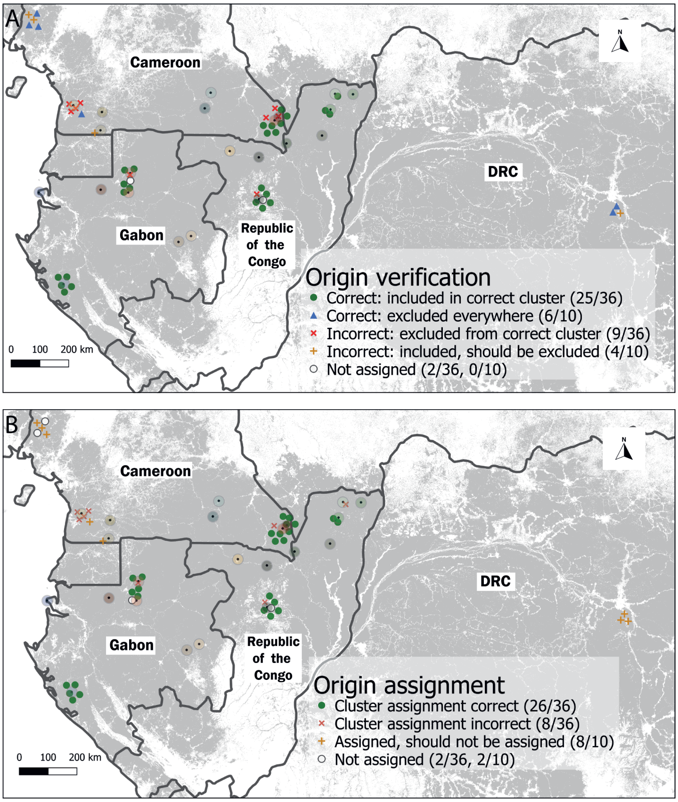


Figure 3.5: Blind sample origin and results of (A) the sample origin verification (does this sample come from cluster X?) and (B) origin assignment (where does this sample come from?), both at the cluster level. To avoid overlap, blind sample symbols were scattered around their site of origin and this does not reflect distance to that site. Origin verification was based on the eight RF models per cluster (models V1-V8, Table 3.1). Origin assignment was based on RF model A4. The total number of samples that was correct or incorrect is indicated in the legend, compared to the total per group: 36 samples originated from <20 km of a site (category 1+2), 10 blind samples originated from >20 km of a site or were another species (category 3). For percentages per species and blind sample category see Table S3.2. Primary tropical forest extent from Global Forest Watch in light grey (Turubanova et al., 2018).

The scale at which tracing performed best was not the site level (assignment accuracy of 74.4%) but the cluster level (assignment accuracy of 88%), which were between 50 and 800 km apart (Fig 3.2 and S3.1). On Borneo, site-level assignment accuracy was already high (87.6%), with distinct chemical patterns across provinces. Even though accuracy increased at the cluster level on Borneo as well (92.6%), these clusters overlapped geographically and are thus unlikely to improve tracing results. This difference between Central Africa and Borneo could arise from the spatial scales at which soil characteristics differ, which is especially clear when comparing the soil types in our study areas. Whereas our study area in Central Africa almost exclusively consists of ferralsol, the area on Borneo consists largely of nitisol which alternates with cambisol and ferralsol even at 10 kms (Poggio et al., 2021).

To put elemental tracing in perspective, we compare its performance to that of other methods, such as genetics and stable isotope ratios. Genetic studies at the population level reported self-assignment success rates between 55% (Ng et al., 2017) and 88% (Chaves et al., 2018), which is comparable to the site level accuracy of 68.0% - 87.6% presented here (Table 3.1). For a Red Meranti species, genetic site-level assignment accuracy in Western Malaysia was lower than our chemical site-level assignment accuracy of Red Meranti on Indonesian Borneo (60.60% vs 87.6%, Ng et al., 2017). Overall, genetic studies that were applied at the same scales reported similar success rates to this study (Low et al., 2022). Compared to the genetic literature, much less work has been done on other chemical methods for tracing purposes. Site-level assignment success based on stable isotope ratios varies between studies, mostly illustrating that the method has more potential at large geographical scales such as the state or country level (Vlam et al., 2018b; Paredes-Villanueva et al., 2022; Watkinson et al., 2020). Therefore our results suggest a higher potential for elemental tracing compared to stable isotope ratios at finer spatial scales.

In addition to performance of the RF classification models, the blind sample tests in Central Africa provided a proof of principle for multi-elemental tracing. Verification accuracy was 72% (does this sample come from cluster X?) and assignment accuracy was 70% (where does this sample come from?; Table S2). Yet, the blind sample exclusion showed that methods to identify samples from outside the area or from species not covered by the reference database had a low success rate (63%, Table S3.2). Additionally, the 25% of false negatives in the blind sample verification analysis (exclusion of timber that should be included; Table S3.2) are especially sensitive as they can lead to erroneous accusations of timber fraud. Better exclusion criteria are needed, which can be achieved by increased sampling effort but also by combining it with other methods such as wood anatomy for species identification.

Blind sample test results based on other methods are limited, but those available showed a comparable accuracy. In Central Africa, higher successes were reported for genetic tracing of Tali (92%, Vlam et al., 2018b) as well as when assigning batches of samples of *E. cylindricum* (86%, Jolivet and Degen, 2012). Assignment of those same Tali blind samples based on stable isotope ratios was unsuccessful however (Vlam et al., 2018b). Other studies verifying blind sample origin based on stable isotope ratios had mixed results. The predicted area of origin of oak samples in the US encompassed the true origin for 78% of the blind samples (Watkinson et al., 2020). Furthermore, two investigations on tropical timbers reported a country-level verification accuracy of 87% (Förstel et al., 2011) and 50%-70% (Degen et al., 2015), and 20%-50% (Degen et al., 2015) at the region-level. Based on those comparisons, multi-elemental tracing of timber is a novel method that seems to achieve sub-country level origin identification at an accuracy that ranges between isotopic and genetic analyses.

The next steps for timber tracing with multi-element analysis

Based on these results we identified key points to be addressed in order to make multi-element analysis operational. Origin verification analysis must be improved, specifically the exclusion of samples from a specific origin. This can be achieved by increased sampling within clusters, to better cover the spread in chemical composition per cluster, as well as by sampling additional trees from different species and origins, to improve cluster specificity. New opportunities could also emerge with the addition of more species to the reference dataset. With site differences larger than differences between species (Fig 3.1), data collection efforts could be reduced by combining the information of multiple species into a single reference dataset. This is a great advantage over other tracing methods that require a species-specific reference set, such as genetic tracing methods, but should be tested with more species.

Lastly, the statistical analyses can be developed further to enable large-scale application of the method. Adding samples to the reference database in a more scattered sampling design allows for interpolating spatial analyses, similar to isoscapes in stable isotope tracing (Bowen, 2010). That would open up opportunities to move forward from site- or cluster-oriented verification of a blind sample towards identifying a potential region of origin.

3.5 Conclusions

This study illustrates a high potential for multi-element analysis to be developed into a timber tracing tool, because (1) wood origin could be identified at a sub-country scale in Central Africa and on Borneo (from 50-800 km), (2) overall cluster verification success is high, even though it varied per cluster (96.0% in the RF self-assignment and 70% for the blind samples) and (3) the spatial variation in Central Africa included non-species specific patterns. This is a first indication that it may be possible to combine multiple species in one reference dataset from which to build classification models. Furthermore, wood chemical composition was associated with soil chemistry, which can help to identify regions where this method may be most promising. However, the methods to exclude samples originating from outside of the reference dataset must be improved to apply elemental tracing in a forensic context. These findings should be confirmed for other sites and species in order to implement multi-element analysis as a timber tracing tool.

3.6 Acknowledgments

This study was supported by the Dutch Research Council (NWO-TTW-OTP-16427). Additional fieldwork support was received from the Alberta Mennega Foundation and from World Forest ID. We thank all the collaborating timber companies and their field teams for facilitating the fieldwork, colleagues at our partner institutes IPB University, University of Dschang, Marien Ngouabi Univeristy, IRET/CENAREST and the National Herbarium of Gabon for help in arranging fieldwork, our colleagues of WFSR for assistance in the laboratory work and our colleagues from the Forest Ecology and Forest Management writing group for feedback on the manuscript. We also thank the Indonesian Ministry of Environment and Forestry for the permission to access the plant sample genetic resources for research purposes with the permit number SK Dirjen: SK.92/KSDAE/ SET.3/KSA.2/4/2022.

Supporting information

S3.1 Blind sample verification and assignment

Blind sample verification

For the blind samples we first tested the origin verification question at the cluster level. All blind samples were classified by the eight verification models (models V1-V8, Table 3.1) to test whether they were included in their cluster of origin as well as excluded from all the others. To take into account the uncertainty in the RF models because of either a small measurement error or uncertainty in the RF models themselves, we subsequently performed two robustness tests. As a first robustness test we randomly constructed 100 new synthetic samples based on every blind sample. For this, the element concentrations of the measured blind sample were taken as mean and for every element a random concentration was drawn from a normal distribution around that mean, with a standard deviation of 10% of the mean. The 100 synthetic blind samples were then classified with the same eight RF models per timber species. If more than 80% of the 100 synthetic samples were classified to the same cluster, this cluster was reported for a particular blind sample. However, if not more than 80% of the samples shared a single outcome, that blind sample was recorded as 'not assigned', because the chances of an incorrect classification were considered too large. For the second robustness test we fitted the eight RF models 100 times, to test for uncertainty in the models themselves. The original 50 blind samples were then classified with every model. Similar to the first robustness test, if more than 80% of the models classified a sample to the same cluster, this cluster was recorded. However, if RF models had different outcomes and none occurred in more than 80% of the models, the blind sample was recorded as 'not assigned'.

Blind sample assignment

To address the other tracing question, origin assignment, we first assessed whether the blind sample should be assigned at all or whether it should be excluded based on its chemical composition. Four steps for outlier detection were applied to identify blind samples that should be excluded. By using four different statistical methods we aimed to reduce the probability of incorrectly excluding a sample, as defining a cutoff value per method can be challenging and this can put too much weight on a specific statistical method. Step 1 was a univariate step, where a sample was labelled as outlier if the chance of finding a specific elemental concentration was below 0.998 based on a normal distribution. Elemental concentrations were box-cox transformed if not normal using the MASS package (Venables & Ripley, 2002). The other three were multivariate steps. In step 2 a PCA was performed using the FactoMineR package (Lê et al., 2008). Outliers were identified visually based on the first two components if they were clearly outside of the cloud of reference sample points. In step 3, a Chord distance analysis was performed using the vegan package (Oksanen et al., 2020). Outliers were identified based on their distance to all reference trees per site, compared to the distances between trees from that site itself. The percentage of trees per site was calculated that had a smaller distance to all trees at that site than to the blind sample. Those 22 site-specific percentages (so % of trees within a site that are chemically more similar to each other than to the blind sample) were averaged and if this was more than 80%, the sample was indicated as outlier. Lastly, in step 4, an isolation tree analysis was performed using the isotree package (Cortes, 2022), where outliers are assumed to have a short branching distance in a classification tree because they can be identified based on few elements only. The isolation tree was developed with the reference data and blind samples were labelled as outlier if their outlier score was >0.5 (Liu et al., 2008). We then excluded blind samples if either two or more of the four steps had identified that sample as outlier or if the value of that sample in one of the four steps was more than 10% above the cutoff value and was thus extremely high.

After exclusion, all blind samples were assigned to their most likely origin using the RF assignment models with the highest accuracy (Table 3.1). For the Azobé blind samples these were the combined Azobé and Tali model for cluster-level assignment and the Azobé only model for site-level assignment (model A4 and A5). For the Tali blind samples these were both the combined models including Azobé and Tali (model A4 and A8). All blind samples were assigned to their most likely cluster and site.

We then applied two types of robustness tests to test the assignment accuracy, in the same way as for the blind sample verification. Using the same synthetic blind samples that were constructed for the verification analysis, we assigned 100 synthetic samples to cluster and site. If not more than 80% of the samples shared a single outcome, that blind sample was recorded as ‘not assigned’. Furthermore, the Random Forest assignment model was run 100 times and if outcome overlapped in less than 80% of the RF models, that blind sample was also recorded as ‘not assigned’.

Table S3.1: Mixed effect model results, showing the model coefficient of timber species (in mg/kg), the coefficient significance (*<0.05, **<0.01, ***<0.001) and element type, ordered by element group. A positive coefficient indicates a higher concentration in Tali, a negative coefficient indicates a higher concentration in Azobé.

Element group	Element	Coefficient	Significance	Element type
Essential elements	K	-9.69	***	Alkali Metal
	Mg	-8.665	***	Alkaline Earth Metal
	Ca	-5.31	***	Alkaline Earth Metal
	Si	0.226	ns	Metalloid
	P	-0.677	***	Nonmetal
	Mn	-4.911	***	Transition Metal
	Fe	0.551	*	Transition Metal
	Ni	0.374	***	Transition Metal
	Cu	-0.126	***	Transition Metal
	Zn	-1.434	***	Transition Metal
	Mo	0.044	***	Transition Metal
Rare earth elements	La	-0.164	***	Lanthanide
	Ce	-0.256	***	Lanthanide
	Pr	-0.091	***	Lanthanide
	Nd	-0.216	***	Lanthanide
	Sm	-0.085	***	Lanthanide
	Eu	-0.036	***	Lanthanide
	Gd	-0.077	***	Lanthanide
	Tb	-0.019	***	Lanthanide
	Dy	-0.041	***	Lanthanide
	Er	-0.022	***	Lanthanide
	Yb	-0.016	***	Lanthanide
	Y	-0.082	***	Transition Metal
Other elements	Li	0.003	ns	Alkali Metal
	Na	0.688	***	Alkali Metal
	Rb	-0.675	***	Alkali Metal
	Cs	-0.092	***	Alkali Metal
	Sr	-0.742	***	Alkaline Earth Metal
	Ba	-0.591	***	Alkaline Earth Metal
	Al	-0.356	ns	Metal
	Ga	-0.055	***	Metal
	Sn	0.019	***	Metal
	Pb	-0.42	***	Metal
	Bi	0.003	ns	Metal
	As	-0.014	***	Metalloid
	Ti	-0.086	**	Transition Metal
	Cr	0.044	**	Transition Metal
	Co	-0.008	ns	Transition Metal
	Zr	-0.006	ns	Transition Metal
	Cd	-0.2	***	Transition Metal
	W	0.021	*	Transition Metal

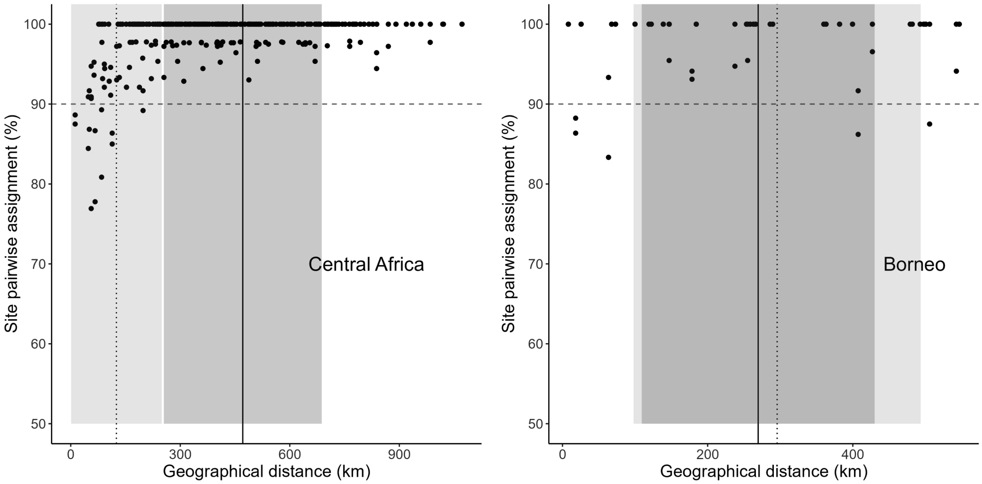


Figure S3.1: Successful site pairwise assignment (%) vs geographical distance (km) between every combination of sites (dots) in Central Africa and Borneo. Successful site pairwise assignment was based on Random Forest models A4 (Central Africa) and A3 (Borneo; Table 3.1), calculated as: 100% minus the percentage of trees that was incorrectly assigned to another site. Every combination of sites is included twice as the % of trees of site A that was assigned to site B and vice versa, the dots at y=100% indicate no trees were assigned incorrectly in that pair of sites. Vertical lines and shaded areas indicate geographical distances (mean and sd) between sites within one cluster (dotted line, light grey) and between sites of separate clusters (solid line, dark grey). Dotted horizontal line indicates 90% successful assignment between a combination of sites (not more than 10% of trees was assigned to the non-origin site).

Table S3.2: Blind sample verification (1), exclusion (2) and assignment success (3). Blind sample origin is depicted in Fig 3.5. Blind samples fit one of three categories: 1 = within our reference sites; 2 = within 20 km of the reference sites but from a different sampling campaign and 3 = more than 20 km away and/or another species.

1. Verification of the cluster of origin

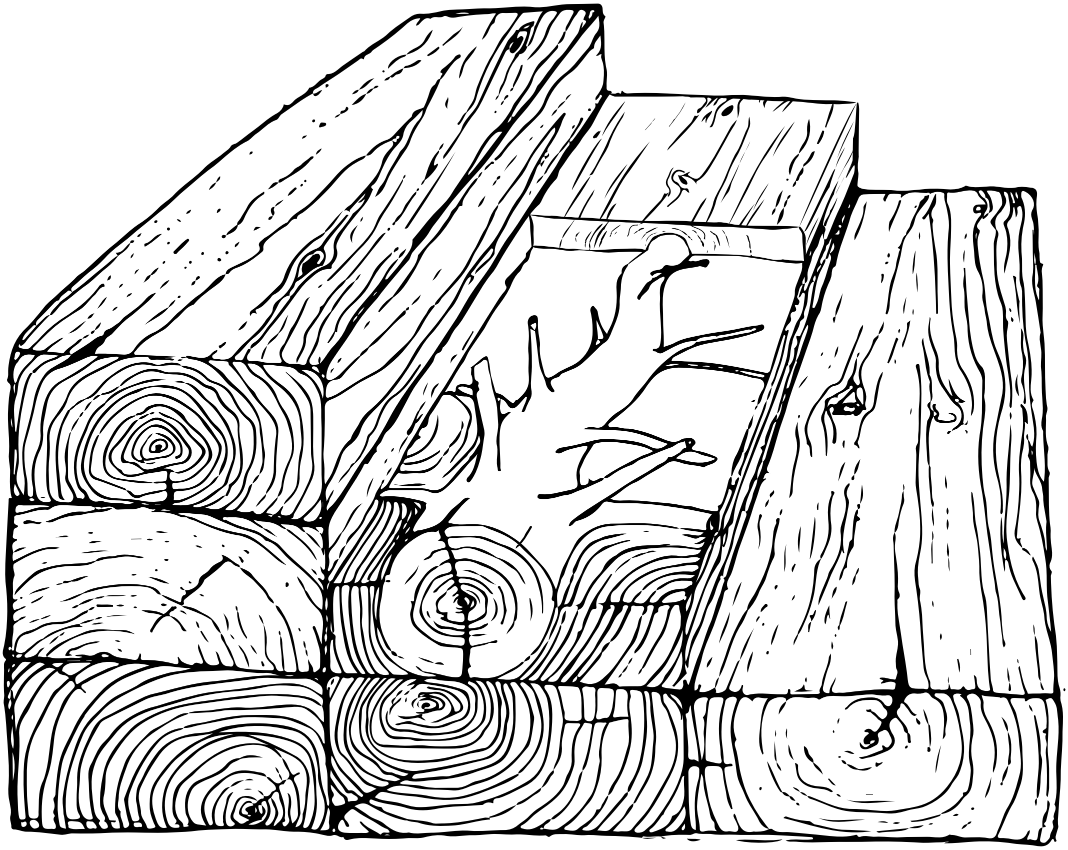
		Cluster						
Species	Blind sample category	Correct: included in correct cluster	Incorrect: excluded from correct cluster	Not assigned	Correct: excluded everywhere	Incorrect: included but should be excluded	Not assigned	Total # trees
Azobé	1	80%	20%	0%				10
Tali	1	77%	15%	8%				13
Azobé	2	38%	50%	12%				8
Tali	2	80%	20%	0%				5
Azobé	3				80%	20%	0%	5
Tali	3				40%	60%	0%	5
% of total		70% (25)	25% (9)	5% (2)	60% (6)	40% (4)	0% (0)	46

2. Exclusion from the reference dataset

Species	Blind sample category	Correct exclusion	Incorrect, should be included	Incorrect, should be excluded	Total # trees
Azobé	1	90%	10%		10
Tali	1	77%	23%		13
Azobé	2	62%	38%		8
Tali	2	40%	60%		5
Azobé	3	0%		100%	5
Tali	3	60%		40%	5
% of total		63% (29)	22% (10)	15% (7)	46

3. Assignment to the most likely cluster of origin

		Cluster			*Site*			
Species	Blind sample category	Correct cluster	Incorrect cluster	Not assigned cluster	Correct site	Incorrect site	Not assigned site	Total # trees
Azobé	1	100%	0%	10%	10%	60%	30%	10
Tali	1	77%	15%	8%	62%	31%	8%	13
Azobé	2	38%	32%	12%	0%	75%	25%	8
Tali	2	60%	20%	0%	40%	60%	20%	5
Azobé	3							5
Tali	3							5
% of total		72% (26)	22% (8)	6% (2)	30% (11)	53% (19)	17% (6)	46



Chapter 4

Evaluating the potential of stable isotopes for coarse-scale timber tracing: a first pantropical isoscape for wood δ^{18}O

Mart Vlam, Laura E. Boeschoten, Arnoud Boom, Kathelyn Paredes-Villanueva, Peter van der Sleen, Pieter Zuidma

Invited for resubmission at Biological Conservation

Abstract

Independent verification of timber origin is needed to enforce legislation aimed at combatting illegal tropical timber trade. A potentially effective technique is stable isotope analysis of wood, but the scarcity of reference data currently hampers its operationalization. This can be overcome by creating isoscapes, predictive maps of isotopic signature based on relations between isotope values and climatic data. Here we develop the first pantropical isoscape for wood $\delta^{18}O$ and assess its potential for timber tracing.

We established a pantropical database of 2514 $\delta^{18}O$ measurements from 27 commercially harvested tree species. We tested effects of $\delta^{18}O$ in rainfall, potential evapotranspiration (PET), temperature and precipitation on wood $\delta^{18}O$ and used these to develop an isoscape based on Quantile Regression Forests. The isoscape was created at 0.5° resolution and 27 species-specific isoscapes were created by extracting areas based on species distribution models.

The Quantile Regression Forest model explained 82% of pantropical variation in wood $\delta^{18}O$, the most important variables were $\delta^{18}O$ in precipitation, temperature annual range and PET. The pantropical isoscape had a wood $\delta^{18}O$ range of 3.9‰ (10th-90th percentile), showed increasing isotopic differences for geographic distances up to 2000 km and revealed pronounced isotopic gradients for South America and Asia.

Our isoscape provides a proof of concept of using $\delta^{18}O$ to verify timber origin at large spatial scales, though precision remains low. Necessary future steps include improving regional representation, rigorous testing of species differences and conducting blind sample tests. We conclude that wood $\delta^{18}O$ measurements have the potential to be used for coarse scale forensic tracing.

Keywords

geographic origin, isoscape, stable isotope, timber forensics, tropical timber

4.1 Introduction

An estimated 50-90 % of the traded volume of timber from major producing tropical countries is illegally sourced (Hoare & Uehara, 2022), causing substantial environmental and socio-economic damage. To combat illegal timber trade, legislation has been implemented that prohibits the production, selling and import of illegally sourced timber. Enforcement of this legislation is based on the verification of required permits along the timber value chain, but this paper trail is sensitive to fraud. Forensic techniques that allow verifying the origin of tropical timber are therefore needed as an independent control mechanism (Dormontt et al., 2015). Forensic tracing tools for tropical timber are in high demand by authorities, but also increasingly by timber producers, timber traders and certification schemes (Boekhout van Solinge et al., 2016; Gasson et al., 2021).

Two main types of fraud in tropical timber trade exist: false reporting of traded species and false reporting of geographic origin. Forensic techniques to detect species fraud are well-developed and widely applied in enforcement (Gasson, 2011). In contrast, forensic techniques to detect false claims of geographic origin of tropical timber – based on chemical or genetic wood characteristics – are upcoming but still under development (Low et al., 2022; Gasson et al., 2021). One of the promising techniques is the measurement of stable isotopes in wood tissue. Relative abundance of stable isotopes of δ^{13}C, δ^{18}O, δ^{2}H, δ^{15}N, δ^{34}S and δ^{87}S in wood of tropical trees are associated with climatic, edaphic or other environmental conditions experienced by trees (van der Sleen et al., 2017; Watkinson et al., 2020; van der Sleen et al., 2022). As a result, these isotopic signatures may contain information on the geographic origin of traded timber and thus assist in verifying the claimed origin of timber (English et al., 2001; Gori et al., 2018; Watkinson et al., 2022a).

In recent years, isotopic tracing has been successfully used to verify the geographic origin of important plant-based commodities produced in temperate (West et al., 2007) and tropical climate zones (Ehleringer et al., 2000; Driscoll et al., 2020; Wang et al., 2020). It has also been applied to wood from temperate (Horacek et al., 2009; Kagawa & Leavitt, 2010; Gori et al., 2018; Watkinson et al., 2020) and tropical trees (Vlam et al., 2018b; Watkinson et al., 2022a). Overall, these results suggest a potential for isotopic timber tracing at large spatial scales. Yet, the application of isotopic tracing of tropical timber is strongly limited by the scarcity of reference data and limited geographic coverage of existing datasets (Gasson et al., 2021).

Two recent developments may help overcoming these limitations. First, advances in isotope theory have resulted in a much better understanding of the processes and mechanisms responsible for the discrimination and fractionation of stable isotopes in trees (van der Sleen et al., 2017; Ehleringer et al., 2010). As a result, the biotic and abiotic conditions influencing isotopic signatures in wood of tropical trees are now better understood. Second, the availability of global data products with climatic variables, soil conditions and species distributions has increased rapidly in recent years. These products contain information on environmental conditions, such as precipitation, drought indices, δ^{18}O in rainfall, δ^{13}C in plant material and N deposition, which are known to influence isotopic fractionation in wood. Together, these developments may overcome the scarcity of reference data, improve predictions of isotopic values, constrain expected isotopic signatures and help creating spatial models of isotopic signatures: isoscapes.

Here we present a unique database of isotopic measurements in tropical tree species for 27 species in 20 countries, which we link to global data products, to evaluate the potential of isotopic tracing for major tropical timbers. We focus on δ^{18}O, which has shown potential to trace the geographic origin of tropical commodities and temperate timber. Our main question is whether wood δ^{18}O values can be used to trace back the origin of tropical timber at a coarse spatial scale, i.e., at country level.

First we tested if wood δ^{18}O is determined by the isotopic signature of rainfall (root uptake), total rainfall ("amount effect") and drought (evapotranspiration). This is based on physiological insights into the main determinants of δ^{18}O in wood of tropical trees (Fig 4.1; Barbour (2007)): δ^{18}O in source water and δ^{18}O enrichment through transpiration in leaves. During water uptake by roots, no fractionation of oxygen isotopes occurs, thus water in the xylem has the same δ^{18}O as the soil water surrounding the roots, i.e. the source water (Dawson & Ehleringer, 1991). Although CO_2 is the substrate for photosynthesis, oxygen in CO_2 exchanges completely with that in leaf water prior to photosynthesis (Epstein et al., 1976). The ^{18}O/^{16}O ratio in source water is thus an essential determinant of wood ^{18}O/^{16}O ratio (Epstein et al., 1976). Additionally, the enrichment of ^{18}O in leaf water due to transpiration is a crucial process determining the ^{18}O/^{16}O ratio in wood. During transpiration, leaf water is enriched with ^{18}O because water molecules containing ^{18}O evaporate at a lower rate (Barbour, 2007). The degree to which leaf water and – as a result – wood cellulose is enriched with ^{18}O depends on factors driving transpiration: potential evapotranspiration (PET), temperature and soil water availability. It is therefore expected that PET and temperature have a positive and water availability a negative effect on tree-ring δ^{18}O values (Roden & Ehleringer, 2000; Cernusak et al., 2016).

As a second step, we used the resulting statistical model to create a δ^{18}O isoscape, which predicts and maps the range and distribution of wood δ^{18}O-isotopic values expected within the tropical forest biome. Finally, we used the δ^{18}O isoscape to evaluate the spatial scale at which δ^{18}O-isotopic tracing is expected to differentiate wood samples within species.

4.2 Methods

Data collection

To set up a database of δ^{18}O measurements in tropical timber species we collected wood samples from two sources: (1) live trees in natural old-growth forests (wood chunks or cores) and (2) xylarium and herbarium specimens. The largest share of samples (>95%) were obtained from live trees. Our database also includes δ^{18}O measurements from previously published work by our lab (van der Sleen et al., 2014; Van der Sleen et al., 2015; Paredes-Villanueva et al., 2022). We did not add published wood isotope data collected by others, because differences

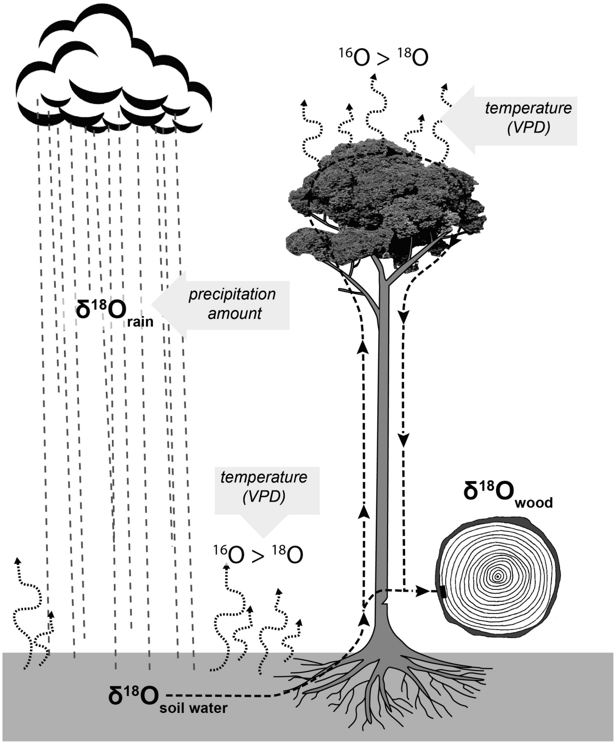


Figure 4.1: Schematic representation of main processes and environmental drivers of $\delta^{18}O$ signature in wood of tropical trees. The first process is the uptake of soil water by trees with a particular $\delta^{18}O$ signature, which is mainly determined by the $\delta^{18}O$ signature of rain water, and thus also on precipitation amount. The second process is transpiration, which strongly enriches leaf water with ^{18}O and thus increases $\delta^{18}O$ in wood.

in sample preparation, e.g. cellulose extraction, prohibit comparison. Our database consisted of 2514 $\delta^{18}O$ measurements obtained from 643 trees, belonging to 27 species and 10 families (Table 4.1; full overview in Table S4.1). The timber of all 27 species is traded internationally, including major timbers such as Azobé (*Lophira alata* Banks ex Gaertn.f.), Okan (*Cylicodiscus gabunensis* Harms.) and Meranti (*Shorea* spp). Our aim has been to cover a large number of traded timber species across tropical forests from all continents and not to sample species at equal proportions. The distribution of samples per species is therefore highly variable (1-178 trees), with 7 species represented by only one sample (Table S4.1).

Isotope measurements

From each wood sample, about 25 mg of wood was obtained using a sharp knife. In all cases measurements per tree were at least duplicated to reduce the potential effect of outliers. Next, cellulose extraction was performed following a modified version of the Jayme-Wise protocol

Table 4.1: Summary of the database of tropical wood δ^{18}O values used in this study. Columns show the number of countries from which trees were included, the number of unique trees, total number of wood samples taken across all trees, and the numbers of species, genera and plant families per continent included in the study. Minimum and maximum latitudes are also shown.

Continent	Countries	Trees	Samples	Species	Genera	Families	Latitudinal range (°)	
Africa	8	449	1283	15	9	6	-5.71	7.61
South-East Asia and Australia	2	10	452	5	2	2	-2.00	15.64
Central and South America	10	184	758	7	2	2	-21.06	22.00
Total	20	643	2514	27	13	10	-21.06	22.00

(Wieloch et al., 2011). For cellulose extraction procedure and isotopic composition determination, see Vlam et al. (2018b). If multiple isotopic measurements were available from the same tree, e.g. from annual measurements in tree rings, these were first averaged per tree. In all cases, care was taken to include as much as possible wood formed in multiple years, by using for example wood samples with large dimensions.

Temporal and local variation in δ^{18}O wood values

Two sources of variability may induce uncertainty and noise in large-scale wood δ^{18}O patterns: inter-annual variation and local spatial variation. To assess their potential influence, we quantified both types of variability.

First, inter-annual variation in wood δ^{18}O values can be substantial (van der Sleen et al., 2017) and may create uncertainty in spatial patterns of δ^{18}O wood values. Our database contains annual δ^{18}O measurements over several decades for 5-7 trees belonging to three species, spread across continents (Tables S4.1; van der Sleen et al. (2014); Van der Sleen et al. (2015)), and allows estimating the possible implications of temporal variation in δ^{18}O. The effect of temporal variation on the uncertainty in δ^{18}O values decreases with the number of years included in wood samples. To assess the effect of the number of years included, we calculated the 10th-to-90th quantile range of wood δ^{18}O values per tree when those values were averaged over 5, 10 or 20 years. We used this range as a proxy for uncertainty in wood δ^{18}O values.

Second, local spatial variation in wood δ^{18}O values among nearby trees (i.e., at distances less than 10 km) may also induce uncertainty. For six species in our database, wood δ^{18}O values were available from 2-8 trees, sampled at the same forest site, for six species. For each of these species we calculated the range of wood δ^{18}O values across individuals, similar to the approach for temporal variation. If this range is large, local spatial variability will create uncertainty in the generation of large-scale δ^{18}O patterns.

Quantile Regression Forests: creating a pantropical wood δ^{18}O isoscape

We identified relevant environmental co-variables from stable isotope theory and prior studies and obtained those variables at the growth locations of all trees. We then predicted the relationship between wood isotopic composition and δ^{18}O in precipitation, elevation, mean maximum temperature, mean precipitation, mean PET, 19 bioclimatic variables (Fick & Hijmans, 2017), species and genus, by applying Quantile Regression Forests (Meinshausen, 2006). Thus, a total of 24 environmental variables and species and genus were used as explanatory variables. Species and genus were later omitted from the final model, because only those variables were used that contain spatial information to generate maps. Random Forest Regression, the underlying method for this analysis, is not influenced by co-linearity and non-normal data distributions and has been successfully applied for other isoscapes (Brlík et al., 2022). The confidence interval around the predicted median was quantified based on all individual prediction trees, from which the interquantile range was calculated. This provided an estimation of prediction uncertainty. Based on these Quantile Regression Forests, we predicted the spatial distribution of wood δ^{18}O in the study area.

Multiple regression and mixed effect models of wood δ^{18}O measurements

To verify and visualize relations between the three most important variables in the Quantile Regression Forest and wood δ^{18}O, we constructed multiple regression models with these explanatory variables and reported its R^2 value. Next, to assess how predicted wood δ^{18}O values may vary with species identity, we constructed a mixed effect model on a subset of the database that included nine species represented by at least five individuals and sampled at various locations. We only included those species, because the random effects of taxonomic identity could be evaluated. As for the multiple regression, explanatory variables in this model were restricted to the three most important variables of the Quantile Regression Forest. We calculated the standard deviation of the random effect 'species' to obtain an estimate of inter-specific variation in wood δ^{18}O. These statistical analyses were conducted in R (R Core Team 2021), using the lme4, lmerTest and usdm packages (Naimi et al., 2014; Bates et al., 2015; Kuznetsova et al., 2017).

Creating species-specific isoscapes

Using the same pantropical wood δ^{18}O isoscape we created species-specific isoscapes, based on distributional ranges for all 27 species represented in our database (Table S4.1). We first established distribution models for all species using occurrence records (GBIF.org, 2021) and species distribution modelling (Phillips et al., 2006). We checked the geographic range of occurrence records and cropped these based on locations that were improbable climatically or those that most likely represented records outside natural forests (e.g., plantations). Next, to reduce the effect of spatial bias due to local high density of observations, we selected only one record per 0.5° grid cell (the spatial resolution of our isoscape). Species distribution models using Maxent were developed using 19 bioclimatic variables (Worldclim; downloaded at 5-minute resolution) to identify climatic determinants of species occurrences. Next, all grid cells with a probability of occurrence >0.2 were identified to get a species distribution map. A relatively low threshold was chosen, because we did not want uncertainty in species range models to restrict the distributional area and hence provide the impression of a limited potential for isotopic tracing. Finally, the species distribution was overlaid with the pantropical isoscape to get a species-specific wood δ^{18}O isoscape. These analyses were conducted in R, using the dismo and terra packages (Hijmans et al., 2020; Hijmans, 2021b).

Using wood δ^{18}O isoscapes to evaluate the potential for tracing

To evaluate the potential of δ^{18}O for tracing of tropical timbers, we conducted two analyses on the continental and species-specific wood δ^{18}O isoscapes. First, histograms of area, expressed as number of grid cells, were produced to obtain the distribution of δ^{18}O values in the isoscapes. We also calculated the 10 and 90 percentile to quantify the expected range of δ^{18}O values. Second, we created distance matrices based on the isoscapes for δ^{18}O and geographic distance. Thus, for all pairs of cells with a δ^{18}O value, we calculated their geographic distance and the difference in predicted wood δ^{18}O. We then plotted these distances and calculated the general trend between these distances using a Mantel test (999 permutations). A general positive tendency (positive Mantel r) indicates that δ^{18}O differences increase with geographic distance, thus suggesting tracing potential at distances included in the database. Apart from the general tendency, the form of the relation between δ^{18}O differences and geographic distances is also of importance: the geographic range at which the strongest increase in δ^{18}O difference occurs is an indication of the geographic range at which wood δ^{18}O will likely have the highest potential. We therefore also conducted a LOWESS regression to appreciate the tendency of the data. Distance analyses were conducted in R using the terra package; Mantel tests using the vegan package (Oksanen et al., 2020).

4.3 Results

Temporal and local spatial variation in wood δ^{18}O

The wood δ^{18}O values in our database varied between 14.7‰ and 32.6‰, with a mean of 26.2‰, a standard deviation of 2.2‰ and a 10th-to-90th quantile range of 5.5‰ (23.0 to 28.5‰). This interquantile range is large compared to the measurement errors of the mass spectrometer used, which are in the order of 0.2‰.

Our analyses on the implication of interannual variability revealed that estimates of wood δ^{18}O from samples including only a few growth years can be quite uncertain. For instance, the 10th-to-90th quantile range of temporal variability in two of the three tested species may be close to 2‰ (*Cariniana ianeirensis* and *Toona ciliata*; Figure 4.2). Yet, these ranges drop strongly, to values around 0.5‰ when 20-year bulk samples are taken. A second analysis was conducted to quantify local variability in wood δ^{18}O, obtained from trees belonging to the same species and sampled within a 0.1 x 0.1° grid cell. Results show that across species, the 10th-to-90th quantile range of local spatial variability in wood δ^{18}O is typically around 0.7‰, but can be as small as 0.1‰ or or as large as 2.0‰. These analyses show that temporal variability in wood δ^{18}O may easily outnumber that caused by large-scale spatial variability if samples contain wood produced during few years only, and that local variability in wood δ^{18}O can be substantial compared to large-scale spatial variability.

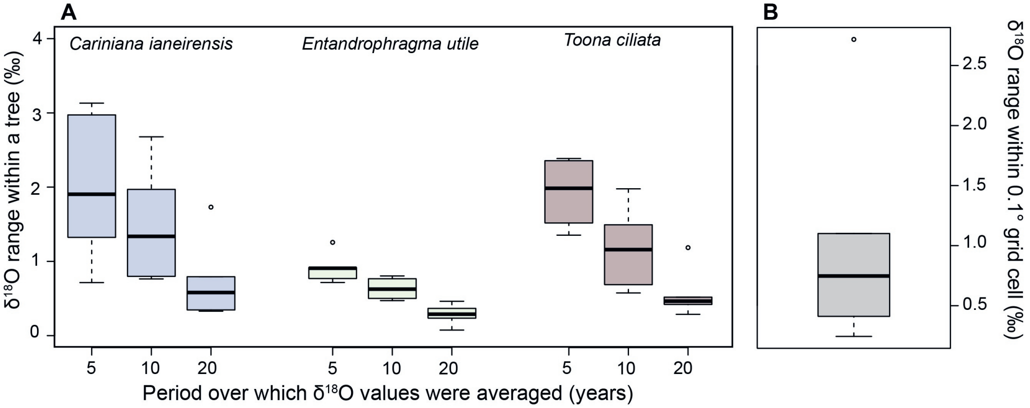


Figure 4.2: Temporal and local spatial variation in tropical wood δ^{18}O values. (A) Inter-annual variability in wood δ^{18}O for tree species from the Americas (light blue), Africa (light green) and Asia (light red). Each boxplot shows the distribution of ranges in wood δ^{18}O obtained for 5-7 trees. The plotted ranges are a proxy of variability in estimates of wood δ^{18}O within a tree when taking bulk wood samples of 5, 10 or 20 years. For each tree, the range was obtained by calculating all possible δ^{18}O values (averaged over 5, 10 or 20 years) and getting the 10th and 90th percentile of these averages. (B) An estimate of the local variability in wood δ^{18}O. Shown is the distribution of wood δ^{18}O ranges (max-min) for trees belonging to the same species and sampled within a 0.1° grid cell.

Determinants of pantropical variation in wood δ^{18}O

Using Quantile Regression Forest, we developed a single pantropical isoscape predicting wood δ^{18}O values. The three most important predictors in this model were mean δ^{18}O in precipitation, temperature annual range (BIO 7) and potential evapotranspiration (PET; Fig S4.1B). This model explained 82% of the variance, with an out of bag prediction error of 1.06‰. When species and genus were included as variables in the Quantile Regression Forest, genus was the second most important variable in the model (Fig S4.1A), indicating a strong species effect in wood δ^{18}O values. We also compared the sign and magnitude of the effect of these three environmental variables in a multiple regression model and found that precipitation δ^{18}O and PET were significant (Fig 4.3). For temperature annual range, the relation was not significant. The mixed effect model explained 58% of the variance (Rscript2 0.58). The mixed effect model for nine well-replicated species revealed a significant random effect for species identity (likelihood ratio test, $p<0.001$), with a standard deviation for species of 1.6‰. Species-level random intercepts ranged from -3.6 to 1.8‰, suggesting additive effects of species on the overall regression model for wood δ^{18}O (Fig S4.2).

