Chemotaxonomic Species and Origin Identification of Wood in Pulp and Paper

DISSERTATION

submitted in fulfilment of the requirements for the

degree of Doctor rerum naturalium (Dr. rer. nat.)

at the Faculty of Mathematics, Informatics and Natural Science

Department of Biology

Institute of Wood Science

University of Hamburg

by

Max Lauritz Flaig

Hamburg 2024

Dissertation examiners

Prof. Dr. Bodo Saake

University of Hamburg
Institute of Wood Science
Chemical Wood Technology

PD Dr. habil. Gerald Koch

Johann Heinrich von Thünen Institute
Institute of Wood Research
Wood Anatomy

Disputation examiners

Prof. Dr. Elisabeth Magel

University of Hamburg
Institute of Wood Science
Wood Biology

Head of examination commission

Prof. Dr. Linnea Hesse

University of Hamburg
Institute of Wood Science
Wood Physics

Date of oral defense

06/02/2025

"Humanity is waging a war on nature. This is suicidal.

Making peace with nature is the defining task of the 21st century.

It must be the top, top priority of everyone, everywhere."

Antonio Gutteres (2020)

Acknowledgements

I would like to thank my main doctoral supervisor Prof. Dr. Bodo Saake very much for his supervision of my entire thesis and for his support in every way, be it technical problems in the lab, writing, focusing on the essentials, scheduling and motivation. I would also like to thank PD Dr. habil. Jürgen Odermatt, who passed away far too early. Without him, I would never have started my PhD. Thank you for the project idea itself, for all the preparatory work that made my work easier, for the passion for chemotaxonomy and GC-MS analysis that you passed on to me and for your trust in me as a doctoral candidate – a great compliment.

I have always enjoyed the friendly cooperation and mutual support of my colleagues and wood science students at the University of Hamburg. In particular, I thank Birte Buske and Dr. Jens Berger for many professional discussions and their active support in the laboratory and with the analytical work. For their lab support I thank Christina Ganser, Frederik Schüler and Hannes Mader. I would also like to thank my colleagues at the Thünen Institute for their collegial environment, friendly working atmosphere and support. In particular, I thank PD Dr. habil. Gerald Koch for reviewing my dissertation and Dr. Andrea Olbrich and Dr. Stephanie Helmling for their excellent collaboration on the project and the joint publications.

Dr. Philip Wenig developed the analysis software and personally assisted with data processing. Thank you very much for all your help.

Dr. Othar Kordsachia helped me with the pulping and bleaching, Alina Wassink and Nils Grützmann did some preliminary lab work on my project. Thank you very much.

I thank my friends Flock, Marv, Höft, Lars, Marie, Jabba, Tim, Leo, Schmu and Laurin for their understanding and for some welcome distractions such as beautiful holidays together. I would like to thank my siblings Hennes, Judith and Philine and my parents Martin and Kirsten for their unconditional support, motivation and love. I would also like to express my deepest gratitude to my partner Heidrun for her love, understanding, patience and all the emotional and practical support she has given me – an extraordinarily great help!

Preface

The present work was carried out between February 2019 and November 2024 in the Department of Chemical Wood Technology, Institute of Wood Science at the University of Hamburg. The research results were collected as part of the project “Detection of Tropical Wood in Paper – Chemotaxonomy and Anatomy for the Identification of Mixed Tropical Hardwood”. The project was funded by the Deutsche Bundesstiftung Umwelt (DBU) under the grant number AZ 31 759-31 and AZ 34 295/01. Project and cooperation partners were the University of Hamburg (UHH), the Thünen-Institute for Wood Research (TI), the Technical University of Darmstadt (TUDa), the ISEGA Forschungs- und Untersuchungsgesellschaft mbH, the World Wide Fund for Nature (WWF) and DIE PAPIERINDUSTRIE (VDP).

The main part of this cumulative dissertation is based on the following three publications:

I. **Flaig, M.L.**, Berger, J., Wenig, P., Olbrich, A. and Saake, B. (2023). Identification of tropical wood species in paper: A new chemotaxonomic method based on extractives. *Holzforschung* 77: 860–878. https://doi.org/10.1515/hf-2023-0048 (Supplementary material under the same link)

II. **Flaig, M.L.**, Berger, J., Helmling, S., Olbrich, A., Schaffrath, H.J., Zahn, D. and Saake, B. (2024). Chemotaxonomic and anatomic wood species identification in bleached pulp: Blind test and method validation. *Holzforschung* 78: 487–502. https://doi.org/10.1515/hf-2024-0025 (Supplementary material under the same link)

III. **Flaig, M.L.**, Berger, J. and Saake, B. (2025). Tracing the geographic origin of wood in pulp and paper by GC-MS analysis of extractives. *Holzforschung* 79: 15–29. https://doi.org/10.1515/hf-2024-0080

The publications are embedded in a comprehensive framework of background knowledge, problem definition, state of the art, further unpublished research and results as well as a discussion of the results and outlook for the future.

Persistent identifier (PID): urn:nbn:de:gbv:18-ediss-127138

Declaration of authorship contributions to the publications I–III

I. **Flaig, M.L.**, Berger, J., Wenig, P., Olbrich, A. and Saake, B. (2023). Identification of tropical wood species in paper: A new chemotaxonomic method based on extractives. *Holzforschung* 77: 860–878. https://doi.org/10.1515/hf-2023-0048 (Supplementary material under the same link)

- Research concept: 70 %
- Execution of the experiments: 80 %
- Data evaluation: 100 %
- Writing of the manuscript: 90 %

Confirmation of the author's contributions to Pub I

Author	Supervisor & dissertation examiner
Max L. Flaig	Prof. Dr. Bodo Saake

II. **Flaig, M.L.**, Berger, J., Helmling, S., Olbrich, A., Schaffrath, H.J., Zahn, D. and Saake, B. (2024). Chemotaxonomic and anatomic wood species identification in bleached pulp: Blind test and method validation. *Holzforschung* 78: 487–502. https://doi.org/10.1515/hf-2024-0025 (Supplementary material under the same link)

- Research concept: 60 %
- Execution of the experiments: 70 %
- Data evaluation: 70 %
- Writing of the manuscript: 90 %

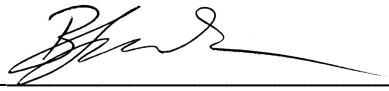
Confirmation of the author's contributions to Pub II

Author	Supervisor & dissertation examiner
Max L. Flaig	Prof. Dr. Bodo Saake

III. **Flaig, M.L.**, Berger, J. and Saake, B. (2025). Tracing the geographic origin of wood in pulp and paper by GC-MS analysis of extractives. *Holzforschung* 79: 15–29. https://doi.org/10.1515/hf-2024-0080

- Research concept: 80 %
- Execution of the experiments: 50 %
- Data evaluation: 100 %
- Writing of the manuscript: 100 %

Confirmation of the author's contributions to Pub III

Author	Supervisor & dissertation examiner
	
Max L. Flaig	Prof. Dr. Bodo Saake

Further publications

IV. Berger, J., **Flaig, M.L**., Olbrich, A., Helmling, S. and Saake, B. (2021). Nachweis von Tropenholz in Papier – Chemotaxonomie und Anatomie zur Identifizierung von Mixed Tropical Hardwood. Abschlussbericht zum DBU-Projekt AZ 34 295/01, 1–86. Universität Hamburg. (link)

V. **Flaig, M.L.**, Berger, J., Wenig, P., Olbrich, A. and Saake, B. (2023). A chemotaxonomic method for identifying mixed tropical hardwood (MTH) species in pulp and paper products. *ISWFPC23 Conference Proceedings Vol. I – Oral Presentations*, 21: 37–40. 4.–7. July, Venice, Italy. (link)

VI. **Flaig, M.L.**, Berger, J., Olbrich, A., Helmling, S. and Saake, B. (2024). Enhancing Wood Species Identification in Pulp and Paper: Synergies of Anatomy and Chemotaxonomy. *EWLP24: 17th European Workshop on Lignocellulosics and Pulp 26–30 August 2024, Laboratory of Natural Materials Technology, Åbo Akademi University, Finland: Book of abstracts*, 17: 100–100. 26.–30. August, Turku/Åbo, Finland. (link)

Affidavit

I hereby declare and affirm that this doctoral dissertation is my own work and that I have not used any aids and sources other than those indicated. If electronic resources based on generative artificial intelligence (gAI) were used in the course of writing this dissertation, I confirm that my own work was the main and value-adding contribution and that complete documentation of all resources used is available in accordance with good scientific practice. I am responsible for any erroneous or distorted content, incorrect references, violations of data protection and copyright law or plagiarism that may have been generated by the gAI.

Hamburg, 13.11.2024 

Place, Date Max L. Flaig

Since I am not a native English speaker, I used two gAI systems in the present work, as follows:

DeepL

- To translate German passages I had written myself into English
- To proofread and optimise language, grammar and spelling

ChatGPT (5 % of the total text)

- To summarise/shorten my own paragraphs, without changing the content
- To optimise/rephrase passages in scientific language, without changing the content

I confirm that I have stated all uses in full. I am aware that missing or incorrect information may be considered as an attempt to deceive.

Hamburg, 13.11.2024 

Place, Date Max L. Flaig

I, the undersigned, declare that this bound copy of the dissertation is identical with the electronic copy of the dissertation submitted (via the Docata upload) and the printed bound copy of the dissertation submitted to the Faculty of Mathematics, Informatics and Natural Sciences (doctoral office biology) for archiving.

Hamburg, 13.11.2024 

Place, Date Max L. Flaig

Abstract

This work presents the development and proof of concept of a novel chemical method for wood taxa and origin identification in pulp and paper. In order to conserve nature and halt the ongoing and severe crossings of planetary boundaries such as land-system change and biosphere integrity, it is essential to combat deforestation and the loss of genetic diversity within global forests. Therefore, a variety of national and international laws have been enacted with the objective of protecting the environment. However, there is a need for science-based monitoring tools to ensure that these laws are effectively enforced in practice.

The method is based on the differentiation of wood extract profiles that remain in pulp (chemotaxonomic fingerprinting) and involves cryo-ball mill grinding and Soxtherm extraction of bleached pulp or paper samples with n-hexane, followed by thermodesorption of the dried extracts and gas chromatography - mass spectrometry (TD-GC-MS) analysis. The generated chromatograms undergo a series of pre-processing steps, including ion filtering, smoothing and peak deconvolution and are then added to the reference database (DB). The investigation of optimal configuration settings for the DB was also part of this work. The subsequent identification of samples of unknown composition is conducted through querying against the DB on three comparison parameters: marker peak scores (MPS), reverse similarity index (RSI), and matched chromatogram area (MCA). The methodology has been validated on different measuring devices. However, there is currently a partial dependence of the database query results on the measuring system. Furthermore, the novel extractant-based approach was evaluated in an independent external blind test and compared to the established anatomical identification method. The blind test comprised 38 exemplary species of mainly tropical hardwood, which were processed in unknown mixtures into 15 test sheets and provided to four participating institutes for analysis. The results demonstrate that the new chemotaxonomic method yielded outcomes that were comparable to those obtained through the morphological-anatomical method, which is currently employed for all assessments. However, both approaches show distinct strengths and weaknesses in the recognition of individual wood taxa, suggesting that a combination of both methods could potentially enhance the overall hit rate.

Another study included in this work successfully differentiated the origins of the wood genera *Gonystylus* and *Rubroshorea* and the wood species *Tectona grandis* in fully bleached pulps. Certain compounds responsible for the differentiation of specific geographical origins were found for every taxon. The results of the origin differentiation could be particularly useful for the verification of harvest origin declarations at border crossings of pulp and paper products, as well as for the enforcement of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Currently, this is the only method available for this purpose. A variety of optimisation possibilities of the chemotaxonomic database approach are discussed, including sample preparation techniques such as advanced tuning of the GC-MS instruments used, data processing methods such as ion filter settings and database

configuration. The outlook section outlines future prospects for further method validation on industrial samples and the potential use of machine learning for automation.

Kurzfassung

Diese Arbeit zeigt die Entwicklung und den Proof of Concept einer neuen chemischen Methode zur Identifizierung von Holztaxa und deren Provenienz in Zellstoff- und Papierprodukten. Um die Natur zu erhalten und die gravierende Überschreitung der planetaren Grenzen, vor allem den Landsystemwandel und die Integrität der Biosphäre, aufzuhalten, ist es unerlässlich, die illegale Entwaldung und den Verlust der genetischen Vielfalt in Wäldern zu bekämpfen. Dazu wurden verschiedene nationale und internationale Gesetze geschaffen. Zu deren praktischer Durchsetzung werden wissenschaftlich fundierte Überwachungsinstrumente benötigt.

Die Identifizierungsmethode basiert auf der Unterscheidung von in gebleichten Zellstoffen verbleibenden Holzextraktstoffprofilen (chemischer Fingerabdruck) und umfasst das Mahlen mit einer Kryomühle, die Extraktion mit n-Hexan, die Thermodesorption der eingedampften Extrakte und die gaschromatografische und massenspektrometrische Analytik (TD-GC-MS).