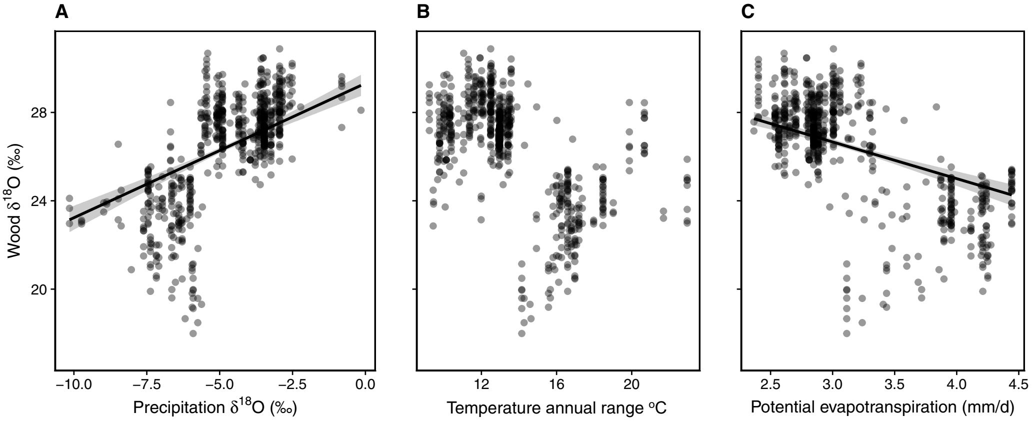


Figure 4.3: Determinants of pantropical variation in wood δ^{18}O. Wood δ^{18}O values are shown as a function of three variables with the highest importance in the Quantile Regression Forest: δ^{18}O in rainwater (A), temperature annual range (BIO7) (B) and potential evapotranspiration (C). Lines represent predicted effects in a multiple regression model and 95% confidence interval.

Pantropical and continental wood δ^{18}O isoscapes

While we found an effect of species on wood δ^{18}O, we chose to produce one pantropical isoscape rather than species-specific isoscapes because species effects could only be estimated for a subset of our database and the sample sizes per species were generally very low. Thus, we used the environmental variables only (Fig S4.3) to create a pantropical wood δ^{18}O isoscape across all species (median and inte range, Fig S4.4), using Quantile Regression Forest.

Within the tropical moist forest range, wood δ^{18}O in our pantropical isoscape varies between 19.6 and 29.4‰, with a mean of 25.9‰ and a 10th-90th percentile range of 3.9‰ (23.9 – 27.8‰). Interquantile ranges increase up to 10‰ in South America and Asia. Across the Andes in South America for example, δ^{18}O in precipitation reaches -20‰ (Fig S4.3A).

The isoscapes of Central and South America and Asia (including Australia) were relatively comparable in spread and magnitude of predicted wood δ^{18}O values (Fig 4.4A, C). The 10th-90th percentile ranges in wood δ^{18}O for these continents were 4.0‰ and 2.8‰, respectively. The isoscape of Africa had a somewhat reduced range in δ^{18}O (10th-90th percentile range of 1.9‰), with comparatively higher values (Fig 4.4B). In South America, a continental-scale gradient in wood isotopic values was observed, with more ^{18}O-depletion in wood of trees growing in the West of the continent compared to the East. This continental gradient mimics that of δ^{18}O in precipitation (Fig S4.3A). In Asia, a continental-scale gradient was also present, with higher wood δ^{18}O values in India and Myanmar, and lower values in Sumatra, Borneo and Papua (Fig 4.4C). Like in South America, this gradient in Asia is mimics that of δ^{18}O in precipitation, which is higher in western Asia. In Africa, no gradual continental-scale shift in isotopic signature is present: spatial variation occurs at a regional scale, with areas richer in heavy isotopes in both west (Ivory Coast and Ghana) and east (Democratic Republic of the Congo) and more depleted areas in between (Liberia and Cameroon) as well as in Madagascar.

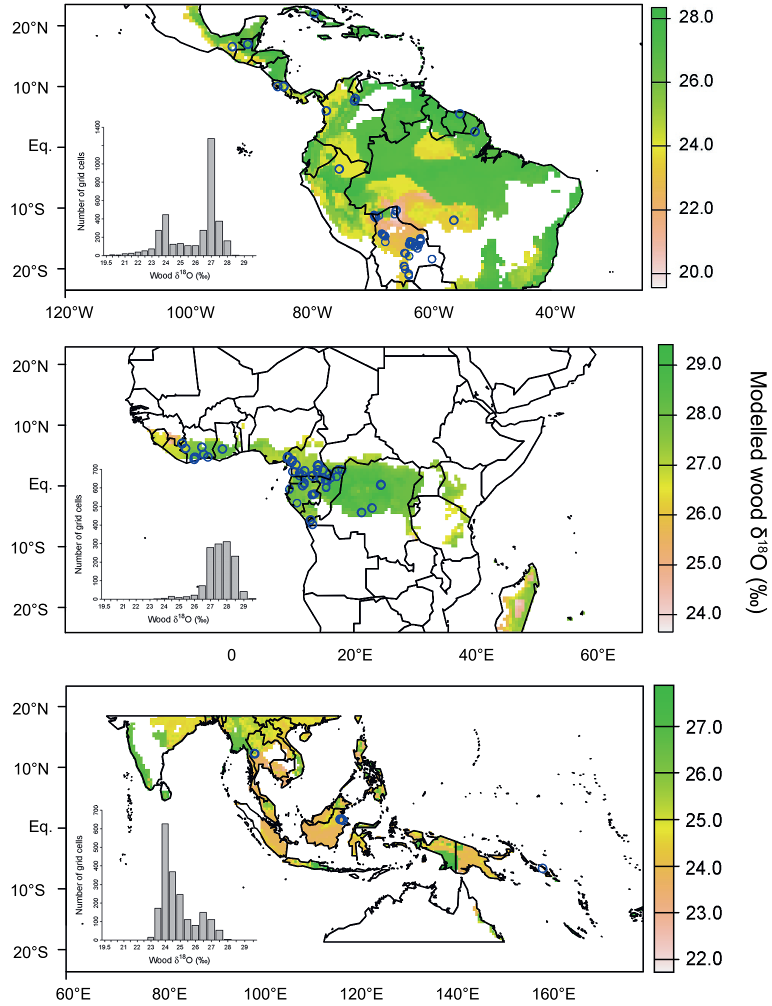


Figure 4.4: Continental-level wood δ^{18}O isoscapes based on the pantropical Quantile Regression Forest model. The isoscape is shown per continent to appreciate continental differences in δ^{18}O ranges, and only for those grid cells with tropical moist forest cover. Insets show histograms of predicted δ^{18}O values for 0.5° grid cells. Blue circles show wood δ^{18}O sample locations.

Species-level δ^{18}O isoscapes

Using species distribution models, we extracted species-level isoscapes of predicted wood δ^{18}O. It should be stressed that these species-level isoscapes only differ among species because of differences in distributional range among species, as all are extracted from the same pantropical isoscape (Fig S4.4A) that does not include a species effect. The resulting 27 species-level isoscape maps were used to calculate variation per species range in wood δ^{18}O (Fig 4.5).

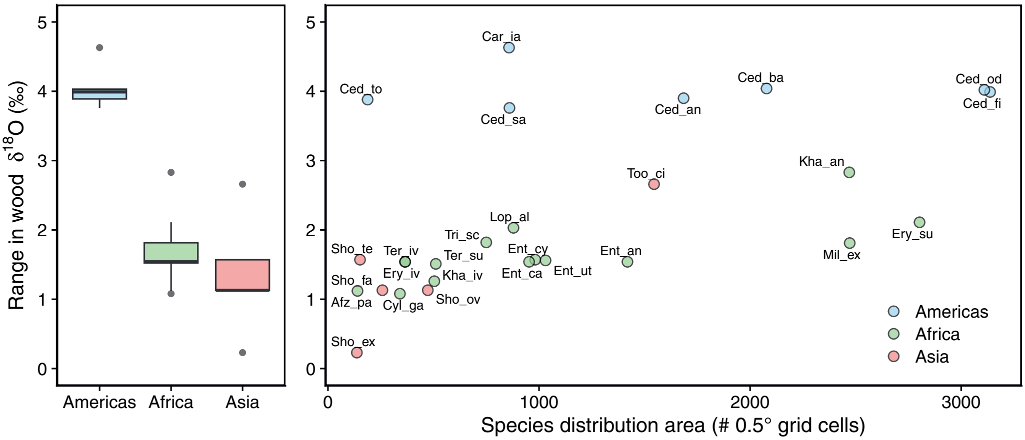


Figure 4.5: Spatial variation in wood δ^{18}O and its relation to tropical timber species distributions. Left: Large-scale spatial variation in wood δ^{18}O within distributional areas of 27 species, grouped by tropical region (colour). Right: Association between species-specific wood δ^{18}O ranges and their distributional areas for 27 species (represented as points; codes explained in Table S4.1). Distributional areas are based on species distribution models and expressed as the number of 0.5° grid cells. The range in wood δ^{18}O is the difference between 10th and 90th percentile of the δ^{18}O values extracted from the pantropical isoscape within the species distribution area.

Across all 27 study species, the predicted ranges of wood δ^{18}O (10th-90th percentile) differed quite strongly among species: from around 1.1‰ for several African species to a maximum of 4.6‰ for a Cedar species in the Americas (Table S4.1). Overall, isotopic ranges were significantly larger for neotropical species (median = 4.0‰) than for African (1.5‰) or Asian (1.1‰) species (Kruskal-Wallis, chi-squared = 15.5, $p<0.001$; Fig 4.5A). Among-species variation in isotopic ranges can be caused by the extent of the species distribution and the spatial variation in explanatory variables within that area (Fig S4.3). The impact of species distribution areas on isotopic ranges seems to be small, as can be observed when plotting these for all species (Fig 4.5B). This scatter plot may help assessing the potential for wood δ^{18}O tracing: species with higher wood isotopic ranges tend to have a higher potential and species combining a high range with a small geographic area have a higher potential for tracing at smaller spatial scales.

Mantel tests of the isoscapes yielded positive correlations between geographic and δ^{18}O distances for all continents (Fig S4.5). It is especially interesting to focus on the scales of autocorrelation.

That is, from which distance the isotopic difference between two points in space ‘levels off’. At those scales, trees that grow further away do not necessarily differ more in δ^{18}O in wood and spatial patterns in the kriging map are more important than absolute distances. For the Americas this is after about 2000 km. In Africa, the δ^{18}O in wood level off after about 1000 km, in Asia the difference in δ^{18}O still seems to increase after 10.000 km.

4.4 Discussion

We present here the first pantropical isoscape for wood. Although our database of wood δ^{18}O values is relatively small, underrepresents various regions and lacks proper species replication, the Quantile Regression Forest model prediction was good. The model explains 82% of pantropical variation in measured wood δ^{18}O by including environmental variables obtained from gridded climate data. The resulting tropical wood δ^{18}O isoscape represents a proof of concept for isotopic timber tracing at large spatial scales.

Temporal and local spatial variability in wood δ^{18}O

Our annual δ^{18}O measurements for some species showed strong temporal variability (Fig 4.2), which may induce substantial uncertainty in estimating tree-level δ^{18}O values and creating isoscapes. This is consistent with results from previous studies, which reported inter-annual variability in δ^{18}O as high as 5‰ (Brienen et al., 2012). The main driver of this variation is the δ^{18}O signature of source water, determined by the “amount effect” and recycling of rainfall (Fig 4.1; Kurita et al. (2009); Brienen et al. (2012)). The uncertainty in tree-level δ^{18}O induced by temporal variability can be strongly reduced by taking samples of wood formed during >20 years and in any case >10 years. For species with average radial growth rates (1.5 mm/y; Locosselli et al. (2020)) this translates to samples measuring at least 3 cm in radial direction (3 mm annual diameter growth); for faster growing species we advise samples of 5 cm in radial direction. In addition, we advise taking and mixing 2-3 wood samples per tree to get a more robust tree-level estimate.

Another source of uncertainty in tree-level δ^{18}O values are gradual changes in wood δ^{18}O over time. These may be caused by increasing rooting depth or shifting transpiration with tree size, or by changes in δ^{18}O of precipitation over time (van der Sleen et al., 2017). To minimize this uncertainty, we advise taking reference samples from recently formed wood and large individuals. The final source of uncertainty in large-scale wood δ^{18}O isoscapes is local spatial variation within a species (Fig 4.2). Such variation can be caused by differences in rooting depth, water availability and crown size of trees (van der Sleen et al., 2017), or – at scales of 10s of km – by steep gradients in isotopic signature of rainfall (Förstel et al., 2011). When taking reference samples, the impact of local variation can be reduced by sampling multiple (5-10) individuals from a location and averaging δ^{18}O values.

Environmental drivers of pantropical variation in wood $\delta^{18}O$

Our Quantile Regression Forest model identified three major drivers of pantropical variation in wood $\delta^{18}O$: $\delta^{18}O$ in precipitation, temperature annual range and potential evapotranspiration. The positive effect of $\delta^{18}O$ in precipitation was expected because of the strong signature of source water $\delta^{18}O$ on wood $\delta^{18}O$ (van der Sleen et al., 2017) and is consistent with studies on spatial variation in $\delta^{18}O$ in oak wood (Watkinson et al., 2020), larch wood (Horacek et al., 2009) and coffee beans (Driscoll et al., 2020). These results are also in accordance with the strong effect of fluctuations in $\delta^{18}O$ in precipitation on temporal variation in wood $\delta^{18}O$ of tropical tree species (Brienen et al., 2012; Baker et al., 2015).

The importance of temperature annual range as an explanatory variable may be explained by two large groups of samples in our database. The first group consists of 'continental samples' - i.e. predominantly *Cedrela* spp. samples from Bolivia – being collected at generally higher elevations and at large distances from the oceanic source of rainfall. The other group of samples - the samples collected in Africa - being collected closer to the coast and at lower elevations. Both higher elevations and distance to the sea are generally associated with lower $\delta^{18}O$ values in precipitation (IAEA/WMO, 2020), which is reflected by lower $\delta^{18}O$ values in wood.

The negative effect of potential evapotranspiration (PET) on wood $\delta^{18}O$ is rather puzzling. The opposite effect was expected, because higher PET is often associated with higher actual transpiration, causing stronger ^{18}O enrichment of leaf water and thus increasing wood $\delta^{18}O$ values. Explaining the negative association between PET and wood $\delta^{18}O$ is complicated because the calculation of PET (Penman–Monteith equation; (Harris et al., 2020)) includes multiple climatic variables. In addition, the relation between PET and wood $\delta^{18}O$ within a species may be quite different from that across species, due to inter-specific differences in transpiration responses to atmospheric conditions.

Species differences

The $\delta^{18}O$ signature of tropical tree species growing in the same area can vary quite strongly, due to inter-specific differences in $\delta^{18}O$ values of water sources, e.g. caused by rooting depth, leaf phenology and transpiration (Watkinson et al., 2022a). Given these two major sources of interspecific wood $\delta^{18}O$, it is remarkable that we obtained a robust model explaining 82% of pantropical variation in wood $\delta^{18}O$ from a database representing 27 species and a rather unbalanced sampling design. This result suggests that large-scale climatic drivers potentially have a stronger effect on tropical wood $\delta^{18}O$ than species differences.

Species differences in wood $\delta^{18}O$ also became apparent in our mixed effect model on a subset of replicated species, as we found mean values between species to vary with a standard deviation of almost 1.6‰. This value provides a first indication of the possible magnitude of species differences in large-scale isoscapes.

A better quantification of species effects on tropical wood $\delta^{18}O$ is needed and can be obtained by sampling multiple species with sufficient spatial replication (at least 10 sites per species) and across large climatic gradients. Our pantropical and species-level isoscapes will be helpful in deciding where sampling efforts should be focused to better quantify species differences and in designing sampling strategies for reference sampling initiatives such as World Forest ID (Gasson et al., 2021).

Using global gridded data to establish isoscapes: opportunities and limitations

In spite of recent initiatives to collect large amounts of samples from tropical timber species (Gasson et al., 2021) and unite references databases (Global Timber Tracking Network; https://globaltimbertrackingnetwork.org/products/reference-database/), it is very unlikely that isotopic reference databases will have a sufficient resolution and coverage of species and distributional areas to produce species-specific isoscapes based on reference samples only. The establishment of statistical relations between isotopic measurements and environmental variables, such as climate and soil, combined with the availability of gridded data for such variables, offers the opportunity to reduce the "burden of proof" on reference samples. This allows for making well-informed – yet cautious – inferences on expected isotopic signatures for regions or species not included in the reference database. It is an approach that has been applied previously in studies on the provenance of various commodities, including coffee (Driscoll et al., 2020) and temperate wood (Gori et al., 2018; Watkinson et al., 2020). Our study is one of the first to do so for tropical timber (Watkinson et al., 2022b).

The creation of isoscapes based on statistical relations with environmental information is greatly facilitated by the increasing availability, spatial resolution and quality of gridded data products for climatic drivers of wood isotopic values. Yet, there are also some limitations to this approach. First, the reliability of the generated isoscapes strongly depends on the quality, spatial resolution and uncertainty of the gridded data, which may vary strongly among data products. Cautious and critical use of these products is therefore required. Ideally, uncertainty of the gridded data would be propagated to the resulting isoscape, but techniques to do so have not been developed until now, as far as we are aware. Second, the low spatial flexibility (i.e., coarse spatial structure) of some of the gridded data products may reduce the association with isotopic data in case these vary strongly at local scales. Thus, the spatial stiffness of gridded data may reduce the spatial resolution of the isoscape, even if the map has a high resolution, and hamper reliable predictions of fine spatial structure. Finally, it may not always be beneficial or necessary to establish isoscapes based on gridded data. For instance, if the reference dataset is large or covers the regions of interest, simple statistical comparisons and assignment tests may suffice (Horacek et al., 2009). Such direct comparisons may also be an appropriate approach if isotopic variation is poorly explained by gridded climate data in statistical models or there is reason to assume that the isoscape does not represent local isotopic variation.

What is the potential for tropical timber tracing based on wood δ^{18}O?

Assessing the potential of wood δ^{18}O measurements for tropical timber tracing based on this study is difficult, because we have not conducted assignment tests or bind sample tests. Nevertheless, our study does provide a first proof of concept for the application of wood δ^{18}O at large spatial scales for tropical timbers. Across continents we consistently find increases in isotopic differences with increasing distance up to several thousands of kms.

The spatial resolution of isotopic tracing can be further increased by combining δ^{18}O isoscapes with other stable isotopes such as δ^{13}C, δ^{2}H, δ^{15}N and δ^{34}S (Chapter 5, this thesis). Combining isotopes may increase predictive power and improve tracing potential and reliability (Watkinson et al., 2020). Additionally, adding samples that cover the entire isotopic range of a species or continent will improve the Quantile Regression Forest models and thus the predictions of δ^{18}O. To be able to do so on a large scale, it is essential that data is comparable across labs to enable a shared database. Currently, this is complicated by differences in sample preparation and measurement protocols.

Our species-level analyses of the wood δ^{18}O isoscape can be used to provide some cautious clues on species characteristics that may increase the potential for δ^{18}O. First, species occurring in regions with a large isotopic variation such as the West-East δ^{18}O gradient in South America or Asia likely are more suitable than those growing in Africa. Second, species with a large distributional range or from which timber is derived from a large geographic range likely have a higher potential, although this seems not to be the case in the Americas. Third, tropical timbers from plantations and natural forests can be told apart based on wood δ^{18}O, especially if plantations and natural forest sources are located on different continents. This is particularly beneficial for timbers with strongly restricted harvesting in natural forests because of the conservation status of the species, such as teak and mahogany.

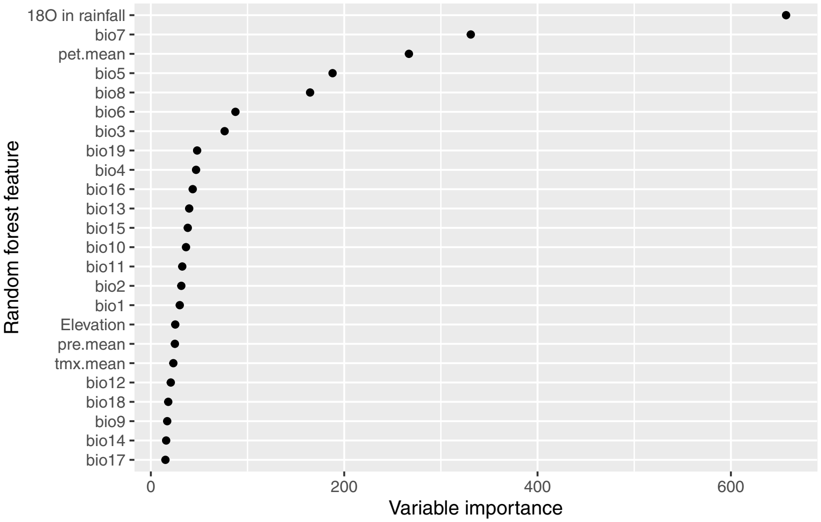
Acknowledgements

This work was supported by the European Research Council (grants ERC-StG-242955 and ERC-PoC-665275) and the Dutch Research Council (I-STW-Take-off-14071 NWO-TTW-OTP-16427). We thank managers of wood collections for their assistance in providing wood samples, in particular Naturalis Biodiversity Center and the Tervuren Xylarium.

Supporting information

Table S4.1: Overview of studied tropical timber species and results of species-level isoscapes. Area refers to geographic extent of the species distributional area, quantified as the number of 0.5° grid cells covered by the species. δ^{18}O range is the difference between the 10th and 90th percentile of predicted δ^{18}O within a species' distribution area.

Species	Sp Code	Family	Trade name	Area (0.5° cells)	δ^{18}O range (‰)	10th δ^{18}O (‰)	90th δ^{18}O (‰)	# trees in dataset
Africa								
Afzelia pachyloba	Afz_pa	Fabaceae	Doussie	140	1.12	26.93	28.05	1
Cylicodiscus gabunensis	Cyl_ga	Fabaceae	Okan	341	1.08	26.97	28.05	1
Entandrophragma utile	Ent_ut	Meliaceae	Sipo	1031	1.56	26.97	28.53	5
Entandrophragma cylindricum	Ent_cy	Meliaceae	Sapele	982	1.57	26.96	28.53	4
Entandrophragma angolense	Ent_an	Meliaceae	Tiama	1419	1.54	26.96	28.5	5
Entandrophragma candollei	Ent_ca	Meliaceae	Kosipo	953	1.54	26.96	28.5	8
Erythrophleum ivorense	Ery_iv	Fabaceae	Tali	365	1.54	26.5	28.04	178
Erythrophleum suaveolens	Ery_su	Fabaceae	Tali	2803	2.11	26.14	28.25	87
Khaya anthotheca	Kha_an	Meliaceae	African mahogany	2470	2.83	25.41	28.24	2
Lophira alata	Lop_al	Ochnaceae	Azobé	879	2.03	26.5	28.53	144
Milicia excelsa	Mil_ex	Moraceae	Iroko	2472	1.81	26.5	28.31	2
Terminalia superba	Ter_su	Combretaceae	Limba	511	1.51	26.5	28.01	3
Terminalia ivorensis	Ter_iv	Combretaceae	Framire	366	1.54	26.5	28.04	4
Triplochiton scleroxylon	Tri_sc	Malvaceae	Abachi	750	1.82	26.51	28.33	2
Americas								
Cariniana ianeirensis	Car_ia	Lecythidaceae	Yesquero blanco	858	4.63	22.76	27.39	6
Cedrela angustifilia	Ced_an	Meliaceae	Cedar	1685	3.9	23.51	27.41	2
Cedrela balansae	Ced_ba	Meliaceae	Cedar	2078	4.04	23.18	27.522	10
Cedrela fissilis	Ced_fi	Meliaceae	Cedar	3137	3.99	23.42	27.41	116
Cedrela odorata	Ced_od	Meliaceae	Cedar	3110	4.02	23.53	27.55	41
Cedrela saltensis	Ced_sa	Meliaceae	Cedar	860	3.76	23.49	27.25	8
Cedrela tonduzii	Ced_to	Meliaceae	Cedar	188	3.88	23.53	27.41	1
Asia								
Shorea ovalis	Sho_ov	Dipterocarpaceae	Meranti	473	1.13	23.72	24.85	1
Shorea exelliptica	Sho_ex	Dipterocarpaceae	Meranti	137	0.23	23.71	23.94	1
Shorea faguetiana	Sho_fa	Dipterocarpaceae	Meranti	258	1.13	23.72	24.85	1
Shorea tenuiramulosa	Sho_te	Dipterocarpaceae	Meranti	152	1.57	23.72	25.29	1
Toona ciliata	Too_ci	Meliaceae	Toon	1545	2.66	23.85	26.51	6



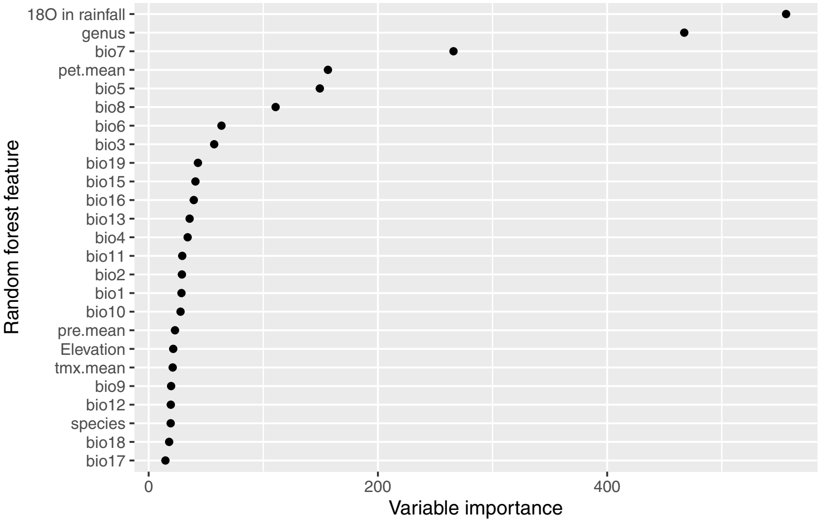


Figure S4.1: Variable importance of the most important variables for the Quantile Regression Forest. Including only environmental predictor variables (top) and with species and genus added as predictor variables (bottom). Bioclim variables bio1-bio11 refer to temperature-related variables, bio12-bio19 refer to rainfall related variables. Please refer to https://worldclim.org/data/bioclim.html for the individual labels.

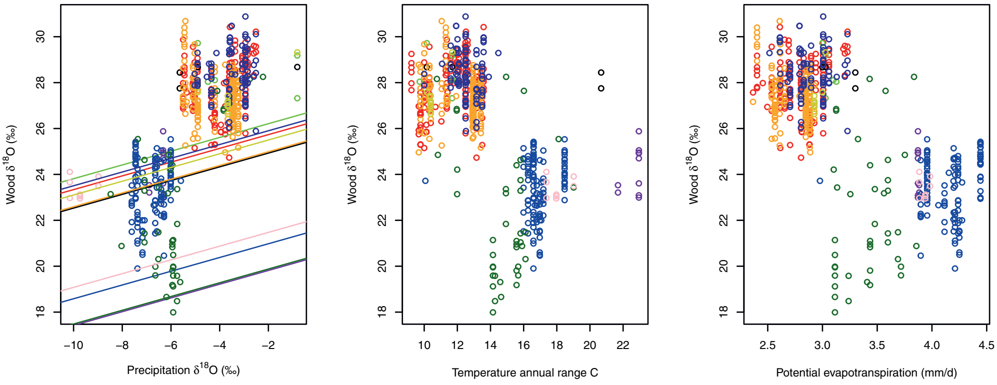


Figure S4.2: Results of the mixed effect model of wood δ^{18}O for a subset of the dataset restricted to species represented by at least five individuals and spatially separated measurements. This thus represents a subset of the data shown in Figure 3. Predictions for the three explanatory variables are shown as lines (precipitation δ^{18}O, temperature annual range and potential evapotranspiration), one colour per species. Colour coding: red = *Lophira alata*, green = Entandrophragma angolense, black = Entandrophragma candollei, orange = Erythrophleum ivorense, darkblue = Erythrophleum suaveolens, yellow = Terminalia superba, darkgreen = Cedrela odorata, pink = Cedrela saltensis, purple = Cedrela balansae, blue = Cedrela fissilis.

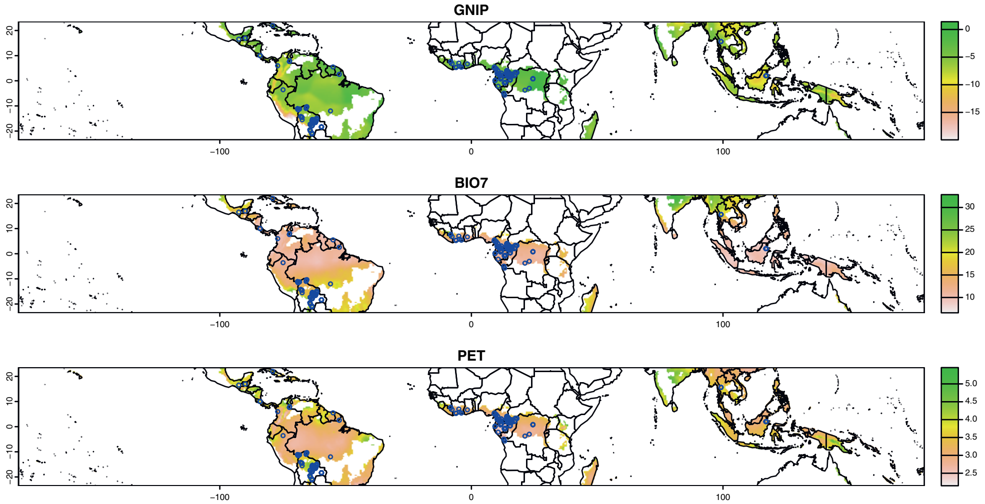


Figure S4.3: Pantropical distributions of values for the three explanatory variables that entered the Quantile Regression Forest model for tropical wood δ^{18}O and sample locations (circles). (A) Annual δ^{18}O in precipitation from the GNIP database (GNIP, in ‰), (B) Temperature annual range (BIO7, in), (C) Daily potential evapotranspiration (PET, in mm/day).

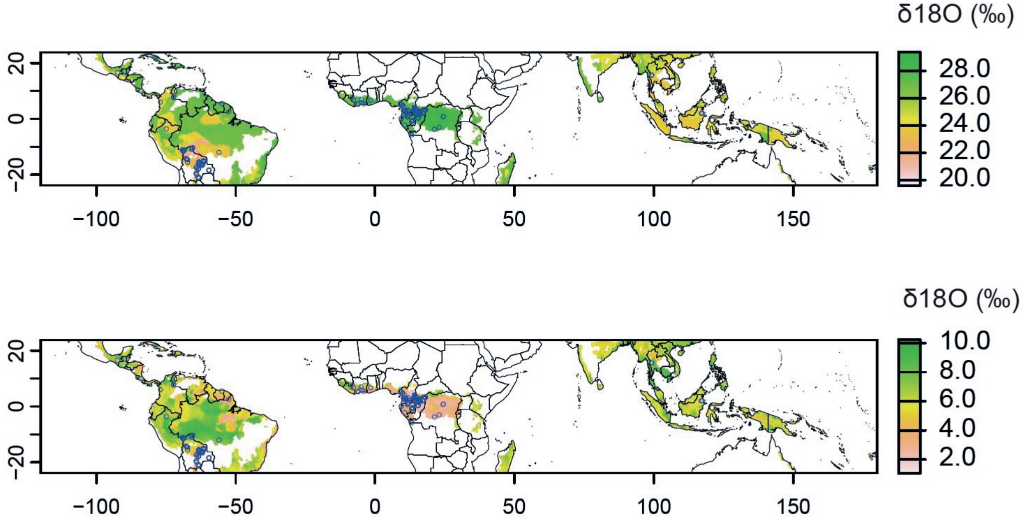


Figure S4.4: Pantropical wood δ^{18}O isoscape. (A) Median wood δ^{18}O isoscape, obtained by Quantile Regression Forest. (B) Inte range (10-90%) of wood δ^{18}O isoscape, obtained by Quantile Regression Forest. Blue dots are the sampling locations.

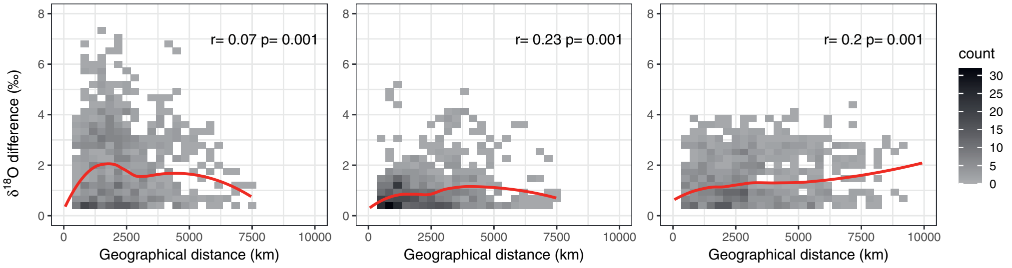
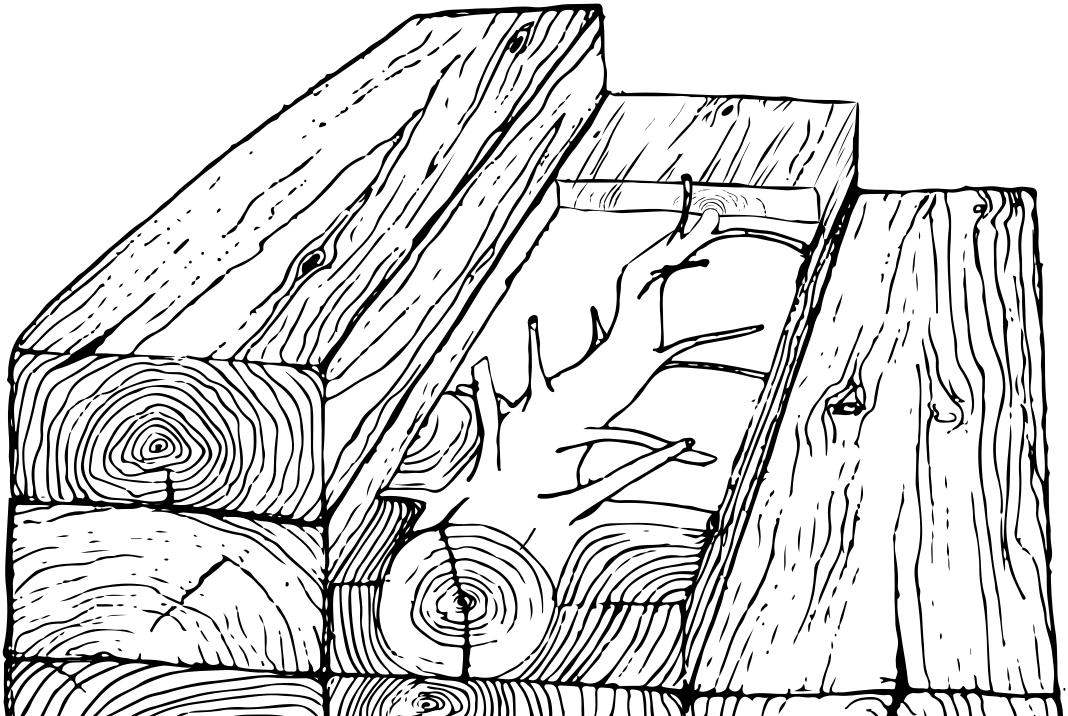


Figure S4.5: Density plot of the relation between pairwise difference in wood δ^{18}O and geographic distance for all 0.5° grid cells of the pantropical isoscape, calculated per continent (left to right: America, Africa, Asia). A random selection of 1000 points is taken per continent, with darker colors indicating more datapoints at that distance. Statistical output of Mantel test is shown (999 permutations): r = Spearman rank correlation and p = p-value. The red LOWESS line is only shown to observe the tendency of data.



Chapter 5

Stable isotope ratios in wood show little potential for sub-country origin verification in Central Africa

Laura E. Boeschoten, Mart Vlam, Ute Sass-Klaassen, Barbara Rocha Venâncio Meyer-Sand, Arnoud Boom, Gaël U.D. Bouka, Jannici C.U. Ciliane-Madikou, Nestor L. Engone Obiang, Mesly Guieshon-Engongoro, Joël J. Loumeto, Dieu-merci M.F. Mbika, Cynel G. Moundounga, Rita M.D. Ndangani, Dyana Ndiade Bourobou, Peter van der Sleen, Steve N. Tassiamba, Martin T. Tchamba, Bijoux B.L. Toumba-Paka, Herman T. Zanguim, Pascaline T. Zemtsa, Pieter A. Zuidema

Published in Forest Ecology and Management, 544 (2023) 121231, https://doi.org/10.1016/j.foreco.2023.121231

Abstract

Origin verification of timber is essential to expose origin fraud and reduce illegal timber trade. A promising forensic method for origin verification is based on stable isotope ratios in wood, but large-scale studies that test local and regional variation to apply the method at a sub-country scale are lacking.

We investigated the isotopic variation in wood in Central Africa for two economically important timbers (Azobé and Tali). We measured wood $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$ and $\delta^{34}S$ at 17 locations in the main timber exporting countries (Cameroon, Republic o/t Congo, Gabon), for the first time quantifying both local and regional variation as well as species differences in the same study. We applied a novel statistical technique, Quantile Regression Forests, to develop spatial predictions of isotopic composition based on gridded covariables (isoscapes). This is a powerful method to develop isoscapes based on non-normally distributed and correlating covariables.

We found limited potential for origin assignment with stable isotope ratios in Central Africa: local variability was too pronounced in all isotopes, that regional variation was not large enough to identify origin. This led to low site classification success (Azobé: 32.1%, Tali: 20.5%) and large uncertainty in the isoscapes. The limited origin signal can be attributed to low variation in the driving environmental factors in the region. The isotope ratios did differ between the three countries and country level assignment was considerably higher (Azobé: 79.4%, Tali: 61.7%). Lastly, we found a significant species and site effect, stressing species-specific reference datasets might be needed for future isotope tracing studies.

These results show that the four isotopes applied here have limited potential for sub-country origin assignment of timber in this region (Cameroon, Republic of the Congo, Gabon). Nevertheless, the variation at large spatial scales resulted in more promising country-level assignment. These findings contribute to our understanding of forensic timber tracing methods and can help to identify the most appropriate method for a tracing question at hand.

Keywords

Illegal logging, Isoscapes, Sulphur isotopes, Quantile Regression Forest, Tali, Azobé

Highlights

- Local variation in wood stable isotope ratios was high in Central Africa
- This resulted in low accuracy when assigning site of origin based on stable isotopes
- High local variation also resulted in high uncertainty in isoscapes
- Isotopic variation was more pronounced at large spatial scales (country-level)
- Species specific variation indicated the need for individual datasets for tracing

5.1 Introduction

Illegal timber trade harms people, ecosystems, and local economies and is especially prevalent in the tropics (Hoare & Uehara, 2022). In the Congo Basin, estimations of illegally exported timber go up to 90% for some countries (Hoare & Uehara, 2022), while law enforcement efforts to reduce this trade are often limited by a lack of available tools to verify the true origin of the wood (Lowe et al., 2016). Origin fraud occurs both within and between countries (Gan et al., 2016), so a verification of both country and region of origin could help to combat the illegal trade. A method that has been applied for origin verification of a variety of products is based on the use of stable isotope ratios. It has been successful for origin verification of for example cocaine (Ehleringer et al., 2000), elephant tusks (Ziegler et al., 2016), coffee (Driscoll et al., 2020) and other foodstuffs (Oulhote et al., 2011). Its application in tropical timber so far is relatively limited however (Low et al., 2022; Dormontt et al., 2015) and has yielded mixed results: some studies could distinguish between sites (Watkinson et al., 2022a,b) but others not (Vlam et al., 2018b; Paredes-Villanueva et al., 2022).

The stable isotope ratios mostly used for timber tracing are the bio-elements δ^{18}O, δ^{2}H, δ^{13}C, δ^{15}N and δ^{34}S, from which an origin-specific isotopic fingerprint can be defined. The stable isotope ratios of a new sample are then compared to a reference database, allowing for the assignment of a sample to its most likely origin (Förstel et al., 2011). Additionally, isotopic composition can be mapped, producing so-called isoscapes (Bowen, 2010). These isoscapes predict the isotopic composition in between measurement points based on relevant drivers of isotopic variation as covariables, allowing for the assessment of isotopic variation at a landscape scale (Gori et al., 2018). Both of these methods have the potential to determine timber origin.

For successful origin assignment, geographical variation in stable isotope ratios is a prerequisite (Bowen et al., 2005). In timber, this variation depends on the availability and uptake of the isotopes by the tree as well as further on fractionation processes within the tree. Both result in geographical isotopic variation as isotope ratios in wood are influenced by climate, geology and/or deposition, although the specific processes differ per isotope. These processes also define the relevant covariables to be used in isoscapes. Important factors that have been linked to geographic variation in isotopes include distance to the ocean (δ^{18}O, δ^{2}H, δ^{34}S), elevation (δ^{18}O, δ^{2}H, δ^{34}S), temperature (δ^{18}O, δ^{2}H, δ^{13}C), rainfall amount as well as rainfall isotopic composition (δ^{18}O, δ^{2}H, δ^{13}C), nitrogen source (δ^{15}N), bedrock (δ^{34}S) and deposition (δ^{15}N, δ^{34}S) (West et al., 2010; Cernusak et al., 2022; Wynn et al., 2014; Lehmann et al., 2022; Allen et al., 2022; van der Sleen et al., 2017; Brlík et al., 2022). Therefore, the method has shown high potential for timber tracing in areas where the natural variation of stable isotopes in wood was large between origins (Förstel et al., 2011; Degen et al., 2015; Horacek et al., 2009; Lee et al., 2015).

For timber origin assignment in West and Central Africa, the use of stable isotopes has shown some potential: one report indicated country-level verification with 55%-75% accuracy for blind samples, including 10 blind samples per test (Degen et al., 2015). At a sub-country scale, timber from two locations in Gabon could be distinguished, mostly based on δ^{34}S (Watkinson et al., 2022a). In contrast, this was not possible in Cameroon based on δ^{15}N, δ^{13}C and δ^{18}O (Vlam et al., 2018b). Studies on animal products in the area did show tracing potential, for example bird feathers have been traced using δ^{34}S (Brlík et al., 2022) and ivory using δ^{13}C, δ^{18}O, δ^{2}H, δ^{15}N and δ^{34}S (Ziegler et al., 2016) although both encompassed a larger study area. Studies on multiple timber species and multiple isotopes with high sampling coverage are lacking however, which is essential to further understand variation in isotopic composition across the area.

The aim of this study was to investigate the isotopic variation in timber from Central Africa in order to test the potential for timber tracing. Additionally, we aimed to improve isotopic origin assignment tests by applying machine learning techniques and by explicitly including prediction uncertainty. To achieve this, we developed an extensive dataset of stable isotope ratios from trees growing at 17 locations across the main timber exporting countries in the region: Cameroon, Gabon and Republic of the Congo. These three countries are the most important sources of timber exports in the Congo Basin, while at the same time illegal practices remain widespread throughout their timber supply chains (Hoare & Uehara, 2022). Increased transparency through scientific origin verification is thus an important step to reduce illegality in the region. We included two important timber species in this study: Azobé (*Lophira alata*) and Tali (*Erythrophleum ivorense* and *E. suaveolens*) to allow for species comparisons. We focused on δ^{18}O, δ^{2}H, δ^{13}C and δ^{34}S as they have been most extensively used for tracing in the area (Watkinson et al., 2022a; Degen et al., 2015; Brlík et al., 2022; Ziegler et al., 2016). δ^{15}N is mostly affected by the source of N_2 and local nitrogen cycling (van der Sleen et al., 2015), which is not expected to show strong geographical variation, and was thus not included here. This is the first study to test the feasibility of isotopic timber tracing using a sampling coverage that allows a quantification of species differences as well as local and regional variation for multiple isotopes.