Die erzeugten Daten durchlaufen eine Reihe von Verarbeitungsschritten, einschließlich Ionenfilterung, Glättung der Massenspuren und Peakerkennung, bevor sie in einer Referenzdatenbank (DB) gespeichert werden. Die optimale Konfiguration der DB wurde ebenfalls in dieser Arbeit untersucht. Zur Identifizierung von Proben unbekannter Zusammensetzung wird eine Abfrage der DB anhand von drei Vergleichsparametern vorgenommen: marker peak scores (MPS), reverse similarity index (RSI) und matched chromatogram area (MCA), durchgeführt. Die Methodik konnte auf unterschiedlichen Messgeräten validiert werden, wobei derzeit eine teilweise Abhängigkeit der Datenbank-Vergleichs-Ergebnisse vom verwendeten Gerät besteht. Darüber hinaus wurde die neu entwickelte Extraktstoff-basierte Methode in einem unabhängigen externen Blindtest mit der etablierten anatomischen Identifizierungsmethode verglichen. Der Blindtest umfasste 38 exemplarische, vorwiegend tropische Laubholzarten, die in unbekannten Mischungen zu 15 Testblättern verarbeitet und vier teilnehmenden Instituten zur Analyse zur Verfügung gestellt wurden. Die Ergebnisse zeigen ähnlich gute Ergebnisse der chemotaxonomischen Methode im Vergleich zur morphologisch-anatomischen Methode, mit welcher derzeit alle offiziellen Gutachten erstellt werden. Beide Ansätze zeigten unterschiedliche Stärken und Schwächen bei der Erkennung einzelner Holztaxa. Durch eine Kombination beider Methoden könnten die Synergien zu einer Verbesserung der Gesamttrefferquote führen.

In einer weiteren Untersuchung wurden Provenienzen der Gattungen *Gonystylus* und *Rubroshorea* sowie der Art *Tectona grandis* in vollständig gebleichten Zellstoffen unterschieden. Es konnten herkunftsspezifische Verbindungen in den Extrakten identifiziert werden. Die Ergebnisse der erfolgreichen Differenzierung könnten vor allem für die Überprüfung von Herkunftsangaben an Grenzübergängen und die Durchsetzung des Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES)-Übereinkommens von Nutzen sein, da dies derzeit die einzige zu diesem Zweck für Zellstoffprodukte verfügbare Methode ist. Des Weiteren wurden verschiedene Möglichkeiten

zur Optimierung der chemotaxonomischen Methode erörtert, sowohl im Hinblick auf die Probenvorbereitung und eine verbesserte Kalibrierung der verwendeten GC-MS Geräte z. B. im s-tune, als auch auf die Datenverarbeitung, wie z. B. Ionenfiltereinstellungen und Datenbankkonfiguration. Ein Ausblick auf weitere Methodenvalidierungen an industriellen Zellstoffen und die mögliche Nutzung von maschinellem Lernen für die Automatisierung der Auswertung werden im Ausblick aufgezeigt.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

30x30	**30** by **30** initiative
3D-RLM	**3D r**eflectance **l**ight **m**icroscopy
AI	**a**rtificial **i**ntelligence
ANOVA	**An**alysis **o**f **Va**riance
APRIL	**A**sia **P**acific **R**esources **I**nternational **L**imited
BLE	**B**undesanstalt für **L**andwirtschaft und **E**rnährung
BMUV	**B**undes**m**inisterium für **U**mwelt, Naturschutz, nukleare Sicherheit und **V**erbraucherschutz
CBD	UN **C**onvention on **B**iological **D**iversity
CITES	**C**onvention on **I**nternational **T**rade in **E**ndangered **S**pecies of Wild Fauna and Flora
COP15	**15**th **C**onference **o**f the **P**arties
CV	**c**oefficient of **v**ariation
CWP	**c**hemical **w**ood **p**ulp
DART-TOF-MS	**d**irect **a**nalysis in **r**eal **t**ime - **t**ime **o**f **f**light - **m**ass **s**pectrometry
DB	reference **d**ata**b**ase
DBU	**D**eutsche **B**undesstiftung **U**mwelt
DNA	**d**eoxyribo**n**ucleic **a**cid
DU	**D**obson **u**nits
E/MSY	**e**xtinctions per **m**illion **s**pecies-**y**ears
EA-IRMS	**e**lemental **a**nalysis - **i**sotope **r**atio **m**ass **s**pectrometry
ESG	**e**nvironmental, **s**ocial, and corporate **g**overnance
EPA	**E**nvironment **P**rotection **A**ustralia
EUDR	**E**uropean **U**nion **D**eforestation **R**egulation
EUTR	**E**uropean **U**nion **T**imber **R**egulation
FAME	**f**atty **a**cid **m**ethyl **e**ster
FAO	**F**ood and **A**griculture **O**rganization of the United Nations
FESEM	**f**ield **e**mission **s**canning **e**lectron **m**icroscopy
FCPF	**F**orest **C**arbon **P**artnership **F**acility

FTIR-ATR	**F**ourier **t**ransform **i**nfra**r**ed - **a**ttenuated **t**otal **r**eflectance
G7ANPE	**G7**-**A**lliance on **N**ature **P**ositive **E**conomies
gAI	**g**enerative **A**rtificial **I**ntelligence
GBF	Kunming-Montreal **G**lobal **B**iodiversity **F**ramework
GDP	**g**ross **d**omestic **p**roduct
GI-TOC	**G**lobal **I**nitiative against **T**ransnational **O**rganised **C**rime
HAC for N&P	**H**igh **A**mbition **C**oalition **for N**ature & **P**eople
HANPP	**h**uman **a**ppropriated **n**et **p**rimary **p**roduction
HMT	**H**er **M**ajesty **T**reasury
HPLC	**h**igh **p**erformance **l**iquid **c**hromatography
IPBES	**I**ntergovernmental Science-Policy **P**latform on **B**iodiversity and **E**cosystem **S**ervices
IPCC	**I**ntergovernmental **P**anel on **C**limate **C**hange
ISEGA	**ISEGA** Forschungs- und Unternehmensgesellschaft mbH
IUCN	**I**nternational **U**nion for **C**onservation of **N**ature
MCA	**m**atched **c**hromatogram **a**rea
MCR-AR	**m**ultivariate **c**urve **r**esolution - **a**lternating **r**egression
MDF	**m**edium **d**ensity **f**ibreboard
MF	**m**atch **f**actor
Mha	**m**illion **h**ect**a**res
MPS	**m**arker **p**eak **s**cores
Mt	**m**illion metric **t**onnes
MTH	**m**ixed **t**ropical **h**ardwood
NGO	**n**on-**g**overnmental **o**rganisation
NIRS	**n**ear-**i**nfra**r**ed **s**pectroscopy
NPI	**N**ature **P**ositive **I**nitiative
NPP	**n**et **p**rimary **p**roduction
originID	**origin id**entification
OSB	**o**riented **s**trand **b**oard
PCx	**p**rincipal **c**omponent **x**
PFTBA	**p**er**f**luoro**t**ri**b**uty**l**amine