As tracing can only be successful if regional variation is larger than local variation, we first quantified the local and the regional isotopic variation within the two species. We then developed isoscapes based on relevant covariables for all four stable isotope ratios to assess the variation in wood isotopic composition across our study area. To develop the isoscapes we made use of Quantile Regression Forests, a novel statistical technique for spatial analysis based on machine learning that has not yet been applied to map isotope variability. The strength of Quantile Regression Forests comes from the fact that it is a powerful method to develop prediction maps based on non-normally distributed, high-dimensional and possibly correlating covariables (the environmental variables driving variation in the stable isotopes). Furthermore, it estimates the conditional quantiles based on all random forest outcomes, which presents an accurate estimation of the uncertainty with a certain prediction (Meinshausen, 2006). Lastly, we tested assignment success by applying classification models based on the stable isotope ratios. Through this research, we aimed to contribute to the development of effective tools for assigning the origin of tropical timber and that could ultimately aid in the reduction of illegal timber trade.

5.2 Methods

Study design

The study was conducted on two valuable tropical timber species from Central Africa: Azobé (*Lophira alata* Banks ex C.F. Gaertn, Ochnaceae) and Tali (*Erythrophleum ivorense* A. Chev. and *E. suaveolens* (Guill. & Perr.) Brenan, Fabaceae). These are among the most important internationally traded species and represent a single-species timber, Azobé, as well as a multi-species timber, Tali. Tali was only identified to species level if leaf, flower and/or fruit material was available as it is hard to distinguish in the field. Both timbers grow in evergreen and moist deciduous forests. Samples were taken from trees at 17 study sites spread across the main timber exporting countries of timber from natural forests in Central Africa: Cameroon (7 sites), Gabon (4 sites) and the Republic of the Congo (6 sites), between September 2019 and April 2022 (Fig 5.1). Samples were taken from both species at all sites except for two locations: at one site only Azobé was sampled (GAB1) and at one site only Tali was sampled (GAB4) due to low occurrence of the other species. All sites were natural forest concessions, accessed in collaboration with the operating forestry companies.

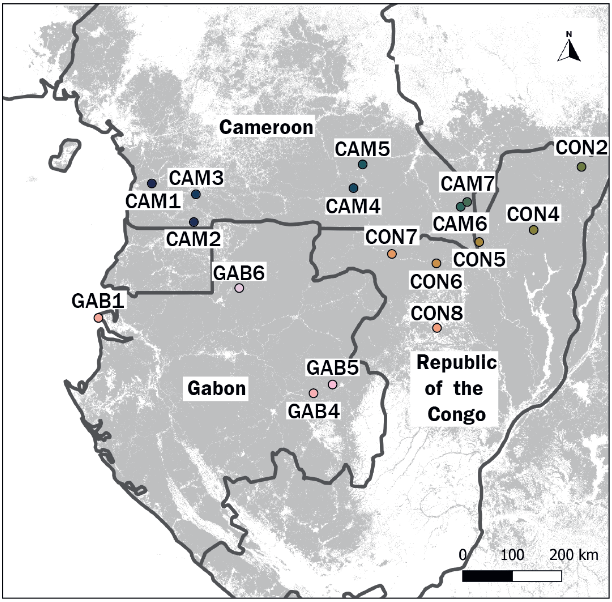


Figure 5.1: Map of the study area and study sites. Primary tropical forest extend from Global Forest Watch in light grey (Turubanova et al., 2018).

Sample collection

At each site, we sampled heartwood from 20-30 trees per timber species. Target trees within one site were located between 100 m and 5 km apart. Trees were either standing or recently felled and were of at least 30 cm diameter at breast height (DBH). We sampled trees of different sizes at all sites to minimize the effect of ontogenetic stage on any isotopic signal, as can be the case for δ^{13}C (Brienen et al., 2017). The heartwood sample was collected from each tree as an increment core (Haglöf Increment borer 350 mm x 5,15 mm; n = 27), with a FAMAG plug cutter of 15 mm diameter (n = 165), as a wood chunk (n = 11) or as a wood powder sample obtained with an electrical drill (n = 85). All samples were taken at least 14 cm into the tree. The samples were stored in plastic straws or paper envelopes and properly ventilated to prevent mould, no additional drying steps were performed. Additionally, GPS-coordinates and DBH were recorded.

Stable isotope analysis

Between four and 10 trees per site were selected for stable isotope analysis, depending on the isotope: we measured δ^{34}S in four trees per species per site (total of 126), δ^{2}H in six or 10 (total of 254), δ^{18}O in six or 10 (total of 246) and δ^{13}C in six or 10 (total of 240) trees per species per site. To define the number of species per site per isotope we balanced between measurement costs and the expected local variation per isotope, local variation of δ^{34}S was expected to be lower than the other three isotopes for example as we assumed less fractionation takes place within the tree. This resulted in a geolocated database of 143 Azobé and 145 Tali trees in total. Measurements of the different isotopes were additive, meaning that all four isotopes were measured in the same tree in as many trees as possible (106 in total). Only if the amount of material was too limited, another tree of that site was analysed.

For the analyses, a subsample of heartwood from every tree was cut in radial direction including at least 3-5 cm to include wood formed during multiple years. δ^{13}C, δ^{18}O and δ^{2}H were measured in cellulose, following Vlam et al. (2018b) for cellulose extractions. δ^{13}C and δ^{18}O were measured at the Leicester Environmental Stable Isotope Laboratory, using an IRMS (Sercon Hydra 20-20). δ^{2}H was measured in duplo at the laboratory of WSL, Birmensdorf, applying an on-line equilibration technique with a DeltaPlus XP IRMS (Finnigan MAT, Schuler et al. (2022)). δ^{34}S was measured in whole wood at Agroisolab GmbH, where the wood was powderized and extracted as described by Watkinson et al. (2022a), after which measurements were performed with an IRMS (Isoprime). The stable isotope compositions were expressed in ‰, relative to an international reference standard (V-SMOW for δ^{18}O and δ^{2}H, V-PDB for δ^{13}C and CDT for δ^{34}S).

Statistical analysis

Linear models for origin and species differences

Statistical analyses were performed in R version 4.1.0 (R Core Team, 2021). Differences between countries and species as well as site and species were tested by two-way ANOVAs including Tukey post-hoc tests. Additionally, the association between the water isotopes (δ^{2}H and δ^{18}O) in wood and rainfall was tested by multiple regression, including species and rainfall isotopic composition as explanatory variables. The association between tree DBH and δ^{13}C was tested by linear regression including species and DBH as explanatory variables, to test the effect of ontogenic stage of the tree on the δ^{13}C ratio. Before all tests, box-cox or log transformations were applied to remove skewedness of the data after adding a constant value to move values above zero (MASS package, Venables & Ripley (2002)).

Individual isoscapes based on Quantile Regression Forests

We identified relevant environmental co-variables from stable isotope theory and prior studies and obtained those variables at the growth locations of all trees (Table S5.1). We then predicted the relationship between wood isotopic composition and the co-variables by applying Quantile Regression Forests, one model per isotope per species. Random Forest Regression, the underlying method for this analysis, is not influenced by co-linearity and non-normal data distributions and has been successfully applied for other isoscapes (Brlík et al., 2022). Another important advantage over other mapping techniques, such as co-kriging, is that no 'one best model' needs to be chosen based on the available data. Therefore, it remains more flexible, as additional data can be easily incorporated once it becomes available and the models can simply be updated by re-training the models. Furthermore, through applying a Quantile Regression Forest, the confidence interval around the predicted median is quantified as the values of all individual prediction trees are stored to calculate the interquantile range. This provides an estimation of prediction uncertainty, an essential step in origin assignment methods. Based on these Quantile Regression Forests, we predicted the spatial distribution of the isotopes in the study area by applying them to gridded layers of all covarables. This resulted in eight isoscapes of the predicted median (2 species, 4 isotopes) and their respective interquantile ranges.

Site and country classification models

Lastly, we applied Random Forest classification models to assign trees to their most likely site (models 1-3, Table 5.1) and country of origin (models 4-6) based on the four isotopes. As we did not measure all isotopes in all trees, the missing isotope ratios were imputed with the mean of that site per species. To ensure a fully independent test set without imputed data, 25% of the trees in which all four isotopes were measured was set apart randomly first and these samples were not included in the Random Forest model development. Assignment success was then evaluated using 50-fold cross-validation (so taking a random test set 50 times), calculated as: 100% minus the percentage of incorrect assigned trees. Furthermore, we calculated the geographical distance between the real and the assigned origin of all test samples in the 50 cross validations (distHaversine from the geosphere package, Hijmans (2022)).

5.3 Results

Species and sites differences

We measured the wood isotopic composition of 288 trees from two tree species at 17 sites, from an area that spanned about 900 km west to east and 450 km north to south across Cameroon, Republic of the Congo and Gabon. δ^{18}O varied between 24.7 and 30.9‰ (n=246; Fig. 5.2), δ^{2}H between -25.8 and 36.6‰ (n=254), δ^{13}C between -31.1 and -21.9‰ (n=240) and δ^{34}S between 4.4 and 10.9‰ (n=126). We found significant differences between species in all four isotopes, except for δ^{18}O which did not differ between Azobé and Tali when comparing the countries (two-way ANOVA with Tukey post-hoc, $p>0.05$, Fig 5.2a): Azobé had overall higher δ^{2}H ratios and lower δ^{13}C and δ^{34}S ratios than Tali, corresponding to more heavy δ^{2}H stable isotopes and less δ^{13}C and δ^{34}S isotopes relative to their lighter counterparts. Isotopic ratios also differed between countries (Fig 5.2a) and individual sites (Fig 5.2b). The δ^{18}O ratio in Azobé for example was significantly lower in Cameroon compared to Republic of the Congo and Gabon and was lower in site CAM1 specifically than in nine of the other sites.

In addition to the species and site effect, the interaction between the two was significant for three isotopes (δ^{18}O, δ^{13}C, δ^{34}S), indicating the two species did not share a geographical pattern across sites. Again, this was clear at site CAM1 for example, where δ^{18}O was lowest of all sites in Azobé but not in Tali. The interaction was not significant in δ^{2}H, indicating shared variation between sites in the two timbers. However, even if some sites were relatively unique in terms of stable isotopes, many sites also showed a large within-site variation that overlapped strongly with the isotopic signal of other sites. As a result, many sites were not significantly different neither from other sites within the same country, nor from sites of the other countries.

Isotopic variation across the study area: individual isoscape models

We developed eight isoscapes predicting stable isotope ratios across the study area (panels a and d in Fig 5.3), one per species for each isotope. All isoscapes showed a gradient from the ocean inwards, reflecting distance to the coast as one of the top predictors in all models (Fig S5.1). Sand content in the soil was another important variable, as well as elevation and several precipitation-related bioclimatic variables (bio 12-19) and temperature-related bioclimatic variables (bio 1-11, see Fick & Hijmans (2017) for the individual labels). Specifically for δ^{34}S, the SO_2 deposition occurred among the top variables as well.

Contrary to what we expected, the modelled stable isotopes in rainwater were not among the most important variables for δ^{18}O and δ^{2}H. δ^{2}H did largely follow the expected pattern of an increased light isotope concentration from the coast inland, and thus lower δ^{2}H for Azobé and Tali from the coast inland (Fig 5.3.2a and d). However, this was much less clear in δ^{18}O (Fig 5.3.1a and d), where the highest δ^{18}O values were found more inland, especially for Azobé. These patterns were also reflected in the associations between wood and modelled rainwater isotopic compositions for δ^{18}O and δ^{2}H individually: the water stable isotope ratios were significantly associated between rainwater and wood (Fig S5.2), but model fit was low for both isotopes (adj. R^2 of 0.04 for δ^{18}O and 0.19 for δ^{2}H). This indicates a strong local deviation in the isotopic signal in wood from the modelled isotopic signal in rainwater.

The predicted values at the locations of our measured trees all correlated well with the observed values (Fig 5.3c and f). However, the range of predicted values was much smaller than the observed range. Even though the median prediction showed a geographical pattern based on

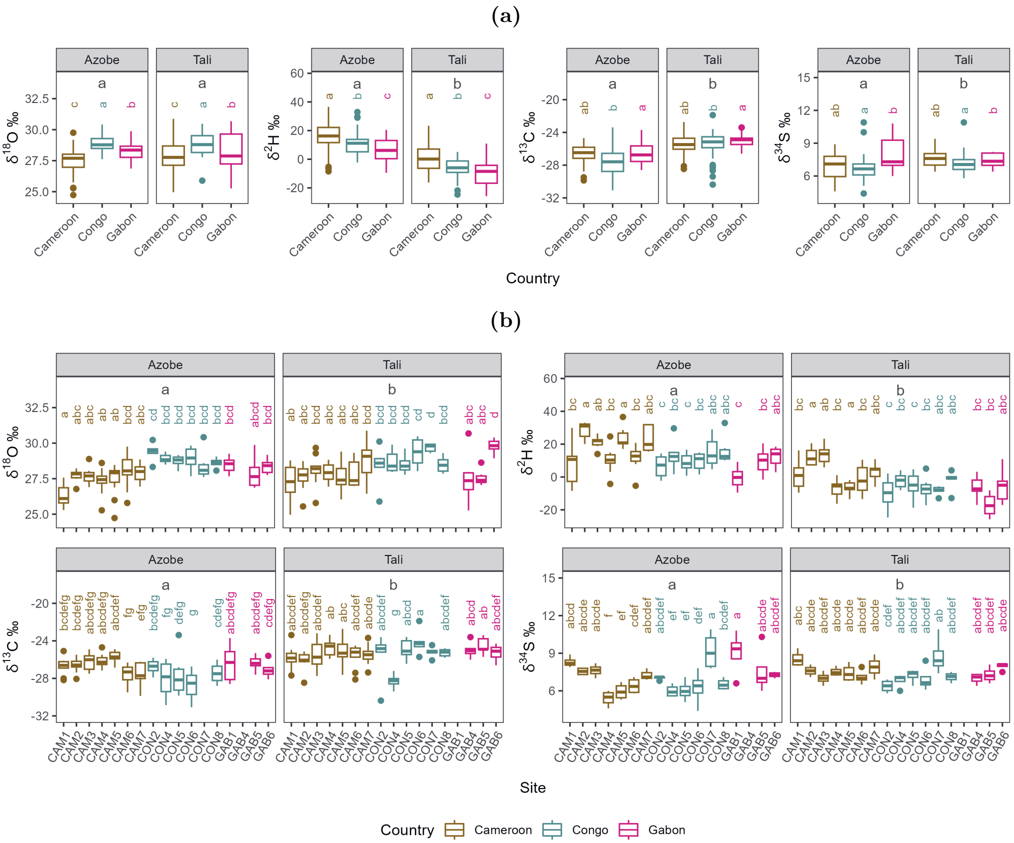


Figure 5.2: Stable isotope ratios in the wood of Azobé and Tali across the countries (a) and study sites (b), expressed relative to the respective international standards in ‰. Letters indicate significant differences between species (grey, horizontal) and between sites or countries (colored, vertical, two-way ANOVA with Tukey post-hoc; n=246 for δ^{18}O, n=254 for δ^{2}H, n=240 for δ^{13}C, n=126 for δ^{34}S). None of the species * country interactions were significant so in (a) letters indicate shared country differences between Azobé and Tali. The species * site interaction was significant in δ^{18}O, δ^{13}C and δ^{34}S, therefore in (b) the letters indicate a unique species * site effect that is not shared between Azobé and Tali. In δ^{2}H, the interaction was not significant so the letters indicate shared site differences between Azobé and Tali.

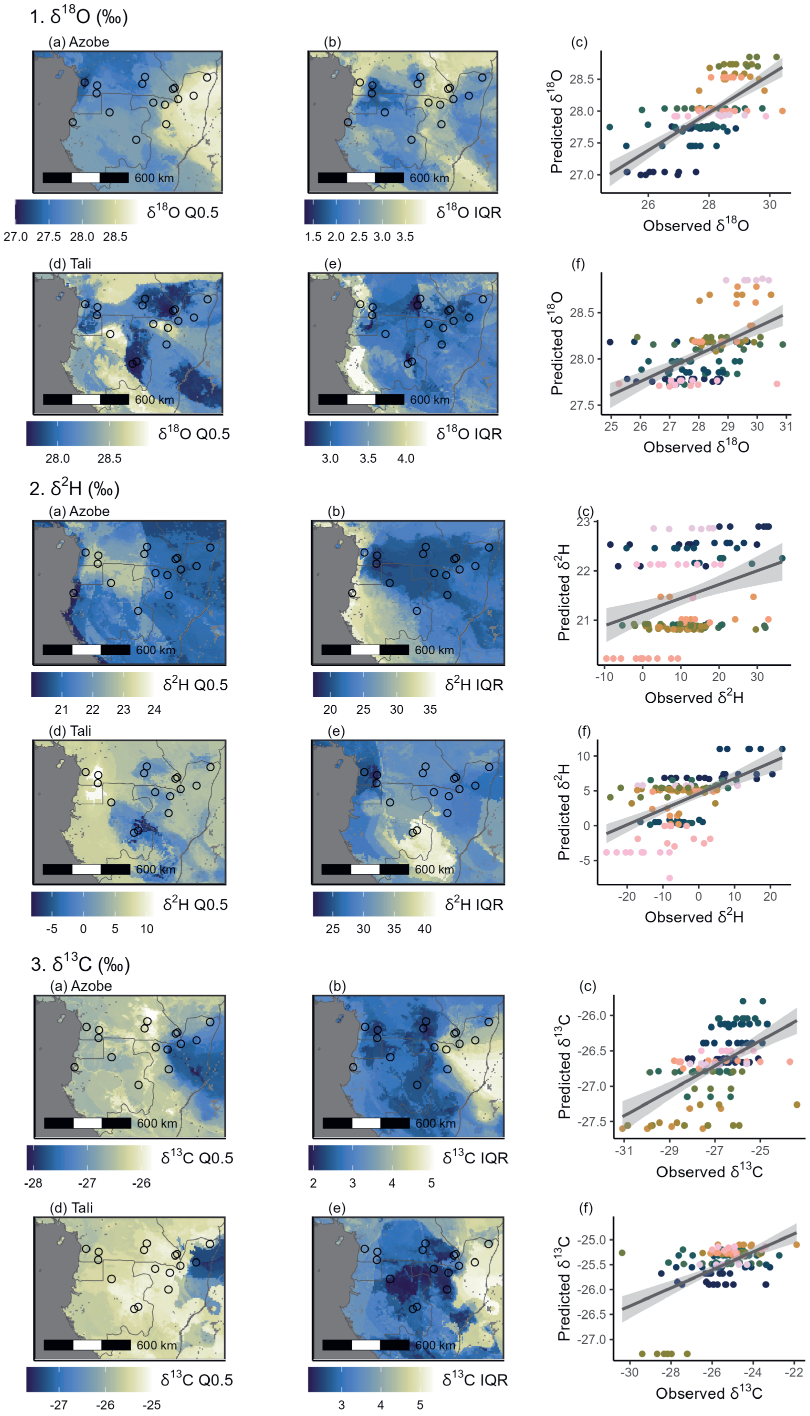
the combination of covariables, the interquantile range around the predicted median of every model was large: sometimes the minimal uncertainty around the prediction was as large as the predicted variation across the study area. In Azobé for example, the δ^2H ratio was predicted between 20 and 23 ‰ across the whole study area (Fig 5.3.2c), but the minimum IQR range was as high as 20 ‰ (Fig 5.3.2b). Assigning origin based on these isoscapes will thus result in a large potential area of origin when the model uncertainty is taken into account.

Origin assignment with multiple stable isotope ratios

Site-level classification success based on the four stable isotopes was low: correct assignment was highest for the Random Forest model predicting Azobé with 32.1% (Table 5.1). Correct site assignment for Tali was 20.5%. Furthermore, there was no shared isotopic signal among the two species that could lead to increased correct assignment when combining both species, also reflected in the significant interaction term in the two-way ANOVAs (Fig 5.2b). Still these results indicate that not all assignments were random, as that would have led to a correct assignment of (1/17 sites)*100=6%. In all models, some individual sites were predicted well, such as GAB1, CAM1 and CAM5 for Azobé. The same holds true for Tali, where the sites CAM1 and GAB5 were predicted well, and the combined model where CAM1 and GAB1 were predicted well. This was also reflected in their specific isotopic ratios (Fig 5.2b).

Country-level classification performed better, the Random forest model yielded a correct assignment of 79.4% for Azobé and 61.7% or Tali (based on random assignment this would be (1/3 countries)*100%=33.3%). None of the three countries was predicted much better or worse than the others. These country-level differences were also found in the stable isotope ratios: both country and species were significantly different (Fig 5.2a).

Lastly, we compared the distance between the predicted site and the real origin of a sample based on the individual random forest models (Fig 5.4). A decrease in assignments further away from the real origin would indicate that the samples that were assigned incorrectly were still assigned relatively close-by. Such a finding would be important as it shows at what scale the assignments based on isotopic composition has the most distinguishing potential. However, we did not find that trees were assigned more to close-by sites than to sites that were further away. About 31% and 23% of the samples (Azobé and Tali) were assigned within 25km of their respective origin, representing the correctly assigned samples. When not assigned correctly, samples were assigned close-by as well as far away (up to 800 km). In comparison, the 17 study sites themselves were on average 395 km apart (sd of 205 km), with a minimum of 16 km and a maximum of 1013 km.


1. δ18O (‰)
(a) Azobe
(b)
(c)
600 km
δ18O Q0.5
27.0 27.5 28.0 28.5
δ18O IQR
1.5 2.0 2.5 3.0 3.5
Predicted δ18O
27.0 27.5 28.0 28.5
Observed δ18O
26 28 30
(d) Tali
(e)
(f)
δ18O Q0.5
28.0 28.5
δ18O IQR
3.0 3.5 4.0
Predicted δ18O
27.5 28.0 28.5
Observed δ18O
25 26 27 28 29 30 31
2. δ2H (‰)
(a) Azobe
(b)
(c)
δ2H Q0.5
21 22 23 24
δ2H IQR
20 25 30 35
Predicted δ2H
21 22 23
Observed δ2H
-10 0 10 20 30
(d) Tali
(e)
(f)
δ2H Q0.5
-5 0 5 10
δ2H IQR
25 30 35 40
Predicted δ2H
-5 0 5 10
Observed δ2H
-20 -10 0 10 20
3. δ13C (‰)
(a) Azobe
(b)
(c)
δ13C Q0.5
-28 -27 -26
δ13C IQR
2 3 4 5
Predicted δ13C
-27.5 -27.0 -26.5 -26.0
Observed δ13C
-31 -29 -27 -25
(d) Tali
(e)
(f)
δ13C Q0.5
-27 -26 -25
δ13C IQR
3 4 5
Predicted δ13C
-27.0 -26.5 -26.0 -25.5 -25.0
Observed δ13C
-30 -28 -26 -24 -22

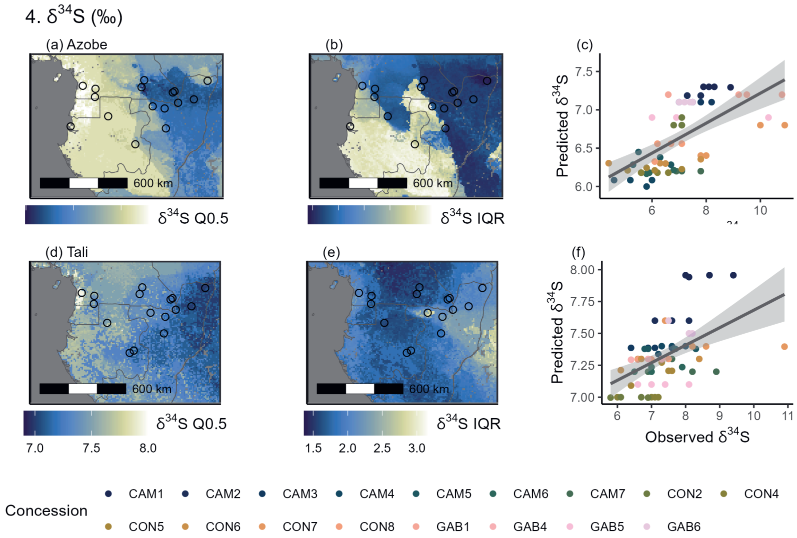


Figure 5.3: Isoscapes based on Quantile Regression Forests for $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$ and $\delta^{34}S$ (1-4) in Azobé (top) and Tali (bottom), in the study region across Cameroon, Republic of the Congo and Gabon. Isotope ratios are expressed relative to their international standards in ‰. Panels indicate the predicted median isotope ratio (a, d), the interquantile range of the predictions (IQR 0.1-0.9, b, e) and the observed versus predicted isotope ratios of the reference trees (c, f). IQR was calculated as the 0.9 quantile minus the 0.1 quantile. Circles indicate the study sites. For variable importance per Quantile Regression Forest please refer to Fig S5.1.

Table 5.1: Random forest assignment success at the site and the country level. Trees were assigned to their respective origin based on reference data of Azobé only (1, 4), Tali only (2, 5) or including both species (3, 6). Classification variables are shown in decreasing order of importance.

		Site	Classification variables in order of importance		Country	Classification variables in order of importance
Azobé	1)	32.1%	$\delta^{34}S$, $\delta^{2}H$, $\delta^{18}O$, $\delta^{13}C$	4)	79.4%	$\delta^{18}O$, $\delta^{34}S$, $\delta^{2}H$, $\delta^{13}C$
Tali	2)	20.5%	$\delta^{34}S$, $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$	5)	61.7%	$\delta^{18}O$, $\delta^{34}S$, $\delta^{2}H$, $\delta^{13}C$
Combined	3)	18.1%	$\delta^{34}S$, $\delta^{18}O$, $\delta^{2}H$, Species, $\delta^{13}C$	6)	63.7%	$\delta^{34}S$, $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$, Species

5.4 Discussion

This is the first study to test the feasibility of isotopic timber tracing with a sampling coverage that allowed quantifying species differences as well as local and regional variation. However, we found large local variation in all studied stable isotope ratios which led to low assignment accuracy at a sub-country scale in the main tropical timber production area in Africa (Cameroon,

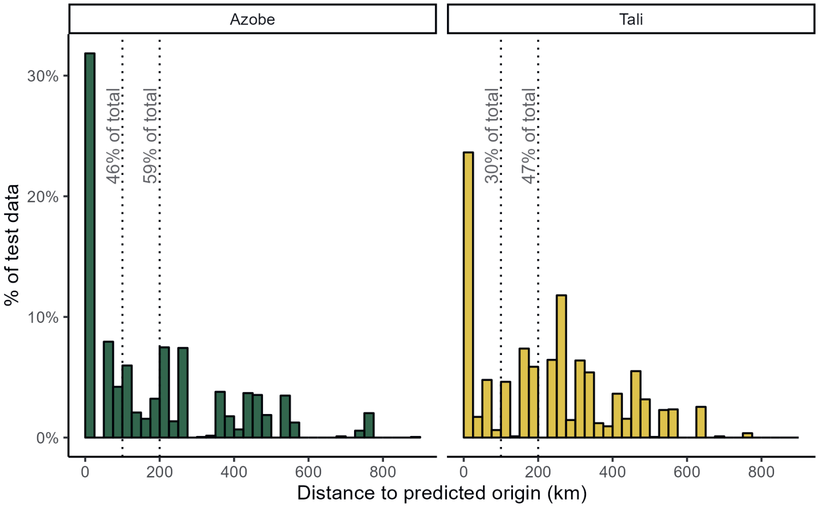


Figure 5.4: Histogram of the distance between the predicted site and the real origin (km), based on Random Forest models 1 (Azobé) and 2 (Tali, Table 5.1). Distance to predicted origin was calculated for every sample in the test set per cross validation, performed for 50 random test sets. The percentage of Azobé and Tali trees that was assigned within 100 km and within 200 km of the real origin is indicated.

Republic of the Congo and Gabon): For the two studied timbers, Azobé and Tali, a few individual sites could be separated based on δ^{18}O, δ^{2}H, δ^{13}C and δ^{34}S, but for most sites the local variation was too large to distinguish between sites. As a result, isoscapes of the area did not provide sufficient isotope variability, and consequently classification models were unable to provide a satisfactory assignment accuracy for a forensic context. Nevertheless, the accuracy of country-level classification was higher.

Stable isotope tracing within countries in Central Africa

Our results showed that for all isotopes, site variation was almost as large as the variation in median isotope ratio across the entire study area (Fig 5.2). The regional variation in wood stable isotopes was thus too limited to show sufficient differences between sites, which can be explained by a lack of spatial variation in driving environmental conditions. In Central Africa, differences in temperature, rainfall and seasonality are relatively small (Fick & Hijmans, 2017), variation in the stable isotopes in rainfall is limited within the study region (Bowen, 2023), elevation differences are mostly very gradual and not more than a few 100 meters, and the bedrock mostly homogeneous (Persits et al., 1997). As a result, strong drivers for geographical variation in the isotopic composition of wood are limited. This shows that tracing with stable isotope analysis on a sub-country scale is more challenging in Central Africa compared to other areas where the gradients in isotope ratios are more defined, such as in Central Europe (Gori et al., 2018; Boner et al., 2007; Hobson et al., 2004) or across South America (Ehleringer et al., 2000; Bowen, 2023). The latter regions have strong variation in temperature, rainfall, the isotopic composition in rainfall, elevation and/or bedrock material, which can drive the geographical isotopic variation in wood.

We did find a more unique isotopic fingerprint at some individual sites, such as CAM1 in the coastal area of Cameroon (Fig 5.2). The method could thus be used to distinguish specific combinations of sites. This is in line with previous work on timber in Central Africa, showing isotopic differences between two sites in Gabon of which one was close to the coast and the other was about 400 km inland (Watkinson et al., 2022a). δ^{34}S was found to be most important for site distinction there, which is especially driven by to distance to coast (Fig S5.1 and Nehlich (2015)). However, a pairwise site comparison is not always the most relevant question for tracing purposes. More often, a claimed origin needs to be verified, for which the claimed origin has to exhibit unique biochemical characteristics that are not found elsewhere. Based on our isoscape results we foresee that origin verification tests at the sub-national level will not improve by expanding the dataset with additional sites. On the contrary, as data collection increases to fill the sampling gaps in the area, the isotopic fingerprint of individual sites might become less unique. As a result, assignment accuracy between sites may go down further.

One of the factors most often described as a source of isotopic variation in δ^{18}O and δ^{2}H in wood is the isotopic composition in rainwater, so we tested the association between the two. However, we did not find strong positive correlations between isotope ratios in cellulose and those modelled in rainwater within our study area (Fig S5.2), contrasting what has been found on a global scale (Lehmann et al., 2022; Barbour et al., 2001). This lack of association could again arise from the low spatial variability in climatic conditions and origin of rainfall within the study region. The modelled δ^{18}O and δ^{2}H in rainwater are a function of rainfall amount and temperature (Bowen, 2023), both of which show little variation across the area. As a result, the range in modelled δ^{18}O in rainfall is 4‰ and the range in δ^{2}H is 30‰, which is very small compared to the global variation of 61.7‰ for δ^{18}O and 486.1‰ for δ^{2}H (Bowen, 2023). This variation in modelled rainwater isotopes is even small compared to the modelled variation within the African continent, which is 15.5‰ for δ^{18}O and 108.2‰ for δ^{2}H. This lack of a strong gradient in the modelled isotopes in rainfall also explains their limited importance in the random forest models, both for the isoscapes (Fig 5.3) and the classification models (Table 5.1).

Causes of local and temporal variation in isotopic ratios

In addition to naturally occurring variation, the observed local variation could also be partially attributed to methodological uncertainty and choices made in sampling design (Horacek et al., 2018). One source of uncertainty is measurement precision, as the variation may be due to measurement errors instead of actual patterns. The measurement precision of all four isotopes was high however and did not substantially contribute to the observed variation within species and sites: measurement errors were on average 0.2‰ for δ^{18}O, 3‰ for δ^{2}H, 0.1‰ for δ^{13}C and 0.3‰ for δ^{34}S. Therefore, although it may add some variation, it does not explain the large variation within our research sites.

Another source of variation are the year-to-year fluctuations in stable isotopes within a wood sample, which can cause variation between sampled trees if not enough years are pooled (Van der Sleen et al., 2015). To minimize this effect, we sampled 3-5 cm of the heartwood, thereby including multiple growth years. Tali and especially Azobé do not produce distinct annual rings but it is assumed that this sample size corresponds to at least 5 years of growth as these trees grow between 0.2 - 1 cm in diameter a year (Engone Obiang et al., 2012; Groenendijk et al., 2014). In all trees the sample was taken from the outer 10 cm of diameter growth of the heartwood, assuming limited additional variation in δ^{13}C due to changing atmospheric CO_2 composition as a result of the Suez effect (Cernusak & Ubierna, 2022). Therefore we are confident this methodological source of variation has a limited effect on our results.

Variation in δ^{13}C could also result from differences in tree age through variation in height and the amount of light it receives (Fonti et al., 2018). We only sampled trees of >30 cm DBH to exclude juvenile trees, but we still found a slight positive trend in δ^{13}C with increasing DBH (Fig S5.3), especially for Azobé. Trees of varying sizes were sampled at all sampling sites, which could thus partially explain the local variation in δ^{13}C. The Suez and juvenile effect highlight the complexity of applying δ^{13}C measurements for timber tracing, as the year of wood formation cannot be easily verified after processing at a sawmill. No trends were found between DBH and any of the other stable isotope ratios.

Lastly, we deliberately chose to measure the δ^{18}O, δ^{2}H and δ^{13}C isotope ratios in cellulose instead of whole wood, even though isotopic composition between the two is assumed to be correlated (Helle et al., 2022). Whole-wood measurements include more wood constituents such as lignin, extractives and resin, which have different isotopic signatures. As a result, the isotopic signal in whole wood is also affected by differences in wood composition such as higher lignin content (Helle et al., 2022). Cellulose, on the other hand, has a more stable isotopic signal as variation due to different biosynthetic pathways of wood constituents has been removed. This is especially relevant when aiming to identify origin in areas with low isotope variability. If measurements would be done on whole wood, a low signal of natural variability could then be masked by differences in wood composition and complicate tracing (Ferrio & Voltas, 2005). Furthermore, measuring cellulose rather than whole wood reduces differences between species as a result of wood composition, therefore facilitating a better species comparison.

Species differences

All four stable isotope ratios were different between Azobé and Tali (Fig 5.2). This species difference is in line with isotopic studies on other tropical tree species in Bolivia Paredes-Villanueva et al. (2022) and Ghana (Förstel et al., 2011) and also suggested by data from the United States and Gabon (Watkinson et al., 2022b,a). Species differences can for example result from differences in fractionation processes due to differences in water use or differences in wood constituents. More importantly however, we also found a significant interaction between site and species for δ^{18}O, δ^{13}C and δ^{34}S. This indicates the two species responded differently to the environmental variables at these sites. Species specific differences such as rooting depth and water-use efficiency could have had a strong influence on the local isotopic signal, therewith suppressing a shared geographical pattern (van der Sleen et al., 2017; Wynn et al., 2014). For tracing purposes this complicates matters, since it means that multiple species cannot easily be combined into one reference dataset for origin classification: accuracy went down in the combined models (Table 5.1). This is especially important when defining sampling designs, which are often based on the assumption that multiple species share their isotopic patterns and are therefore aimed at a low amount of samples per species per site. This also warrants a check of the two species that are sold as Tali and regarded as one in this study. The differences between Azobé and Tali samples were clearly larger than any variation within the Tali samples themselves, therefore we think that combining them here did not have a strong effect on the outcomes of our study. This lack of isotopic geographical correspondence among species should be tested further in areas with a stronger isotopic gradient (f.e. in terms of rainfall, elevation, etc), as any environmental signal may be reflected more clearly in the wood isotopic composition there.

Potential for stable isotope tracing in Central Africa

The potential for tracing at a sub-country level with stable isotopes in Central Africa has been a subject of debate in scientific literature in recent years. Site-level differentiation based on δ^{18}O, δ^{15}N and δ^{13}C was not possible for three sites in south Cameroon (Vlam et al. (2018b) and see Horacek et al. (2018); Vlam et al. (2018a)), but two sites in Gabon could be distinguished mostly based on δ^{34}S and δ^{2}H (Watkinson et al., 2022a). Compared to previous studies, which were based on samples obtained from 2-3 sites within Cameroon or Gabon, our study is much more elaborate (17 sites) and covers a considerably larger area (across Cameroon, Gabon and Republic of the Congo). While the sets of isotopes differ among these studies (Vlam: δ^{18}O, δ^{15}N and δ^{13}C; Watkinson: δ^{18}O, δ^{15}N, δ^{13}C, δ^{2}H and δ^{34}S ; this study: δ^{18}O, δ^{13}C, δ^{2}H and δ^{34}S), two to four isotopes overlap among them. With the four isotopes that have showed the highest potential for origin verification, we found limited potential for sub-country timber tracing with stable isotopes in Central Africa. This does not imply that timber tracing with stable isotopes has no potential at the sub-country scale in other areas, for example in the US (Watkinson et al., 2020; Kagawa & Leavitt, 2010), Central Europe (Gori et al., 2018; Horacek et al., 2009) or the East Mediterranean (Rich et al., 2016).

Despite the high local variability in the isotopic signal in wood hindering provenancing at the site level in the study area, we did find country-specific differences. Stable isotope ratios can thus still offer larger-scale origin identification, also in Central Africa, which can be relevant depending on the provenancing question and required scale of tracing. For example, this is relevant when international trade routes are coming together, as it is often a moment where fraud occurs (Lowe et al., 2016). These findings are in line with results on three timber species in West and Central Africa for example, where country verification success based on stable isotope ratios was between 55% and 75% (Degen et al., 2015).

Still, logically, country verification can only be achieved when natural variation in stable isotope ratios is present between origins. As the drivers of isotopic variation do not always correspond to country borders, a clear isotope signal per country might not always be found. To improve our scientific understanding of the relevant drivers of variation, the established databases here could be extended to cover more countries in West Africa as well as the Democratic Republic of the Congo. This will allow for a comparison of local and continent-wide variation, further testing the scale at which isotopic composition can aid timber tracing in Africa. However, this is less relevant for origin verification in practice as by far the largest share of timber from natural forests in Africa originates from the three countries in this study (International Tropical Timber Organization, 2021).

5.5 Conclusion

The urgency for accurate timber tracing methods is high, especially in tropical regions such as Central Africa. However, our study shows limited potential to trace origin based on stable isotope ratios within the three most important timber exporting countries in the region (Cameroon, Republic of the Congo, Gabon). Isotopic composition may have more potential on larger scales (such as between countries) or in other regions, but here the regional variation was not strong enough to identify sub-country origin as individual sites already showed high local variability in isotopic composition. Future studies on timber tracing in Central Africa should consider at what scale tracing is most relevant, depending on the types of fraud that are occurring, and investigate methods that can identify such types of fraud. This also includes small-scale fraud such as timber movement close to the border (Nellemann, 2012), for which other methods such as genetic analysis (Jolivet & Degen, 2012; Vlam et al., 2018b) or multi-element analysis (Boeschoten et al., 2022) might have more potential. On the other hand, stable isotope ratios could still be a cost-effective method to identify timber fraud across larger scales and in parts of the world where environmental gradients show strong spatial variability.

5.6 Acknowledgements

This study was supported by the Dutch Research Council (NWO-TTW-OTP-16427). Additional fieldwork support was received from the Alberta Mennega Foundation, FSC International and from World Forest ID. Additional analysis support was received from the Ecology Fund of the KNAW. We thank all the collaborating timber companies and their field teams for facilitating the fieldwork, colleagues at our partner institutes University of Dschang, Marien Ngouabi Univeristy, IRET/CENAREST and the National Herbarium of Gabon for help in arranging fieldwork, our colleagues Gerard Heuvelink and David Rossiter of ISRIC for helpful suggestions on the statistical analyses and our colleagues from the Forest Ecology and Forest Management writing group for feedback on the manuscript.

Supporting information

Table S5.1: Covariables of the Quantile Regression Forests for isotope-specific isoscapes of Azobé and Tali and for which isoscape they were used.

Covariable	Isoscape of	Covariable units	Original resolution	Source
19 bioclimatic variables	$\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$	°C (temperature) and mm (rainfall)	0.083°(∼9.3 km)	Fick & Hijmans (2017)
Modelled $\delta^{18}O$ and $\delta^{2}H$ in rainwater	$\delta^{18}O$, $\delta^{2}H$	‰	0.083°(∼9.3 km)	Annual means from Bowen (2023)
Sand content	$\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$	%	0.083°(∼9.3 km)	Poggio et al. (2021)
Elevation	$\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$, $\delta^{34}S$	m.a.s.l.	0.083°(∼9.3 km)	Hijmans et al. (2022)
Distance to the ocean	$\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$, $\delta^{34}S$	km	0.25°(∼27.8 km)	Calculated with geosphere package, Hijmans (2022)
SO_2 deposition	$\delta^{34}S$	mol m^{-2}	1 km	Copernicus Sentinel. (2023)
Soil extractable S	$\delta^{34}S$	‰	0.3 km	Hengl et al. (2021)
Bedrock type	$\delta^{34}S$	Nominal (class)	Discrete	Persits et al. (1997)

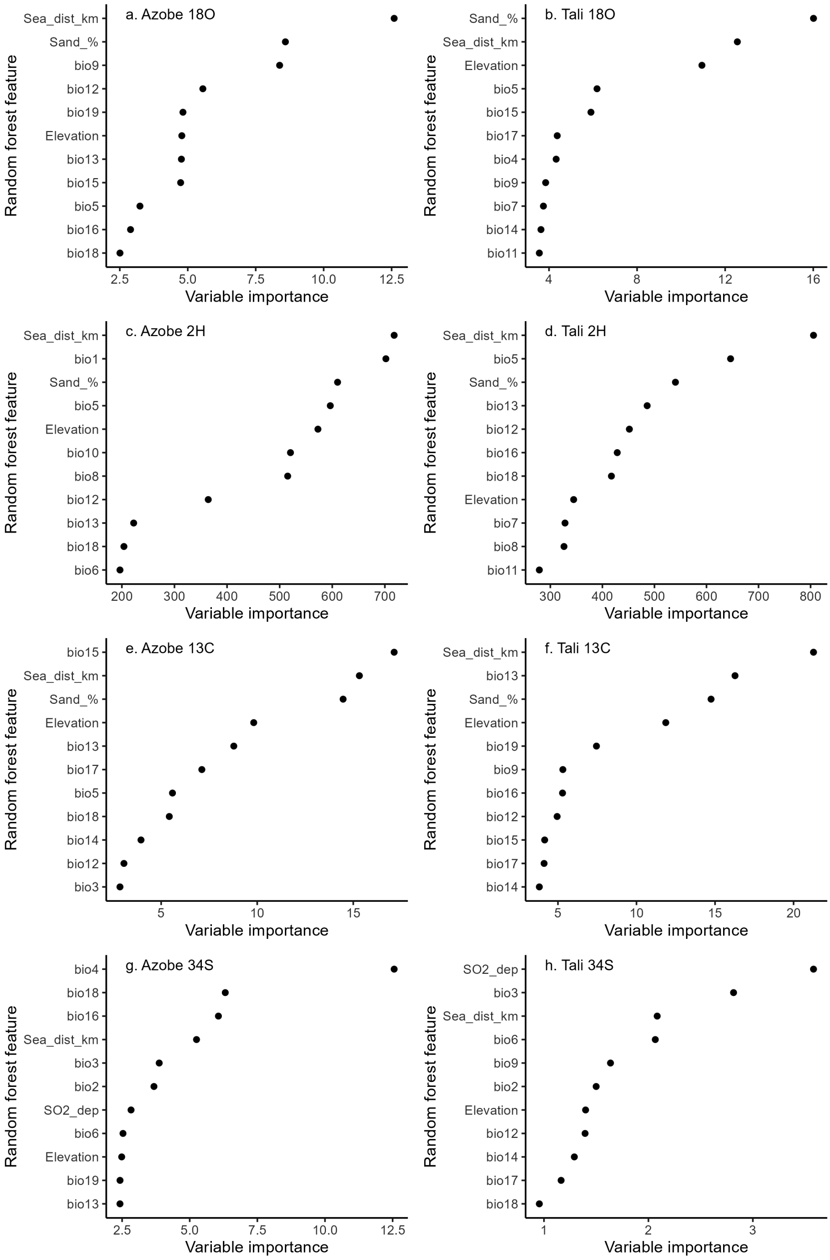


Figure S5.1: Variable importance of the most important variables for the Quantile Regression Forests per isotope per timber species, as applied to develop the maps in Figure 5.3. Bioclim variables bio1-bio11 refer to temperature-related variables, bio12-bio19 refer to rainfall related variables. Please refer to https://worldclim.org/data/bioclim.html for the individual labels.

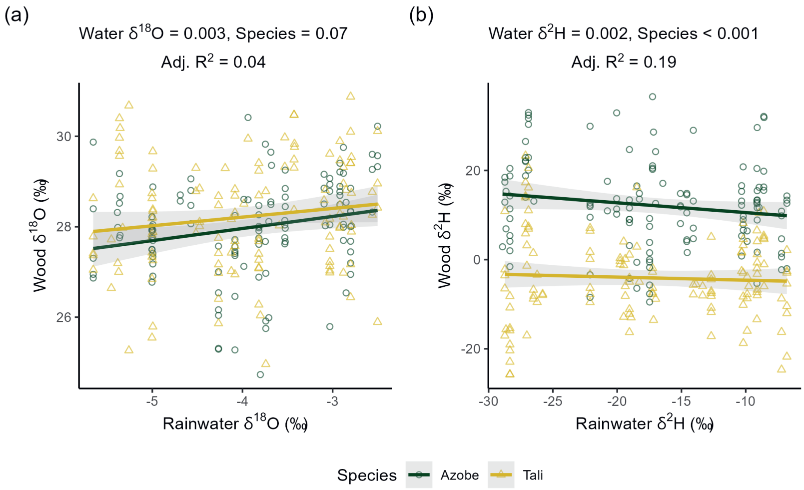


Figure S5.2: Multiple linear regressions between the measured water stable isotopes δ^{18}O (a) and δ^{2}H (b) in wood and the modelled isotopes in rainwater. Both regressions were significant (p<0.01), variable significance and adj. R^{2} are indicated above the plots. Dots indicate the individual trees, lines indicate the regression per species with shaded 95% CI.

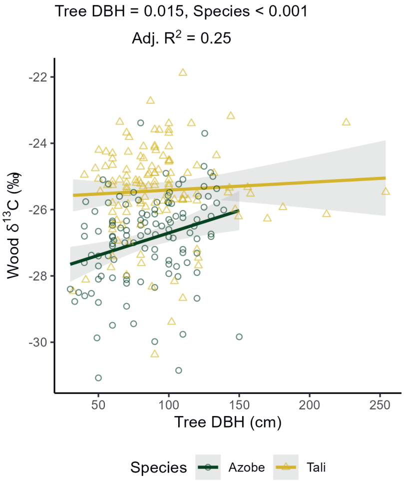
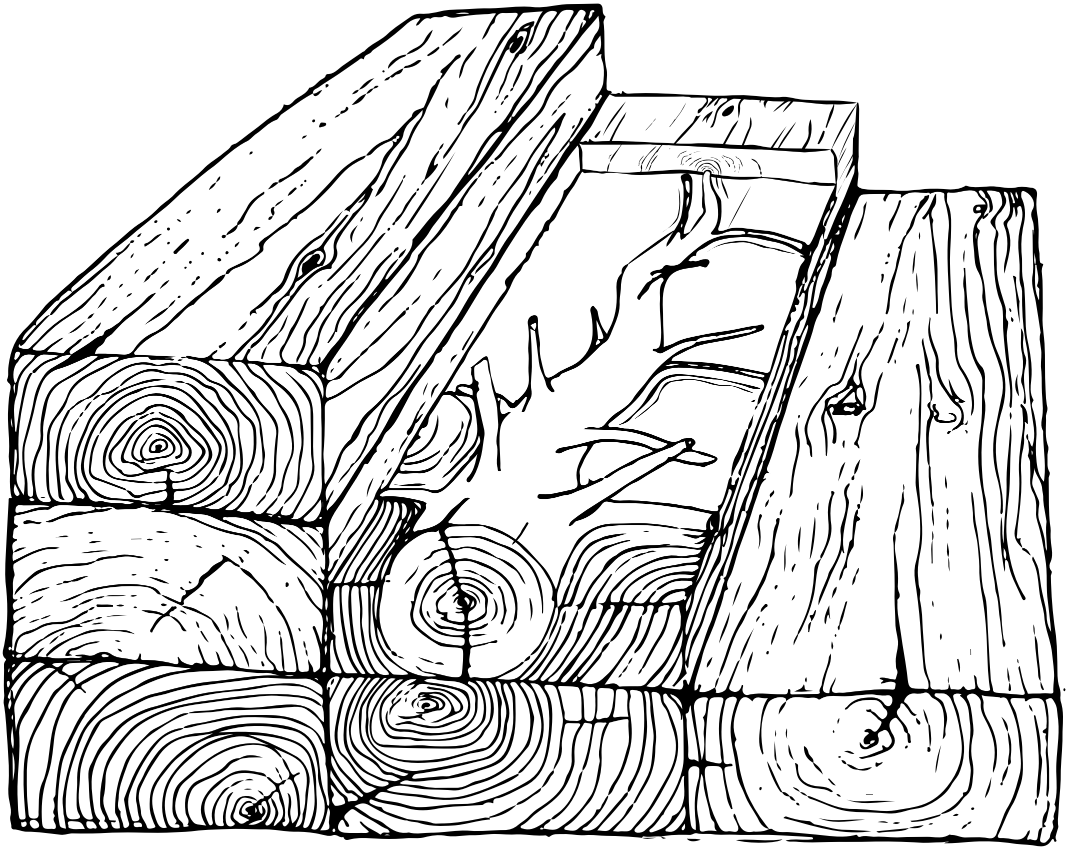


Figure S5.3: Multiple linear regression between tree DBH (cm) and δ^{13}C in wood. The regression was significant (p<0.01), variable significance and adj. R^{2} are indicated above the plot. Dots indicate the individual trees, lines indicate the regression per species with shaded 95% CI.



Chapter 6

Combining genetic and chemical methods boosts accuracy of timber tracing at small spatial scales

Laura E. Boeschoten*, Barbara Rocha Venâncio Meyer-Sand*, Arnoud Boom, Gaël U.D. Bouka, Jannici C.U. Ciliane-Madikou, Nestor Laurier Engone Obiang, Mesly Guieshon-Engongoro, Arjen de Groot, Joël J. Loumeto, Dieu-merci M.F. Mbika, Cynel G. Moundounga, Rita M.D. Ndangani, Dyana Ndiade Bourobou, Ute Sass-Klaassen, Marinus J. M. Smulders, Steve N. Tassiamba, Martin T. Tchamba, Bijoux B.L. Toumba-Paka, Mart Vlam, Herman T. Zanguim, Pascaline T. Zemtsa, and Pieter A. Zuidema

* L.E. Boeschoten and B. Rocha Venâncio Meyer-Sand contributed equally to this paper

Abstract

Improving transparency in the timber supply chain is crucial to halt illegal timber trade and promote sustainable timber trade. To support law enforcement, independent verification of timber origin is needed, which can be based on genetic and chemical wood properties. For forensic cases, these methods need to achieve a high tracing accuracy (above c. 90%) at small scales (below c. 100 km). So far, no single method has yielded both these goals for large geographic areas. Combining tracing methods may improve verification accuracy, but this has never been tested.

In this study, we collected wood samples of Azobé (*Lophira alata*), an economically important timber, from 13 georeferenced sites in three Central African countries. For a varying number of trees per site, we (1) detected genome-wide chloroplast Single Nucleotide Polymorphisms (cpSNP's), (2) measured three stable isotope ratios (δ^{18}O, δ^{2}H and δ^{34}S) and (3) analysed the multi-elemental composition. Based on the three methods, we developed Random Forest models for site classification and tested their individual as well as combined assignment accuracy.

The accuracy with which samples were assigned to their site of origin varied between 43% and 73% when applying single methods. Combining cpSNP's and multi-element analysis boosted assignment success to 83% at site level and yielded correct assignment of 90% within 100 km. Combining all three methods further increased accuracy to 94% within 100 km, and even yielded a 90% accuracy at 50 km distance. Methods complemented each other, such that a low accuracy of one was compensated by a high accuracy of another.

Our study shows that by combining tracing methods unprecedented accuracy levels in assigning timber origin can be attained at small spatial scales across three countries. It further demonstrates the potential of combining methods to boost tracing accuracy and resolution for timber, and likely can also be applied to illegal wildlife and agricultural commodities.

Keywords

Origin verification, machine learning, illegal logging, cpDNA, stable isotope ratio, wood chemistry

6.1 Introduction

Illegal timber trade is a global problem that contributes to deforestation, biodiversity loss and climate change (Hoare et al., 2020; Kleinschmit et al., 2016). To address this, various regulations and certification schemes have been implemented in order to reduce illegal trade Lowe et al. (2016); May (2017). However, verifying the claimed origin of wood products remains a challenge and fraudulent practices are still widespread, especially in tropical countries (Hoare & Uehara, 2022). Forensic methods that can verify timber origin are therefore crucial to ensure the legality and sustainability of timber products.

Multiple forensic methods are currently under development to enable independent verification of the origin of wood. These methods use intrinsic wood properties to verify the origin, such as genetic or chemical variations within species (Dormontt et al., 2015; Low et al., 2022). Three promising tracing methods for tropical timber include genetic differentiation, stable isotope ratios and multi-element analysis. Genetic methods remain the most applied method for origin testing of timber to date (Low et al., 2022). The population genetic approach is based on the principle that individuals that are geographically closer are also genetically more closely related. This relationship arises from historical demographic processes, which have led to spatial variation and consequently to the link between genetic differentiation and geographical distance. Second, stable isotope ratios also have a long history of origin identification research in wood as well as in other commodities (Ehleringer et al., 2000; Gori et al., 2018). For timber, the bio-elements δ^{18}O, δ^{2}H, δ^{13}C, δ^{15}N and δ^{34}S are most used for tracing (Ch 4+5, this thesis, Watkinson et al. (2020, 2022a); Gori et al. (2018)). The drivers of geographical variation in these isotope ratios in wood are mostly related to climate, geology and/or deposition (West et al., 2010; Cernusak et al., 2022; Wynn et al., 2014; Lehmann et al., 2022; Allen et al., 2022; van der Sleen et al., 2017). The third method is based on multi-element analysis (Boeschoten et al., 2022, 2023). A large number of elements (such as Mg, Ca, La) is measured simultaneously using mass spectrometry and based on this elemental composition, an origin-specific chemical fingerprint is defined. Geographical variation in elemental composition is caused by variation in physical and chemical soil properties (Boeschoten et al., 2022).

While each of these methods holds great potential for identifying wood origin (Vlam et al., 2018b; Ng et al., 2017; Watkinson et al., 2020; Boeschoten et al., 2023; Dormontt et al., 2020), their effectiveness in tracing origins at short distances remains challenging. The accuracy of tracing wood origins at small spatial scales (<100 km) has not consistently reached a reliable accuracy to be used as a tracing method at those distances. At short distances, high tracing accuracies based on genetic or isotopic methods have been limited to studies with a low number of sites, or to some specific species and/or some regions (Vlam et al., 2018b; Watkinson et al., 2022a), whereas in other studies only large geographical regions could be distinguished (Ng et al. (2017); Watkinson et al. (2020), Chapter 4, this thesis). This was also the case for multi-elemental analysis: some sites that were 50 km apart could be distinguished well, but a few sites that were >200 km apart showed a non-discriminatory chemical composition (Boeschoten et al., 2023).

As different mechanisms underly the variation used to distinguish between different origins, each method has its own strengths and weaknesses. As a result, regions where one of the methods performs poorly may coincide with those where one of the others yields high tracing accuracy. Therefore, tracing accuracy is expected to improve when combining multiple methods (Dormontt et al., 2015; Low et al., 2022), but this has never been tested for tropical timber. In archaeological wood research, combining multiple methods (anatomical proxies as well as strontium isotope analysis) indeed yielded higher tracing precision at small spatial scales (Akhmetzyanov et al., 2019; D'Andrea et al., 2023). In other commodities, the combination of stable isotopes and multi-element analysis also yielded good tracing results (Wang et al., 2020; Drivelos & Georgiou, 2012). Furthermore, the combination of stable isotopes with genetic analysis has been applied to assign migratory birds to their breeding grounds and the combination also greatly reduced the area to which birds were assigned (Chabot et al., 2012).

The objective of this study was to test the potential for origin verification of timber of three established forensic methods (genetics, stable isotope ratios and multi-element analysis) individually as well as combined. We did so by answering the following questions: 1) what is the natural variation underlying these forensic methods and is there a clear spatial structure as a basis for tracing purposes? 2) what is the assignment accuracy at a site level of individual as well as combined forensic methods? and 3) what method or combination of methods reaches the highest accuracy at small spatial scales (<100km)? To answer these questions we focused on an economically important tropical timber species from Central Africa, traded as Azobé (*Lophira alata*). We applied all three methods to the same trees at 13 locations across the main timber exporting countries in the region (Cameroon, Gabon, and Republic of the Congo). To specifically test small-scale origin assignment, we tested tracing performance at the site level, as some sites were as close as 10 km. With this research we aimed to improve the understanding of how these tracing methods performed individually when applied to the same trees, and we aimed to test their potential for combined use.

6.2 Methods

Sample collection

The study was conducted on the commercial Central African timber species Azobé (*Lophira alata* Banks ex C.F. Gaertn, Ochnaceae). Azobé occurs from Guinea to the Democratic Republic of Congo (GBIF.org, 2021), but almost all timber export is originating from Cameroon, Gabon, Republic of the Congo and Democratic Republic of Congo. Sampling was conducted in 13 logging concessions in Cameroon, Gabon and Republic of the Congo between September 2019 and April 2022 (see Fig 6.1). All sites were natural forest concessions. Sampling was conducted in collaboration with the operating forestry companies.

At each site, we sampled heartwood as well as bark of 20-30 trees per timber species. Target trees within one site were located between 100 m and 5 km apart. Sampled trees were either standing or recently felled and were of at least 30 cm diameter at breast height (DBH). We sampled trees of different sizes at all sites, to allow conclusions independent of ontogenetic stage. The size variation of the sampled trees was comparable across all sites. The heartwood sample was collected from each tree as an increment core (Haglöf Increment borer 350 mm x 5,15 mm; n = 27), with a FAMAG plug cutter of 15 mm diameter (n = 322), as a wood chunk (n = 20) or as a wood powder sample obtained with an electrical drill (n = 133). All samples were taken at least 14 cm into the tree. The heartwood samples were stored in plastic straws or paper envelopes and properly ventilated to prevent mould. Additionally, three cambium samples per tree were taken with punches of 2.5 cm diameter. The bark samples were stored in plastic bags with silica. The silica was refreshed up to 1 week after sampling to ensure the samples were dried fully. Additionally, GPS-coordinates and DBH were recorded.

Lab methodology

Genetic analysis

DNA was isolated from leaf, cambium or sapwood material from between 17 and 28 trees per site. Genomic material was isolated with an optimised cetyltrimethyl ammonium bromide (CTAB) protocol (Dumolin et al., 1995) with additional cleaning steps (Supplementary Note S6.1). DNA purity of all extracts was checked with Nanodrop (Thermo Fisher Scientific, Schwerte, Germany), DNA concentrations were measured with the Qubit™ kit (Thermo Fisher Scientific, Schwerte, Germany) following the manufacturer's instructions, and 1.5% agarose gel was used to check the fragments length range. The DNA isolates were used to prepare three libraries, 300 bp or more insert size, with the 'RIPTIDE High Throughput Rapid Library Prep Kit' (Twist Bioscience, South San Francisco, USA). The libraries were sequenced with Illumina Novaseq6000 PE150 (Novogene, Cambridge, United Kingdom). The Illumina sequences of the trees were mapped to the annotated chloroplast genome (MZ274135.1, Mascarello et al. (2021)) using Bowtie2 (Langmead & Salzberg, 2012)). The variant call considered all mapped reads without filtering and was performed using NGSEPcore (Tello et al., 2023). A variant call file containing only biallelic loci was generated, and further analyses were carried out using R version 4.1.0 (R Core Team, 2021). The detected variants underwent further filtering, including a minimum sequencing depth of three reads, a maximum depth of 250, and individuals with more than 50% missing data as well as SNPs with more than 25% missing data were removed (SNPfiltR package, DeRaad (2022)). This resulted in a genetic dataset of 322 trees. Haplotypes were defined in a non-restrictive manner with inclusion of sequences of varying lengths within the same haplotype, while ambiguities resulted in certain sequences being assigned to different haplotype (Pegas package, Paradis (2010)).

Stable isotope analysis

Between four and 10 trees per site were selected for stable isotope analysis, depending on the isotope: we measured δ^{34}S in three to four trees per species per site (total of 51), δ^{2}H in six or 10 (total of 101) and δ^{18}O in six or 10 (total of 105) trees per site. Trees were selected for isotopic measurements such that multiple isotopes were measured for the same set of trees as much as possible. This resulted in a geolocated database of 295 Azobé trees in total, with between 1 and 3 isotopes measured per tree. δ^{13}C was measured in 12 sites, but was not found to add to the assignment accuracy. As measurements were missing at one site, this isotope was not included in the final model.

Stable isotope measurements are explained in detail in Chapter 5. In short, a subsample of heartwood from every tree was cut in radial direction including at least 3-5 cm to include wood formed during multiple years. δ^{18}O and δ^{2}H were measured in cellulose, following (Vlam et al., 2018b) for cellulose extractions. δ^{34}S was measured in whole wood. The stable isotopes ratios were determined by IRMS, expressed in per mill (‰) relative to an international reference standard (V-SMOW for δ^{18}O and δ^{2}H and CDT for δ^{34}S).

Multi-element analysis

Twenty trees per site were selected for multi-element analysis. The wood chemical composition was measured following Boeschoten et al. (2022). In short, a 1.0 g subsample was cut from 3-5 cm of heartwood and dissolved in 70% HNO3 by heating in a microwave digestion system (Mars 6, CEM Cooperation, USA). The lowest detection limit per element was calculated as three times the intensity of that element in a blank standard. If elements were found in quantities below the detection limit in more than 100 samples (the equivalent of about a quarter of the trees), they were excluded. This resulted in a multi-elemental composition of 41 elements, measured in 398 Azobé trees.

Statistical analysis

All other statistical analyses were performed in R version 4.2.3 (R core team, 2023). To test the variation between sites for each of the forensic methods, a db-RDA was performed based on 1-proportion of shared alleles for the cpDNA (PopGenReport package, Gruber & Adamack (2022)) and on Chord distances for stable isotope ratios and multi-element analysis (vegan package, (Oksanen et al., 2020)).

Site classification

Site classification models were developed based on Random Forest analysis (ranger package, (Wright & Ziegler, 2017)). Random Forest models cannot be developed with missing values in the dataset, therefore imputation steps were performed to fill the gaps in the dataset. The cpSNP's dataset (haplotypes) contained the highest number of trees per site (up to 28). Missing data in the SNP dataset were randomly distributed across cpDNA and individual trees. Therefore, the most common allele was used at the population of origin to impute missing data, which may have resulted in a slight over-estimation of the accuracy of the Random Forest assignment: due to the imputation, some trees may have ended up with the same haplotype while in fact they were very similar but not identical, but the effect on site assignment was regarded to be minimal.

In the dataset of stable isotope ratios, less individuals per site were measured than for the genetic analysis (between 4 and 10 per isotope), because it was expected that local variation was smaller in the isotope ratios. To match the size of the genetic dataset, isotopic data was imputed for the non-measured trees. We first randomly selected a subset of trees for which all measurements were performed for δ^2H and δ^{18}O and we excluded those from the dataset as test set, to minimize the effect of the imputation on further assessment of the assignment models (50% of all fully measured trees, so 44 test trees, spread across all 13 sites). For the remaining dataset, missing values were imputed with the site means.

For the multi-element analysis, 20 trees were measured per site. Elemental concentrations for the non-measured trees were imputed with site means to match the genetic dataset, like the stable isotope ratios. Again, only trees that were measured directly were used as test trees to calculate assignment accuracy.

We then developed seven types of Random Forest models based on different reference datasets: one per forensic method (SNP, ISO, EL), one for each pairwise combination of method (SNP-ISO, SNP-EL, ISO-EL), and one with all three methods included (SNP-ISO-EL). We tested their assignment accuracy as % of correctly assigned test trees. Test trees did not contain any imputed data. The model development was repeated 25 times per Random Forest model type (with different random subset of test trees), resulting in 7 x 25 assignment models. For each model, the sites to which test trees were assigned were saved and distances between actual and assigned origin were calculated. Assignment accuracy and mismatches were visualized in donut charts, based on the 25 x 44 test trees per Random Forest model category.

6.3 Results

Genetic analysis

After filtering, a total of 238 cpSNP's for 322 individual trees composed the genetic dataset. Three main genetic clusters were observed based on proportion of shared haplotypes among sites (Fig S6.1). The first main cluster consisted of 10 sites, spanning from West Cameroon, to North-West Congo and Centre-East Gabon. This cluster showed a clear substructure with sites CAM1 and CAM3 in West Cameroon being the most genetically distant from the other nine sites (Fig 6.1A). The second cluster, composed of GAB1, was the most genetically distinct cluster. The clear distinction of GAB1 indicates a possible species effect, as the site is in the region where a cryptic *L. alata* species occurs (L. alata1 - West Gabon, Ewédjè et al. (2020)). To a lesser extent, the third cluster was also distinct from the other two clusters. It was composed of two sites in North Congo (CON2 and CON4), and such a cluster was not previously detected using nuclear microsatellite markers (Ewédjè et al., 2020).

We used a Random Forest classification model to assign trees to their most likely site of origin based on the 238 cpSNP's. Assignment success to the correct site was 50.4% ± 3.7% (Fig 6.1A, 6.2), but it showed large variation across sites. Whereas no trees were correctly assigned at some sites (CAM2 and CAM4, Fig 6.1A), at other sites correct assignment was 100% (GAB1, CAM5 and CON4). Furthermore, the genetic groups corresponded to a clear spatial structure: incorrect assignments often occurred between groups of sites that were geographically close, such as between CON2 and CON4. This was also clear from the distances to predicted origin: 72.4% of the trees was assigned within 100 km of the sample site, 89.2% within 300 km, and no trees were assigned to sites more than 500 km away (Fig 6.2). This reflects a high precision.

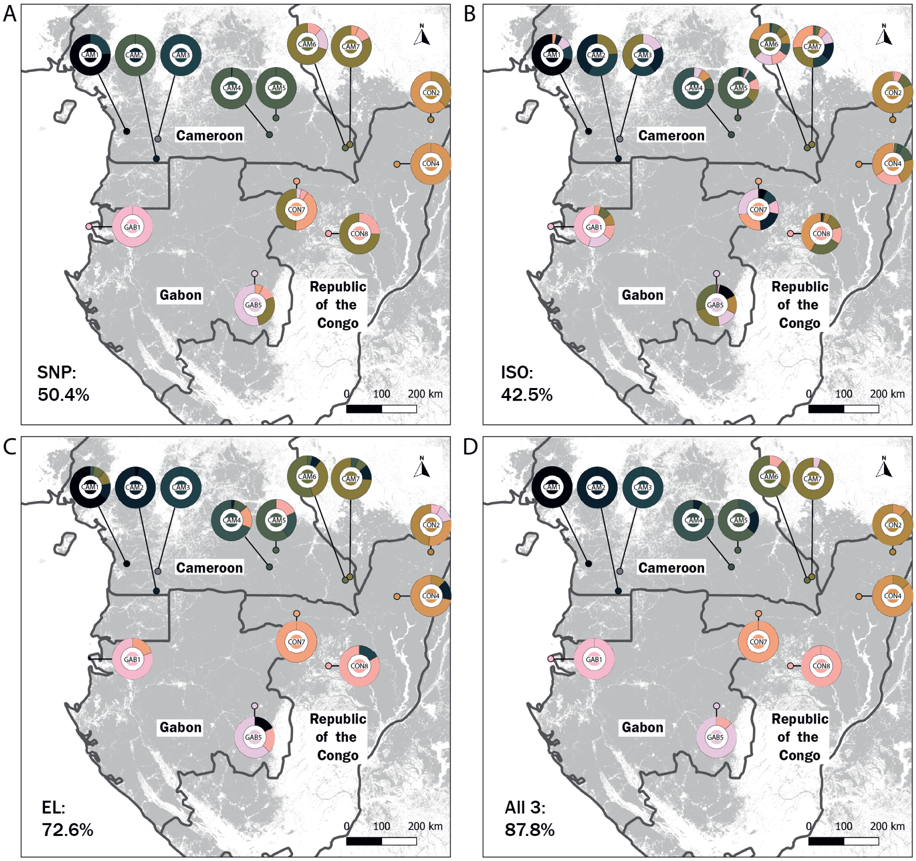


Figure 6.1: Confusion charts of the assignments of trees to their respective origin based on the different forensic methods: A) cpSNP's (SNP), B) three stable isotope ratios (ISO), C) multi-element analysis (EL) and D) all three combined. Mean assignment accuracy is indicated in the Figure. Colours in the inner circle indicate the colour of the unique site. Colours in the outer circle indicate to which site the trees of that location were assigned. Primary tropical forest extent from Global Forest Watch is indicated in light grey (Turubanova et al., 2018).

Stable isotope analysis

Overall, stable isotope ratios ranged from 24.7 - 30.4 ‰ for δ^{18}O, -9.5 - 36.55 ‰ for δ^{2}H and 4.6 - 10.9 ‰ for δ^{34}S . Local variation in isotopic ratios was high, which resulted in a standard deviation within the sites that was similar to the variation in isotope ratios across all sites (Table S6.1). However, even with this high isotopic overlap among sites, site was significant

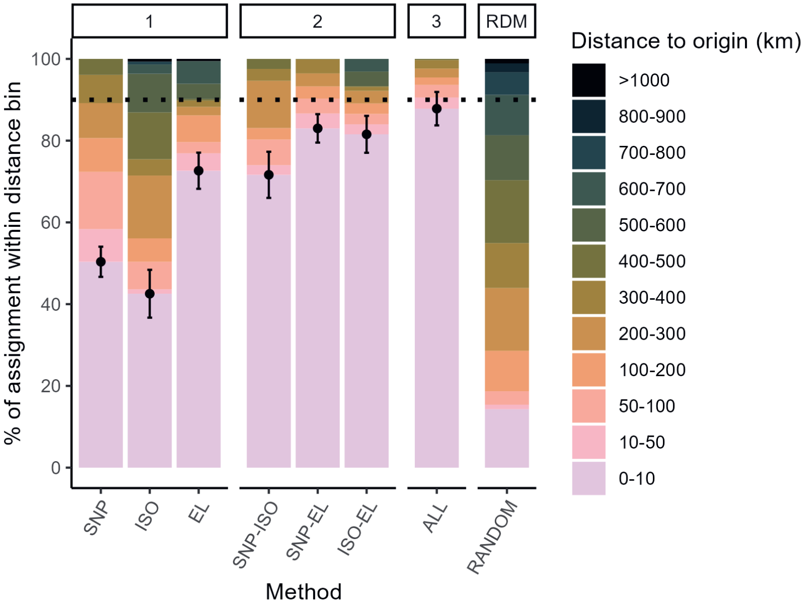


Figure 6.2: Distance to predicted origin (km), as % of the total number of test trees per distance category (bin) per method. The lowest bar represents correct site origin assignment, which is the assignment accuracy, also represented by a dot with an error bar that indicates the variation (st. dev.) in assignment accuracy, estimated using different test and training datasets of the Random Forest models. SNP = cpSNP's, ISO = stable isotope ratios, EL = multi-element analysis, ALL = all three methods, RDM = the occurrence of pairwise distances in the reference dataset, representing the result of random assignments. The dotted line indicates 90% accuracy.

in the db-RDA ($p<0.001$, Fig S6.2). Especially δ^2H was a strong predictor for the first axis, which explained 60.6% of the total variation.

Despite the significant differences among sites, assignment accuracy was rather low. Site-level Random Forest classification success was 42.5% ± 5.9% (Fig 6.1B, 6.2). Even though assignment success for some individual sites was around 75% (CAM1, CAM4 and CON2), it was much lower for most other sites. Furthermore, there was no clear spatial structure in the assignments. Trees were assigned to many different sites (up to eight out of 12), only 50.4% of the trees was assigned within 100 km and misassignments were as far as 1000 km from the original origin (low precision, Fig 6.2).

Multi-element analysis

Concentrations of the 41 measured elements varied between 0.001 µg/kg (Yb) and 4.0 g/kg (K) in the wood samples. Differences between sites were clear, as site was a significant predictor in the db-RDA ($p<0.001$, Fig S6.3).

The Random Forest classification model based on the multi-elemental concentration performed better than that of the other methods: 72.6% ± 4.4% of trees was assigned to the correct site of origin (Fig 6.1C, 6.2) and 79.6% of the trees was assigned within 100 km of the sample site. Important elements for the classification based on variable importance were W, Ba, Mo, K and Cr. Sites with a high assignment accuracy were CAM2, CAM3 and CON7, where all trees were assigned correctly. However, across sites, trees that were not assigned to the correct site could be assigned to sites as far as 1000 km away (low precision, Fig 6.2). This contrasts with the stronger spatial structure in the cpSNP's data, for which distances to actual origin of mis-assigned individuals were much lower (high precision, Fig 6.2).

Combining methods

All the pairwise combinations of forensic methods increased assignment accuracy compared to the single methods (Fig 6.1D, 6.2), indicating that the methods complement each other. Especially combinations including multi-element analysis resulted in a high accuracy (83.0% and 81.5%). These were also the most consistent methods across different Random Forest models, indicated by a low standard deviation (Fig 6.2).

The combination of all three methods improved site assignment accuracy even further, to a correct assignment of 87.8% ± 4.1%. It also yielded good precision: 90.6% of all trees was assigned within 50 km of the sample site and 93.6% of within 100 km. Furthermore, no trees were assigned more than 500 km away (Fig 6.2).

The complementarity of the methods was reflected in the individual site confusion charts (Fig 6.1), as the sites with the highest accuracy were not the same across the three methods. For example, three in west Cameroon (CAM1, CAM2 and CAM3) showed 100% correct site assignments based on the combination of methods, whereas none of the individual methods could distinguish all three well.

6.4 Discussion

This is the first study that compared and combined genetic and chemical tracing methods to assign the origin of a tropical timber, Azobé. The combination of three methods (genetic differentiation based on chloroplast SNPs, stable isotope ratios, and multi-element analysis) resulted in a high accuracy for origin identification at small spatial scales: 87.8% of the trees was assigned to the correct site of origin and 93.6% was assigned within 100 km of its origin. This was considerably higher than the assignment accuracy of the individual methods (which ranged from 42.5% to 72.6%).

Combining methods improves tracing accuracy at small spatial scales

There was clear chemical and genetic variation within the study region (Figs S6.1-S6.3). As a result, site assignments of all three tracing methods performed better than random (Fig 6.2). Multi-elemental analysis performed best at these 13 locations (72.6%), followed by assignments based on cpSNP's (50.4%) and stable isotope ratios (42.5%). However, the genetic profiles were more region-specific than the other two methods: no trees were assigned more than 500 km away based on cpSNP's (Fig 6.2). In comparison, some trees (6.1%) were assigned more than 600 km away based on multi-element analysis, which would result in assignment to another country in this context.

Most importantly, the three methods showed distinct spatial patterns: the uniquely distinguished sites differed between the three methods (Fig 6.1). It was clear that this was thanks to the distinct clusters and/or sites that were observed: the genetic clusters enhanced the spatial resolution of the chemical tracing for example. As a result, combining the three methods led to increased performance (87.8% correct at the site level, 93.6% within 100 km) and a reduction of the misassignment distances to less than 500 km (Fig 6.2). This complementarity can be explained by the different drivers of each method, thus resulting in a mosaic of spatial variation.

The reported combined assignment accuracy is comparable to the highest reported accuracies in other studies applying genetic analysis. Within Cameroon, blind samples of Tali were assigned to three sites as close as 14 km with 92% accuracy (Vlam et al., 2018b) and a batch of *E. cylindricum* samples was correctly verified to a concession of 2.500 km2 with 86% accuracy (Jolivet & Degen, 2012). Yet, not all genetic studies found such differences at small spatial scales. Population assignment of *Shorea platyclados* in Malaysia yielded only 60.6% correct assignment at distances between 10 and 300 km (Ng et al., 2017) and for *Cedrela odorata* in Bolivia the assignment accuracy was only 66.3% at distances between 268 and 501 km (Paredes-Villanueva et al., 2019).

Compared to previous studies on individual chemical methods, the accuracy of the combined assignment exceeds any previously reported assignment at small spatial scales. As a comparison, *Aucoumea klaineana* timber from two locations in Gabon that were roughly 400 km apart differed in isotopic composition (δ^2H, δ^{18}O, δ^{13}C, δ^{15}N, δ^{34}S), but no assignment tests were performed (Watkinson et al., 2022a). Contrastingly, Tali timber (*Erythrophleum suaveolens* and *E. ivorense*) from three sites in Cameroon that were between 14 and 216 km apart could not be distinguished isotopically (δ^{18}O, δ^{13}C, δ^{15}N, Vlam et al. (2018b)). Overall, isotopic tracing has proven particularly successful across large environmental gradients, such as those caused by elevation (tested for *Picea abies*, Gori et al. (2018)) and over larger distances (tested for *Quercus* spp., Watkinson et al. (2020), Chapter 5, this thesis). However, region of origin could not even be distinguished based on δ^{13}C and δ^{18}O in four *Cedrela* species across a 500 m elevation gradient, spanning 400 km in the Bolivian Andes (Paredes-Villanueva et al., 2022). For multi-element analysis, the only existing other study was conducted on Tali for an overlapping set of sites, yielding similar assignment accuracy as for Azobé here (68.8% vs 72.6%, Boeschoten et al. (2023)).

In contrast to previous studies, that mainly achieved high accuracy within a limited spatial extent or across countries, our study showcases the tracing potential with high accuracy across three countries while focusing on small spatial scales. To further improve small scale timber tracing for Azobé in Central Africa, we suggest improvements of the individual methods. Specifically for the genetic analysis, the development of a few nuclear DNA markers with high individual discriminating power in addition to the chloroplast DNA may add more genetic sub-structure to differentiate at the site level (Duminil et al., 2013; Blanc-Jolivet & Liesebach, 2015). Additionally, for all three methods it will be essential to cover the sampling gaps within the study region. Even though the current sampling was extensive, large parts of exploited forests remain unmeasured, such as in central Gabon or in the Democratic Republic of the Congo. Therefore, genetic and chemical profiles in these regions remain unknown.

Combining methods by stacking predictive maps

When sampling effort increases and more sites are added to the database, we also propose a statistical improvement for combining these different data types into a standardized approach for origin verification, illustrated in Figure 3. This would move tracing methods forward from the level of site verification to identifying a region of origin with a certain level of certainty, which strongly reduces the large dependency on reference datasets from all potential areas of origin (Truszkowski et al., 2023). The idea is based on prediction maps of isotopic composition, known as isoscapes ((West et al., 2010), Chapter 4+5, this thesis). Based on the same principle, interpolated maps could also be made for each individual element based on relevant covariables such as soil pH and clay concentrations (Boeschoten et al., 2022). Furthermore, the frequency at which different genotypes occur at the different sites can be mapped based on Kriging statistics (Bucci et al., 2000; Manel et al., 2003). By stacking all these different maps, the potential area of origin can then be defined for any blind sample with a specific accuracy. This also requires the development an uncertainty map along with each prediction map, therewith explicitly including uncertainty in the assignment. This approach does not require every method to be applied to the same individual trees as maps are developed separately per method, making combining methods more flexible.

To develop these stacked maps, the first requirement is to include more sites for all methods. Traditional statistical interpolation techniques such as kriging require a minimum of at least 30-50 datapoints and these should be scattered throughout the study area, so that the distance at which spatial autocorrelation occurs can be estimated (Webster & Oliver, 2007). In the context of our study in Central Africa, this translates to at least 30-50 sampled sites that can accurately capture the genetic and chemical variations across the region. Interpolating the chemical data is relatively straightforward due to the continuous nature of the data, facilitating interpolation, as well as due to the availability of known gridded covariables that can enhance the model (West et al., 2010). However, the categorical genetic data requires the calculation of haplotype frequencies per site for the development of an interpolated map. This requires a higher number of measured trees per site. While the strictness of the minimum site requirement may vary for other interpolation methods like Quantile Regression Forests, the spatial autocorrelation between sampling points needs to be estimated in all cases to determine prediction uncertainty (Veronesi & Schillaci, 2019). Therefore, additional sampling should cover a variety of distances, to be able to have a good indication of the spatial structure found for each of the methods. The additional sites should cover the areas in which Azobé is harvested most, so that the resulting maps are relevant to be used in practice.

The main advantage of this approach lies in the potential of combining databases that don't fully overlap. The reason is that maps are produced for individual methods and only combined when assigning a (blind) sample. Therefore, the selection and distribution of sites can vary across tracing methods, depending on needs, limitations, strengths and costs of each individual method. This allows much more flexibility in the combination of methods compared to the method used in this study, which required all data to be collected at the same sites. This approach may also allow combining maps produced by different tracing projects or organizations.

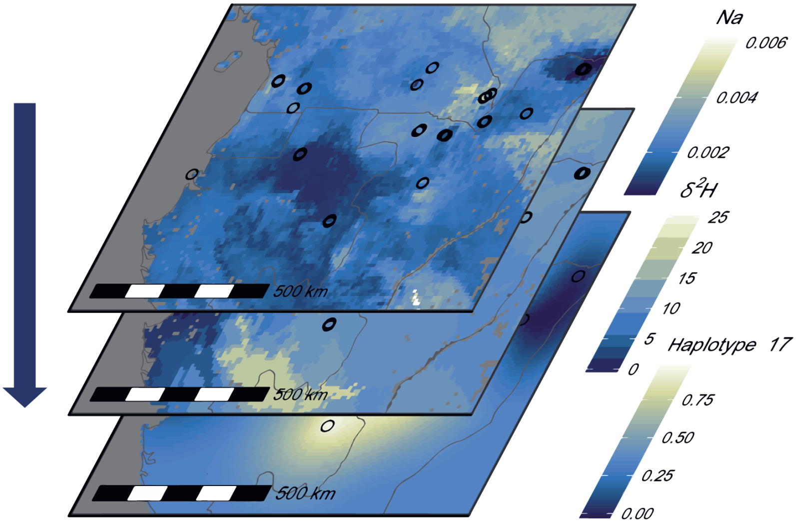


Figure 6.3: Examples of spatially interpolated maps of the study region. Individual maps based on properties of sampled trees may be integrated with maps of gridded covariables such as soil properties (for element maps) or environmental variables (for isotope maps). Prediction maps can be based on machine learning methods such as Quantile Regression Forest (Chapters 4+5, this thesis), or more classical spatial interpolation statistical techniques such as Kriging (Takoutsing & Heuvelink, 2022).

Implementing combined forensic tracing methods for other species

Our study reveals promising results of combining tracing methods for one species in one region. Evidently, the combination of tracing methods should be tested for other species as well as in other geographic regions. We anticipate that the complementarity of these tracing methods will be highest where the environmental gradients driving variation in wood genetic and chemical properties are smallest. This likely also explains why in our study region the combination of methods worked well. In Central Africa, no strong geographical barriers that would genetically separate tree populations are present, environmental variation driving wood isotopic signatures is limited (Chapter 4, this thesis), and variations in bedrock or soil type are minimal in some areas as well (Boeschoten et al., 2023). These factors limit the potential for the individual tracing methods, but offer opportunities to reach the required accuracy levels by combining methods. Contrastingly, in regions where significant physical barriers or environmental gradients are present, such as mountain ranges, a single method may already be sufficiently reliable to address tracing issues with high accuracy, making the investment in multiple forensic methods unnecessary. Clearly, these hypotheses on the possible roles of geographic barriers and environmental gradients will need to be tested further.

6.5 Conclusion

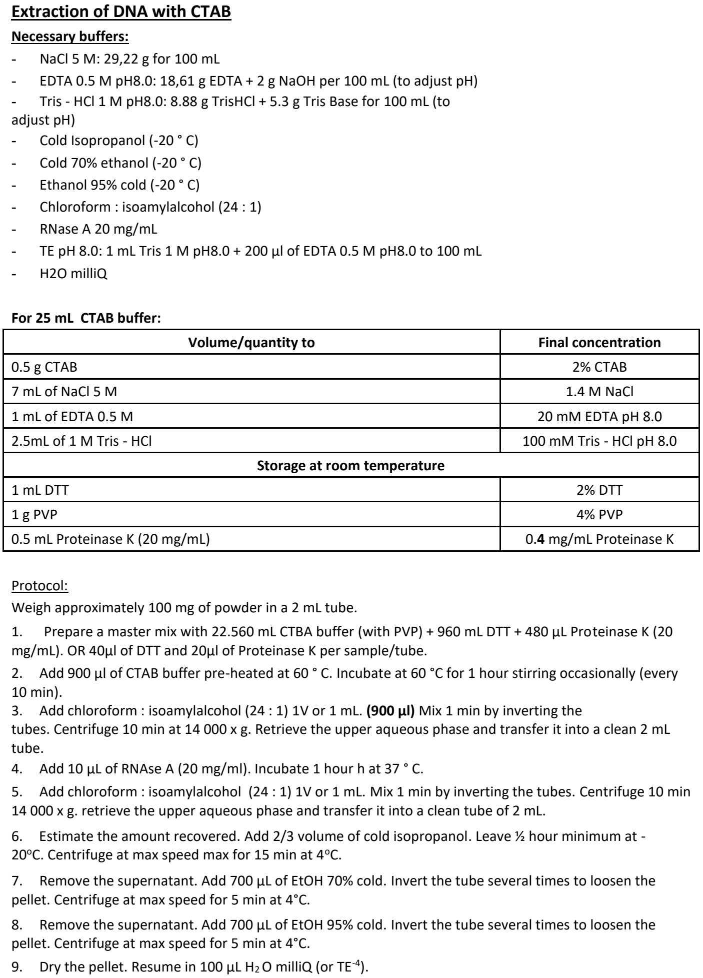
The combination of three forensic tracing methods (population genetics, here with cpSNP's, stable isotope ratios and multi-element analysis) yielded highly reliable results to trace tropical timber origin at small spatial scales in Central Africa. The combined methods attained an unprecedented accuracy level in assigning the origin of timber at small spatial scales, and did so across three countries. This shows that the combination of methods is reliable and highlights its potential to be applied to address timber legality questions. Especially in areas with limited geographical barriers and environmental variation, such as in Central Africa, the combination of methods can reach highly accurate tracing results.