PPCDAm	**P**lan for the **P**revention and **C**ontrol of **D**eforestation in the legal **Am**azon
Pub I–III	**pub**lications **I–III**
REDD+	**R**educing **E**missions from **D**eforestation and Forest **D**egradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries
RSI	**r**everse **s**imilarity **i**ndex
RT	**r**etention **t**ime
SNIP	**s**tatistics-sensitive **n**on-linear **i**terative **p**eak-clipping
speciesID	**species id**entification
taxonID	**taxon id**entification
TD-GC-MS	**t**hermal **d**esorption - **g**as **c**hromatography - **m**ass **s**pectrometry
THM	**t**hermally assisted **h**ydolysis and **m**ethylation
TI	**T**hünen **I**nstitute for Wood Research
TIC	**t**otal **i**on **c**ount
TMAH	**t**etra**m**ethyl**a**mmonium **h**ydroxide
TUDa	**T**echnische **U**niversität **Da**rmstadt
UHH	**U**niversität **H**amburg
UKTR	**U**nited **K**ingdom **T**imber **R**egulation
UNFCCC	**U**nited **N**ations **F**ramework **C**onvention on **C**limate **C**hange
UNEP	**U**nited **N**ations **E**nvironment **P**rogramme
UNODC	**U**nited **N**ations **O**ffice on **D**rugs and **C**rime
WWF	**W**orld **W**ide **F**und for Nature

1 Introduction

1.1 Background

1.1.1 Pulp and paper production

According to the food and agriculture organization of the United Nations (FAO), the total global roundwood production in 2022 was 3 983 million metric tonnes (Mt), of which roughly 18 % was consumed by the pulp and paper industry to produce 414 Mt of paper and paperboard (FAO 2022a). Chemical wood pulp (CWP) is of particular interest for this work, because it is the most produced of all pulp products (chemical, semi-chemical and mechanical pulp) and a very large amount of fresh wood is used for its production. As a result, CWP poses one of the greatest global deforestation threats and is regulated by various local and international laws, as explained in section 1.1.3 (Political commitments and binding legislation).

CWP production has grown steadily since the 1960s from less than 38 Mt, as shown in Figure 1, to an annual worldwide production quantity of 158.1 Mt in 2022 according to the latest data (FAO 2022a). Historically, species identification was mainly requested for solid wood. However, with increased production volumes and the implementation of the European Timber Regulation (EUTR) (EU 2010), which also regulates pulp and paper products, the market and its demand for wood identification services have shifted. As a result, pulp, paper and fibre materials have become more important. According to wood anatomists at the Thünen Institute for wood research (TI) – one of the world's leading wood identification institutes – the number of requests for expert opinions on the identification of wood taxa in fibre and paper products has increased (Olbrich 2024; Koch et al. 2017).

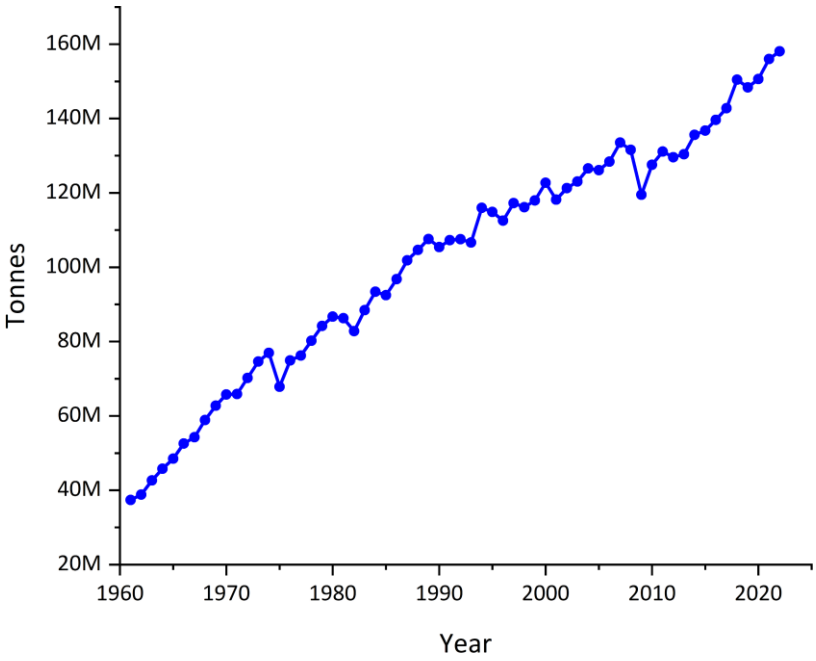


Figure 1: World production quantity of chemical wood pulp in Mt from 1961 to 2022. Data source: FAO (2022a).

The biggest proportional increase in CWP production by world region is seen in Asia (Figure 2). According to FAO data, Asia produced about 17.9 Mt in the year 2000, representing a market share of 14.7 % (Figure 2). By 2022, Asian production had more than doubled to 38.7 Mt, representing 24.5 % of world production – a proportional increase of 10 % in global share compared to 2000. The main driver of the increase in Asian CWP production was China, which alone accounted for almost 12 % of the world production in 2022 and is estimated to potentially grow to one-third of global pulp and paper production in less than ten years (Elhardt et al. 2018). The total world production quantity in 2000 was 122.7 Mt, which is 35.3 Mt less than in 2022. Global paper consumption is unsustainably high, still growing and unevenly distributed (Martin and Haggith 2018). In 2020, the entire continent of Africa consumed less than 2 % of the world's paper and board, while China and the USA alone consumed almost 46 %, with a global average consumption of 51 kg per person (FAO 2022c).

Due to the strong growth of Asian pulp and paper production as seen in Figure 2 (FAO 2022a), this work puts a spotlight on mixed tropical hardwood (MTH) from South East Asia as listed in the first column of Table 2.

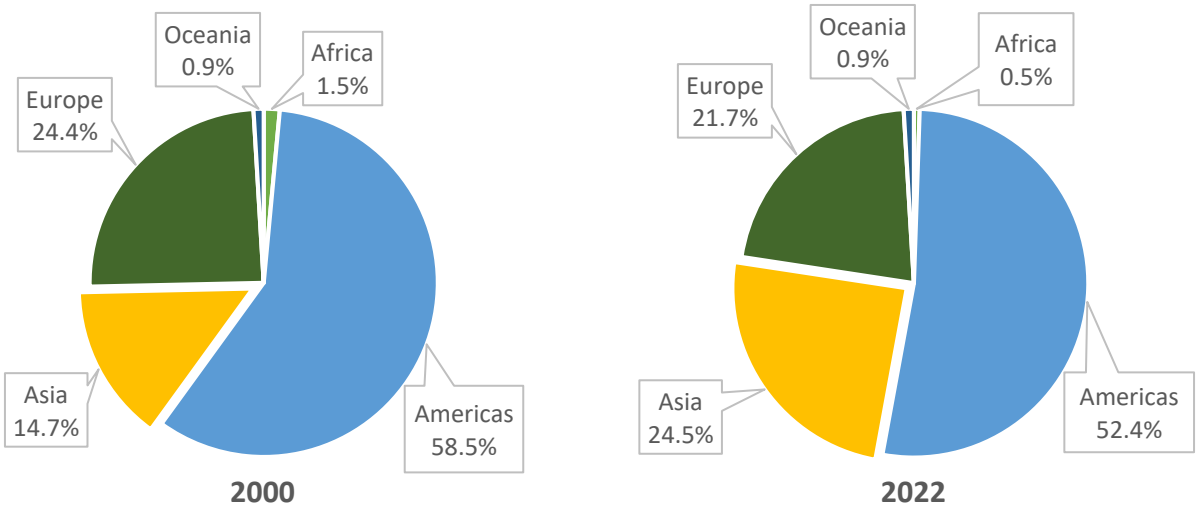


Figure 2: Chemical wood pulp production quantity share in the year 2000 (left) and 2022 (right) by world regions. Data source: FAO (2022a).