6.6 Acknowledgements

This study was supported by the Dutch Research Council (NWO-TTW-OTP-16427). Additional fieldwork support was received from the Alberta Mennega Foundation, FSC International and from World Forest ID. Additional analysis support was received from the Ecology Fund of the KNAW. We thank all the collaborating timber companies and their field teams for facilitating the fieldwork and all involved colleagues at our partner institutes University of Dschang, Marien Ngouabi Univeristy, IRET/CENAREST and the National Herbarium of Gabon for help in arranging fieldwork.

Supporting information

S6.1 Modified DNA isolation protocol



Extraction of DNA with CTAB

Necessary buffers:

- NaCl 5 M: 29,22 g for 100 mL
- EDTA 0.5 M pH8.0: 18,61 g EDTA + 2 g NaOH per 100 mL (to adjust pH)
- Tris - HCl 1 M pH8.0: 8.88 g TrisHCl + 5.3 g Tris Base for 100 mL (to adjust pH)
- Cold Isopropanol (-20 ° C)
- Cold 70% ethanol (-20 ° C)
- Ethanol 95% cold (-20 ° C)
- Chloroform : isoamylalcohol (24 : 1)
- RNase A 20 mg/mL
- TE pH 8.0: 1 mL Tris 1 M pH8.0 + 200 µl of EDTA 0.5 M pH8.0 to 100 mL
- H2O milliQ

For 25 mL CTAB buffer:

Volume/quantity to	Final concentration
0.5 g CTAB	2% CTAB
7 mL of NaCl 5 M	1.4 M NaCl
1 mL of EDTA 0.5 M	20 mM EDTA pH 8.0
2.5mL of 1 M Tris - HCl	100 mM Tris - HCl pH 8.0
Storage at room temperature	
1 mL DTT	2% DTT
1 g PVP	4% PVP
0.5 mL Proteinase K (20 mg/mL)	0.**4** mg/mL Proteinase K

Protocol:

Weigh approximately 100 mg of powder in a 2 mL tube.

1. Prepare a master mix with 22.560 mL CTBA buffer (with PVP) + 960 mL DTT + 480 µL Proteinase K (20 mg/mL). OR 40µl of DTT and 20µl of Proteinase K per sample/tube.
2. Add 900 µl of CTAB buffer pre-heated at 60 ° C. Incubate at 60 °C for 1 hour stirring occasionally (every 10 min).
3. Add chloroform : isoamylalcohol (24 : 1) 1V or 1 mL. **(900 µl)** Mix 1 min by inverting the tubes. Centrifuge 10 min at 14 000 x g. Retrieve the upper aqueous phase and transfer it into a clean 2 mL tube.
4. Add 10 µL of RNAse A (20 mg/ml). Incubate 1 hour h at 37 ° C.
5. Add chloroform : isoamylalcohol (24 : 1) 1V or 1 mL. Mix 1 min by inverting the tubes. Centrifuge 10 min 14 000 x g. retrieve the upper aqueous phase and transfer it into a clean tube of 2 mL.
6. Estimate the amount recovered. Add 2/3 volume of cold isopropanol. Leave ½ hour minimum at -20ºC. Centrifuge at max speed max for 15 min at 4ºC.
7. Remove the supernatant. Add 700 µL of EtOH 70% cold. Invert the tube several times to loosen the pellet. Centrifuge at max speed for 5 min at 4°C.
8. Remove the supernatant. Add 700 µL of EtOH 95% cold. Invert the tube several times to loosen the pellet. Centrifuge at max speed for 5 min at 4°C.
9. Dry the pellet. Resume in 100 µL H_2O milliQ (or TE^{-4}).

Table S6.1: Mean and standard deviation of the three isotope ratios, averaged within sites as well as across sites, expressed relative to the respective international standards in ‰.

	δ^{18}O	δ^{2}H	δ^{34}S
Mean of the averages within sites	28.0	14.4	7.2
SD within sites	0.7	7.7	0.8
Mean across all sites	27.9	12.7	7.2
SD across sites	1.1	10.7	1.4

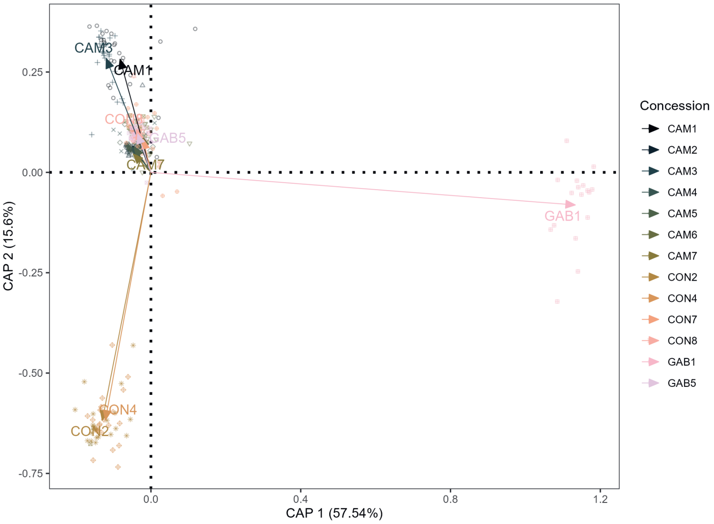


Figure S6.1: Db-RDA analysis of the genetic distances in the cpDNA, calculated based on 1-proportion of shared alleles across 13 sites, using 238 cpDNA loci for 322 trees. The coloured dots indicate the trees, coloured by site.

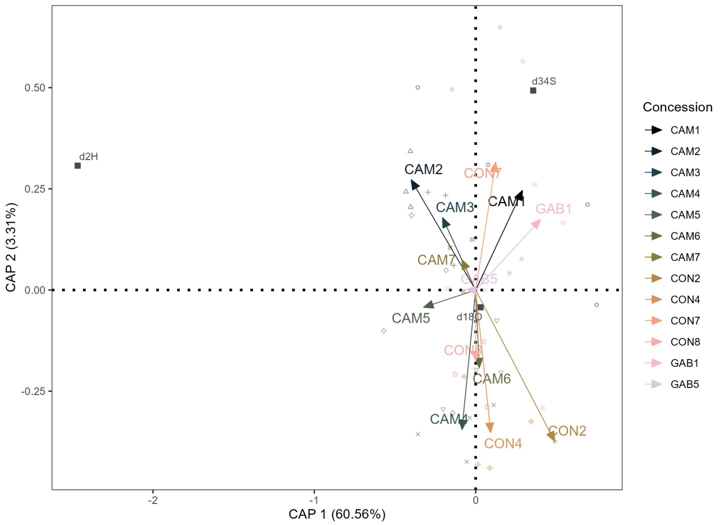


Figure S6.2: Db-RDA analysis of isotopic composition of the 45 trees across 13 sites in which all three isotope ratios were measured (d18O, d2H and d34S), based on Chord distances. The coloured dots indicate the trees, coloured by site. The grey dots indicate the elements.

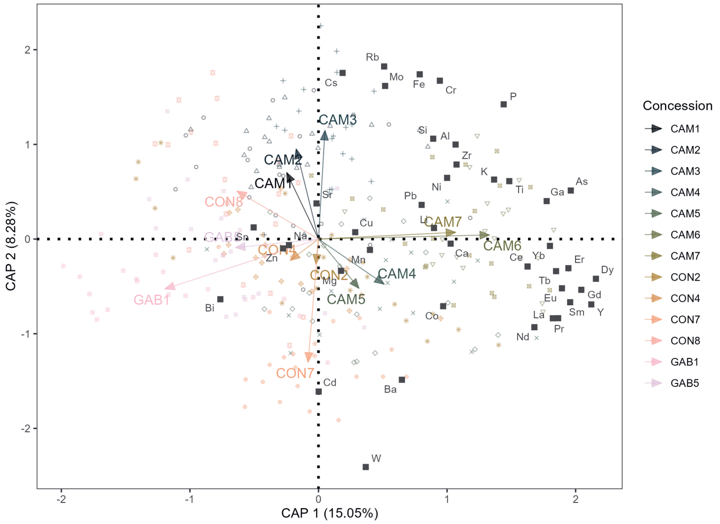
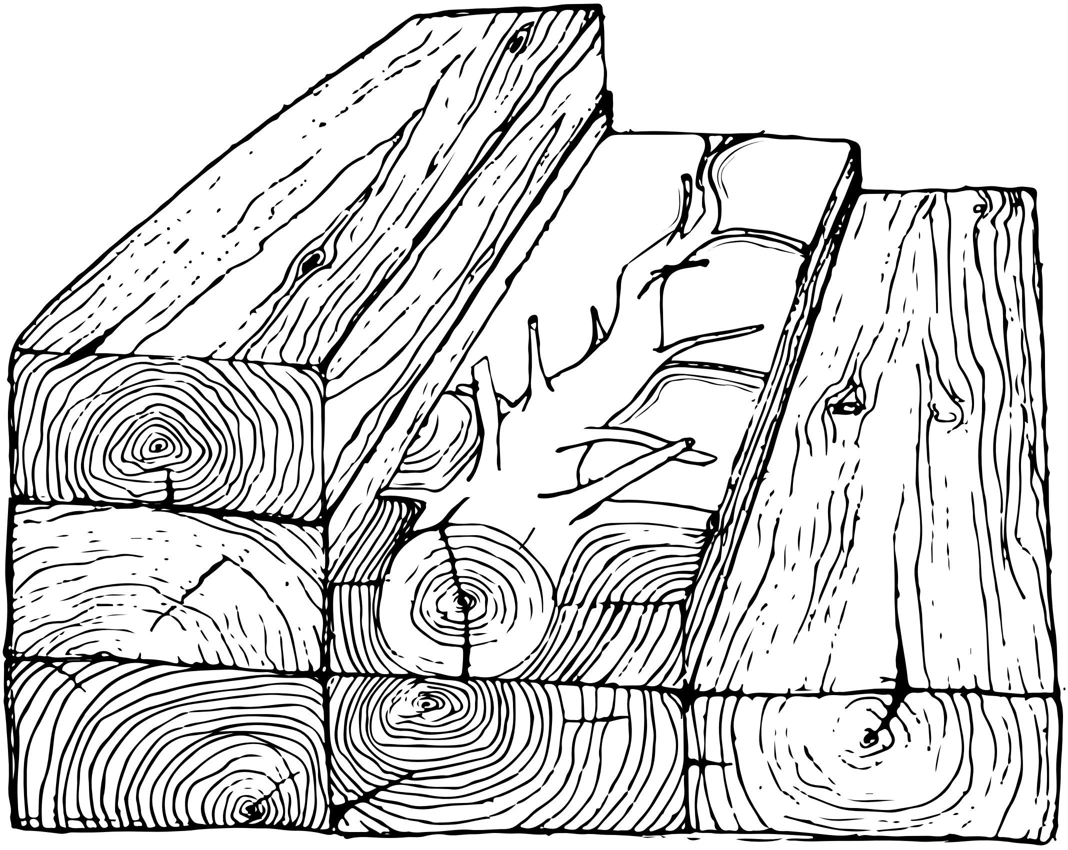


Figure S6.3: Figure S3 Db-RDA analysis of the multi-elemental composition of 259 trees across 13 sites, based on Chord distances. The coloured dots indicate the trees, coloured by site. The grey dots indicate the elements.



Chapter 7

General discussion

7.1 Forensic methods for timber tracing

Transparency in the timber supply chain is critical to combat illegal logging and promote sustainable timber trade (Lowe et al., 2016). Illegal practices remain widespread, with an estimated 4-30% of the global trade made up of illegal timber, despite attempts to halt this through regulations and law enforcement (Hoare & Uehara, 2022; Wiedenhoeft et al., 2019). In tropical regions, illegal exports are estimated to be even higher, ranging from 50-90% (Hoare & Uehara, 2022). Given the complexity of the timber supply chain and the many borders that timber crosses, a global response is needed to enhance transparency and promote legal and sustainable forest management (Uehara, 2023). This will require action and cooperation within and between timber producing, processing, and consuming countries (Bisschop, 2012; Zeitlin & Overdevest, 2021; Boekhout van Solinge et al., 2016).

The main types of timber fraud involve false reporting of species and/or origin (Gan et al., 2016), with origin fraud being the focus of this thesis. Origin fraud occurs in many ways and at various scales; a schematic overview is presented in Figure 7.1, categorized based on the scale at which it takes place. Fraud of the first category (A) occurs within the legal timber operations, where trees are harvested that should not have been cut within a concession. Fraud of the second category (B) is related to plantation establishment, for example by mixing illegal timber with plantation products. Fraud of the third category (C) includes all cases where timber was logged illegally elsewhere within the country and laundered by mixing it with legal timber or falsifying paperwork to make it seem legal. Lastly, the fourth category (D) includes laundering practices across borders. The only way to detect these types of false reported origin after the wood has entered the complex global trade network is through independent origin verification, as it cannot be checked by paperwork alone (Lowe et al., 2016). This is where forensic science can make a difference: methods for timber tracing can be used to verify the paperwork accompanying timber, to improve transparency along the timber supply chain and to detect illegal practices (Dormontt et al., 2015).

The toolbox for forensic timber tracing includes methods that are based on chemical, genetic and structural variation in wood from different origins, as those cannot be falsified (Table 7.2, Dormontt et al. (2015); Low et al. (2022)). All methods are based on a reference database that includes samples of known origin. To then trace a sample of unknown origin, it is compared to the reference database and based on the similarity with the reference samples an origin claim can be accepted or rejected.

However, barriers remain to implement these forensic methods and make them fully operational. The **first main barrier** is related to the reference databases: for many traded species no georeferenced databases are available (Low et al., 2022; Lowe et al., 2016). This can be improved by an increased effort to extend the databases through collecting and analyzing additional samples (Gasson et al., 2021). However, we can never measure every tree and it might still take years before sufficient reference databases are available to trace origin of priority species (Low et al., 2022). Therefore, the dependency on extensive reference databases can also be alleviated by developing methods that are non-species specific and/or that allow interpolation in between sampling points (Truszkowski et al., 2023). Both would greatly reduce the sample number that is required for accurate origin identification.

The **second main barrier** is related to the spatial scale at which the methods operate. The scale at which tracing can be successful depends on geographic variation in chemical, genetic and structural characteristics of the wood within a tree species, as this determines how specific the 'fingerprint' of a location is. For the best results, this variation needs to be small at the local scale, so that noise is limited within a specific origin. At the same time, the variation needs to be large across different areas, i.e. larger distances, to differentiate between origins. The natural variation in wood characteristics thus determines the types of fraud that can be detected with a specific method (Dormontt et al., 2015; Low et al., 2022). Origin verification at small spatial scales (10s of km) is especially important as it can identify the largest range of fraudulent practices (Gan et al., 2016; Nellemann, 2012), but tracing results have not been consistently high at those scales. Especially for chemical methods such as stable isotope ratios, the spatial resolution is limited as variation in some isotopes occurs at large spatial scales, sometimes continental (Vlam et al., 2018b; Paredes-Villanueva et al., 2022). This limits the potential of catching timber fraud within a country.

In this thesis I contributed to removing these barriers in timber tracing, with a specific focus on chemical methods. I did so through showing the potential of a new timber tracing method, multi-element analysis, which proved to have great potential at small spatial scales (barrier 2, Chapters 2 and 3). Furthermore, the elemental concentrations showed non-species specific patterns, which significantly reduces the amount of reference samples needed for accurate tracing (barrier 1). Additionally, I contributed to the reference database of stable isotope tracing and developed isoscapes that show potential at a continental and a country level (barrier 1, Chapters 4 and 5). Lastly, by combining multiple forensic methods for the first time (multi-element analysis, stable isotopes and genetic analysis) it was clear that the combination of tracing methods can really improve tracing accuracy at small spatial scales (barrier 2, Chapter 6).

Other methods that have been investigated for timber tracing include DART-TOFMS (Musah et al., 2015), NIRS (Silva et al., 2018) and dendrochronology (Sass-Klaassen et al., 2008; Kuhl et al., 2023), among others. Furthermore, origin specificity might be increased by combining multiple proxies, such as by combining more isotopes (Watkinson et al., 2022a) or combining stable isotopes with other methods (Wang et al., 2020). As the last review on the applicability of these methods has been published eight years ago (Dormontt et al., 2015), I will discuss the strengths and weaknesses of these methods for origin verification of timber here (see Table 7.2 for the overview).

A distinction in timber tracing is to be made between origin verification and origin identification (Chapter 3, Deklerck (2023)). The goal of verification is to confirm whether timber originated from a specific location ('Did this timber come from location X?'), while identification aims to determine the specific source of the timber ('Where did this timber come from?'). While these questions are clearly related, they do require different statistical techniques. For practical

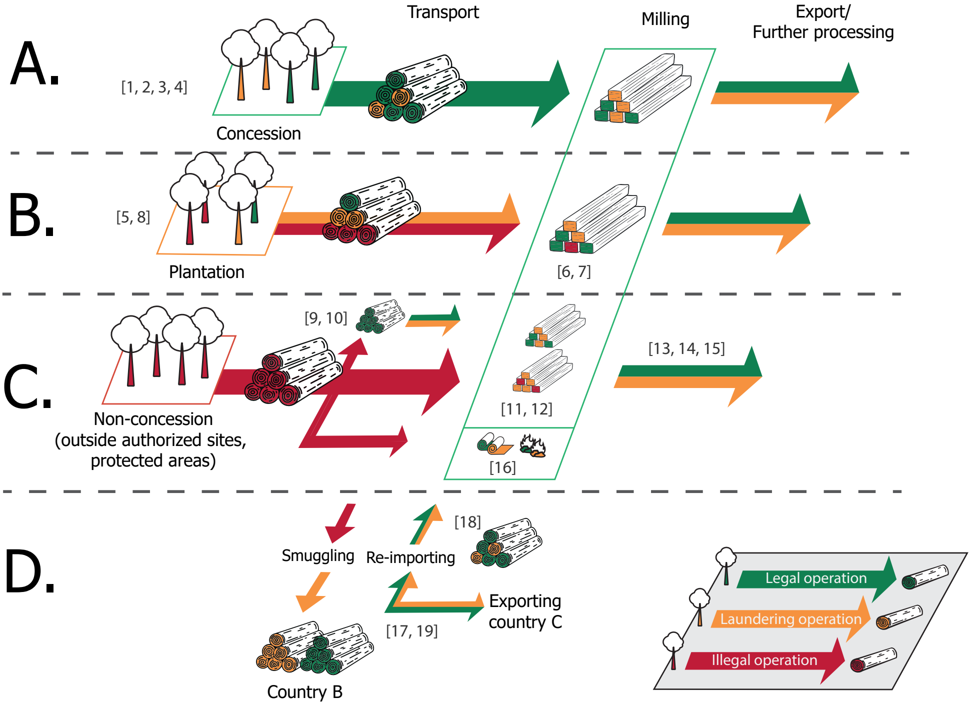


Figure 7.1: Examples of laundering practices throughout the timber supply chain. Color indicates legality of the timber: green is legal, red is illegal and orange is laundered timber that has become 'legal'. The numbers correspond to Table 7.1, representing examples of laundering practices from literature. Sustainable and/or legal timber originates from well-managed concessions, is transported to a log yard and subsequently to a mill after which it is processed before or after export. The four rows correspond to the categories at which fraud occurs: A) represents examples of fraud occurring within the legal forest operation itself, B) represents examples of fraud related to plantation establishment, C) includes all types of fraud where timber was illegally harvested elsewhere in the country and mixed with legal logs as well as through further processing to obscure origin, and D) includes laundering practices that involve exporting timber to other countries and claiming origin elsewhere. This is an extended version of Figure 1 of Chapter 1 and it is based on Nelleman et al. (2016); Lowe et al. (2016).

Table 7.1: Illegal timber laundering practices as reported in literature. Fraud categories correspond to Figure 7.1: A is illegal logging within a legal concession, B is related to plantation establishment, C includes all fraud where timber was logged illegally elsewhere in the country and D includes laundering practices across borders. ME = multi-element analysis, ISO = stable isotope ratios, DART = DART-TOFMS, DNA = DNA fingerprinting and population genetics, DC = Dendrochronology and Comb = a combination of ME and DNA. Potential of the different tracing methods to catch that type of fraud is indicated as none (empty), uncertain (+/-), low (+), medium (++) and high (+++).

Nr	Example	Cat	Potential for tracing						
			ME	ISO	DART	NIRS	DNA	DC	Comb
1	Exceeding cutting quotas on-site [1, 2]	A							
2	Cutting wide corridors along new roadways [1]	A							
3	Getting ‘permits’ through bribes and then operating as ‘legal’ forest operation [3]	A							
4	Cutting beyond concession borders [3]	A	+				+		++
5	Cutting down existing forest to ‘establish’ a legal plantation [1]	B	++		+/-	+/-	++	++	++
6	Obtaining a permit for plantation production for mills and funnelling illegal timber through the non-productive plantation permit [1]	B	++	+	+/-	+/-	++	+++	++
7	Laundering by mixing illegal wood with legally produced plantation products [1]	B	++	+	+/-	+/-	++	+++	++
8	Selling illegal timber as part of legal land clearing operations for palm oil/soy plantations or ranching establishment [1]	B	++		+/-	+/-	++		+++
9	Mixing logs by transporting illegal timber to a legal forest operation [1]	C	++	+	+/-	+/-	++		++
10	Using permits or logging concessions in one area to cut in a different area [1]	C	++	+	+/-	+/-	++		+++
11	Mixing illegally logged timber with legal logs at a sawmill or pulp mill [1]	C	++	+	+/-	+/-	++		+++
12	Over-stating the percentage of wood extracted in mills per logs processed, leaving room for illegal timber [1]	C	++	+	+/-	+/-	++		+++
13	Using existing export permits to export illegally logged timber originating from another part of the country [1]	C	++	+	+/-	+/-	++		+++
14	Multiple re-use of existing export permits and certificates for export in harbours [1]	C	++	+	+/-	+/-	++		+++
15	Using forged permits or permits obtained illegally through hacking of government sites or bribery. [1, 4]	C	++	+	+/-	+/-	++		+++
16	Processing roundwood into paper or charcoal before exporting [3]	C	+	+			+	+	
17	Exporting illegal logs cross-border from origin country A, and exporting as ‘originating’ from country B [1, 5]	D	+++	++	+	+	++	+	+++
18	Exporting logs illegally from origin country A to country B, then re-importing to a mill in country A as “legal” import from B [1]	D	+++	++	+	+	++	+	+++
19	Exporting illegal logs from country A to country B, and through trading it multiple times on paper the logs are owned by a company from country C by the time it arrives [1, 5]	D	+++	+++	+	+	++	+	+++
20	False declaration of tree species on customs papers, especially for endangered or rare species [1]	A-D	+/-	+/-	+++	+++	+++	+++	

From [1] Nellemann (2012), [2] Brancalion et al. (2018), [3] Nelleman et al. (2016), [4] Global Witness (2019), [5] Gan et al. (2016).

application of forensic methods in timber importing countries, origin verification will be relevant most often as it involves verifying the claimed origin of a piece of wood. However, for understanding trade routes as well as for the development of forensic tracing methods, it is also relevant to test whether origin can be identified, so whether a sample can be assigned to a specific origin. For example, origin identification can help a great deal in uncovering illegal logging practices at the source, which is not possible with verification alone. Moreover, testing the potential for origin identification is a crucial step in the development of forensic methods as it gives insights into how variation is spread across the different origins (Chapter 3). Therefore, in this thesis I focused more on origin identification then verification, as potential for identification suggests good potential for verification as well.

7.2 The available methods for forensic timber tracing

Multi-element analysis

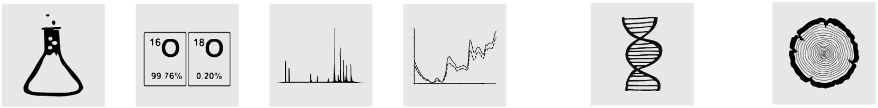
Multi-element analysis was first described for timber tracing in this thesis. The novel method is based on the quantification of a large set of chemical elements in the wood. The organic material is digested, after which up to 60 elements are measured simultaneously by for example Inductively Coupled Plasma Mass Spectrometry (ICP-MS). The elements in the wood then provide a specific chemical profile or ‘fingerprint’ per area, which is used for origin identification. The method had been applied for origin verification of commodities such as cocoa, cabbages and capers (Acierno et al., 2020; Pepi & Vaccaro, 2018; Bong et al., 2012) and was successful at continental and regional spatial scales (Joebstl et al., 2010; Baroni et al., 2015).

Elemental composition of wood was found to be linked to soil physical and chemical properties, indicating a high timber tracing potential in areas where spatial differences in soil properties occur (Chapter 2, Štulc et al. (2023); Hevia et al. (2017)). Furthermore, the potential of multi-element analysis was confirmed for origin identification of timber at a sub-country scale in Central Africa and Borneo (Chapter 3). The scale at which variation in chemical composition was detected in two important source areas for tropical timbers makes this a valuable new tool for origin identification, as distinct chemical patterns were measured already at distances of 50 km.

Strengths

The main strength of multi-element analysis includes the spatial variation at small spatial scales (Chapter 2, 3). Differences were found at small spatial scales (50-100 km), even more so on Borneo than in Central Africa, thus showing high specificity for timber tracing (barrier 2). Furthermore, the variation in Central Africa included non-species specific patterns. This is a first indication that multiple species can be combined in one reference database from which to build classification models, thus partly alleviating the dependency on species-specific reference databases (barrier 1). Additionally, a main advantage of the method is the speed and relative ease of use: over 60 elements can be measured at once, providing a large basis

Table 7.2: Simplified overview of the discussed forensic methods for timber tracing, based on Dormontt et al. (2015) and Low et al. (2022). The comparison includes: lab processing time, costs per sample (lab time as well as expertise), potential scales at which tracing can work (depends on the level of variation between origins), the state of knowledge on the mechanisms that drive spatial variation, methodological advancement (how far the methods have been developed for origin verification of timber), requirements and availability of reference databases and the upscaling potential per method (whether the method can be upscaled by the development of non-species specific databases or through spatial interpolation).



	Chemical methods				Genetic methods		Optical methods
	1. Multi-element analysis	**2. Stable isotope ratios**	**3. DART-TOFMS**	**4. NIRS**	**5. DNA fingerprinting**	**6. Population genetics**	**7. Dendro-chronology**
Lab processing time	1 day	Several days	Minutes	Minutes	Several days	Several days	Several days
Approximate costs per sample	\$100-\$200	\$100-\$200	\$1-\$100	\$1-\$100	\$100-\$300	\$100-\$300	<\$100
Potential scale for tracing	50-200 km	250-1000 km	?	?	< 1m	15-100 km	100-500 km
Knowledge on mechanisms	Moderate	Moderate	Limited	Limited	Strong	Strong	Strong
Methodological advancement	Limited	Moderate	Limited	Limited	Strong	Strong	Moderate
Reference dataset requirements	(Species specific) multi-element database of the region	(Species specific) isotope database of the region	Species specific DART-TOFMS database of the region	Species specific NIRS database of the region	DNA of the tree stump and from other individuals	Species specific genetic database of the region	Tree ring series from reference trees of the region
Availaibilty of reference datasets	Very limited	Limited	Moderate	Limited	Limited	Moderate	Limited
Upscaling potential	Strong	Strong	Limited	Limited	None	Limited	Moderate

for origin differentiation. Some of these elements even provided a chemical basis for origin identification after the wood was charred (Štulc et al., 2023), suggesting the method could be applied to charcoal for which to date no origin verification methods are available other than the challenging visual tree-ring analysis (February, 2000).

Limitations

Nevertheless, there are still challenges that need to be overcome before the method can be widely implemented. Most importantly, the methods for in- or exclusion of samples must be improved when compared to a reference database, as it was shown that blind samples from other geographical areas or other species were chemically not distinct enough to exclude them in the assignment tests (Chapter 3). This is problematic in a court case, as it could lead to false acquitting of origin fraud. Including a species identification step before chemical analysis, which is relatively easy to implement for many species, can already improve the exclusion of other timber species before measurement. This was not included in this thesis as the goal was to test how the method would operate by itself. Improving the exclusion of samples from origins not covered by the reference database is more complicated however. It is expected that this will improve as the geographical range of the reference database is expanded, because that gives a better representation of which chemical compositions should be excluded (the 'outgroup').

Stable isotope ratios

Stable isotope ratios have a longer history of research for origin tracing. The stable isotopes are taken up and included in wood relative to their availability in the surrounding environment, which in turn is influenced by climate and geology (van der Sleen et al., 2017; Lehmann et al., 2022; Wynn et al., 2014; Allen et al., 2022). Isotopes have been applied for origin verification in a variety of commodities, sometimes in combination with multi-element analysis (Baroni et al., 2015; Griboff et al., 2021; Wang et al., 2020). In timber, the stable isotope ratios mostly used for tracing are the bio-elements δ^{18}O, δ^{2}H, δ^{13}C, δ^{15}N and δ^{34}S, from which an origin-specific isotopic fingerprint can be defined. As the factors influencing the isotope ratios differ per isotope, combining multiple isotopes maximizes the potential differences between different origins (Ehleringer et al., 2000; Watkinson et al., 2020). Some studies applied the heavy isotope ^{87}Sr/^{86}Sr for provenancing of timber as well (Rich et al., 2016), but practical issues related to the low Sr concentrations in wood hamper its broad implementation.

Strengths

A major strength of stable isotope ratios is that we have a strong understanding of the drivers of the naturally occurring variation in timber. Stable isotope ratios in wood have been investigated for multiple applications, such as for studying drought responses or climate change effects (van der Sleen et al., 2017; Lehmann et al., 2022; Wynn et al., 2014). Based on previous knowledge on variation in these drivers, locations can be targeted specifically where high potential for timber tracing is expected. Additionally, these mechanisms can be used as a basis for developing interpolated maps (isoscapes), which alleviates the dependency on origin-specific reference databases (barrier 1, Watkinson et al. (2020), Chapter 4 and 5).

Limitations

However, the reported success of origin assignment for timber based on stable isotopes varies. Some studies reported high accuracy and potential (Watkinson et al., 2020; Degen et al., 2015; Watkinson et al., 2022a; Gori et al., 2018), but others reported the opposite (Vlam et al. (2018b); Paredes-Villanueva et al. (2022), Chapter 5). This difference can partly be attributed to the isotopes that were measured and the scale at which they were applied, as in some areas the underlying variation in those isotopes might not have been enough to find differences between origins. In Central Africa for example, the gradients in rainfall and temperature are limited, thus limiting the applicability of stable isotopes for tracing. Therefore we concluded that stable isotope ratios are most promising for origin identification at large spatial scales (100s of km), because local variation can be too large to find strong differences at small spatial scales (Chapter 5). They can for example be used for country differentiation (Chapter 4), as in Degen et al. (2015), or across large states as in Watkinson et al. (2020).

DART-TOFMS

A third chemical method discussed here is Direct Analysis in Real Time - Time of Flight Mass Spectrometry (DART-TOFMS). With DART-TOFMS, the metabolomic profile is measured in a wood sliver, which can then be compared to a reference database to define similarities between the sample and the reference set. DART-TOFMS is used more and more for species identification of timber, as it can provide a fast and accurate species identity analysis given an appropriate reference database is available that includes the species in question (Deklerck et al., 2019; Musah et al., 2015). Differences between species are mostly driven by the presence of unique secondary metabolites, which determine species identification. Recently, studies have also shown potential of DART-TOMFS for the differentiation of wood origin (Finch et al., 2017; Deklerck et al., 2020; Price et al., 2021): the species metabolomic profile differed between continents (Price et al., 2021) and states (Finch et al., 2017).

Strenghts

Strengths of the method are mostly related to the ease of measurement: the method is cheap and fast, and only a small wood sample is needed to measure the metabolomic profile (Deklerck, 2023). This makes it an ideal method for forensic purposes. It is already applied at a large scale for species identification and thus reference databases are currently being developed, for example as part of the World Forest ID project (Gasson et al., 2021).

Limitations

However, the application for origin differentiation remains experimental and strongly depends on the specificity of the reference samples (Deklerck et al., 2020; Deklerck, 2023). The drivers of spatial variation in these metabolomic profiles are poorly studied and represent the main limitation of DART-TOFMS. The within-species variation can be caused by both genotype and local environment and especially the influence of climate and soil on these profiles is not understood (Deklerck et al., 2020). As a result, the method can only be applied on a case-by-case basis and it's applicability for tracing at small spatial scales is not yet known. Furthermore, the lack of understanding hampers spatial interpolation of spectra, so the potential of the method will still largely depend on an extensive reference database (Deklerck, 2023). To move the method forward, more research is needed to understand which molecules might be driven by climate. This will improve our understanding of which areas should be targeted to further expand reference databases because high chemical variation is expected in those areas. Furthermore, it can allow for spatial interpolation of metabolites as proposed by Deklerck (2023).

NIRS

The last chemical method for origin identification I discuss here is Near Infrared Spectroscopy (NIRS). The method is based on variation in phytochemical properties as well, characterised by wood absorption spectra when a sample is exposed to near infrared electromagnetic energy. Rather than the relative abundances of molecules, as with DART-TOFMS, the near infrared spectrum of a sample is obtained. These NIR spectra have been applied mostly to species identification (Bergo et al., 2016), but they have also been linked to timber origin (Sandak et al., 2011; Silva et al., 2018; Li et al., 2019).

Strengths

Strengths of the method are similar to DART-TOFMS: measurements can be done quickly and they are not expensive (Sandak et al., 2016). Furthermore, operators can be trained easily and measurements can be taken with a hand-held device, which facilitates large-scale implementation (Sandak et al., 2016). As the method is non-destructive, it is ideal for tracing cases where no physical sample can be extracted, such as for valuable instruments or art panels.

Limitations

A main limitation of NIRS for origin verification is the high variation that occurs within a tree already. The obtained spectra can vary between the tree tip and the base, between inner and outer wood and when compression wood is present (Sandak et al., 2016). This complicates the characterisation of the spectrum of an individual tree, let alone the characterisation of a site of origin. Additionally, the method requires calibration against reference data of known composition and this can lead to variation between measurements, if not done carefully (Sandak et al., 2016). Lastly, like with DART-TOFMS it is not yet understood what is driving geographical variation in these spectra. Overall, the potential for origin tracing was tested on a case-by-case basis and this will need to be expanded for further implementation of the method.

Genetic analysis

Genetic analyses have the longest history in timber tracing research and have been applied to the largest number of tree species (Low et al., 2022). The analyses rely on small variations in the genetic code that accumulate over generations, resulting in larger differences between the DNA of individuals related distantly than those related closely. At the individual level, genetic analyses are referred to as DNA fingerprinting, where the DNA of a piece of wood is compared to the DNA of the corresponding tree stump in the forest (Lowe et al., 2010). This provides strong evidence of the timber's claimed origin, but it is expensive as it requires sampling and analysis of individual trees. Alternatively, genetic differences at the population level have also been used for origin verification (Jolivet & Degen, 2012; Ng et al., 2022; Vlam et al., 2018b). By analyzing the genetic differences between groups of individuals (populations), the spatial genetic structure can be identified, and genetic markers can be used to assign individuals to different populations. The population genetics approach has been quite successful, with reported assignment accuracies ranging from 55% to 92% across various spatial scales (Ng et al., 2017; Chaves et al., 2018; Vlam et al., 2018b).

Strengths

The strength of genetic analysis lies in its potential for high precision, as high-quality reference databases can differentiate blind samples from sites as close-by as 14 km (barrier 2, Vlam et al. (2018b)). This still depends on the species, the geographic setting and the reference database however, as we found limited potential at small spatial scales when applying cpDNA for tracing of Azobé (Chapter 6). Another important strength of genetic tracing is that reference databases are robust over time, especially compared to chemical methods which are driven by soil and/or climate. Differences in a genetic profile at a certain location will not change much at a timescale of 50-100 years, whereas changes in soil and especially climate can be expected at those timescales and these may drive changes in the chemical proxies described here (elemental composition and stable isotope ratios). Lastly, genetic data especially lends itself for exclusion tests because it is categorical: the presence or absence of a specific mutation can lead to the inclusion or exclusion of a sample from a potential population of origin. This differs from the previously described chemical methods, where a population is always characterized by a range of values. Therefore, the chemical data requires a statistical analysis for outliers to exclude samples from a certain origin and this is not always straightforward (Chapter 3).

Limitations

However, limitations to the broad implementation of the method remain and these are strongly linked to its strengths. Characterizing local genetic profiles to reach the desired accuracy is time-consuming and expensive, as many trees need to be sampled and analyzed: a minimum of 20-30 trees is advised to describe a population and the DNA extraction takes several days. Therefore, there will always be a trade-off between sampling effort and accuracy of the method (Jolivet & Degen, 2012). Furthermore, upscaling the method by using statistical interpolation is challenging due to the categorical nature of the data as well as the difficulty in finding relevant covariables (Chapter 6). To develop an appropriate genetic reference database for forensic use, each relevant population must thus be sampled, as it not clear whether inferences can be made from locations where no data is available. Another challenge in genetic analysis is related to DNA extraction from wood, which can be challenging due to the fragmentation of DNA in the heartwood (Degen et al., 2013; Jolivet & Degen, 2012). The use of chloroplast gene markers has shown potential in alleviating this issue, as there are more copies of the cpDNA in the cells compared to nuclear DNA, making it easier to obtain sufficient amounts of DNA from degraded samples (Tnah et al., 2009; Deguilloux et al., 2003; Ng et al., 2022). However, the chloroplast may not always provide enough genetic diversity, complicating the distinction between trees from close-by regions.

Overall, genetic analysis is a useful tool especially for small-scale tracing, mostly because other methods are more cost-effective at large scales. To move the methods forward, it will be essential to define priority timbers or areas where genetics can provide added value compared to other tracing methods. In that way, costs for sample collection and analysis can be optimized.

Dendrochronology

Lastly, dendrochronology is a method of timber tracing that involves analysing the unique pattern of tree rings. By analyzing the width, density, and other characteristics of these rings, the age and origin of a piece of wood can be identified (Akhmetzyanov et al., 2019). The method relies strongly on the yearly growth patterns of trees, where the width of each ring is influenced by environmental factors such as temperature, rainfall, and soil conditions. It has been particularly useful for determining the age and origin of wood from historic buildings or archaeological sites (Mundo et al., 2022; Kuhl et al., 2023; Sass-Klaassen et al., 2008), but it could also be used to trace the origin of illegally traded timber.

Strengths

A main advantages of dendrochronology is that not much specialised laboratory equipment is needed for the measurements. It does not require specific extractions, such as for genetic analysis, nor is it affected by chemical changes over time that may occur in wood. This gives it a big advantage over other methods for archaeological wood for example, especially if wood has been water-logged a long time (Domínguez-Delmás et al., 2020), as well as for treated wood such as charcoal (Rybníček et al., 2022). Another large advantage is its ability to provide a date for the origin of a piece of wood in addition to origin. This is especially useful in an archaeological context, where it can provide additional pieces of the puzzle to understand local history, as well as for climatology when the archaeological tree ring records are used for climate reconstruction (Hellmann et al., 2017). Knowing the age of the analysed wood sample can be helpful for illegally traded wood as well. In the light of international trade regulations such as in the new European anti-deforestation regulation, only timber that was logged after a certain date is banned and this thus requires an approximation of the logging date.

Limitations

The main limitation of dendrochronology in the context of this thesis is that it can only be applied to tree species that produce annual rings, which is rare in the tropics where illegal logging rates are highest (February, 2000; Hoare & Uehara, 2022). Even though variation in other anatomical features could be used as well to develop a chronology (Akhmetzyanov et al., 2019), those patterns have hardly been studied in the tropics and will first require a better understanding of the variables that might provide useful spatial information. Additionally, dendrochronology can only trace timber back if enough rings are preserved to 1) date the wood and 2) to link it to reference databases. This makes it less useful for sawn wood and other wood products such as charcoal and wood chips for example. Lastly, the need for reference samples to identify the geographic origin of a piece of wood is especially pressing for dendrochronology. Similar to the other tracing methods, researchers must have access to a database of tree-ring patterns from different regions in order to match the pattern of a sample to a specific location. However, reference chronologies require a lot of samples per site and many hours of cross-dating to make them reliable and very few are available for tropical timbers (Zuidema et al., 2022). An important step forward for dendrochronology for origin verification would therefore be to develop machine learning tools that can predict chronologies at non-sampled locations (Bodesheim et al., 2022). These models can be trained by existing sites and the local environmental information, which is then used to build chronologies at new locations. This would greatly reduce the sampling burden as well as the time investment to upscale these methods.

7.3 What about combining methods?

Based on these different tracing tools, it can be concluded that there is clear potential for forensic origin verification of timber. However, small scale tracing remains challenging and no method has yet shown consistently high accuracies across species and study regions when aiming to distinguish origin within 100 km. To improve small scale tracing when individual methods are not sufficiently accurate, combining methods might be the outcome (Dormontt et al., 2015). As the processes that lead to geographical variation in each of the methods are spatially heterogeneous, combining them may enforce the strengths of the individual methods. This allows for an optimization of the unique fingerprint of a certain origin (Vlam et al., 2018b; Dormontt et al., 2015; Lowe et al., 2016).

The additionality of multiple proxies for tracing purposes has already been demonstrated in other commodities and in other scientific disciplines. For example, the combination of stable isotopes and multi-element analysis has proven successful for origin identification of a large range of commodities such as honey, meat products and vegetables (Wang et al., 2020; Drivelos & Georgiou, 2012). Similarly, in studies that trace the provenance of archaeological wood, the integration of multiple methods (anatomical proxies as well as strontium isotope analysis) also reached a higher tracing precision than the individual methods would have done (Akhmetzyanov et al., 2019; D'Andrea et al., 2023). Moreover, the combination of genetic analysis with stable isotopes even led to improved tracing of the breeding grounds of birds (Chabot et al., 2012). Tracing can thus become more reliable by exploiting the strengths of multiple proxies.