1.1.2 Loss of biodiversity and forests

1.1.2.1 Species extinction rates, population declines and planetary boundaries

Not only tree species in the tropical forests of South East Asia are at risk of extinction. Biodiversity loss in general is a major global environmental problem that is progressing at an alarming rate. Estimates of the total number of species worldwide vary from 5 ± 3 million (Costello et al. 2013) to 11 million (Chapman 2009), with high uncertainties for insects, fungi and species living in the deep sea (Scheffers et al. 2012). Pimm et al. (2014) estimate that there are more than 450 000 land plant species, including ~352 000 *Magnoliophyta* (Chapman 2009) and 58 497 described tree species (BGCI 2021).

In this context, it is important to understand the background rate of extinction, a term often used in conservation biology and described by Pimm et al. (1995) as a reference to the natural average extinction rates without any human influence over long periods of time. Extinctions are counted per million species-years (E/MSY), since fossil species are assumed to last 10^6–10^7 years. An average fossil species for instance has a background extinction rate of 0.1 to 1 E/MSY. With approximately 5 to 10 million species currently on the planet, this would translate to a background rate of 0.5 to 10 extinctions per year globally. Pimm et al. (2014) conclude that the modern average extinction rate of ~100 E/MSY is 1 000 times higher than the overall average pre-human background rate of ~0.1 E/MSY for animals and plants combined. For example, even the conservatively calculated current average loss rates of different vertebrate species by Ceballos et al. (2015) are still up to 100 times higher than the benchmark used: an also very conservatively estimated background rate of 2 mammalian extinctions per 10 000 species per 100 years – correctly given as 2 E/MSY. Ceballos et al. (2015) confidently conclude that the total modern and human-induced extinction rate is well above the natural background rate and still accelerating. They predict the “sixth mass extinction” in 4.5 billion years of history on earth.

According to the Red List of Threatened Species of the International Union for Conservation of Nature (IUCN), over 42 000 of the total 163 000 species that have undergone assessment so far are currently threatened with extinction (IUCN 2024). This represents a proportion of more than 25 %. The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) estimates that ~1 million of the ~8 million animal and plant species on Earth are already threatened with extinction (IPBES 2019).

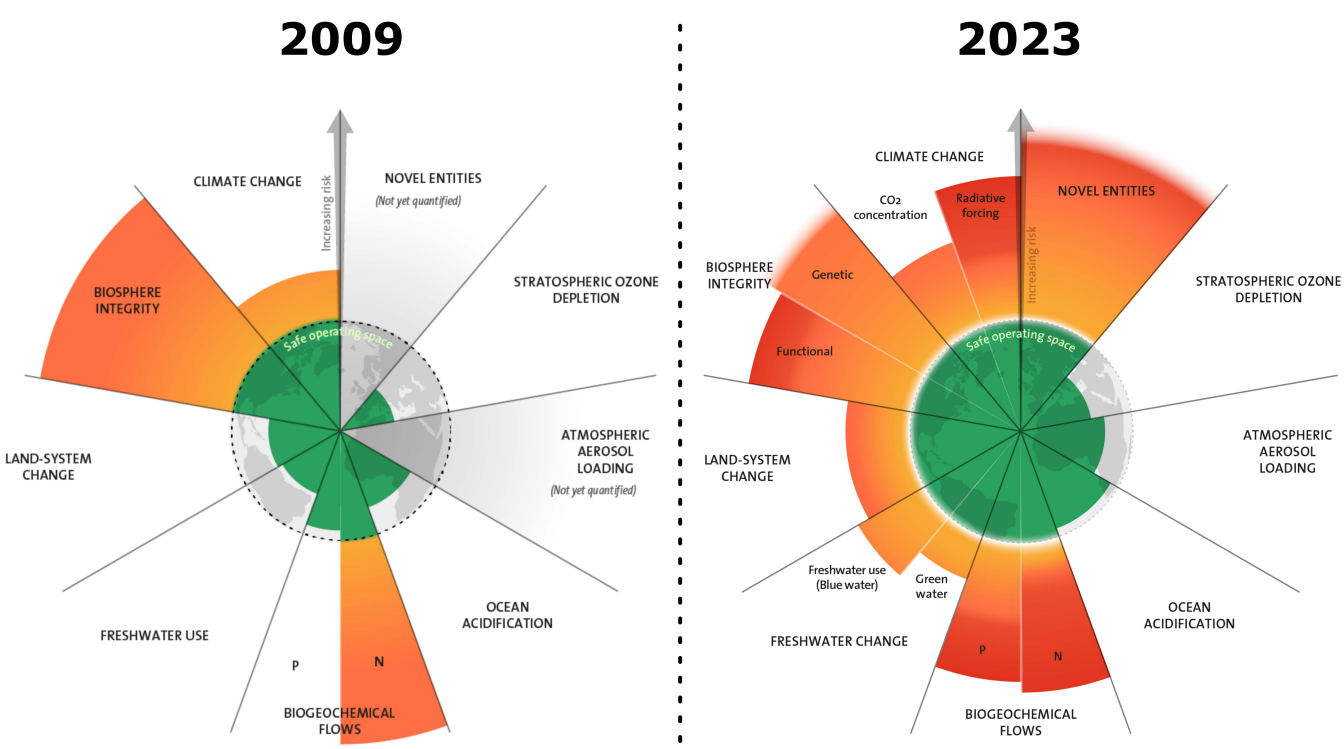


Figure 3: The planetary boundaries framework: status in 2009 (Rockström et al. 2009) and status in 2023 (Richardson et al. 2023). Licenced under CC BY-NC-ND 3.0. Credit: © Azote for Stockholm Resilience Centre, Stockholm University, based on Richardson et al. (2023), Steffen et al. (2015), and Rockström et al. (2009).

Richardson et al. (2023) also used 1 E/MSY as "pre-industrial Holocene base value" to compare with the change in genetic diversity as one half of the biosphere integrity planetary boundary shown in Figure 3. The figure illustrates the progression of planetary boundaries from their status in 2009, as outlined by Rockström et al. (2009), to their status in 2023, as presented by Richardson et al. (2023). Consequently, the content and colour grading of the two periods are not identical. The figure is structured into zones of different risk levels. The green inner safe operating zone of each boundary has a measurable threshold value up to which there is no danger. The safe zone gets filled with green colour until the border to the zone of increasing risk (light orange) is crossed. After exceeding the increasing risk zone, the high risk zone begins. For the genetic diversity within the biosphere integrity boundary, the "upper end of the zone of increasing risk" is set at 100 E/MSY. The "current value of [the] control variable" is estimated to be > 100 E/MSY (Ceballos et al. 2015; Rounsevell et al. 2020; Cowie et al. 2022), which is why Richardson et al. (2023), Steffen et al. (2015) and already Rockström et al. in (2009) assessed this planetary boundary as far transgressed, already approaching the high risk zone, coloured dark orange in Figure 3. The colour at the end of the genetic boundary for 2023 in Figure 3 is blurred due to high uncertainties in the measures of biodiversity loss. However, it is still certainly well outside the safe zone, which means that this system of the Earth has never been so disrupted before the appearance of humans (Richardson et al. 2023). This giant extinction crisis also includes forests and trees, which do play an important role in many

ecosystems. Approximately 50 % of all terrestrial animal and plant species depend on trees to provide them with shelter and habitat. A great variation of trees is specifically important in this context as they grow in different constellations, densities, heights etc. and thus create different habitats. According to a report by the Botanic Gardens Conservation International (BGCI 2021) trees are not only important for the biodiversity and ecology, they are also important for the economy, for culture and for their role as a carbon sink and hence in preventing the climate crisis.

According to BGCI (2021) approximately 30 % of all tree species are threatened with extinction and 142 species are already extinct in their natural wild habitat. This is due to invasive pests, illegal logging and overexploitation, which reduce the number of trees and the habitat they need to recover. Climate change itself is also a major threat to many tree species, as it alters their environment and weakens them by, for example, reducing or increasing water availability (BGCI 2021).

The tropical regions of Central and South America, tropical South East Asia and Africa have the highest total number of tree species, while Madagascar has the highest proportion of threatened species. As the temperate zones have less species diversity, they also have the lowest proportion of threatened species (BGCI 2021).

For all conifers assessed, the proportion of threatened species is even higher at 34 % according to IUCN (2024) than the total number of 30 % for conifers and broadleaf trees combined (BGCI 2021). The Living Planet Report by the World Wildlife Fund for Nature (WWF) (2022) states a 69 % decline in wildlife populations between 1970 and 2018 in their global “Living Planet Index”. This number is based on nearly 32 000 populations representing 5 230 species of mammals, birds, amphibians, reptiles and fish that have been monitored over time and tracked for changes in their relative abundances. The index has been stress-tested by excluding some species or populations to ensure it is independent of extreme increases or decreases. Both declining and increasing trends were found, but the average index development over the period observed was -69 % (WWF 2022).

In total, there are nine planetary boundaries affecting the stability and resilience of the Earth as entire system – with a major impact on whether the Earth remains in the stable pre-industrial Holocene condition in which modern humanity evolved (Richardson et al. 2023), or continues to drift into a new unknown state – the Anthropocene (Waters et al. 2016). All boundaries, described by Richardson et al. (2023) are interconnected and must be seen as a complex system, with the probability of impacts on the Earth and humanity increasing as the scientifically based thresholds for safe zones are crossed. Crossing a single boundary increases the risk of huge and irreversible environmental changes, but does not cause the whole system to collapse. This is the reason why the focus on climate change alone is said to be insufficiently targeted and not sustainable. More importantly, the interactions between all nine, but especially between biodiversity as a regulator of ecosystem functions and climate change, need to be understood and addressed together (Richardson et al. 2023).

Figure 3 shows that by 2023, six of the nine planetary boundaries have already been crossed from the safe operating zone (green) to either the increasing risk zone (light orange) or the high risk zone (red). The boundaries defined by Richardson et al. (2023) are listed and briefly explained below:

1. **Biosphere integrity** consists of two dimensions: First, the genetic diversity described above, with a maximum safe operating extinction rate of < 10 E/MSY to preserve the necessary genetic complexity of the biosphere (Richardson et al. 2023), of which an estimated 10 % is already lost (Exposito-Alonso et al. 2022). Second, planetary function, strongly interrelating with the first – biodiversity. Functional integrity refers to how much energy based on biomass, produced by plants through photosynthesis, is available to the ecosystems, measured in net primary production (NPP) per year. How much of an ecosystem's energy flow remains available to support its natural processes depends on the human appropriation of NPP (HANPP) through harvesting for food and fuel through agriculture, as well as land-use changes like deforestation and urbanization, which reduce the capacity of ecosystems to produce biomass (Haberl et al. 2014). The variability of the pre-industrial Holocene average NPP was 1.9 %. The upper limit of the zone of increasing risk is 20 % HANPP. The current global HANPP value is 30 % – far into the high risk zone (Richardson et al. 2023).
2. **Novel entities** are defined as "restricted to truly novel anthropogenic introductions to Earth system. These include synthetic chemicals and substances (e.g., microplastics, endocrine disruptors, and organic pollutants); anthropogenically mobilized radioactive materials, including nuclear waste and nuclear weapons; and human modification of evolution, genetically modified organisms and other direct human interventions in evolutionary processes" (Richardson et al. 2023).
3. **Stratospheric ozone depletion**: The Earth's ozone layer is being damaged by industrially released volatile halocarbon compounds. However, the layer is very important for protecting life on Earth from various types of UV radiation and must be kept above a certain level, a maximum of 5 % lower than the 290 Dobson units (DU) that characterize the pre-industrial state benchmark. It is currently in the safe zone (Richardson et al. 2023).
4. **Atmospheric Aerosol Loading**: High levels of anthropogenic aerosols affect many things, such as temperature changes in the northern and southern hemispheres (Ocko et al. 2014) or changes precipitation, e.g. leading to less monsoon rainfall over land (Steffen et al. 2015; Westervelt et al. 2020), which in turn affects the biosphere integrity (Richardson et al. 2023).
5. **Ocean acidification** is currently still in the safe operating range, but is worsening due to high and constant CO_2 emissions (Richardson et al. 2023).
6. **Biogeochemical flows**: Global nitrogen (N) and phosphorus (P) cycles are critical for aquatic systems. Too much input to freshwater lakes and oceans causes eutrophication, algal blooms (Bennett et al. 2001) and very low oxygen levels, known as anoxia, which are critical for many aquatic plants and animals and for entire

underwater ecosystems (Zillén et al. 2008). Both N and P limits are being exceeded globally by industry and intensive fertilization in agriculture (Richardson et al. 2023).