To test whether combining methods also improves tracing results for tropical timbers, we were the first to compare and combine methods by developing and testing reference databases based on samples from the same trees (Chapter 6). We combined multi-element analysis, stable isotope ratios and cpSNP analysis and tested their individual as well as their combined assignment accuracy for 13 locations in Central Africa. The combination of multi-element analysis and cpSNPs proved to be the most complementary and improved tracing at a sub-country scale (100 km; multi-element accuracy of 80% and cpSNP accuracy of 72%, combined accuracy of 91%). This finding is especially important in Central Africa, where the absence of significant geographical barriers and limited environmental variation hampers the potential of individual tracing methods. Tracing with multi-elemental analysis was limited to chemical clusters, due to the similarity in elemental concentrations between some sites (Chapter 3), limited variation in stable isotopes was found (Chapter 5) and also genetic variation was too little to achieve high accuracy for origin identification (Chapter 6). However, when all three methods were combined, we achieved exceptional accuracy at smaller spatial scales, with 94% of the trees being accurately assigned within a 100 km radius. This accuracy is similar to the highest reported accuracy for genetic tracing from previous studies at similar spatial scales (Vlam et al., 2018b; Jolivet & Degen, 2012). Importantly, this result was found across 13 sites in three Central African countries, thus presenting an unprecedented accuracy for such a large study area.

The potential for timber tracing based on a combination of chemical and genetic methods was thus clear from this study on Azobé. Our results demonstrated the added value of combining methods: in areas where an individual method failed to distinguish between sites, other methods did show enough variation between those sites to identify timber origin accurately. The greatest advantage of this combined approach thus lies in its ability to overcome the limitations of individual methods when their accuracy alone is inadequate. I expect that this combination of methods will also prove valuable for small-scale tracing of other timber species in Central Africa, where natural variation in the individual methods only occurs at large spatial scales. On the other hand, in regions where substantial geographical barriers or environmental gradients are already present, such as large variation in elevation or differences in soil characteristics, a single method may provide sufficient reliability for tracing purposes. This could for example be investigated by measuring the genetic and isotopic composition of the Red Meranti trees in Kalimantan of Chapter 3, for which the multi-elemental composition has been measured already. Since soil characteristics were already highly variable at small spatial scales, tracing based on multi-elemental analysis alone already reached a site-level assignment of 93% for nine sites. Adding other forensic methods may not be needed in this case.

The main disadvantage of combining different tracing methods is related to the time investment and associated costs to develop a comprehensive reference database. Currently, the reference databases of the individual methods are already limited, let alone a database with multiple methods applied to the same trees. Still, I would advocate for an expansion of reference databases to advance the application of this combination of methods, until at least 40-50 sites are sampled per method in the areas where relevance is high due to high illegality. This minimum number of locations will allow for a statistically strong spatial interpolation of the individual variables, such as the stacked individual maps as mentioned in Chapter 6.

Applying such spatial interpolation methods can facilitate the integration of reference databases of different sizes and locations, eliminating the need for all measurements to be conducted on the same set of trees. This approach will enable the use of existing databases as well, as there are already reference databases available for each method (Low et al., 2022). By developing the stacked interpolated maps, it would become possible to integrate these existing databases without the need to start a novel reference database specifically for the combination of methods. This strategy thus optimizes the use of available resources while maximizing the benefits of combining multiple forensic methods.

In conclusion, combining forensic methods has the potential to improve tracing accuracy when the individual methods alone are insufficient to trace origin at the relevant scale (barrier 2). Furthermore, an integrated reference database that includes multiple forensic methods also helps to better understand when only one of the methods might already be enough for the tracing question at hand. However, the development of reference databases for multiple methods is time-consuming and expensive. Therefore, the expansion of the individual reference databases should be directed towards the most relevant areas for tracing and further implementation of spatial interpolation methods can help to overcome the sampling burden (barrier 1).

7.4 From scientific development to real-life application

For forensic tracing methods to be relevant, it is essential to link them back to the problems encountered in the timber trade and to confirm they can actually make a difference. Illegal timber is sold and mixed with legal timber in many ways (Figure 7.1, Brancalion et al. (2018); Bisschop (2012); May (2017); Gan et al. (2016)) and further scientific development can only be useful if it can be applied for actual cases of timber fraud. Twenty examples of timber fraud, divided in four scale categories, are included in Table 7.1 and Figure 7.1. It is clear that illegal practices within or very close to a legal concession (Category A) will not be caught by the forensic methods discussed in this thesis: extensive local monitoring is the only way to stop these sorts of fraud and origin cannot be identified after the wood has entered the global timber supply chain. This underlines the need for collaboration amongst all levels of the timber supply chain, from the supply to the demand side. However, for all laundering of timber that does take place outside of a concession, e.g. at least 100 km away, the methods discussed here can make a difference even if the wood has already been traded across multiple countries. These are also the types of fraud that occur in the largest volumes, such as selling illegal timber from the establishment of a plantation (nr. 7) or mixing logs at mills (nr. 11).

To distinguish between real plantation timber and illegal timber from natural forests that is claimed to originate from a plantation (Category B), dendrochronology has high potential as long as the trees produce visible growth rings. Trees with visible growth rings have distinct growth patterns that differ between natural forest and plantations due to the different growing conditions. For example, trees growing in plantations grow faster and will show larger rings as they are less subjected to competition from other trees (Dié et al., 2015; Worbes, 1999).

To distinguish between legal and illegal timber originating from the same region or country (Category C), multi-element analysis and SNP analysis have the highest potential. When combined, the precision with which timber can be traced is even higher and origin could be pinpointed at the concession level (Chapter 6). Nevertheless, DART-TOFMS, NIRS and stable isotopes can also give an indication of origin at a sub-country level, provided enough variation is present at small spatial scales. This was clear for example for stable isotope ratios in timber from the Solomon islands (Watkinson et al., 2022b) or from the Alps (Gori et al., 2018).

To combat fraud taking place across larger distances, e.g. country borders (Category D), all the methods can help to verify origin, given that an accurate reference database is available. Depending on the region, the methods will vary in their potential to trace origin as the available natural variation will be more or less pronounced depending on the physical and environmental gradients. For effective law enforcement, the best forensic tracing method is likely to be a trade-off between the required accuracy and the investment costs (time and money), prioritizing methods that are shown to work in the area, are cost-effective and fast. For instance, when aiming to determine the country of origin of a wood sample, constructing a reference database with genetic profiles or reference chronologies of relevant reference countries can be an immense task. Characterising the isotopic fingerprint of a country on the other hand might be easier, as the isotope ratios within a country can be interpolated based on relevant covariables. Therefore, if there is sufficient variation within the potential area of origin of the sample (identified by assessing the species distribution and potentially the main exporting countries), stable isotope ratios are likely to be the preferred approach (Chapter 4 and 5).

However, the best tracing method still depends on what kind of tracing is required: origin verification or origin identification. If it is a question of origin verification, reliable in- and exclusion criteria are essential, whereas the question of origin identification requires a reference database of all potential origins. Genetic tracing especially lends itself for verification questions due to the nature of the data: forensic tests can be based on presence or absence of specific mutations that occur at that location (Vlam et al., 2018b), even though a small chance remains this specific mutation was not sampled. As long as reference samples can be taken at the claimed origin, the verification tests are rather straightforward. On the contrary, stable isotope ratios as well as multi-element analysis do not provide a site-specific fingerprint in the same way: they provide a range of values and exclusion must be based on outlier statistics (Chapter 3). At the same time, the continuous nature of the data makes them more appropriate for spatial interpolations and therefore for questions of origin identification: a potential region of origin can be defined based on an interpolated map, where areas may be included at which no samples were taken yet. This is much harder for genetic data if no clear geographical boundaries are present that would split genetic groups (Chapter 6).

Overall, forensic methods clearly have potential to be relevant for the types of timber laundering that have been reported: many types of fraud can be detected by one or more of the methods that are currently available. Only timber fraud taking place at the smallest spatial scale, within or right next to a legal concession, will be extremely hard to control after the trees have left the forest. Deciding on the best method or combination of methods for a tracing question will thus depend on the tracing question, the species and the location and whether reference databases are already available, the ease of use of the methods as well as the available resources for the investigation.

7.5 Multi-element analysis: what's next?

My thesis contains the first test of multi-elemental analysis applied to timber tracing. We found chemical differences between origins at the sub-country level in Central Africa and on Borneo, likely related to differences in soil composition between locations. Three main advantages of the method became apparent in comparison with other forensic methods: 1) two timber species could be combined into one classification model, therewith alleviating the dependency on species-specific reference databases (barrier 1), 2) the chemical composition of the wood could be linked to soil variables of which global databases are available, potentially allowing for interpolation of the chemical profiles (barrier 1), and 3) up to 60 elements could be measured simultaneously of which 40 were found in the wood, therewith providing a large base for origin verification at small spatial scales (barrier 2). However, additional research is needed to develop the method further and test its robustness and applicability beyond the study species and areas included in this thesis. I suggest the following research lines.

The first priority to develop the method further should be to better understand the link between the chemical composition of the wood and the soil physical and chemical properties. This could be achieved by extending the analysis as done in Chapter 2 of this thesis: measuring physical and chemical soil properties and linking those measurements to wood elemental composition, so that the most important drivers of variation in wood elemental composition can be determined. Furthermore, measuring the total nutrient content in the soil, not just the extractable content, could provide valuable insights into the coupling between soil and wood chemistry. We found no strong links between the soil reactive pools and the wood elemental concentrations, possibly because trees accumulate elements over a long time. Therefore, total elemental content may be a better indicator of element availability to trees. The locations for these analyses should be carefully selected so that contrasting soil conditions are included, as this will add the most to our understanding of geographical variation. A better understanding of soil drivers of wood composition could also be achieved in an experimental setup, for example by growing young trees under different soil conditions and comparing uptake and incorporation of chemical elements in the wood. This will require more resources however and the applicability to mature trees in a natural system remains uncertain.

This research line should receive the highest priority as its expected outcomes would move the method forward in three ways: 1) The knowledge will help to identify areas where the method is expected to have high potential, because the wood composition is expected to vary at relevant scales. If for example soil clay content is found to be a main driver of wood chemical composition, multi-element analysis should be preferred over other forensic methods where soil clay content varies at small spatial scales. 2) Knowing the drivers of chemical variation in wood will also inform the development of spatial models as it determines which covariables are most relevant to include. Prediction models can then be developed in the same way as isoscapes, as suggested in Chapter 6, to estimate the chemical composition in areas where no wood samples were measured directly. 3) These measurements will also help to understand if elemental composition might change over time and through climate change. Ideally the reference databases are robust to future changes in wood composition due to potential changes in soil and climate, so that measurements will not have to be repeated. For example, changing rainfall patterns or rising temperatures might have an effect on the soil characteristics and available elements for trees (Chen et al., 2018; Dalal et al., 2021; Quan et al., 2014). By including at least 3-5 cm of wood and thus including multiple years of growth we aimed to ensure the robustness of the elemental composition measurements (Chapter 2), but this needs to be tested further to ensure the reference database does not have to be updated over time. To do so, we need to understand what drives the chemical composition in the wood.

The second priority for future research is to expand the reference database with more species and sites. This will enable testing the performance and the robustness of the method in other timbers and landscapes and also improve our understanding of the drivers of variation in elemental composition. Furthermore, by including additional species at the same locations, we can test with more species whether there are elements that are important for origin identification across sites. We already planned to start with this research line in collaboration with the World Forest ID project (Gasson et al., 2021). Additional samples from other species and sites in Central Africa will be analysed for their multi-elemental composition to extend the database and address these questions. Priority should be given to species linked to high occurrence of illegality, for example by focusing on the target species identified by the GTTN (Cramm & Van Brusselen, 2019). Furthermore, the current database should be extended to the Democratic Republic of the Congo, to cover the major timber exporting countries in Central Africa. Additionally, adding more timber species from sites on Borneo, preferably including sites in Malaysia, would enable investigating any signal across species in another context than Central Africa.

As expanding the reference database is costly, the use of available resources could be optimized by establishing a minimum number of samples per site that is required for accurate tracing. To do so, it is crucial to understand what sample number can characterise the chemical variation within a certain origin. In this study we started with 20 trees per site, but a smaller amount of samples might result in similar assignment accuracy, thus reducing measurement costs. Therefore, I assessed what the effect was of the number of trees in the reference database on the resulting assignment accuracy (Figure 7.2). I did this both at the research site level and the regional chemical cluster level, as we found both were relevant scales of tracing with multi-element analysis (Chapter 3).

I tested the assignment accuracy of various Random Forest models, using a subset of samples per site from the 22 sites and 8 clusters of Azobé and Tali in Cameroon, Republic of the Congo and Gabon (see Chapter 3 for method details). Assignment accuracy clearly increased when the number of reference samples for the Random Forest classification models increased (Figure 7.2). However, accuracy did not increase a lot when more than 10-12 samples per origin were included in the reference set, neither for site classification nor for cluster classification. Neither did the models get more precise after including 10-12 samples: the interquartile range (IQR) as indicated by the size of the boxes did not get consistently smaller when more trees were added to the reference sets. Therefore, in this region in Central Africa, it can be concluded that a minimum of 10-12 samples per origin can provide accurate site and cluster assignments. For defining clusters, it is likely that more samples than the minimum of 10 will still be sampled anyway as most clusters likely consist of multiple sites. Nevertheless, this analysis shows that cluster assignment will not improve much further when more samples would be added per cluster.

Third, testing the effect of post-harvesting treatment, such as charring, chemical treatments or

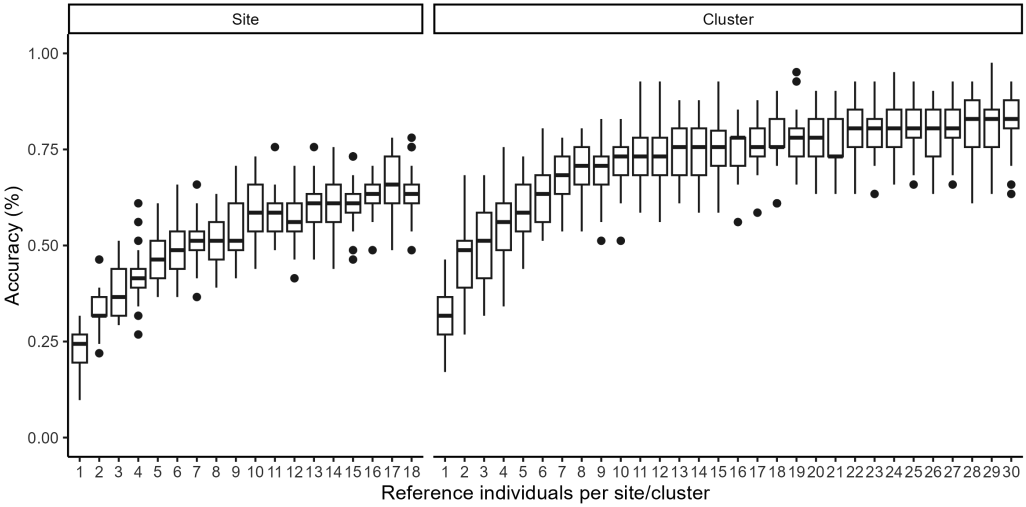


Figure 7.2: Assignment accuracy (%) of Random Forest models based on multi-elemental composition (25 models per number of reference trees) with increasing number of reference trees per site or cluster to develop the model, predicting site of origin (left) and cluster of origin (right). Accuracy was evaluated by assigning a randomly selected subset of test trees to their site or cluster of origin based on those models. The test set consisted of 5% of the original database per site (= 2 test trees) or cluster (= between 2 and 14 test trees; depending on the size of the cluster). Accuracy was calculated in this way to allow for SMOTE resampling of the cluster database, due to the different sample sizes of the cluster (see Chapter 3). The Random Forest models were base on the same reference databases as models A8 (for the site level) and A4 (for the cluster level) in Table 1 of Chapter 3.

heat treatments will provide useful insights into the robustness of the method for application on traded wood products. In this thesis we aimed to test the method on wood samples taken directly from trees, to limit external variation to the reference database. However, a study on charcoal showed promising results that some elements were not affected by charring (Štulc et al., 2023), which would greatly expand the applicability of the method. Experimentally testing the effect of different treatments would help to assess the potential impact of wood processing on the multi-elemental composition of the wood, and thus the applicability of the method on anything else than untreated wood.

Fourth, a relevant research line would be to test whether there is a genotype effect on variation in the chemical composition of wood within a tree species. As we have developed a unique database as part of the Timtrace project where both genetic variation and chemical variation are defined for the same trees, it can be tested whether genetic distance and chemical distance between individual trees are correlated. As Chapter 6 showed the methods are complementary for origin assignment tests, it is not expected that these distances are strongly correlated: the additional accuracy when these methods were combined would not have been so profound if this was the case. Nevertheless, when aiming to develop interpolated maps of elemental concentrations in wood based on relevant covariables, we must rule out the effect of genetic variation on chemical composition as this will add noise to the maps. Therefore this research line can provide useful information for the link between chemical and genetic variation.

Lastly, a more practical research line will be to establish the lowest required element concentrations that need to be measured for accurate tracing results. ICP-MS is a strong method for measuring multi-elemental composition at low concentrations, but it may not be widely accessible. This limitation could hinder widespread implementation. Alternative techniques, such as ICP-OES and ICP-AES, are more commonly available but lack the sensitivity of ICP-MS (Drivelos & Georgiou, 2012). It should be investigated whether they may still provide the necessary chemical information to differentiate between timber origins. Along the same line of research, ring tests must be set up to compare measurements across laboratories, to ensure consistency in the application.

In conclusion, tracing timber origin based on the multi-elemental composition of the wood has the potential to be a valuable tool for origin verification at a sub-country scale. However, further testing and method development is required to demonstrate its robustness and applicability in different contexts.

7.6 Towards reducing illegal logging

Taking another step back and considering the application of forensic timber tracing methods, some key points should be addressed to make them operational and achieve their aim: preservation of forests and good forest management. For all methods, the statistical analyses can still be improved to develop the most accurate verification and assignment models. This is on the one hand a matter of scalability: as mentioned in the introduction, it will take years until relevant reference datasets are available for the main traded timbers (Low et al., 2022). Therefore, we need to optimize sampling schemes, tailor them to the specific geographical setting, and improve the spatial interpolation of the different methods (Truszkowski et al., 2023). Furthermore, the statistical analysis to exclude a sample from a specific origin with a degree of certainty will require more attention when is it based on continuous data, as various outlier statistics did not perform well enough to identify species/individuals from outside of the reference dataset (Chapter 3).

Forensic methods also need to be embedded into a system of checks along the timber supply chain, as there will be no one-size-fits-all solution to reduce illegal logging. Origin verification methods can be combined with other methods, such as external tagging (Shirmohammadi, 2023) and blockchain technology (Munoz et al., 2021; Komdeur & Ingenbleek, 2021). These systems can all become more reliable if they are verified independently by forensic tracing. Furthermore, an especially valuable combination would be to link these tracing methods to deforestation alerts based on remote-sensing information (Reiche et al., 2021). In that way, the location to which timber is assigned based on the forensic tests could be checked for recent disturbance, making origin assignment claims even stronger if disturbances indeed took place in that area.

Additionally, while global initiatives are essential to reduce illegal timber trade (Bisschop, 2012; Zeitlin & Overdevest, 2021; Boekhout van Solinge et al., 2016), reducing illegal logging will ultimately be achieved at the local level. Once the tree has arrived at its point of sale, often on another continent, it may be too late to stop further logging. However, by applying forensic verification methods at the point of export, a more direct link can be established to the location of illegal logging and relevant measures can be taken to stop it. Additionally, data shows that domestic markets account for the largest share of consumption of illegal wood (Gan et al., 2016). These forensic methods must thus become available close to the source, to improve the transparency within local timber trade routes. They can be combined with other local measures to stop illegal logging, including simple efforts such as putting gates to restrict access to forests, improving checkpoints and installing wood tracking operations (Thompson & Magrath, 2021; Hoare et al., 2020). To ensure long-term application of laboratories and methods, this will require both equipment and capacity building in timber producing countries. However, willing governments that want to invest in transparency in local timber markets are crucial for structural improvements. An example could be Gabon, who have pledged to make all forest concessions FSC certified in order to improve their sustainability and to create a marketing asset for the sale of Gabonese timber (Karsenty, 2018). The increased tax revenue from the illegal timber could be a good incentive to move towards stronger law enforcement, but corruption and weak governance are still widespread in many producer countries (Hoare & Uehara, 2022), complicating the implementation of strong preventive methods.

Focusing solely on producing countries to tackle illegal logging is not enough however, because demand-side factors also play an important role in the illegal trade. Changes in wood consumption are necessary to have a lasting impact, as people will always find ways to obtain a product that is in high demand. For instance, the demand for natural teak from Myanmar for superyacht decks and African hardwood for furniture in the US that is processed in China will continue to drive illegal logging (Fuller et al., 2019; Chin-A-Fo & Kuijpers, 2023). Therefore, addressing the root cause of the problem, the demand for illegally-sourced timber, is crucial to stop the illegal trade as well (Challender & MacMillan, 2014).

Lastly, it is clear from this thesis that there is still research to be done in the field of forensic timber tracing. To reach its full potential and to overcome the identified barriers as outlined here, collaboration among institutions will be key. This collaboration should involve institutions in timber importing countries with existing reference databases of the different methods, aiming to avoid double work and to integrate already developed databases with future data. Additionally, the collaborations should involve institutions in timber producing countries that can expand reference databases and have an impact locally. Currently, a lot of different institutions have established individual reference databases, but it is not clear what data is available and whether the information can be shared among parties. A coordination and integration of data can thus give a strong push to the tracing field as a whole, ensuring the utilization of available data and potentially leading to more accurate tracing results. Initiatives have been taken to establish a network among research institutes, such as the Global Timber Tracking Network (https://globaltimbertrackingnetwork.org/) and the World Forest ID program (Gasson et al., 2021). The Timtrace project, in which this thesis was produced, is a great example of a project where different forensic methods were combined and compared, to understand which method or combination of methods is most appropriate for a certain tracing question. Furthermore, it involved institutions in producer and consumer countries who collaborated on method development. More projects or collaborations like this will move the field forward, as integrating databases does not require additional sampling or laboratory analysis and can thus be accomplished relatively easily. However, for effective implementation, these collaborations must be embedded within larger institutions rather than relying solely on project-based endeavors, as is the case with the Timtrace project. Therefore, securing stable funding sources must be a main priority to ensure the practical implementation of these findings.

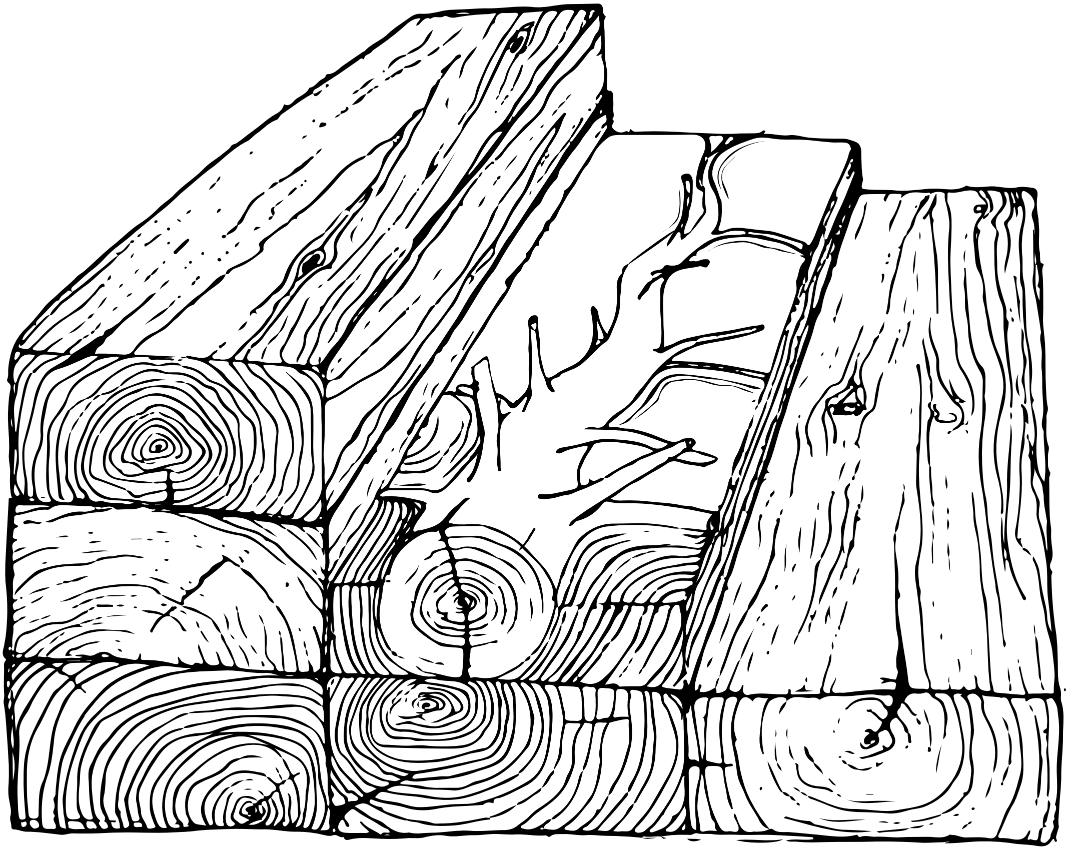
7.7 Conclusion

The main objective of this thesis was to contribute to the development of effective chemical timber tracing tools, which can verify the origin of tropical timber and, ultimately, combat illegal timber trade. Based on this discussion, the potential of forensic tracing is clear: even if the methods themselves can still be optimized further, the available methods can expose many of the reported laundering routes (Table 7.1), from a continental level down to a 100 km resolution.

Despite the progress made, the widespread implementation of forensic tracing methods is still hindered by the dependency on widespread reference databases (barrier 1), which are often lacking for many species and regions (Low et al., 2022). Due to the high costs associated with developing accurate reference databases, it is unlikely that forensic tracing will transition from its current case-by-case application to a more widespread implementation in the near future (Low et al., 2022). Shifting the focus towards forensic methods that can be spatially interpolated more easily, such as multi-element analysis and stable isotope ratios, can be an important way to reduce this dependency on reference databases (as described in the discussion of Chapter 6). The spatial scale at which these two chemical methods can achieve high tracing accuracy may already be sufficient for many tracing questions at hand (Figure 7.1, Table 7.1). In cases where a higher spatial resolution is required (barrier 2), the reference database can then be expanded by incorporating additional methods such as genetic analysis.

In addition to addressing these two research barriers, the use of existing databases must also be optimized. This can be achieved through stronger collaborations between institutions working on forensic timber tracing, as through collaboration these databases can be shared and used more efficiently. Institutions in timber producer countries must be involved in the collaboration as well, as this moves forensic tracing methods closer to the source and enables more direct action to stop illegal logging. Lastly, advancements in the applied statistical methods will also enable better use of existing data, especially when this is aimed towards combining databases from different forensic methods.

In conclusion, timber tracing methods are ready to be used by competent authorities, customs, certification bodies, non-governmental organisations and timber traders. The need for these uses is expected to increase due to new policies on illegal trade and deforestation such as the EUDR, in which origin verification plays a crucial role. However, there is still room for improvement in the success and widespread application of timber tracing. This can be accomplished by further developing the individual methods and expanding reference databases. When doing so, it will be crucial to focus on the most suitable methods for each region and timber species to maximize their effectiveness and realize the full potential of timber tracing.



References

Acierno, V., de Jonge, L., & van Ruth, S. (2020). Sniffing out cocoa bean traits that persist in chocolates by PTR-MS, ICP-MS and IR-MS. *Food Research International*, *133*, 109212. doi:doi: 10.1016/j.foodres.2020.109212.

Akhmetzyanov, L., Buras, A., Sass-Klaassen, U., den Ouden, J., Mohren, F., Groenendijk, P., & García-González, I. (2019). Multi-variable approach pinpoints origin of oak wood with higher precision. *Journal of Biogeography*, *46*, 1163–1177. doi:doi: 10.1111/jbi.13576.

Allen, S. T., Sprenger, M., Bowen, G. J., & Brooks, J. R. (2022). Spatial and Temporal Variations in Plant Source Water: O and H Isotope Ratios from Precipitation to Xylem Water. (pp. 501–535). doi:doi: 10.1007/978-3-030-92698-4_18.

Amais, R. S., Moreau, P. S., Francischini, D. S., Magnusson, R., Locosselli, G. M., Godoy-Veiga, M., Ceccantini, G., Ortega Rodriguez, D. R., Tomazello-Filho, M., & Arruda, M. A. (2021). Trace elements distribution in tropical tree rings through high-resolution imaging using LA-ICP-MS analysis. *Journal of Trace Elements in Medicine and Biology*, *68*. doi:doi: 10.1016/j.jtemb.2021.126872.

Andrews, J. A., Siccama, T. G., & Vogt, K. A. (1999). The effect of soil nutrient availability on retranslocation of Ca, Mg and K from senescing sapwood in Atlantic white cedar. *Plant and Soil*, *208*, 117–123. doi:doi: 10.1023/A:1004512317397.

Antoniadis, V., Levizou, E., Shaheen, S. M., Ok, Y. S., Sebastian, A., Baum, C., Prasad, M. N., Wenzel, W. W., & Rinklebe, J. (2017). Trace elements in the soil-plant interface: Phytoavailability, translocation, and phytoremediation–A review. *Earth-Science Reviews*, *171*, 621–645. doi:doi: 10.1016/j.earscirev.2017.06.005.

Baker, J. C., Hunt, S. F., Clerici, S. J., Newton, R. J., Bottrell, S. H., Leng, M. J., Heaton, T. H., Helle, G., Argollo, J., Gloor, M., & Brienen, R. J. (2015). Oxygen isotopes in tree rings show good coherence between species and sites in Bolivia. *Global and Planetary Change*, *133*, 298–308. doi:doi: 10.1016/j.gloplacha.2015.09.008.

Barbour, M. M. (2007). Stable oxygen isotope composition of plant tissue: A review. doi:doi: 10.1071/FP06228.

Barbour, M. M., Andrews, T. J., & Farquhar, G. D. (2001). Correlations between oxygen isotope ratios of wood constituents of Quercus and Pinus samples from around the world. *Australian Journal of Plant Physiology*, *28*, 335–348. doi:doi: 10.1071/pp00083.

Baroni, M. V., Podio, N. S., Badini, R. G., Inga, M., Ostera, H. A., Cagnoni, M., Gautier, E. A., García, P. P., Hoogewerff, J., & Wunderlin, D. A. (2015). Linking soil, water, and honey composition to assess the geographical origin of Argentinean honey by multielemental and isotopic analyses. *Journal of Agricultural and Food Chemistry*, *63*, 4638–4645. doi:doi: 10.1021/jf5060112.

Barton, K. (2020). MuMIn: Multi-Model Inference. R package version 1.43.17.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. R package version 1.1-27.1. *Journal of Statistical Software*, *67*. doi:doi: 10.18637/jss.v067.i01.

Bergo, M. C., Pastore, T. C., Coradin, V. T., Wiedenhoeft, A. C., & Braga, J. W. (2016). NIRS identification of Swietenia macrophylla is robust across specimens from 27 countries. *IAWA Journal*, *37*, 420–430. doi:doi: 10.1163/22941932-20160144.

Bisschop, L. (2012). Out of the woods: the illegal trade in tropical timber and a European trade hub. *Global Crime*, *13*, 191–212. doi:doi: 10.1080/17440572.2012.701836.

Blanc-Jolivet, C., & Liesebach, M. (2015). Tracing the origin and species identity of Quercus robur and Quercus petraea in Europe: A review. doi:doi: 10.1515/sg-2015-0017.

Bodesheim, P., Babst, F., Frank, D. C., Hartl, C., Zang, C. S., Jung, M., Reichstein, M., Mahecha, M. D., & Stifel Center Jena, M. (2022). Predicting spatiotemporal variability in radial tree growth at the continental scale with machine learning. *Environmental Data Science*, *1*, 9–10. doi:doi: 10.17871/BACI.248.

Boeschoten, L. E., Sass-Klaassen, U., Vlam, M., Comans, R. N., Koopmans, G. F., Meyer-Sand, B. R. V., Tassiamba, S. N., Tchamba, M. T., Zanguim, H. T., Zemtsa, P. T., & Zuidema, P. A. (2022). Clay and soil organic matter drive wood multi-elemental composition of a tropical tree species: Implications for timber tracing. *Science of The Total Environment*, *849*, 157877. doi:doi: 10.1016/j.scitotenv.2022.157877.

Boeschoten, L. E. et al. (2023). A new method for the timber tracing toolbox: applying multi-element analysis to determine wood origin. *Environmental Research Letters*, *18*, 054001. doi:doi: 10.1088/1748-9326/acc81b.

Bolson, M., De Camargo Smidt, E., Brotto, M. L., & Silva-Pereira, V. (2015). ITS and trnH-psbA as Efficient DNA Barcodes to Identify Threatened Commercial Woody Angiosperms from Southern Brazilian Atlantic Rainforests. *PLoS ONE*, *10*. doi:doi: 10.1371/journal.pone.0143049.

Bonde, N., Tyers, I., & Wazny, T. (1995). Where does the timber come from? Dendrochronological evidence of the timber trade in Northern Europe. *Archaeological sciences*, *34*, 201–204.

Boner, M., Sommer, T., Erven, C., & Forstel, H. (2007). Stable isotopes as a tool to trace back the origin of wood. In *"Fingerprinting methods for the identification of timber origins"* (pp. 47–57).

Bong, Y. S., Shin, W. J., Gautam, M. K., Jeong, Y. J., Lee, A. R., Jang, C. S., Lim, Y. P., Chung, G. S., & Lee, K. S. (2012). Determining the geographical origin of Chinese cabbages using multielement composition and strontium isotope ratio analyses. *Food Chemistry*, *135*, 2666–2674. doi:doi: 10.1016/j.foodchem.2012.07.045.

Bowen, G. (2023). Gridded maps of the isotopic composition of meteoric waters. .

Bowen, G. J. (2010). Isoscapes: Spatial Pattern in Isotopic Biogeochemistry, . doi:doi: 10.1146/annurev-earth-040809-152429.

Bowen, G. J., Wassenaar, L. I., & Hobson, K. A. (2005). Global application of stable hydrogen and oxygen isotopes to wildlife forensics. *Oecologia*, *143*, 337–348. doi:doi: 10.1007/s00442-004-1813-y.

Brancalion, P. H. S., de Almeida, D. R. A., Vidal, E., Molin, P. G., Sontag, V. E., Souza, S. E. X. F., & Schulze, M. D. (2018). Fake legal logging in the Brazilian Amazon. *Science Advances*, *4*, eaat1192. doi:doi: 10.1126/sciadv.aat1192.

Branco, P., Ribeiro, R. P., & Torgo, L. (2016). UBL: an R Package for Utility-Based Learning.

Bridge, M. (2012). Locating the origins of wood resources: A review of dendroprovenancing. doi:doi: 10.1016/j.jas.2012.04.028.

Brienen, R., Schöngart, J., & Zuidema, P. (2016). Tree rings in the tropics: insights into the ecology and climate sensitivity of tropical trees. In *Tropical tree physiology: Adaptations and responses in a changing environment* (pp. 439–461). Springer.

Brienen, R. J., Helle, G., Pons, T. L., Guyot, J. L., & Gloor, M. (2012). Oxygen isotopes in tree rings are a good proxy for Amazon precipitation and El Niño-Southern Oscillation variability. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 16957–16962. doi:doi: 10.1073/pnas.1205977109.

Brienen, R. J. W. et al. (2017). Tree height strongly affects estimates of water-use efficiency responses to climate and CO2 using isotopes. *Nature Communications*, *8*, 288. doi:doi: 10.1038/s41467-017-00225-z.

Brlík, V., Procházka, P., Hansson, B., Stricker, C. A., Yohannes, E., Powell, R. L., & Wunder, M. B. (2022). Animal tracing with sulfur isotopes: Spatial segregation and climate variability in Africa likely contribute to population trends of a migratory songbird. *Journal of Animal Ecology*, . doi:doi: 10.1111/1365-2656.13848.

Brun, C. B., Peltola, P., Åström, M. E., & Johansson, M.-B. (2010). Spatial distribution of major, trace and ultra trace elements in three Norway spruce (Picea abies) stands in boreal forests, Forsmark, Sweden. *Geoderma*, *159*, 252–261. doi:doi: 10.1016/j.geoderma.2010.07.018.

Bucci, G., Vendramin, G. G., & Ellstrand, N. C. (2000). Delineation of genetic zones in the European Norway spruce natural range: Preliminary evidence. *Molecular Ecology*, *9*, 923–934. doi:doi: 10.1046/j.1365-294X.2000.00946.x.

Bukata, A. R., & Kyser, T. K. (2008). Tree-ring elemental concentrations in oak do not necessarily passively record changes in bioavailability. *Science of the Total Environment*, *390*, 275–286. doi:doi: 10.1016/j.scitotenv.2007.09.005.

Buras, A., Hirsch, F., Schneider, A., Scharnweber, T., van der Maaten, E., Cruz-García, R., Raab, T., & Wilmking, M. (2020). Reduced above-ground growth and wood density but increased wood chemical concentrations of Scots pine on relict charcoal hearths. *Science of the Total Environment*, *717*, 137189. doi:doi: 10.1016/j.scitotenv.2020.137189.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach.* (2nd ed.). New York: Springer New York.

Cernusak, L. A. et al. (2022). Do 2H and 18O in leaf water reflect environmental drivers differently? *New Phytologist*, (pp. 41–51). doi:doi: 10.1111/nph.18113.

Cernusak, L. A. et al. (2016). Stable isotopes in leaf water of terrestrial plants. doi:doi: 10.1111/pce.12703.

Cernusak, L. A., & Ubierna, N. (2022). Carbon Isotope Effects in Relation to CO2 Assimilation by Tree Canopies. (pp. 291–310). doi:doi: 10.1007/978-3-030-92698-4_9.

Chabot, A. A., Hobson, K. A., van Wilgenburg, S. L., McQuat, G. J., & Lougheed, S. C. (2012). Advances in linking wintering migrant birds to their breeding-ground origins using combined analyses of genetic and stable isotope markers. *PLoS ONE*, *7*. doi:doi: 10.1371/journal.pone.0043627.

Challender, D. W., & MacMillan, D. C. (2014). Poaching is more than an enforcement problem. *Conservation Letters*, *7*, 484–494. doi:doi: 10.1111/conl.12082.

Chaves, C. L., Degen, B., Pakull, B., Mader, M., Honorio, E., Ruas, P., Tysklind, N., & Sebbenn, A. M. (2018). Assessing the Ability of Chloroplast and Nuclear DNA Gene Markers to Verify the Geographic Origin of Jatoba (Hymenaea courbaril L.) Timber. *Journal of Heredity*, *109*, 543–552. doi:doi: 10.1093/jhered/esy017.

Chen, X., Deng, Q., Lin, G., Lin, M., & Wei, H. (2018). Changing rainfall frequency affects soil organic carbon concentrations by altering non-labile soil organic carbon concentrations in a tropical monsoon forest. *Science of the Total Environment*, *644*, 762–769. doi:doi: 10.1016/j.scitotenv.2018.07.035.

Chin-A-Fo, H., & Kuijpers, K. (2023). Superyacht builders close their eyes to illegal teak trade.

Colin-Belgrand, M., Ranger, J., & Bouchon, J. (1996). Internal nutrient translocation in chestnut tree stemwood: III. Dynamics across an age series of Castanea sativa (Miller). *Annals of Botany*, *78*, 729–740. doi:doi: 10.1006/anbo.1996.0183.

Copernicus Sentinel. (2023). Sentinel-5P NRTI SO2: Near Real-Time Sulphur Dioxide.

Cortes, D. (2022). isotree: Isolation-Based Outlier Detection.

Cramm, M., & Van Brusselen, J. (2019). *List of priority tree species for the development of reference data for determination of tree species and their geographic origin. Update of the GTTN priority tree species list of 2013.*. Technical Report, Global Timber Tracking Network, European Forest Institute Joensu.

Dalal, R. C., Thornton, C. M., Allen, D. E., & Kopittke, P. M. (2021). A study over 33 years shows that carbon and nitrogen stocks in a subtropical soil are increasing under native vegetation in a changing climate. *Science of the Total Environment*, *772*. doi:doi: 10.1016/j.scitotenv.2021.145019.

D'Andrea, R., Corona, C., Poszwa, A., Belingard, C., Domínguez-Delmás, M., Stoffel, M., Crivellaro, A., Crouzevialle, R., Cerbelaud, F., Costa, G., & Paradis-Grenouillet, S. (2023). Combining conventional tree-ring measurements with wood anatomy and strontium isotope analyses enables dendroprovenancing at the local scale. *Science of the Total Environment*, *858*. doi:doi: 10.1016/j.scitotenv.2022.159887.

Dawson, T. E., & Ehleringer, J. R. (1991). *Streamside trees that do not use stream water*. Technical Report, .

Degen, B., Bouda, H.-N., & Blanc-Jolivet, C. (2015). *ITTO project PD 620/11 Rev.1 (M): Development and implementation of a species identification and timber tracking system with DNA fingerprints and isotopes in Africa*. Technical Report, .

Degen, B., Ward, S. E., Lemes, M. R., Navarro, C., Cavers, S., & Sebbenn, A. M. (2013). Verifying the geographic origin of mahogany (Swietenia macrophylla King) with DNA-fingerprints. *Forensic Science International: Genetics*, *7*, 55–62. doi:doi: 10.1016/j.fsigen.2012.06.003.

Deguilloux, M. F., Pemonge, M. H., Bertel, L., Kremer, A., & Petit, R. J. (2003). Checking the geographical origin of oak wood: Molecular and statistical tools. *Molecular Ecology*, *12*, 1629–1636. doi:doi: 10.1046/j.1365-294X.2003.01836.x.

Deklerck, V. (2023). Timber origin verification using mass spectrometry: Challenges, opportunities, and way forward. *Forensic Science International: Animals and Environments*, *3*, 100057. doi:doi: 10.1016/j.fsiae.2022.100057.

Deklerck, V., Lancaster, C. A., Acker, J. V., Espinoza, E. O., den Bulcke, J. V., & Beeckman, H. (2020). Chemical fingerprinting of wood sampled along a pith-to-bark gradient for individual comparison and provenance identification. *Forests*, *11*. doi:doi: 10.3390/f11010107.

Deklerck, V., Mortier, T., Goeders, N., Cody, R. B., Waegeman, W., Espinoza, E., Van Acker, J., Van den Bulcke, J., & Beeckman, H. (2019). A protocol for automated timber species identification using metabolome profiling. *Wood Science and Technology*, *53*, 953–965. doi:doi: 10.1007/s00226-019-01111-1.

DeRaad, D. A. (2022). ¡scp¿snpfiltr¡/scp¿ : An R package for interactive and reproducible SNP filtering. *Molecular Ecology Resources*, *22*, 2443–2453. doi:doi: 10.1111/1755-0998.13618.

DeWalle, D. R., Swistock, B. R., Sayre, R. G., & Sharpe, W. E. (1991). Spatial Variations of Sapwood Chemistry with Soil Acidity in Appalachian Forests. *Journal of Environmental Quality*, *20*, 486–491. doi:doi: 10.2134/jeq1991.00472425002000020024x.

Dié, A., De Ridder, M., Cherubini, P., Kouamé, F. N., Verheyden, A., Kitin, P., Toirambe, B. B., Van Den Bulcke, J., Van Acker, J., & Beeckman, H. (2015). Tree rings show a different climatic response in a managed and a non-managed plantation of teak (Tectona grandis) in

West Africa. *IAWA Journal*, *36*, 409–427. doi:doi: 10.1163/22941932-20150111.