7. **Freshwater change** is divided into “blue water” and “green water” increases and decreases (Richardson et al. 2023). “Blue water” represents flowing ground and surface water, such as lakes and rivers, and represents the integrity of aquatic ecosystems (Porkka et al. 2024). “Green water” is defined as the water available to plants in the wet soil between their roots, representing terrestrial ecosystem and hydroclimate interactions (Wang-Erlandsson et al. 2022).
8. **Climate change**: Various sources such as Richardson et al. (2023), Forster et al. (2023) and IPCC (2021) state that the primary drivers of climate change are greenhouse gas and aerosol emissions, as well as changes in surface albedo, which is a measure of reflectivity. The two controlling variables for the planetary boundary are the atmospheric CO_2 concentration of no more than 350 ppm for safe operation and the radiative forcing, which measures the change in energy flux at the top of the atmosphere. The safe operating maximum for radiative forcing is +1 W/m^2, which represents an increase in energy entering the Earth’s system leading to warming (incoming solar radiation minus outgoing radiation). The concentration of CO_2 has exceeded the preindustrial Holocene base value (280 ppm) and the planetary safe operating boundary of 350 ppm, reaching a current level of 417 ppm. This value is situated just below the upper end of the zone of increasing risk, which is 450 ppm. Meanwhile, the current radiative forcing of +2.91 W/m^2 has already exceeded the upper limit of the increasing risk zone (+1.5 W/m^2) and entered the high risk zone (Richardson et al. 2023; Forster et al. 2023; IPCC 2021).
9. **Land system change** is based on the three main forest biomes tropical, temperate and boreal (Steffen et al. 2015). The control variable is the remaining percentage of the original area covered by forest. On a global scale, Richardson et al. (2023) present the average of 75 % as the planetary boundary of the safe zone and 54 % as the upper limit of the increasing risk zone, which is currently not exceeded with a total of 60 %. For the different biomes, the individual upper limits of high risk have been set at 60/30/60 per cent for tropical/temperate/boreal. FAO (2020) and Richardson et al. (2023) present a mixed picture of the current state of tropical forests, with tropical America doing well at 84 % forest cover, but tropical Africa and especially Asia already in the high-risk zone at 54 % and 38 % respectively. Temperate forests are all in the increasing risk zone, with the Americas at 51 %, Europe at 34 % and Asia at 38 %. For boreal forests, the Americas have just entered the high risk zone at 57 % and Eurasia is still assessed as being at increasing risk at 70 % (Richardson et al. 2023; FAO 2020).

The following section 1.1.2.2 (Habitat loss) describes in detail the loss of crucial habitats especially the forests.

1.1.2.2 Habitat loss

Loss of habitat and geographic area, whether terrestrial or marine, leads to fewer individuals of both flora and fauna, and therefore less genetic diversity and ecosystem function, spiralling back to the planetary boundary of biosphere integrity (Exposito-Alonso et al. 2022; Richardson et al. 2023). Loss of genetic diversity in turn increases the risk of extinction in the wild, as shown by Spielman et al. (2004), who found that heterozygosity was 35 % lower in more than three-quarters of 170 threatened taxa than in closely related but nonthreatened taxa. Heterozygosity is an important measure of genetic diversity within a population. Populations with lower heterozygosity have reduced reproductive fitness and less genetic variation, which is disadvantageous for surviving disease and adapting to environmental changes, leading to a higher risk of future extinction (Spielman et al. 2004). An example of a tropical tree species studied for heterozygosity in dependence on degradation of its habitat is *Palaquium obovatum* (Phang et al. 2024). Populations in Singapore and Malaysia were sampled across different tree generations and separated habitat fragments. The results show that near-wild populations from fragmented geographical areas suffer from genetic erosion, measured as loss of heterozygosity (Phang et al. 2024).

The IPBES (2019) Global Assessment Report on Biodiversity and Ecosystem Services shows that the majority of natural landscapes (about 75 %) have been altered by human activities: These significant changes are mainly due to agriculture, urbanisation and resource extraction. Many impacts on ecosystems and declines in biodiversity have been observed. The same is true for marine areas, 66 % of which have been affected, and for wetlands, 85 % of which have been lost. The average global rate of forest degradation and loss has slowed since 2000, but this is partly due to the reforestation of formally natural forest areas, mostly in temperate and boreal regions, with monocultures, which obviously do not have the same ecosystem and biodiversity value. Since 1990, a total global forest area loss of about 420 Mha has been reported. In tropical and subtropical areas with high biodiversity, for instance, 32 Mha of primary and recovering forest area was lost between 2010 and 2015 (IPBES 2019). Agricultural deforestation disturbed/cleared 17 %/11 % of the Amazon between 1995 and 2017 (Bullock et al. 2020). The total global annual rate of deforestation between 2015 and 2020 was about 10 Mha, according to the “Global Forest Resources Assessment 2020: Main report” by the FAO (2020). This compares to a global rate of forest growth of only 5 Mha per year over the same period (Figure 4). The previous period (2010–2015) had both higher losses of 12 Mha and higher gains of 7 Mha per year (FAO 2020).

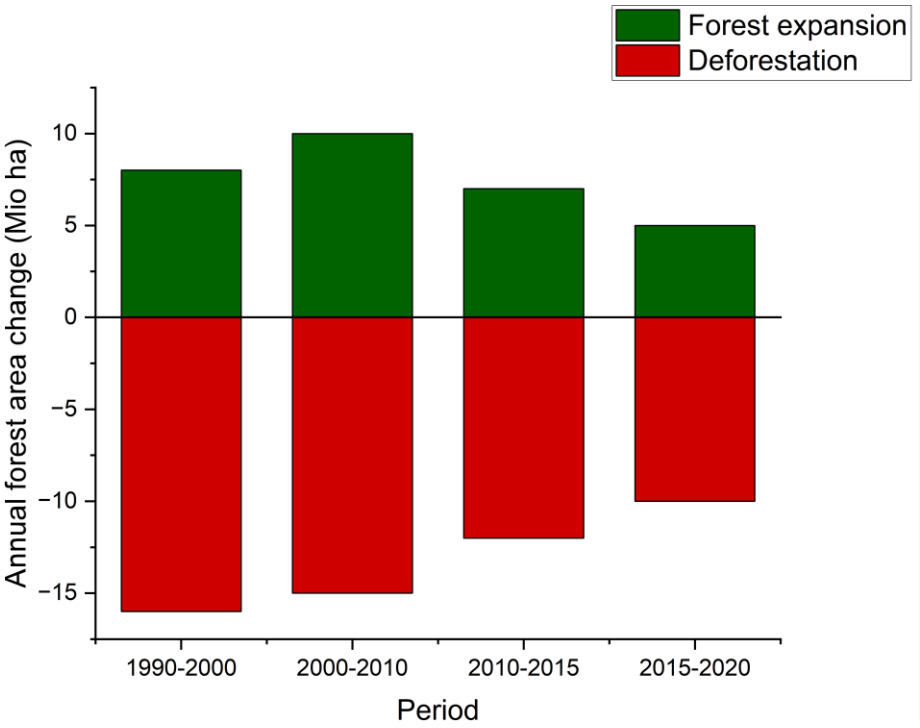


Figure 4: Global forest area gains and losses from 1990 to 2020. Data source: FAO (2020). Graph adapted from the original by © FAO (2020) (https://doi.org/10.4060/ca9825en), licensed under CC BY-NC-SA 3.0 IGO.