Domínguez-Delmás, M., Rich, S., Traoré, M., Hajj, F., Poszwa, A., Akhmetzyanov, L., García-González, I., & Groenendijk, P. (2020). Tree-ring chronologies, stable strontium isotopes and biochemical compounds: Towards reference datasets to provenance Iberian shipwreck timbers. *Journal of Archaeological Science: Reports*, *34*. doi:doi: 10.1016/j.jasrep.2020.102640.

Dormontt, E., Jardine, D., van Dijk, K.-J., Dunker, B., Dixon, R., Hipkins, V., Tobe, S., Linacre, A., & Lowe, A. (2020). Forensic validation of a SNP and INDEL panel for individualisation of timber from bigleaf maple (Acer macrophyllum Pursch). *Forensic Science International: Genetics*, *46*, 102252. doi:doi: 10.1016/j.fsigen.2020.102252.

Dormontt, E. E. et al. (2015). Forensic timber identification: It's time to integrate disciplines to combat illegal logging. *Biological Conservation*, *191*, 790–798. doi:doi: 10.1016/j.biocon.2015.06.038.

Doumenge, C., & Sene, V. O. (2012). Lophira alata Banks ex C.F.Gaertn.

Driscoll, A. W., Howa, J. D., Bitter, N. Q., & Ehleringer, J. R. (2020). A predictive spatial model for roasted coffee using oxygen isotopes of α-cellulose. *Rapid Communications in Mass Spectrometry*, *34*. doi:doi: 10.1002/rcm.8626.

Drivelos, S. A., & Georgiou, C. A. (2012). Multi-element and multi-isotope-ratio analysis to determine the geographical origin of foods in the European Union. *TrAC Trends in Analytical Chemistry*, *40*, 38–51. doi:doi: 10.1016/j.trac.2012.08.003.

Duminil, J., Brown, R. P., Ewédjè, E. E. B., Mardulyn, P., Doucet, J. L., & Hardy, O. J. (2013). Large-scale pattern of genetic differentiation within African rainforest trees: Insights on the roles of ecological gradients and past climate changes on the evolution of Erythrophleum spp (Fabaceae). *BMC Evolutionary Biology*, *13*. doi:doi: 10.1186/1471-2148-13-195.

Dumolin, S., Demesure, B., & Petit, R. J. (1995). *Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method*. Technical Report, .

Durand, S. R., Shelley, P. H., Antweiler, R. C., & Taylor, H. E. (1999). Trees, chemistry, and prehistory in the American Southwest. *Journal of Archaeological Science*, *26*, 185–203. doi:doi: 10.1006/jasc.1998.0315.

Ehleringer, J. R., Casale, J. F., Lott, M. J., & Ford, V. L. (2000). Tracing the geographical origin of cocaine: Cocaine carries a chemical fingerprint from the region where the coca was grown. *Nature*, *408*, 311–312. doi:doi: 10.1038/35042680.

Ehleringer, J. R., Thompson, A. H., Podlesak, D. W., Bowen, G. J., Chesson, L. A., Cerling, T. E., Park, T., Dostie, P., & Schwarcz, H. (2010). A framework for the incorporation of isotopes and isoscapes in geospatial forensic investigations. In *Isoscapes: Understanding movement, pattern, and process on Earth through isotope mapping* (pp. 357–387).

English, N. B., Betancourt, J. L., Dean, J. S., & Quade, J. (2001). Strontium isotopes reveal distant sources of architectural timber in Chaco Canyon, New Mexico. *Proceedings of the National Academy of Sciences*, *98*, 11891–11896. doi:doi: 10.1073/pnas.211305498.

Engone Obiang, N. L., Ngomanda, A., White, L. J. T., Jeffery, K. J., Chézeaux, , & Picard, N. (2012). A growth model for azobé, Lophira alata, in Gabon. *BOIS ET FORÊTS DES TROPIQUES, N° 314 (4)*, .

Epstein, S., Yapp, C. J., & Hall, J. H. (1976). The determination of the D/H ratio of non-exchangeable hydrogen in cellulose extracted from aquatic and land plants. *Earth and Planetary Science Letters*, *30*, 241–251. doi:doi: 10.1016/0012-821X(76)90251-X.

Ewédjè, E.-E. B. K., Jansen, S., Koffi, G. K., Staquet, A., Piñeiro, R., Essaba, R. A., Obiang, N. L. E., Daïnou, K., Biwolé, A. B., Doucet, J.-L., & Hardy, O. J. (2020). Species delimitation in the African tree genus Lophira (Ochnaceae) reveals cryptic genetic variation. *Conservation Genetics*, . doi:doi: 10.1007/s10592-020-01265-7.

February, E. (2000). Archaeological charcoal and dendrochronology to reconstruct pastenvironments of southern Africa. *South African Journal of Science*, *96*, 111–115.

Ferrio, J. P., & Voltas, J. (2005). Carbon and oxygen isotope ratios in wood constituents of Pinus halepensis as indicators of precipitation, temperature and vapour pressure deficit. *Tellus, Series B: Chemical and Physical Meteorology*, *57*, 164–173. doi:doi: 10.1111/J.1600-0889.2005.00137.X.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*, 4302–4315. doi:doi: 10.1002/JOC.5086.

Finch, K., Espinoza, E., Jones, F. A., & Cronn, R. (2017). Source Identification of Western Oregon Douglas-Fir Wood Cores Using Mass Spectrometry and Random Forest Classification. *Applications in Plant Sciences*, *5*, 1600158. doi:doi: 10.3732/apps.1600158.

Fonti, M., Vaganov, E., Wirth, C., Shashkin, A., Astrakhantseva, N., & Schulze, -D. (2018). Age-Effect on Intra-Annual δ13C-Variability within Scots Pine Tree-Rings from Central Siberia. *Forests*, *9*, 364. doi:doi: 10.3390/f9060364.

Förstel, H., Boner, M., Höltken, A. M., Fladung, M., Degen, B., & Zahnen, J. (2011). *Final project report WWF Germany - Fighting illegal logging through the introduction of a combination of the isotope method for identifying the origins of timber and DNA analysis for differentiation of tree species Project*. Technical Report, WWF Germany Berlin.

Fuller, T. L. et al. (2019). Assessing the impact of China's timber industry on Congo Basin land use change. *Area*, *51*, 340–349. doi:doi: 10.1111/area.12469.

Gan, J., Cerutti, P. O., Masiero, M., Pettenella, D., Andrighetto, N., & Dawson, T. (2016). Quantifying Illegal Logging and Related Timber Trade. In *IUFRO World Series Volume 35 Illegal Logging and Related Timber* (pp. 38–58). volume 35.

Gasson, P. (2011). How precise can wood identification be? Wood anatomy's role in support of the legal timber trade, especially cites. *IAWA Journal*, *32*, 137–154. doi:doi: 10.1163/22941932-90000049.

Gasson, P. E., Lancaster, C. A., Young, R., Redstone, S., Miles-Bunch, I. A., Rees, G., Guillery,

R. P., Parker-Forney, M., & Lebow, E. T. (2021). WorldForestID: Addressing the need for standardized wood reference collections to support authentication analysis technologies; a way forward for checking the origin and identity of traded timber. *PLANTS, PEOPLE, PLANET*, *3*, 130–141. doi:doi: 10.1002/ppp3.10164.

Gazel, J. (1956). Carte geologique a 1:1 000 000 du Cameroun. Planche 1-A nord.

GBIF.org (2021). GBIF Occurrence Download https://doi.org/10.15468/dl.wpwm8v. Date accessed: 2021-11-25.

Global Witness (2019). *Buyers beware. European companies buying timber from Industrie Forestière du Congo risk falling foul of EU laws*. Technical Report, .

Gonzalvez, A., Armenta, S., & de la Guardia, M. (2009). Trace-element composition and stable-isotope ratio for discrimination of foods with Protected Designation of Origin. *TrAC Trends in Analytical Chemistry*, *28*, 1295–1311. doi:doi: 10.1016/j.trac.2009.08.001.

Gonzálvez, A., Llorens, A., Cervera, M. L., Armenta, S., & de la Guardia, M. (2009). Elemental fingerprint of wines from the protected designation of origin Valencia. *Food Chemistry*, *112*, 26–34. doi:doi: 10.1016/j.foodchem.2008.05.043.

Gori, Y., Stradiotti, A., & Camin, F. (2018). Timber isoscapes. A case study in a mountain area in the Italian Alps. *PLOS ONE*, *13*, e0192970. doi:doi: 10.1371/journal.pone.0192970.

Griboff, J., Horacek, M., Wunderlin, D. A., & Monferrán, M. V. (2021). Differentiation Between Argentine and Austrian Red and White Wines Based on Isotopic and Multi-Elemental Composition. *Frontiers in Sustainable Food Systems*, *5*, 1–9. doi:doi: 10.3389/fsufs.2021.657412.

Groenenberg, J. E., Römkens, P. F., Zomeren, A. V., Rodrigues, S. M., & Comans, R. N. (2017). Evaluation of the Single Dilute (0.43 M) Nitric Acid Extraction to Determine Geochemically Reactive Elements in Soil. *Environmental Science and Technology*, *51*, 2246–2253. doi:doi: 10.1021/acs.est.6b05151.

Groenendijk, P., Sass-Klaassen, U., Bongers, F., & Zuidema, P. A. (2014). Potential of tree-ring analysis in a wet tropical forest: A case study on 22 commercial tree species in Central Africa. *Forest Ecology and Management*, *323*, 65–78. doi:doi: 10.1016/j.foreco.2014.03.037.

Gruber, B., & Adamack, A. (2022). PopGenReport.

Harris, I., Osborn, T., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*.

Heineman, K. D., Turner, B. L., & Dalling, J. W. (2016). Variation in wood nutrients along a tropical soil fertility gradient. *The New phytologist*, *211*, 440–454. doi:doi: 10.1111/nph.13904.

Helle, G., Pauly, M., Heinrich, I., Schollän, K., Balanzategui, D., & Schürheck, L. (2022). Stable Isotope Signatures of Wood, its Constituents and Methods of Cellulose Extraction. (pp. 135–190). doi:doi: 10.1007/978-3-030-92698-4_5.

Hellmann, L., Tegel, W., Geyer, J., Kirdyanov, A. V., Nikolaev, A. N., Eggertsson, , Altman, J., Reinig, F., Morganti, S., Wacker, L., & Büntgen, U. (2017). Dendro-provenancing of Arctic driftwood. *Quaternary Science Reviews*, *162*, 1–11. doi:doi: 10.1016/j.quascirev.2017.02.025.

Hengl, T. et al. (2021). African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Scientific Reports*, *11*. doi:doi: 10.1038/s41598-021-85639-y.

Herbauts, J., Penninckx, V., Gruber, W., & Meerts, P. (2002). Radial variations in cation exchange capacity and base saturation rate in the wood of pedunculate oak and European beech. *Canadian Journal of Forest Research*, *32*, 1829–1837. doi:doi: 10.1139/x02-097.

Hevia, A., Sánchez-Salguero, R., Camarero, J. J., Buras, A., Sangüesa-Barreda, G., Galván, J. D., & Gutiérrez, E. (2017). Towards a better understanding of long-term wood-chemistry variations in old-growth forests: A case study on ancient Pinus uncinata trees from the Pyrenees. *Science of the Total Environment*, *625*, 220–232. doi:doi: 10.1016/j.scitotenv.2017.12.229.

Hietz, P., Horsky, M., Prohaska, T., Lang, I., & Grabner, M. (2015). High-resolution densitometry and elemental analysis of tropical wood. *Trees*, *29*, 487–497. doi:doi: 10.1007/s00468-014-1126-7.

Hijmans, R., Ghosh, A., & Mandel, A. (2022). geodata: Download Geographic Data. R package version 0.5-3.

Hijmans, R. J. (2021a). raster: Geographic Data Analysis and Modeling. R package version 3.5-9.

Hijmans, R. J. (2021b). Terra: Spatial Data Analysis.

Hijmans, R. J. (2022). geosphere: Spherical Trigonometry, R package version 1.5-18.

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2020). dismo: Species Distribution Modeling. R package version 1.3-5.

Hillis, W. E. (1999). The Formation of Heartwood and Its Extractives. In *Phytochemicals in Human Health Protection, Nutrition, and Plant Defense* (pp. 215–253). Springer, Boston, MA. doi:doi: 10.1007/978-1-4615-4689-4_9.

Hoare, A. (2015). Tackling Illegal Logging and the Related Trade: What Progress and Where Next? doi:doi: 10.1016/S0966-3274(00)00025-3.

Hoare, A., & Uehara, T. H. K. (2022). *Establishing fair and sustainable forest economies: Lessons learned from tackling illegal logging*. Technical Report September, Royal Institute of International Affairs London. URL: `https://chathamhouse.soutron.net/Portal/Public/en-GB/RecordView/Index/191315`. doi:doi: 10.55317/9781784135386.

Hoare, A., Young, D., Uehara, T., Seidu, K., Birikorang, G., Soh, L. W., & Kamga, J. K. (2020). *Forest sector accountability in Cameroon and Ghana Exploring the impact of transparency and participation Energy, Environment and Resources Programme*. Technical Report, .

Hobson, K. A., Bowen, G. J., Wassenaar, L. I., Ferrand, Y., & Lormee, H. (2004). Using stable hydrogen and oxygen isotope measurements of feathers to infer geographical origins of migrating European birds. *Oecologia*, *141*, 477–488. doi:doi: 10.1007/s00442-004-1671-7.

Horacek, M., Jakusch, M., & Krehan, H. (2009). Control of origin of larch wood: Discrimination between European (Austrian) and Siberian origin by stable isotope analysis. *Rapid Communications in Mass Spectrometry*, *23*, 3688–3692. doi:doi: 10.1002/rcm.4309.

Horacek, M., Rees, G., Boner, M., & Zahnen, J. (2018). Comment on: Developing forensic tools for an African timber: [...], by Vlam et al., 2018. *Biological Conservation*, *226*, 333–334. doi:doi: 10.1016/j.biocon.2018.06.037.

Houba, V. J. G., van der Lee, J. J., Novozamski, I., & Wallinga, I. (1995). Part 5B: Soil Analysis Procedures. In *Soil and Plant analysis*. Wageningen: Agricultural University Wageningen.

Houba, V. J. G., Temminghoff, E. J. M., Gaikhorst, G. A., & van Vark, W. (2000). Soil analysis procedures using 0.01 M calcium chloride as extraction reagent. *Communications in Soil Science and Plant Analysis*, *31*, 1299–1396. doi:doi: 10.1080/00103620009370514.

IAEA/WMO (2020). *Global Network of Isotopes in Precipitation. The GNIP Database.*. Technical Report, International Atomic Energy Agency Vienna, Austria.

International Tropical Timber Organization (2021). *Biennial review and assessment of the world timber situation 2019-2020*. Technical Report, .

Jahanbanifard, M., Beckers, V., Koch, G., Beeckman, H., Gravendeel, B., Verbeek, F., Baas, P., Priester, C., & Lens, F. (2020). Description and evolution of wood anatomical characters in the ebony wood genus Diospyros and its close relatives (Ebenaceae): A first step towards combatting illegal logging. *IAWA Journal*, *41*, 577–619. doi:doi: 10.1163/22941932-bja10040.

Jahanbanifard, M., Gravendeel, B., Lens, F., & Verbeek, F. (2019). Ebony Wood Identification to Battle Illegal Trade. *Biodiversity Information Science and Standards*, *3*. doi:doi: 10.3897/biss.3.37084.

Joebstl, D., Bandoniene, D., Meisel, T., & Chatzistathis, S. (2010). Identification of the geographical origin of pumpkin seed oil by the use of rare earth elements and discriminant analysis. *Food Chemistry*, *123*, 1303–1309. doi:doi: 10.1016/j.foodchem.2010.06.009.

Johnson, C. M., Vieira, I. C., Zarin, D. J., Frizano, J., & Johnson, A. H. (2001). Carbon and nutrient storage in primary and secondary forests in eastern Amazônia. *Forest Ecology and Management*, *147*, 245–252. doi:doi: 10.1016/S0378-1127(00)00466-7.

Jolivet, C., & Degen, B. (2012). Use of DNA fingerprints to control the origin of sapelli timber (Entandrophragma cylindricum) at the forest concession level in Cameroon. *Forensic Science International: Genetics*, *6*, 487–493. doi:doi: 10.1016/j.fsigen.2011.11.002.

Kabata-Pendias, A. (2011). *Trace elements in soils and plants*. (4th ed.). CRC Press.

Kagawa, A., Kuroda, K., Abe, H., Fujii, T., & Itoh, Y. (2007). Stable isotopes and inorganic elements as potential indicators of geographic origin of Southeast Asian timber. In *Proceedings of the International Symposium on Development of Improved Methods to Identify Shorea*

Species Wood and its Origin (pp. 39–44).

Kagawa, A., & Leavitt, S. W. (2010). Stable carbon isotopes of tree rings as a tool to pinpoint the geographic origin of timber. *Journal of Wood Science*, *56*, 175–183. doi:doi: 10.1007/s10086-009-1085-6.

Karsenty, A. (2018). The legal institutionalization of FSC certification in Gabon (commentary).

Klaassen, R. K. W. M. (Ed.) (2018). *Houtvademecum.* (11th ed.). Centrum Hout Almere.

Kleinschmit, D., Mansourian, S., Wildburger, C., & Purret, A. (2016). *Illegal Logging and Related Timber Trade – Dimensions, Drivers, Impacts and Responses. A Global Scientific Rapid Response Assessment Report. IUFRO World Series Volume 35*. International Union of Forest Research Organizations (IUFRO).

Komdeur, E. F., & Ingenbleek, P. T. (2021). The potential of blockchain technology in the procurement of sustainable timber products. *International Wood Products Journal*, *12*, 249–257. doi:doi: 10.1080/20426445.2021.1967624.

Kuhl, E., Zang, C., Esper, J., Riechelmann, D. F. C., Büntgen, U., Briesch, M., Reinig, F., Römer, P., Konter, O., Schmidhalter, M., & Hartl, C. (2023). Using machine learning on tree-ring data to determine the geographical provenance of historical construction timbers. *Ecosphere*, *14*. doi:doi: 10.1002/ecs2.4453.

Kurita, N., Ichiyanagi, K., Matsumoto, J., Yamanaka, M. D., & Ohata, T. (2009). The relationship between the isotopic content of precipitation and the precipitation amount in tropical regions. *Journal of Geochemical Exploration*, *102*, 113–122. doi:doi: 10.1016/j.gexplo.2009.03.002.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*. doi:doi: 10.18637/jss.v082.i13.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. doi:doi: 10.1038/nmeth.1923.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR : An R Package for Multivariate Analysis. *Journal of Statistical Software*, *25*, 1–18. doi:doi: 10.18637/jss.v025.i01.

Lee, S., Park, B. S., Lee, D., Chung, H., & Lee, K.-S. (2015). Spatial variability in hydrogen and oxygen isotopic composition of Korean Red Pine and its implication for tracing wood origin. *Environmental Earth Sciences*, *73*, 8045–8052. doi:doi: 10.1007/s12665-014-3960-8.

Lehmann, M. M., Schuler, P., Cormier, M.-A., Allen, S. T., Leuenberger, M., & Voelker, S. (2022). The Stable Hydrogen Isotopic Signature: From Source Water to Tree Rings. (pp. 331–359). doi:doi: 10.1007/978-3-030-92698-4_11.

Lens, F., Liang, C., Guo, Y., Tang, X., Jahanbanifard, M., da Silva, F. S. C., Ceccantini, G., & Verbeek, F. J. (2020). Computer-assisted timber identification based on features extracted from microscopic wood sections. *IAWA Journal*, *0*, 1–21. doi:doi: 10.1163/22941932-bja10029.

Li, Y., Via, B. K., Young, T., & Li, Y. (2019). Visible-Near infrared spectroscopy and chemometric methods for wood density prediction and origin/species identification. *Forests*, *10*. doi:doi: 10.3390/F10121078.

Liaw, A., & Wiener, M. (2002). randomForest: Classification and Regression by randomForest. R package version 4.6-14.

Lira-Martins, D., Humphreys-Williams, E., Strekopytov, S., Ishida, F. Y., Quesada, C. A., & Lloyd, J. (2019). Tropical tree branch-leaf nutrient scaling relationships vary with sampling location. *Frontiers in Plant Science*, *10*. doi:doi: 10.3389/fpls.2019.00877.

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining* (pp. 413–422). IEEE. doi:doi: 10.1109/ICDM.2008.17.

Locosselli, G. M., Brienen, R. J. W., Leite, M. d. S., Gloor, M., Krottenthaler, S., Oliveira, A. A. d., Barichivich, J., Anhuf, D., Ceccantini, G., Schöngart, J., & Buckeridge, M. (2020). Global tree-ring analysis reveals rapid decrease in tropical tree longevity with temperature. *Proceedings of the National Academy of Sciences*, *117*, 33358–33364. doi:doi: 10.1073/pnas.2003873117.

Low, M. C. et al. (2022). Tracing the world's timber: the status of scientific verification technologies for species and origin identification. *IAWA Journal*, *44*, 63–84. doi:doi: 10.1163/22941932-bja10097.

Lowe, A. J., Dormontt, E. E., Bowie, M. J., Degen, B., Gardner, S., Thomas, D., Clarke, C., Rimbawanto, A., Wiedenhoeft, A., Yin, Y., & Sasaki, N. (2016). Opportunities for Improved Transparency in the Timber Trade through Scientific Verification. *BioScience*, *66*, 990–998. doi:doi: 10.1093/biosci/biw129.

Lowe, A. J., Wong, K. N., Tiong, Y. S., Iyerh, S., & Chew, F. T. (2010). A DNA method to verify the integrity of timber supply chains; Confirming the legal sourcing of merbau timber from logging concession to sawmill. *Silvae Genetica*, *59*, 263–268. doi:doi: 10.1515/sg-2010-0037.

Ma, G., Zhang, Y., Zhang, J., Wang, G., Chen, L., Zhang, M., Liu, T., Liu, X., & Lu, C. (2016). Determining the geographical origin of Chinese green tea by linear discriminant analysis of trace metals and rare earth elements: Taking Dongting Biluochun as an example. *Food Control*, *59*, 714–720. doi:doi: 10.1016/j.foodcont.2015.06.037.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2021). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.2.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. doi:doi: 10.1016/S0169-5347(03)00008-9.

Martin, R. R., Naftel, S. J., Macfie, S. M., Jones, K. W., Feng, H., & Trembley, C. (2006). High variability of the metal content of tree growth rings as measured by synchrotron micro x-ray fluorescence spectrometry. *X-Ray Spectrometry*, *35*, 57–62. doi:doi: 10.1002/xrs.817.

Mascarello, M., Amalfi, M., Asselman, P., Smets, E., Hardy, O. J., Beeckman, H., & Janssens,

S. B. (2021). Genome skimming reveals novel plastid markers for the molecular identification of illegally logged African timber species. *PLoS ONE*, *16*. doi:doi: 10.1371/journal.pone.0251655.

May, C. (2017). Transnational crime and the developing world.

McCune, B., & Grace, J. (2002). Chapter 6: Distance Measures. In *Analysis of Ecological Communities*. Mjm Software Design.

McFerrin, L. (2013). HDMD: Structural Analysis Tools for High Dimensional Molecular Data. R package version 1.2.

Meerts, P. (2002). Mineral nutrient concentrations in sapwood and heartwood: a literature review. *Annals of Forest Science*, *59*, 713–722. doi:doi: 10.1051/forest:2002059.

Meinshausen, N. (2006). Quantile Regression Forests. *Journal of Machine Learning Research*, *7*, 983–999.

Mendez, J. C., Van Eynde, E., Hiemstra, T., & Comans, R. N. (2022). Surface reactivity of the natural metal (hydr)oxides in weathered tropical soils. *Geoderma*, *406*, 115517. doi:doi: 10.1016/j.geoderma.2021.115517.

Momoshima, N., & Bondietti, E. A. (1990). Cation binding in wood: applications to understanding historical changes in divalent cation availability to red spruce. *Canadian Journal of Forest Research*, *20*, 1840–1849. doi:doi: 10.1139/x90-247.

Mundo, I. A., Murray, C., Grosso, M., Rao, M. P., Cook, E. R., & Villalba, R. (2022). Dendrochronological dating and provenance determination of a 19th century whaler in Patagonia (Puerto Madryn, Argentina). *Dendrochronologia*, *74*. doi:doi: 10.1016/j.dendro.2022.125980.

Munoz, M. F., Zhang, K., Shahzad, A., & Ouhimmou, M. (2021). LogLog: A blockchain solution for tracking and certifying wood volumes. In *IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2021*. Institute of Electrical and Electronics Engineers Inc. doi:doi: 10.1109/ICBC51069.2021.9461153.

Musah, R. A., Espinoza, E. O., Cody, R. B., Lesiak, A. D., Christensen, E. D., Moore, H. E., Maleknia, S., & Drijfhout, F. P. (2015). A High Throughput Ambient Mass Spectrometric Approach to Species Identification and Classification from Chemical Fingerprint Signatures. *Scientific Reports*, *5*, 11520. doi:doi: 10.1038/srep11520.

Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling. *Ecography*, *37*, 191–203.

Nehlich, O. (2015). The application of sulphur isotope analyses in archaeological research: A review. *Earth-Science Reviews*, *142*, 1–17. doi:doi: 10.1016/j.earscirev.2014.12.002.

Nelleman, C., Henriksen, R., Kreilhuber, A., Stewart, D., Kotsovou, M., Raxter, P., Mrema, E., & Barrat, S. (2016). *The Rise of Environmental Crime. A Growing Threat to Natural Resources, Peace, Development and Security. A UNEP- INTERPOL Rapid Response Assessment.*. United Nations Environment Programme and RHIPTO Rapid Response– Norwegian Centre for Global Analyses.

Nellemann, C. (2012). *Green Carbon, Black Trade: Illegal logging, tax fraud and laundering in the world's tropical forests. A Rapid Response Assessment. United Nations Environment Programme, GRID- Arendal. www.grida.no*.

Ng, C. H., Lee, S. L., Tnah, L. H., Ng, K. K. S., Lee, C. T., Diway, B., & Khoo, E. (2017). Geographic origin and individual assignment of Shorea platyclados (Dipterocarpaceae) for forensic identification. *PLOS ONE*, *12*, e0176158. doi:doi: 10.1371/journal.pone.0176158.

Ng, C. H., Ng, K. K. S., Lee, S. L., Zakaria, N.-F., Lee, C. T., & Tnah, L. H. (2022). DNA databases of an important tropical timber tree species Shorea leprosula (Dipterocarpaceae) for forensic timber identification. *Scientific Reports*, *12*, 9546. doi:doi: 10.1038/s41598-022-13697-x.

Nurrochmat, D. R., Dharmawan, A. H., Obidzinski, K., Dermawan, A., & Erbaugh, J. T. (2016). Contesting national and international forest regimes: Case of timber legality certification for community forests in Central Java, Indonesia. *Forest Policy and Economics*, *68*, 54–64. doi:doi: 10.1016/j.forpol.2014.09.008.

Oksanen, J. et al. (2020). vegan: Community Ecology Package. R package version 2.5-7.

Oulhote, Y., Le Bot, B., Deguen, S., & Glorennec, P. (2011). Using and interpreting isotope data for source identification. doi:doi: 10.1016/j.trac.2010.10.015.

Paradis, E. (2010). pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*, 419–420. doi:doi: 10.1093/bioinformatics/btp696.

Paredes-Villanueva, K., Boom, A., Ottenburghs, J., Van Der Sleen, P., Manzanedo, R. D., Bongers, F., & Zuidema, P. A. (2022). Isotopic Characterization of Cedrela to Verify Species and Regional Provenance of Bolivian Timber. *Tree-Ring Research*, *78*, 73–89. doi:doi: 10.3959/2021-17.

Paredes-Villanueva, K., Espinoza, E., Ottenburghs, J., Sterken, M. G., Bongers, F., & Zuidema, P. A. (2018). Chemical differentiation of Bolivian Cedrela species as a tool to trace illegal timber trade. *Forestry: An International Journal of Forest Research*, *91*, 603–613. doi:doi: 10.1093/forestry/cpy019.

Paredes-Villanueva, K., de Groot, G. A., Laros, I., Bovenschen, J., Bongers, F., & Zuidema, P. A. (2019). Genetic differences among Cedrela odorata sites in Bolivia provide limited potential for fine-scale timber tracing. *Tree Genetics and Genomes*, *15*. doi:doi: 10.1007/s11295-019-1339-4.

Pepi, S., & Vaccaro, C. (2018). Geochemical fingerprints of “Prosecco” wine based on major and trace elements. *Environmental Geochemistry and Health*, *40*, 833–847. doi:doi: 10.1007/s10653-017-0029-0.

Persits, F., Ahlbrandt, T., Tuttle, M., Charpentier, R., Brownfield, M., & Takahashi, K. (1997). Maps showing geology, oil and gas fields and geological provinces of Africa. . *U.S. Geological Survey Open-File Report 97-470-A*, .

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259. doi:doi: 10.1016/j.ecolmodel.2005.03.026.

Pilgrim, T. S., Watling, R. J., & Grice, K. (2010). Application of trace element and stable isotope signatures to determine the provenance of tea (Camellia sinensis) samples. *Food Chemistry*, *118*, 921–926. doi:doi: 10.1016/j.foodchem.2008.08.077.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, *7*, 217–240. doi:doi: 10.5194/soil-7-217-2021.

Pribyl, D. W. (2010). A critical review of the conventional SOC to SOM conversion factor. doi:doi: 10.1016/j.geoderma.2010.02.003.

Price, E. R., Miles-Bunch, I., Gasson, P. E., & Lancaster, C. A. (2021). Inference of origin of Pterocarpus timber by chemical profiling of ambient ionization mass spectra. *Forensic Science International: Animals and Environments*, *1*, 100032. doi:doi: 10.1016/j.fsiae.2021.100032.

Quan, Q., Wang, C., He, N., Zhang, Z., Wen, X., Su, H., Wang, Q., & Xue, J. (2014). Forest type affects the coupled relationships of soil C and N mineralization in the temperate forests of northern China. *Scientific Reports*, *4*. doi:doi: 10.1038/srep06584.

R Core Team (2021). R: A Language and Environment for Statistical Computing.

R core team (2023). R: A Language and Environment for Statistical Computing.

Reiche, J. et al. (2021). Forest disturbance alerts for the Congo Basin using Sentinel-1. *Environmental Research Letters*, *16*. doi:doi: 10.1088/1748-9326/abd0a8.

Rich, S., Manning, S. W., Degryse, P., Vanhaecke, F., & Van Lerberghe, K. (2016). Provenancing East Mediterranean cedar wood with the 87Sr/86Sr strontium isotope ratio. *Archaeological and Anthropological Sciences*, *8*, 467–476. doi:doi: 10.1007/s12520-015-0242-7.

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

Rieuwerts, J. S. (2007). The mobility and bioavailability of trace metals in tropical soils: a review. *Chemical Speciation & Bioavailability*, *19*, 75–85. doi:doi: 10.3184/095422907X211918.

Roden, J. S., & Ehleringer, J. R. (2000). *Hydrogen and Oxygen Isotope Ratios of Tree Ring Cellulose for Field-Grown Riparian Trees*. Technical Report 4, .

Römkens, P. F., Guo, H. Y., Chu, C. L., Liu, T. S., Chiang, C. F., & Koopmans, G. F. (2009). Prediction of Cadmium uptake by brown rice and derivation of soil-plant transfer models to improve soil protection guidelines. *Environmental Pollution*, *157*, 2435–2444. doi:doi: 10.1016/j.envpol.2009.03.009.

Russell, A. E., Hall, S. J., & Raich, J. W. (2017). Tropical tree species traits drive soil cation dynamics via effects on pH: a proposed conceptual framework. *Ecological Monographs*, *87*, 685–701. doi:doi: 10.1002/ecm.1274.

Rybníček, M., Kyncl, T., Vavrčík, H., & Kolář, T. (2022). Dendrochronology improves understanding of the charcoal production history. *Dendrochronologia*, *75*. doi:doi: 10.1016/j.dendro.2022.125994.

Sandak, A., Sandak, J., & Negri, M. (2011). Relationship between near-infrared (NIR) spectra and the geographical provenance of timber. *Wood Science and Technology*, *45*, 35–48. doi:doi: 10.1007/s00226-010-0313-y.

Sandak, J., Sandak, A., & Meder, R. (2016). Assessing trees, wood and derived products with near infrared spectroscopy: Hints and tips. In *Journal of Near Infrared Spectroscopy* (pp. 485–505). IM Publications LLP volume 24. doi:doi: 10.1255/jnirs.1255.

Sass-Klaassen, U., Vernimmen, T., & Baittinger, C. (2008). Dendrochronological dating and provenancing of timber used as foundation piles under historic buildings in The Netherlands. *International Biodeterioration and Biodegradation*, *61*, 96–105. doi:doi: 10.1016/j.ibiod.2007.07.013.

Scharnweber, T., Hevia, A., Buras, A., van der Maaten, E., & Wilmking, M. (2016). Common trends in elements? Within- and between-tree variations of wood-chemistry measured by X-ray fluorescence — A dendrochemical study. *Science of the Total Environment*, *566-567*, 1245–1253. doi:doi: 10.1016/j.scitotenv.2016.05.182.

Schuler, P., Cormier, M. A., Werner, R. A., Buchmann, N., Gessler, A., Vitali, V., Saurer, M., & Lehmann, M. M. (2022). A high-temperature water vapor equilibration method to determine non-exchangeable hydrogen isotope ratios of sugar, starch and cellulose. *Plant Cell and Environment*, *45*, 12–22. doi:doi: 10.1111/pce.14193.

Shirmohammadi, M. (2023). A Review of Traceability Systems in the Timber Industry. In *Wood Industry - Past, Present and Future Outlook*. IntechOpen. doi:doi: 10.5772/intechopen.106704.

Silva, D. C., Pastore, T. C., Soares, L. F., de Barros, F. A., Bergo, M. C., Coradin, V. T., Gontijo, A. B., Sosa, M. H., Chacón, C. B., & Braga, J. W. (2018). Determination of the country of origin of true mahogany (Swietenia macrophylla King) wood in five Latin American countries using handheld NIR devices and multivariate data analysis. *Holzforschung*, *72*, 521–530. doi:doi: 10.1515/hf-2017-0160.

Siriwat, P., & Nijman, V. (2023). Quantifying the illegal high-value rosewood trade and criminal trade networks in the Greater Mekong Region. *Biological Conservation*, *277*. doi:doi: 10.1016/j.biocon.2022.109826.

Van der Sleen, P., Groenendijk, P., & Zuidema, P. A. (2015). Tree-ring δ18O in African mahogany (Entandrophragma utile) records regional precipitation and can be used for climate reconstructions. *Global and Planetary Change*, *127*, 58–66. doi:doi: 10.1016/J.GLOPLACHA.2015.01.014.

van der Sleen, P., Soliz-Gamboa, C. C., Helle, G., Pons, T. L., Anten, N. P. R., & Zuidema, P. A. (2014). Understanding causes of tree growth response to gap formation: 13C-values in tree rings reveal a predominant effect of light. *Trees*, *28*, 439–448. doi:doi: 10.1007/s00468-

013-0961-2.

van der Sleen, P., Vlam, M., Groenendijk, P., Anten, N. P. R., Bongers, F., Bunyavejchewin, S., Hietz, P., Pons, T. L., & Zuidema, P. A. (2015). 15N in tree rings as a bio-indicator of changing nitrogen cycling in tropical forests: an evaluation at three sites using two sampling methods. *Frontiers in Plant Science*, *6*, 229. doi:doi: 10.3389/fpls.2015.00229.

van der Sleen, P., Zuidema, P. A., & Pons, T. (2022). Stable Isotopes in Tree Rings of Tropical Forests. In *Stable Isotopes in Tree Rings: Inferring Physiological, Climatic and Environmental Responses* (pp. 631–649).

van der Sleen, P., Zuidema, P. A., & Pons, T. L. (2017). Stable isotopes in tropical tree rings: theory, methods and applications. *Functional Ecology*, *31*, 1674–1689. doi:doi: 10.1111/1365-2435.12889.

Smith, K. T., Balouet, J. C., Shortle, W. C., Chalot, M., Beaujard, F., Grudd, H., Vroblesky, D. A., & Burken, J. G. (2014). Dendrochemical patterns of calcium, zinc, and potassium related to internal factors detected by energy dispersive X-ray fluorescence (EDXRF). *Chemosphere*, *95*, 58–62. doi:doi: 10.1016/j.chemosphere.2013.08.017.

Snel, F. A., Braga, J. W., da Silva, D., Wiedenhoeft, A. C., Costa, A., Soares, R., Coradin, V. T., & Pastore, T. C. (2018). Potential field-deployable NIRS identification of seven Dalbergia species listed by CITES. *Wood Science and Technology*, *52*, 1411–1427. doi:doi: 10.1007/s00226-018-1027-9.

Boekhout van Solinge, T., Zuidema, P. A., Vlam, M., Cerutti, P. O., & Yemelin, V. (2016). Organized forest crime: a criminological analysis with suggestions from timber forensics. In *Illegal Logging and Related Timber Trade – Dimensions, Drivers, Impacts and Responses. A Global Scientific Rapid Response Assessment Report* (pp. 81–98). Vienna: International Union of Forest Research Organizations (IUFRO).

Soong, J. L. et al. (2020). Soil properties explain tree growth and mortality, but not biomass, across phosphorus-depleted tropical forests. *Scientific Reports*, *10*. doi:doi: 10.1038/s41598-020-58913-8.

Štulc, A. I. et al. (2023). Multi-elemental and Strontium-Neodymium Isotopic Signatures in Charred Wood: Potential for Wood Provenance Studies. *International Journal of Wood Culture*, (pp. 1–48). doi:doi: 10.1163/27723194-bja10019.

Takoutsing, B., & Heuvelink, G. B. (2022). Comparing the prediction performance, uncertainty quantification and extrapolation potential of regression kriging and random forest while accounting for soil measurement errors. *Geoderma*, . doi:doi: 10.1016/J.GEODERMA.2022.116192.

Tello, D. et al. (2023). NGSEP 4: Efficient and accurate identification of orthogroups and whole-genome alignment. *Molecular Ecology Resources*, *23*, 712–724. doi:doi: 10.1111/1755-0998.13737.

Thompson, S. T., & Magrath, W. B. (2021). Preventing illegal logging. doi:doi:

10.1016/j.forpol.2021.102479.

Tnah, L. H., Lee, S. L., Ng, K. K., Tani, N., Bhassu, S., & Othman, R. Y. (2009). Geographical traceability of an important tropical timber (Neobalanocarpus heimii) inferred from chloroplast DNA. *Forest Ecology and Management*, *258*, 1918–1923. doi:doi: 10.1016/j.foreco.2009.07.029.

Truszkowski, J. et al. (2023). A probabilistic approach to estimating timber harvest location., .

Turubanova, S., Potapov, P. V., Tyukavina, A., & Hansen, M. C. (2018). Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. *Environmental Research Letters*, *13*, 074028. doi:doi: 10.1088/1748-9326/aacd1c.

Tyler, G., & Olsson, T. (2001). Plant uptake of major and minor mineral elements as influenced by soil acidity and liming. *Plant and Soil*, *230*, 307–321. doi:doi: 10.1023/A:1010314400976.

Uehara, T. K. (2023). Trade system change: An interdisciplinary proposal for true sustainability and planetary health. *SocArXiv*, .

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. (4th ed.). New York: Springer.

Veronesi, F., & Schillaci, C. (2019). Comparison between geostatistical and machine learning models as predictors of topsoil organic carbon with a focus on local uncertainty estimation. *Ecological Indicators*, *101*, 1032–1044. doi:doi: 10.1016/j.ecolind.2019.02.026.

Vitousek, P. M. (1984). Litterfall, Nutrient Cycling, and Nutrient Limitation in Tropical Forests. *Ecology*, *65*, 285–298. doi:doi: 10.2307/1939481.

Vlam, M., Boom, A., de Groot, G. A., & Zuidema, P. A. (2018a). Tropical timber tracing and stable isotopes: A response to Horacek et al. *Biological Conservation*, *226*, 335–336. doi:doi: 10.1016/j.biocon.2018.06.036.

Vlam, M., de Groot, G. A., Boom, A., Copini, P., Laros, I., Veldhuijzen, K., Zakamdi, D., & Zuidema, P. A. (2018b). Developing forensic tools for an African timber: Regional origin is revealed by genetic characteristics, but not by isotopic signature. *Biological Conservation*, *220*, 262–271. doi:doi: 10.1016/j.biocon.2018.01.031.

Wang, Z., Erasmus, S. W., Dekker, P., Guo, B., Stoorvogel, J. J., & van Ruth, S. M. (2020). Linking growing conditions to stable isotope ratios and elemental compositions of Costa Rican bananas (Musa spp.). *Food Research International*, *129*. doi:doi: 10.1016/j.foodres.2019.108882.

Watanabe, T., Broadley, M. R., Jansen, S., White, P. J., Takada, J., Satake, K., Takamatsu, T., Tuah, S. J., & Osaki, M. (2007). Evolutionary control of leaf element composition in plants: Rapid report. *New Phytologist*, *174*, 516–523. doi:doi: 10.1111/j.1469-8137.2007.02078.x.

Watkinson, C. J., Gasson, P., Rees, G. O., & Boner, M. (2020). The Development and Use of Isoscapes to Determine the Geographical Origin of Quercus spp. in the United States. *Forests*, *11*, 862. doi:doi: 10.3390/f11080862.

Watkinson, C. J., Rees, G. O., Gwenael, M. C., Gasson, P., Hofem, S., Michely, L., & Boner, M. (2022a). Stable Isotope Ratio Analysis for the Comparison of Timber From Two Forest Concessions in Gabon. *Frontiers in Forests and Global Change*, *4*, 1–16. doi:doi: 10.3389/ffgc.2021.650257.

Watkinson, C. J., Rees, G. O., Hofem, S., Michely, L., Gasson, P., & Boner, M. (2022b). A Case Study to Establish a Basis for Evaluating Geographic Origin Claims of Timber From the Solomon Islands Using Stable Isotope Ratio Analysis. *Frontiers in Forests and Global Change*, *4*, 1–15. doi:doi: 10.3389/ffgc.2021.645222.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists.*. John Wiley & Sons.

West, J. B., Bowen, G. J., Dawson, T. E., & Tu, K. P. (2010). *Isoscapes*. Dordrecht: Springer Netherlands. doi:doi: 10.1007/978-90-481-3354-3.

West, J. B., Ehleringer, J. R., & Cerling, T. E. (2007). Geography and vintage predicted by a novel GIS model of wine δ18O. *Journal of Agricultural and Food Chemistry*, *55*, 7075–7083. doi:doi: 10.1021/jf071211r.

Wiedenhoeft, A. C., Simeone, J., Smith, A., Parker-Forney, M., Soares, R., & Fishman, A. (2019). Fraud and misrepresentation in retail forest products exceeds U.S. forensic wood science capacity. *PLOS ONE*, *14*, e0219917. doi:doi: 10.1371/journal.pone.0219917.

Wieloch, T., Helle, G., Heinrich, I., Voigt, M., & Schyma, P. (2011). A novel device for batch-wise isolation of α-cellulose from small-amount wholewood samples. *Dendrochronologia*, *29*, 115–117. doi:doi: 10.1016/j.dendro.2010.08.008.

Worbes, M. (1999). Annual growth rings, rainfall-dependent growth and long-term growth patterns of tropical trees from the Caparo Forest Reserve in Venezuela. *Journal of Ecology*, *87*, 391–403. doi:doi: 10.1046/j.1365-2745.1999.00361.x.

Wortel, M., & van Erp, H. (2020). Labonderzoek in de strijd tegen fout hout.

Wright, M. N., & Ziegler, A. (2017). ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, *77*, 1–17. doi:doi: 10.18637/jss.v077.i01.

Wynn, P. M., Loader, N. J., & Fairchild, I. J. (2014). Interrogating trees for isotopic archives of atmospheric sulphur deposition and comparison to speleothem records. *Environmental Pollution*, *187*, 98–105. doi:doi: 10.1016/j.envpol.2013.12.017.

Wyttenbach, A., Furrer, V., & Tobler, L. (1995). The concentration ratios plant to soil for the stable elements Cs, Rb and K. *Science of The Total Environment*, *173-174*, 361–367. doi:doi: 10.1016/0048-9697(95)04737-9.

Xing, W., Wu, H., Shi, Q., Hao, B., Liu, H., Wang, Z., & Liu, G. (2015). Multielement stoichiometry of submerged macrophytes across Yunnan plateau lakes (China). *Scientific Reports*, *5*, 1–9. doi:doi: 10.1038/srep10186.

Yavitt, J. B., Harms, K. E., Garcia, M. N., Wright, S. J., He, F., & Mirabello, M. J. (2009).

Spatial heterogeneity of soil chemical properties in a lowland tropical moist forest, Panama. *Australian Journal of Soil Research*, *47*, 674–687. doi:doi: 10.1071/SR08258.

Zeitlin, J., & Overdevest, C. (2021). Experimentalist interactions: Joining up the transnational timber legality regime. *Regulation and Governance*, *15*, 686–708. doi:doi: 10.1111/rego.12350.

Zhu, Y. G., & Smolders, E. (2000). Plant uptake of radiocaesium: A review of mechanisms, regulation and application. *Journal of Experimental Botany*, *51*, 1635–1645. doi:doi: 10.1093/jexbot/51.351.1635.

Ziegler, S., Merker, S., Streit, B., Boner, M., & Jacob, D. E. (2016). Towards understanding isotope variability in elephant ivory to establish isotopic profiling and source-area determination. *Biological Conservation*, *197*, 154–163. doi:doi: 10.1016/j.biocon.2016.03.008.

Zuidema, P. A. et al. (2022). Tropical tree growth driven by dry-season climate variability. *Nature Geoscience*, *15*, 269–276. doi:doi: 10.1038/s41561-022-00911-8.

Summary

Illegal logging undermines sustainable timber production, therewith threatening ecosystems, economies, and human well-being. Global estimations of the trade in illegally harvested timber range from 4% to 30%, making it the third largest transnational crime worldwide, after counterfeiting and drug trafficking. Especially in tropical countries, the share of illegally traded timber is high, threatening the preservation of tropical forests and coinciding with violations of human rights.

Combatting illegal timber trade through law enforcement strongly relies on the ability to independently verify trade claims. However, the verification of trade claims is hindered by the long and complex supply chain of wood: timber is traded through multiple countries by many different stakeholders, which makes it extremely hard to verify the documentation that accompanies timber after the wood has left the forest. To support law enforcement, scientific forensic methods are being developed to independently verify timber origin, based on intrinsic wood properties that cannot be falsified. These methods include analyzing the genetic, structural and chemical variation of wood. The accuracy of these forensic methods relies on the natural variation in these wood properties. If this variation is larger between sites than within sites of potential origin, there is potential for origin verification.

The development of tracing techniques begins with a reference database, which is based on samples from a known origin from which the intrinsic properties are measured. To determine the origin of a new sample, it is compared to the reference database, and an origin claim can then be accepted or rejected based on its similarity to the reference samples. However, important barriers in the implementation of these scientific tools remain. First, the lack of georeferenced databases hampers direct implementation of the methods, even if the methods are shown to perform well (barrier 1). For many traded species, no reference databases are available to support origin verification methods. Second, especially small-scale tracing at distances of 100 km remains challenging (barrier 2). While forensic methods based on genetics and stable isotope ratios have demonstrated promising results for specific locations, they have not been as successful in other areas or species. These challenges highlight the need for further research and development to improve the accuracy and applicability of origin verification methods in timber trade.

In this thesis I aimed to address these barriers in scientific timber tracing and contribute to its operationalization with a specific focus on two chemical methods: multi-element analysis and stable isotope ratios. I tested the potential of these methods for tropical timber tracing and investigated the mechanisms driving geographical variation within tree species. I focused on three economically important tropical timber species: Azobé (*Lophira alata*) and Tali (*Erythrophleum ivorense* and *E. suaveolens*) from Central Africa and Red Meranti (*Shorea* spp.) from Borneo. From these species, wood samples were collected at natural forest concessions in four countries, representing major timber exporting countries in Central Africa (Cameroon, Republic of the Congo, Gabon) and South East Asia (Indonesia). All data was collected within the framework of the Timtrace project.

The first chemical method applied in this thesis was the multi-element analysis of wood. For multi-elemental analysis, about 60 elements can be measured all at once through mass spectrometry (such as Mg, Ca, K, etc.) and these make up the elemental composition of the wood. While the elemental composition of various commodities has been successfully used for their origin verification, the potential of multi-elemental analysis for timber tracing had not been explored.

We started with a small-scale study to investigate the drivers of variation in the elemental composition of Azobé at two locations in Cameroon (*Chapter 2*). It was clear that wood elemental composition was associated with that in the topsoil and that variation in the elemental composition was related to soil clay and organic matter content. In spite of these general patterns, no direct associations were found for many individual elements between the wood and the soil reactive element pool. Nevertheless, the differences between trees suggested that the multi-elemental composition of wood could provide a chemical fingerprint that can distinguish locations with distinct soil properties.

We continued with a large-scale study on tracing with multi-element analysis, including all three timbers across two continents (*Chapter 3*). We tested tracing potential at 22 sites in Central Africa and nine sites on Borneo. We found distinct spatial differences in chemical composition in all three timbers. For Azobé and Tali in Central Africa, tracing accuracy was 86-98% for regional clusters of chemically similar sites. These clusters were 50-800 km apart. Tracing accuracy of Red Meranti on Borneo was 88% at the site level, of which some were as close as 40km. This high accuracy at a small scale may be related to the short distances at which differences in soil type occur on Borneo. Furthermore, we put multi-element analysis to the test in this study by simulating a real-life tracing case. We measured 46 African 'blind' timber samples and assigned them to their most likely origin. Only then, their real origin was revealed and it turned out we had correctly identified the origin of 70-72% of the samples. However, we had failed to exclude 70% of the samples that were obtained from different species or from outside the study area. Overall, these results illustrate a high potential for multi-element analysis to be developed into a timber tracing tool and it can do so at a within-country scale. To reach this potential, reference databases will need to cover wider geographic areas and represent more timbers.

The second chemical method that was applied for timber tracing in this thesis was based on stable isotope ratios in wood. We first tested the potential for large-scale timber tracing, in a pantropical study focusing on one specific stable isotope: δ^{18}O (*Chapter 4*). Our objective was to develop a spatial prediction model (also called isoscape) for δ^{18}O, covering the entire tropical region. Spatial interpolation can greatly reduce the burden of sample collection (barrier 1), so developing such isoscapes has great potential to accelerate the implementation of forensic timber tracing. To develop the isoscape, we made use of a database of δ^{18}O measurements across the tropics. We combined this database with relevant covariables, chosen based on physiological insights into the main determinants of δ^{18}O in wood of tropical trees. These covariables included the isotopic signature of rainfall (root uptake) as well as variables related to precipitation, temperature and potential evapotranspiration. This resulted in the first pantropical isoscape of δ^{18}O. As a second step, we used the δ^{18}O isoscape to evaluate the spatial scale at which δ^{18}O-isotopic tracing was expected to differentiate wood samples within species. This confirmed the potential of δ^{18}O for tracing at a coarse spatial scale (e.g. country level).

In addition to testing the large-scale tracing potential of stable isotopes, we aimed to test their potential for tracing within and across countries in Central Africa. Previous studies were not conclusive on the potential for small-scale tracing in the region. In a study on Tali in Cameroon, no tracing potential was found at 15-200 km, whereas in another study in Gabon, differences were found in the isotopic composition of Okoume (*Aucoumea klaineana*) at a scale of 240 km. However, large-scale studies (more than 2-3 sites) that tested local and regional variation were lacking. Therefore, we tested the potential for tracing within and across countries in Central Africa, based on four stable isotopes that are known to vary at a landscape scale: δ^{18}O, δ^{2}H, δ^{13}C and δ^{34}S (*Chapter 5*). We measured these isotopes in Azobé and Tali at 17 locations in Cameroon, Republic of the Congo and Gabon. Based on this database, we found limited potential for origin assignment with stable isotope ratios in Central Africa: regional variation was not large enough to identify origin, as local variability was pronounced in all isotopes. This led to low site classification success (Azobé: 32.1%, Tali: 20.5%) and high uncertainty in the spatial prediction models. Furthermore, we found a significant species and site effect, stressing species-specific reference databases might be needed in future isotope tracing studies. The isotope ratios did differ between the three countries and country level assignment was considerably higher (Azobé: 79.4%, Tali: 61.7%).

As an exciting last step, we combined the two chemical tracing methods with a third tracing method, based on genetic analysis of the same trees (*Chapter 6*). The main aim was to enhance small-scale tracing accuracy (below c. 100 km), as no single method had achieved this level of accuracy for large geographic areas. By combining tracing methods, we hypothesized that tracing at small scales would improve due to the different mechanisms driving geographic variation within species. Comparing and combining the three tracing methods for Azobé at 13 sites in Central Africa, we found that the methods indeed complemented each other. While the accuracy of individual methods ranged from 43% to 73%, the combined methods achieved a correct site assignment of 88%. Moreover, the trees that were assigned to the wrong origin were not assigned as far away when the methods were combined: the model based on the combination of forensic methods assigned 94% of all trees within 100 km of their origin. Overall, these findings demonstrated a clear added value of combining multiple methods for tracing at small spatial scales, especially in areas where the individual methods are not specific enough by themselves.

Taken together, the studies described in this thesis reveal great scientific potential for forensic timber tracing. Multi-element analysis as well as stable isotope ratios can be applied at the scales at which timber fraud occurs (*Chapter 7*) and can be complemented by other methods to improve accuracy at small spatial scales, such as genetic analysis (barrier 2). However, the dependency on reference databases persists (barrier 1) and will continue to hamper the transition from the current case-by-case application of forensic tracing to a more widespread implementation. As the development of reference databases for all traded species will still take years, we need to look further and evaluate tracing methods for their potential to be upscaled while minimizing the number of trees that need to be measured. Future research should also focus on linking existing databases as well as on improving spatial interpolation methods, so that the use of data that is already collected can be optimized.

In conclusion, timber tracing methods are ready to be used by competent authorities, customs, certification bodies, non-governmental organisations and timber traders. However, there is still room for improvement in the success and widespread application of timber tracing. This can be accomplished by further developing the individual methods and expanding reference databases. When doing so, it will be crucial to focus on the most suitable methods for each region and timber species to maximize their effectiveness and realize the full potential of timber tracing.

Samenvatting

Illegale houtkap staat duurzame houtproductie in de weg en heeft negatieve effecten voor ecosystemen, economieën en het welzijn van mensen. Schattingen van de wereldwijde handel in illegaal gekapt hout variëren tussen de 4% en 30%, waardoor het de op twee na grootste vorm van georganiseerde misdaad wereldwijd is. Alleen in de handel in vervalsde producten en drugssmokkel gaat naar schatting meer geld om. Vooral in tropische landen is het aandeel van illegaal verhandeld hout hoog, waardoor tropische bossen worden bedreigt en mensenrechten worden geschonden.

Voor het bestrijden van de illegale houthandel zijn verschillende internationale wetten in werking gesteld. De handhaving van die wetten is afhankelijk van de mogelijkheid om de gerapporteerde herkomst van hout onafhankelijk te controleren. Het controleren van deze claims wordt echter belemmerd door de lange en complexe handelsketen van hout. Een groot aandeel van het totale verhandelde houtvolume passeert meerdere bedrijven en landen. Hierdoor kan de documentatie moeilijk gecontroleerd worden nadat het hout het bos verlaten heeft. Om handhaving te ondersteunen zijn wetenschappelijke forensische methodes ontwikkeld, die de herkomst van hout onafhankelijk kunnen controleren. Deze wetenschappelijke methodes zijn gebaseerd op intrinsieke kenmerken van het hout die niet kunnen worden vervalst. Ze maken gebruik van genetische, structurele en chemische variatie in het hout. De betrouwbaarheid van deze methodes hangt af van de natuurlijke variatie tussen locaties in die houteigenschappen. Als er meer variatie is tussen locaties dan binnen die locaties, dan biedt dat mogelijkheden om de herkomst te traceren.

De ontwikkeling van traceringsmethodes begint met een referentiedataset. Die dataset is gebaseerd op houtmonsters van bekende herkomst, waarvan de intrinsieke eigenschappen gemeten zijn. Om de herkomst van een nieuw houtmonster te bepalen wordt die vergeleken met de referentiedataset, waardoor een gerapporteerde herkomstclaim kan worden geverifieerd of gefalsificeerd. Er bestaan echter nog twee belangrijke barrières die de grootschalige implementatie van deze forensische technieken momenteel in de weg staan. De eerste barrière is de beperkte beschikbaarheid van referentiedatasets waarin gegevens van hout van bekende herkomst zijn opgeslagen (barrière 1). Voor veel verhandelde houtsoorten zijn er op dit moment geen referentiedatasets beschikbaar. De tweede barrière is de nauwkeurigheid van tracering op korte afstanden, van ˜100 km, die nog beperkt is gebleven (barrière 2). Genetische methodes en metingen van stabiele isotopen ratio's hebben veelbelovende resultaten opgeleverd, soms ook op kleine schaal. Maar die nauwkeurigheid was beperkt tot specifieke locaties en ze zijn minder succesvol gebleken in andere gebieden of voor andere houtsoorten. Om deze barrières te doorbreken is verder onderzoek noodzakelijk, met name om de nauwkeurigheid en precisie van traceringsmethodes te verbeteren.

In dit proefschrift heb ik bijgedragen aan het doorbreken van deze barrières en daarmee het operationaliseren van houttracering, met een specifieke focus op twee chemische methodes: multi-elementanalyse en stabiele isotopen ratio's. Ik heb het potentieel van deze methodes getest voor het traceren van tropisch hout en heb onderzocht welke mechanismen de geografische variatie binnen soorten veroorzaken. Ik heb me gericht op drie economisch belangrijke tropische houtsoorten: Azobé (*Lophira alata*) en Tali (*Erythrophleum ivorense* en *E. suaveolens*) uit

Centraal-Afrika en Rode Meranti (*Shorea* spp.) uit Borneo. We verzamelden houtmonsters in houtkapconcessies in vier landen: Kameroen, Republiek Congo, Gabon en Indonesië. Alle gegevens zijn verzameld in het kader van het Timtrace-project, dat zich richt op het ontwikkelen van methodes voor herkomsttracering van tropisch hout.

De eerste chemische methode die in dit proefschrift werd toegepast, was multi-elementanalyse van het hout. Hierbij worden zo'n 60 elementen gemeten (zoals Mg, Ca, K, etc.) in een massaspectrometer. Tezamen geven die metingen een indruk van de elementaire samenstelling van het hout. Hoewel de multi-elementenanalyse al succesvol was toegepast voor herkomstverificatie van verschillende andere producten, zoals kappertjes en cacao, was het potentieel voor houttracering nog niet onderzocht.

We hebben eerst de factoren onderzocht die variatie in elementaire samenstelling in het hout veroorzaken. Dit deden we op basis van Azobé houtmonsters afkomstig van twee locaties in Kameroen (*Hoofdstuk 2*). Uit onze analyses werd duidelijk dat de elementaire samenstelling van het hout gerelateerd was aan die in de grond. Daarnaast was de elementaire samenstelling in het hout ook gerelateerd aan het kleigehalte en het aandeel organische stof in de bodem. Ondanks deze algemene relaties was er voor veel individuele elementen geen direct verband te vinden tussen de concentratie in het hout en die in de reactieve elementenpool in de bodem. We vonden wel grote verschillen in elementaire samenstelling van hout tussen deze twee locaties. De methode levert dus inderdaad een specifieke chemische vingerafdruk op van locaties met verschillende bodemeigenschappen.

Als vervolgstudie hebben we de elementaire samenstelling gemeten van de drie de onderzochte houtsoorten uit dit proefschrift, op twee continenten (*Hoofdstuk 3*). We hebben de potentie voor tracering getest voor 22 locaties in Centraal Afrika en negen locaties op Borneo. Voor alle drie de houtsoorten vonden we duidelijke verschillen in chemische samenstelling tussen locaties. Voor Azobé en Tali in Centraal-Afrika werd 86-98% van de houtmonsters correct toegewezen aan het juiste regionale chemische cluster, bestaande uit locaties waar het hout een vergelijkbare chemische samenstelling had. Deze clusters lagen 50-800 km uit elkaar. In Borneo konden 88% van de Red Meranti bomen op locatieniveau correct worden toegewezen. Deze hoge nauwkeurigheid op kleinere schaal wordt mogelijk verklaard door de grote variatie van bodemtypes die op korte afstanden van elkaar voorkomen op Borneo. Daarnaast hebben we de multi-elementanalyse ook getest door een daadwerkelijke traceringssituatie te simuleren met zogenaamde 'blinde' houtmonsters, waarvan de herkomst ons onbekend was. We hebben de elementaire samenstelling van 46 nieuwe houtmonsters uit Afrika gemeten en ze toegewezen aan hun meest waarschijnlijke herkomst, op basis van de referentiedataset. Hiermee konden we de herkomst van 70-72% van de houtmonsters correct identificeren, maar we slaagden er niet in om 70% van de houtmonsters uit te sluiten die afkomstig waren van andere soorten of die van buiten het studiegebied kwamen. Deze resultaten lieten zien dat multi-elementanalyse zeker potentie heeft voor de herkomsttracering van hout en dat tracering succesvol kan zijn op kleine schaal (binnen een land). Om dit potentieel te benutten en de methode verder te testen, moeten referentiedatasets een breder geografisch gebied bestrijken en moeten er meer houtsoorten aan worden toegevoegd.

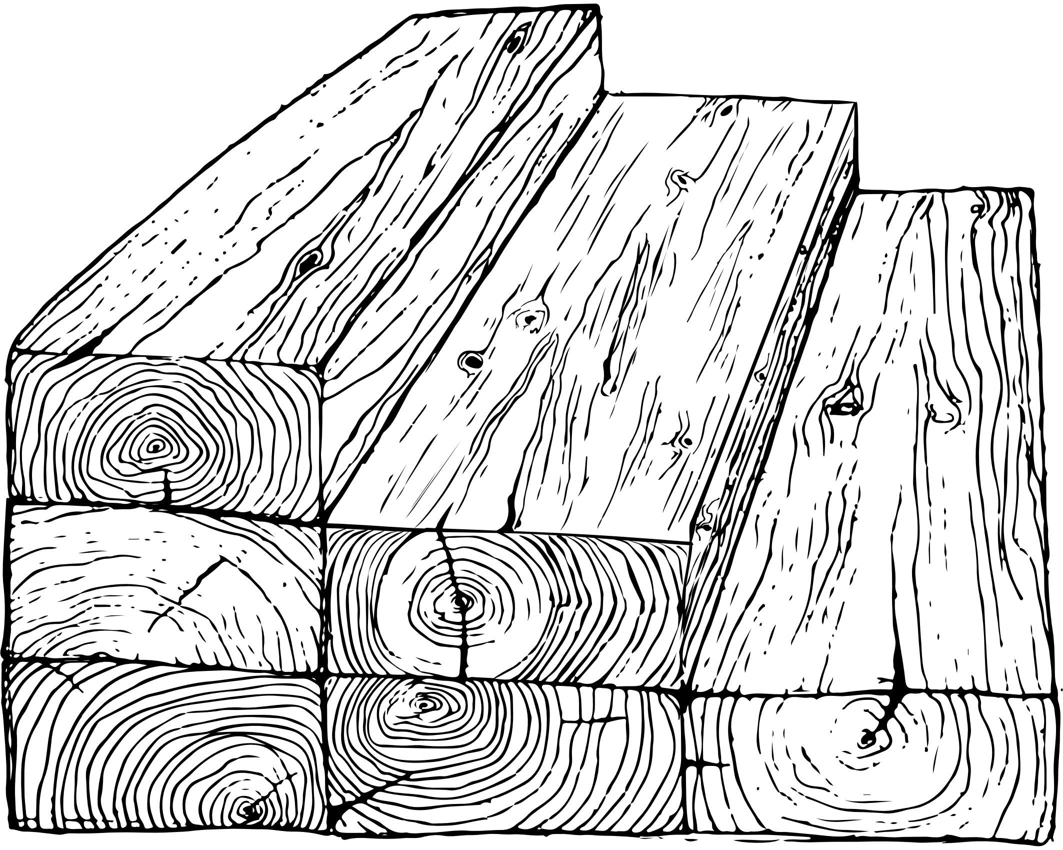
De tweede chemische methode die in dit proefschrift werd toegepast voor houttracering is gebaseerd op stabiele isotopen ratio's in hout. We hebben deze methode eerst toegepast voor houttracering op grote schaal, in een pantropische studie waarbij we ons richtten op één specifieke stabiele isotoop: δ^{18}O (*Hoofdstuk 4*). Ons doel was om een ruimtelijk voorspellingsmodel ('isoscape') voor δ^{18}O te ontwikkelen, dat de gehele tropische klimaatzone bestrijkt. Statistische ruimtelijke interpolatietechnieken kunnen de benodigde hoeveelheid referentiemonsters aanzienlijk verminderen (barrière 1). Het ontwikkelen van dergelijke isoscapes heeft dus het potentieel om de implementatie van forensische houttracering te versnellen. Om de isoscape te ontwikkelen maakten we gebruik van een dataset van δ^{18}O-metingen in de tropen. We combineerden deze dataset met relevante voorspellende variabelen voor δ^{18}O, gekozen op basis van fysiologisch inzichten in de belangrijkste bepalende factoren van δ^{18}O in hout van tropische bomen. We gebruikten de volgende variabelen voor het ontwikkelen van de isoscape: de isotopen samenstelling van neerslag (opname door de wortels), de hoeveelheid neerslag, temperatuur en potentiële verdamping en transpiratie. Als tweede stap hebben we de δ^{18}O-isoscape gebruikt om de ruimtelijke schaal te beoordelen waarop δ^{18}O-isotopische tracering naar verwachting houtmonsters binnen soorten zou kunnen differentiëren. Dit bevestigde het potentieel van δ^{18}O voor tracering op een grove ruimtelijke schaal, bijv. op landniveau.

In aanvulling op het testen van de potentie voor grootschalige tracering, wilden we ook nagaan in hoeverre stabiele isotopen ratio's een rol kunnen spelen voor tracering binnen en tussen landen in Centraal-Afrika. Eerdere studies gaven geen eenduidig beeld van de potentie voor kleinschalige tracering. In een studie naar Tali in Kameroen werd geen potentie voor herkomst tracering gevonden op 15-200 km, terwijl in een studie naar Okoumé (*Aucoumea klaineana*) in Gabon wel verschillen gevonden werden in de isotopenratios op 240 km afstand. Echter ontbrak tot nu toe een grootschalige studie met meer dan 3 locaties. Daarom hebben we de potentie voor traceren met behulp van stabiele isotopen ratio's binnen en tussen landen in Centraal-Afrika getest, op basis van vier stabiele isotopen waarvan bekend is dat ze variëren op landschapsschaal: δ^{18}O, δ^{2}H, δ^{13}C en δ^{34}S (*Hoofdsuk 5*). We hebben deze isotopen gemeten in Azobé en Tali op 17 locaties in Kameroen, de Republiek Congo en Gabon. Op basis van deze dataset vonden we geringe nauwkeurigheid in het toewijzen van herkomst: de regionale variatie in isotopenwaardes was niet groot genoeg om de oorsprong van het hout te identificeren, omdat de lokale variatie binnen een locatie te groot was in alle isotopen. Dit leidde tot een lage nauwkeurigheid in het toewijzen van herkomst (Azobé: 32,1%, Tali: 20,5%) en een hoge onzekerheid in de ruimtelijke voorspellingsmodellen (isoscapes). Bovendien vonden we significante verschillen tussen soorten, wat aangeeft dat soort-specifieke referentiedatasets nodig zullen zijn in toekomstige traceringsstudies met isotopen. De isotopen ratio's verschilden wel tussen de drie landen en de toewijzing van herkomst op landelijk niveau was aanzienlijk hoger (Azobé: 79,4%, Tali: 61,7%).

Als laatste en zeer vernieuwende stap hebben we de twee chemische methodes gecombineerd met een derde methode, gebaseerd op de genetische analyse van dezelfde bomen (*Hoofdstuk 6*). Het belangrijkste doel hiervan was om de nauwkeurigheid te verbeteren voor kleinschalige tracering (op afstanden van ˜100 km). Op die schaal had nog geen enkele individuele methode consequent een hoge nauwkeurigheid laten zien over grote geographische gebieden en in meerdere houtsoorten. Daarom hebben we de hypothese getest dat tracering op kleine schaal verbetert wanneer methodes gecombineerd worden, vanwege de verschillende mechanismen die geografische variatie in houteigenschappen binnen soorten veroorzaken. Door de drie methodes te vergelijken en te combineren voor Azobé op 13 locaties in Centraal-Afrika bleek dat de methodes elkaar inderdaad aanvulden en versterkten. Waar de nauwkeurigheid van de afzonderlijke methodes vrij laag bleef, tussen 43% en 73%, behaalde de combinatie van methodes een nauwkeurigheid van 88% voor toewijzing aan de locatie van herkomst. Bovendien: van de bomen die niet aan hun oorspronkelijke locatie waren toegewezen, werden de meeste (94%) toegewezen aan locaties op minder dan 100 km afstand. Deze resultaten laten duidelijk zien dat het combineren van meerdere methodes een grote toegevoegde waarde heeft voor houttracering op kleine ruimtelijke schaal. Dit is met name waardevol in gebieden waar de afzonderlijke methodes minder nauwkeurig zijn.

De studies in dit proefschrift tonen aan dat forensische houttracering duidelijk potentie heeft. Zowel de multi-elementanalyse als de stabiele isotopen ratio's kunnen worden toegepast op de schalen waarop houtfraude plaatsvindt (*Hoofdstuk 7*) en kunnen worden aangevuld met andere methodes om de nauwkeurigheid te verbeteren, zoals genetische analyses (barrière 2). De afhankelijkheid van referentiedatasets blijft echter nog steeds een belemmering voor de grootschalige implementatie van forensische tracering (barrière 1). Aangezien het ontwikkelen van referentiedatasets voor alle verhandelde houtsoorten nog vele jaren zal duren, zal het belangrijk zijn om traceringmethodes zodanig te ontwikkelen en op te schalen dat hun potentieel wordt geoptimaliseerd met een minimaal aantal houtmonsters. Toekomstig onderzoek zal dus gericht moeten zijn op het koppelen van reeds bestaande datasets en het verbeteren van de statistische methodes voor ruimtelijke interpolatie.

Houttraceringmethodes zijn dus geschikt voor gebruik door autoriteiten, douanes, certificeringsinstanties, niet-gouvernementele organisaties en houthandelaren. Echter, de nauwkeurigheid van deze methodes kan nog worden verbeterd en de toepassing zou algemener moeten worden. Dit kan gerealiseerd worden door de methodes verder te ontwikkelen en door de referentiedatasets uit te breiden. Om het potentieel van houttraceringsmethodes optimaal te benutten, is het cruciaal dat verdere ontwikkeling en verbetering gericht wordt op de meest geschikte forensische methodes per regio en per houtsoort.



Acknowledgements

One of the reasons I became interested in forest ecology was because I realized early on that it takes collaboration to answer the truly interesting questions. A dataset from one forest can be interesting, but by combining it with data from other people and other forests we might be able to see the underlying processes through the trees. This then attracts a dynamic group of researchers with a collaborative mindset, which I am truly grateful to be a part of. Even though a PhD title is awarded to a single person, my PhD journey has been especially inspiring because of the people I shared it with.

First of all I want to thank my truly great supervisors: Pieter, Ute and Mart. I felt supported both personally and professionally, which is the best one can get from their supervisors. Your trust made me feel right for the job while your inquisitive questions always made me look just that little bit further. It feels like you showed me how to read and write. We shared inspiring and creative retreats as well as pragmatic fieldwork discussions, both valued highly. I cherish our collaboration and hope to continue this in the future!

Thanks also to the extended Timtrace Wageningen team: Rene, Arjen and Peter, for all the fun and inspiring meetings we've had. Putting so much brainpower in a room always ensured great new ideas! I am also extremely thankful for working with all the international Timtrace collaborators in Cameroon, Republic of the Congo, Gabon and Indonesia. Thank you for sharing your work and life with us. This was truly indispensable, especially during the covid pandemic. And thanks to our colleagues from World Forest ID: by teaming up for the fieldwork we could maximize the sampling effort.

I want to thank the people at the labs I've worked in/with: the heavy metal lab at Wageningen Food Safety Research where I could always count on help with the ICP-MS. The Dendrolab for wood preparations and cellulose extractions. The soil chemistry lab, where I learned about all things soils. The lab of Plant Ecology and Nature Conservation, for letting me use the microscales. And Arnoud Boom at Leicester University, the isotope lab at the WSL and Agroisolab for stable isotope measurements.

My PhD would have not been the same without the FEMily with its coffee breaks (with and without cake), lively lunches and international dinners. Thanks to the staff of the Forest Ecology and Forest Management group for providing a supportive but inquisitive environment and for showing that collaboration is key: Frans, Frank, Paul, Madelon, Frits, Gert-Jan, Jan, Lourens, Masha, Douglas, Ellen, Leo, Mattieu, Joke, Louise, Liesje, Marina, Peter, Susan. A

special thanks to Marielos, for giving me a desk in the FEM corridor as an MSc student.

And then the hardest part comes. Barbara, I don't even know how to write down in words how to describe my gratitude for having shared these years with you. This journey together has made us grow in ways we probably never thought of. I love you and I'm incredibly proud of you, of us. Thank you. Thank you also to Sophie, we've shared so many coffees, teas, sleepovers and late night talks. Thank you for being your fun, impulsive and recalcitrant self. Don't ever think you cannot be the great researcher you are. Louis, thank you for talking nonsense when no one is listening, for always letting me stay at the farm and for saving me from chopping off my own fingers. Jose, homie, thank you for welcoming me as your office roomie and letting me host parties in your house. Thanks to all FEMmers that made me feel like I fitted in from the start: Fede, Juan Ignacio, Daisy, Heitor, Ursula, Linar, Shanshan, Yanjun, Etienne, Marleen, Ambra, Carlinhos, Meike, Lan, Rodrigo, Maaike. Thanks for the unofficial coffee breaks, smiles bigger than your faces, hotpots and other cooking sessions. And thanks to all FEMmers that I met along the way: Qi, Rens, Jazz, Tomo, Alan, Alejandra, Nicola, Xiaohan, Bo, Johan, Carolina, Marco, Nic, Bianca, Arjen, Catarina, Sylvana, Eva, Sietse, Jente, Aldicir, Danju, Amanda, Lennart, Franciska, Lisa, Arief, Iris, Jorad, Laura. It has been a pleasure to share the hallway with you! A special thanks to my writing group members, Louis, Ambra, Marleen and Nicola, for basically proof-reading most of my thesis. And of course a big thanks to my amazing paranymphs Barbara and Sophie, I'm proud to have you by my side.

Next to this great working environment, I am also very thankful to all the friends that made me forget trees for a while: Lieke, for sharing gardening, triathlons and multivariate analyses, and for showing me what a non-stressful PhD and postdoc can look like. The Vreetfestijn group (including all plus ones!), for all the jokes, for cooking, eating, games and dancing. De Schijf van Vijf, for warm and curing sleep-overs. Judy and Sanne, for hours of reflections on life, ambition and plants. Iris and Eeke, Biology friends of the first hour. Thanks to my Wageningen friends of the Farm and the WURst (the double WK Knotsbal champions!). Sarah for coffees and taking great care of Tijger.

A special thanks also to my family: my mum Everdien, who sets the best example that it is never too late to start new things in life, and my dad Ricolt, who is probably the most content person I know (and who gave me the best Dutch thesis subtitle: "Hout de Dief!"). You are the most supportive parents I could have wished for. Thank you both, for always stimulating me to go further and at the same time providing a safe home to come back to. Thanks to my brother Erik and my sister Emma for being proud and making fun of me at the same time. You are great siblings to hang out with!

Thanks also to Petra, for all the amazing dinners. Food definitely brings people together. Thanks to Bram for sharing your passion and perseverance to promote a more sustainable world. And thanks to Annelot, Shyro and Vesper, again for countless dinners and laughter.

And then my most favourite person, Aram. Thank you. Thank you for believing in me at all times. For being patient with me. Thank you for being open to all of my adventures while still pursuing your own goals. Thank you for how you can truly enjoy things in life.

About the author

Laura Emily Boeschoten was born on May 7, 1994, in Zaltbommel, The Netherlands. She obtained her bachelor Biology at Utrecht University and graduated *cum laude* in 2015. During her BSc she also completed a minor in Ecology at the University of Hull, Scarborough, UK.

She continued with a master Environmental Biology at Utrecht University. She followed the track 'Ecology and Natural Resource Management', and graduated *cum laude* in 2018. It was during her MSc that she carried out tropical fieldwork for the first time, when she did an internship with Dr Marijke van Kuijk and Rens Vaessen MSc, to understand fruit-frugivore relationships in the tropical forest of French Guiana. Excited to learn more about tropical forest ecology, she went to the Smithsonian Tropical Research Institute in Panama for her second MSc internship. Under the supervision of Dr Jeff Hall she studied successful tree planting strategies in an area affected by a highly competitive invasive grass and recurrent fires.

She continued her academic career with a PhD at Wageningen University in 2019. She joined the Forest Ecology and Management group and became a part of the Timtrace project, led by Prof. Pieter Zuidema and also supervised by Dr Ute Sass Klaassen and Dr Mart Vlam. During her PhD, she investigated chemical methods to trace the origin of tropical timber, with the aim of improving the independent verification of timber origin, which resulted in this thesis.

Sustainability is something she cares about a lot. She aims to work on the transition towards a more sustainable world, both in her professional and private life. In her time off she enjoys being outdoors and doing sports, watering the plants in her veggy garden or cooking for and with friends.

Publications

Laura E. Boeschoten, Mart Vlam, Ute Sass-Klaassen, Barbara Rocha Venâncio Meyer-Sand, Arnoud Boom, Gaël U.D. Bouka, Jannici C.U. Ciliane-Madikou, Nestor L. Engone Obiang, Mesly Guieshon-Engongoro, Joël J. Loumeto, Dieu-merci M.F. Mbika, Cynel G. Moundounga, Rita M.D. Ndangani, Dyana Ndiade Bourobou, Peter van der Sleen, Steve N. Tassiamba, Martin T. Tchamba, Bijoux B.L. Toumba-Paka, Herman T. Zanguim, Pascaline T. Zemtsa, Pieter A. Zuidema (2023). Stable isotope ratios in wood show little potential for sub-country origin verification in Central Africa. *Forest Ecology and Management*, 544, 121231. https://doi.org/10.1016/j.foreco.2023.121231

Laura E. Boeschoten, Mart Vlam, Ute Sass-Klaassen, Barbara Rocha Venâncio Meyer-Sand, Ulfa Adzkia, Gaël U.D. Bouka, Jannici C.U. Ciliane-Madikou, Nestor L. Engone Obiang, Mesly Guieshon-Engongoro, Joël J. Loumeto, Dieu-merci M.F. Mbika, Cynel G. Moundounga, Rita M.D. Ndangani, Dyana Ndiade Bourobou, Mohamad M. Rahman, Iskandar Z. Siregar, Steve N. Tassiamba, Martin T. Tchamba, Bijoux B.L. Toumba-Paka, Herman T. Zanguim, Pascaline T. Zemtsa, Pieter A. Zuidema (2023). A new method for the timber tracing toolbox: applying multi-element analysis to determine wood origin. *Environmental Research Letters*, 18, 054001. https://doi.org/10.1088/1748-9326/acc81b

Laura E. Boeschoten, Ute Sass-Klaassen, Mart Vlam, Rob N.J. Comans, Gerwin F. Koopmans, Barbara Rocha Venâncio Meyer-Sand, Steve N. Tassiamba, Martin T. Tchamba, Herman T. Zanguim, Pascaline T. Zemtsa, Pieter A. Zuidema (2022). Clay and soil organic matter drive wood multi-elemental composition of a tropical tree species: Implications for timber tracing. *Science of the Total Environment*, 849, 157877. https://doi.org/10.1016/j.scitotenv.2022.157877

Melita C. Low, Nele Schmitz, **Laura E. Boeschoten**, José A. Cabezas, Mathias Cramm, Volker Haag, Gerald Koch, Barbara R.V. Meyer-Sand, Kathelyn Paredes-Villanueva, Erin Price, Andrew H. Thornhill, Jo Van Brusselen, Pieter A. Zuidema, Victor Deklerck, Eleanor E. Dormontt, Alison Shapcott, Andrew J. Lowe (2022). Tracing the world's timber: the status of scientific verification technologies for species and origin identification. *IAWA Journal*, 44(1), 63-84. http://dx.doi.org/10.1163/22941932-bja10097

Nele Schmitz, Hans Beeckman, Celine Blanc-Jolivet, **Laura E. Boeschoten**, Jez W.B. Braga, José Antonio Cabezas, Gilles Chaix, Simon Crameri, Victor Deklerck, Bernd Degen, Eleanor Dormontt, Edgard Espinoza, Peter Gasson, Volker Haag, Stephanie Helmling, Micha Horacek, Gerald Koch, Cady Lancaster, Frederic Lens, Andrew Lowe, Sandra Martínez-Jarquín, Justyna A. Nowakowska, Andrea Olbrich, Kathelyn Paredes-Villanueva, Tereza C.M. Pastore, Tahiana Ramananantoandro, Andriambelo Radonirina Razafimahatratra, Prabu Ravindran, Gareth Rees, Liz F. Soares, Niklas Tysklind, Mart Vlam, Charlie Watkinson, Elisabeth A. Wheeler, Robert Winkler, Alex C. Wiedenhoeft, Valentina T. Zemke, Pieter A. Zuidema (2020). Overview of current practices in data analysis for wood identification. A guide for the different timber tracking methods. Global Timber Tracking Network, GTTN secretariat, European

Forest Institute and Thünen Institute. https://doi.org/10.13140/RG.2.2.21518.79689

Laura E. Boeschoten, Michiel van Breugel, Mario Bailon, Johana Balbuena, Miguel Nuñez, Arturo Cerezo, Jefferson S. Hall (2020). Framework species approach proves robust in restoring forest on fire prone invasive grass: A case study from Panama. *Journal of Sustainable Forestry*, 40(2), 197-215. https://doi.org/10.1080/10549811.2020.1746915

Papers submitted

Rens Vaessen, Klaske van Wijngaarden, **Laura E. Boeschoten**, Ronja Knippers, Livia Durazzo, Loes Verkuil, Marijke van Kuijk. Fruit and seed traits and vertebrate-fruit interactions of tree species occurring in Guyana, Suriname and French Guiana.(*under review* at Ecology, preprint available at https://doi.org/10.34894/3X8JWB)

Awards

TRACE conference (2022), Erlangen, Germany: Best PhD oral presentation

TRACE conference (2023), Coimbra, Portugal: Best PhD poster

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review/project proposal (9 ECTS)

- Tracing the world's timber: the status of scientific verification technologies for species and origin identification
- Developing chemical tracing techniques for tropical timbers

Post-graduate courses (7.8 ECTS)

- Multivariate analysis; PE&RC (2019)
- Geostatistics; PE&RC (2019)
- Hands on digital soil mapping; ISRIC (2021)
- Forest management across Europe; PE&RC (2021)
- IR-MS for isotope analyses; University of Erlangen (2022)

Laboratory training and working visits (1.5 ECTS)

- Training in inductively coupled plasma mass spectrometry (ICP-MS) to measure trace and rare earth elements; RIKILT (2019)

Invited review of journal manuscript (3 ECTS)

- Dendrochronologia: advances in increment coring system for large tropical trees with high wood densities
- Molecules: isotope ratio mass spectrometry for provenancing

- Restoration Ecology: aboveground biomass accumulation and tree size distribution in seasonal Atlantic Forest restoration sites

Competence, skills and career-oriented activities (2.8 ECTS)

- An introduction to LaTeX; PE&RC (2020)
- Intercultural communication; PE&RC (2021)
- Scientific artwork, data visualisation and infographics with Adobe illustrator; PE&RC (2021)
- Scientific writing; WGS (2022)
- Design and deliver effective scientific presentations; Presenting Scientist (2022)

Scientific integrity/ethics in science activity (0.3 ECTS)

- Ethics in plant and environmental sciences; PE&RC (2020)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.5 ECTS)

- PE&RC First years weekend (2019)
- PE&RC Last years retreat (2022)

Discussion groups/local seminars or scientific meetings (5.4 ECTS)

- Defaunation in the anthropocene (2019)
- STTC Conference Berlin (2019)
- PE&RC/FEM-PEN-WEC R discussion group (2019, 2020)
- FEM journal club (2019-2023)
- Global timber tracing network meeting (2020)
- AVETH Diversity and inclusion seminar series (2020)
- Research seminars at IPB and CIFOR Indonesia (2020)
- 2Emes journees d'etudes en sciences et technologies du bois au Gabon (2021)
- Old forest, new forest (2023)

International symposia, workshops and conferences (9.4 ECTS)

- ATBC; oral presentation; Cartagena, Colombia (2022)
- TRACE; oral presentation; Erlangen, Germany (2022)
- ATBC; oral presentation; Coimbatore, India (2023)
- TRACE; poster presentation; Coimbra, Portugal (2023)

Societally relevant exposure (1.7 ECTS)

- Timtrace user committee meeting (2019-2022)
- Timtrace blog (2019-2023)
- Tracing seminar at WWF the Netherlands (2020)
- NWO Bessensap presentation (2020)
- Politie nieuwsbrief, NWO photo story, chemisch twee-weekblad, houtwereld articles (2021)
- Radio1 interviews (2023)

Lecturing/supervision of practicals/tutorials (2.7 ECTS)

- Resource dynamics (2020, 2021, 2022)
- Restoration ecology (2021, 2022)
- Forest resources (2022)

BSc/MSc thesis supervision (7.5 ECTS)

- Tracing African timber: an evaluation of the potential of multi element analysis to verify the geographic origin of Lophira alata
- The first steps in a successful market entry: an advisory report about the practical implementation of Timtrace
- Combining tracing techniques for timber origin verification

This research was funded by the Dutch Research Council (NWO-TTW-OTP-16427). Additional support for fieldwork was received from the Alberta Mennega Foundation, FSC International and from World Forest ID. Additional support for isotope analysis was received from the Ecology Fund of the KNAW.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design and lay-out by Laura Emily Boeschoten.

Printed by ProefschriftMaken.nl on FSC-certified paper.