In the decade from 2010 to 2020, Africa experienced the highest rate of forest loss at -3.9 Mha per year, followed by South America at -2.6 Mha per year. Interestingly, Asia recorded a positive forest cover increase of 1.2 Mha per year in the same period, which is only half of the 2.4 Mha increase in the previous decade (2000–2010) (FAO 2020). Europe gained only 0.3 Mha of forest per year between 2010 and 2020. As explained above, forest expansion does not necessarily mean recovery of natural forest areas. Between 2010 and 2020, 3 Mha of planted forests were gained, while -8 Mha of naturally regenerating forests were lost. In total, 7 % of the world's forests are planted, of which about 45 % (131 Mha) are plantations. 99 % of South America's planted forests are plantations, of which 97 % consist of introduced species. These plantations account for 2 % of the total forest area in South America. Globally, 66 % of plantations are of native species. In 2020, only 18 % (726 Mha) of the world's forests will be in protected areas, with Europe having the smallest share of protected forests at only 6 % (FAO 2020). Primary forests, which have not been disturbed by humans to the extent that natural processes have been disrupted, and which consist of native trees, cover 27 % (1.11 billion ha) of the total forest area, which has decreased by 81 Mha since 1990 (FAO 2020).

1.1.2.3 Causes of biodiversity loss

There are many reasons for loss of biodiversity. One is habitat destruction and fragmentation (Wilson et al. 2016), due to overexploitation for instance by overfishing, overhunting or deforestation, as described in section 1.1.2.2 (Habitat loss). Deforestation is caused by large-scale logging for timber and fuel, or by the removal of trees for agriculture, especially the

conversion of forests into arable land for soy and cattle farming, urbanisation and infrastructure projects (Johanson 2024; FAO and UNEP 2020; WWF 2022). This is especially predominant in tropical regions such as the Amazon. The Amazon rainforest is home to around 16 000 tree species (ter Steege et al. 2013), many of which are threatened by habitat loss and fragmentation, which is the main threat to the majority of species listed as threatened and endangered by the WWF (2024) and IUCN (2024).

Forests are being fragmented into smaller pieces, isolating species and reducing genetic diversity, and also drying out due to less precipitation. This in turn feeds into another major cause of deforestation and a dangerous threat to tree species and wildlife: forest fires (FAO 2020). Both natural and human-caused fires can devastate large areas of forest, further reducing their habitat. Although some ecosystems, such as savannahs, need natural fires to maintain their function, most fires are human-caused and accelerated by low humidity and rainfall as well as high temperatures and winds (FAO and UNEP 2020; FAO 2020). Between 2003 and 2012, around 67 Mha of forest area burned each year, mainly in South America and Africa (van Lierop et al. 2015). In 2015, a total of 98 Mha of forest area was burnt (FAO 2020). 4 % of tropical forests were affected by these fires, most of which occurred in Africa and South America. The impact of fires extends beyond the loss of biodiversity; they also result in the loss of human life, destruction of built assets and a decline in productivity. On the other hand, boreal and temperate forests have also been affected by severe weather events, mainly in Asia, and insect pests such as the bark beetle outbreak, which has mainly been observed in North America and Europe (Hlásny et al. 2021), and the emergence of diseases, which have been identified as a significant threat in Europe and Asia. A total of 142 Mha of forest was damaged by other events than fire between 2003 and 2012, which accounts for 5 % of the reporting 75 countries' total forest area (van Lierop et al. 2015). In 2015 alone 40.6 Mha of forest were affected (30.2 Mha by insects, 6.6 Mha by diseases, and 3.8 Mha by severe weather events) reported by over 60 countries worldwide, representing a proportion of 2.1 % of the reporting countries' forested area (FAO 2020). As the reporting countries and assessment methods of van Lierop et al. (2015) and FAO (2020) are not identical, the above numbers are not comparable. In addition to agriculture, deforestation and fires, industrial expansion also affects biodiversity. For example, in the energy sector, in addition to oil and gas extraction, renewable energy such as wind energy causes habitat destruction/fragmentation and species mortality, as shown by Jones et al. (2015).

Other causes of biodiversity loss include pollution (Spielman et al. 2004), invasive species (Holmes et al. 2009), climate change and severe weather events. Approximately 0.3 % of reported forest area (37 % of the world's forests) were affected by severe weather events in 2015, for example 2.1 Mha in North and Central America, representing 0.5 % of the region's reported forests (FAO 2020). Climate change is increasingly recognised as a significant factor, threatening species that cannot adapt quickly enough to changing conditions (WWF 2022). Insects are another factor damaging forests, leading to loss of habitat and biodiversity. In 2015, 8.8 Mha of forest in Asia were affected by insect pests. This alone accounted for 3 % of

the total area reported that year (FAO 2020). Diseases, reported by countries covering 42 % of the world's forests, affected 6.6 Mha, representing 0.4 % of the total forest area of the 51 reporting countries (FAO 2020).

1.1.2.4 Impacts on Ecosystem Services

Vital and healthy forests and ecosystems in general with high functional integrity provide more goods like food, wood, fibre, fuel, clean air, water filtration and regulate disease and climate better than degraded ecosystems with high HANPP (FAO 2020; Sarukhán and Whyte 2005). Trees play a critical role in regulating water cycles, including groundwater recharge and flood control. Deforestation reduces this capacity, leading to more frequent floods and droughts. Deforestation can also reduce downstream water quality and emit greenhouse gases. Many plants, especially trees, contribute to soil formation, protection against soil erosion and nutrient cycling (supporting ecosystem services). Loss of tree species reduces soil fertility, which can affect the health of an ecosystem and lead to desertification (FAO 2020). Healthy forests can act as carbon sinks, sequestering CO_2 and mitigating climate change. Loss of biodiversity reduces this capacity, exacerbating global warming, according to the intergovernmental panel on climate change (IPCC) (IPCC 2022). Another important ecosystem service is the pollination of fruits, vegetables and cereal crops by insects and some birds, securing the food supply for humans and livestock. A decline in flowering plants can lead to a decline in bees and other insects, resulting in reduced agricultural yields and, in turn, reduced plant reproduction in these ecosystems (IPCC 2022).

Loss of biodiversity also affects societies. In addition to the provisioning, regulating and supporting services mentioned above, ecosystems also provide cultural services such as aesthetic, spiritual, educational and recreational (Sarukhán and Whyte 2005). For indigenous societies, the loss of marine or terrestrial wildlife, e.g. changes their way of life, not only through the loss of food, but also through the loss of knowledge, language and culture (WWF 2022). Zhang et al. (2022b) provided a broad overview of how human mental and physical health in modern societies is also strongly influenced by the interdependence of a healthy environment and ecosystem functioning. Both affect human health, for example by providing aromatic and medicinal plants and species-rich habitats. Ecosystem services protect us from floods, heat waves, polluted air and water, and also help us to recover from stress and to support physical activity as well as aesthetic value and satisfaction (Figure 5) (Zhang et al. 2022b).

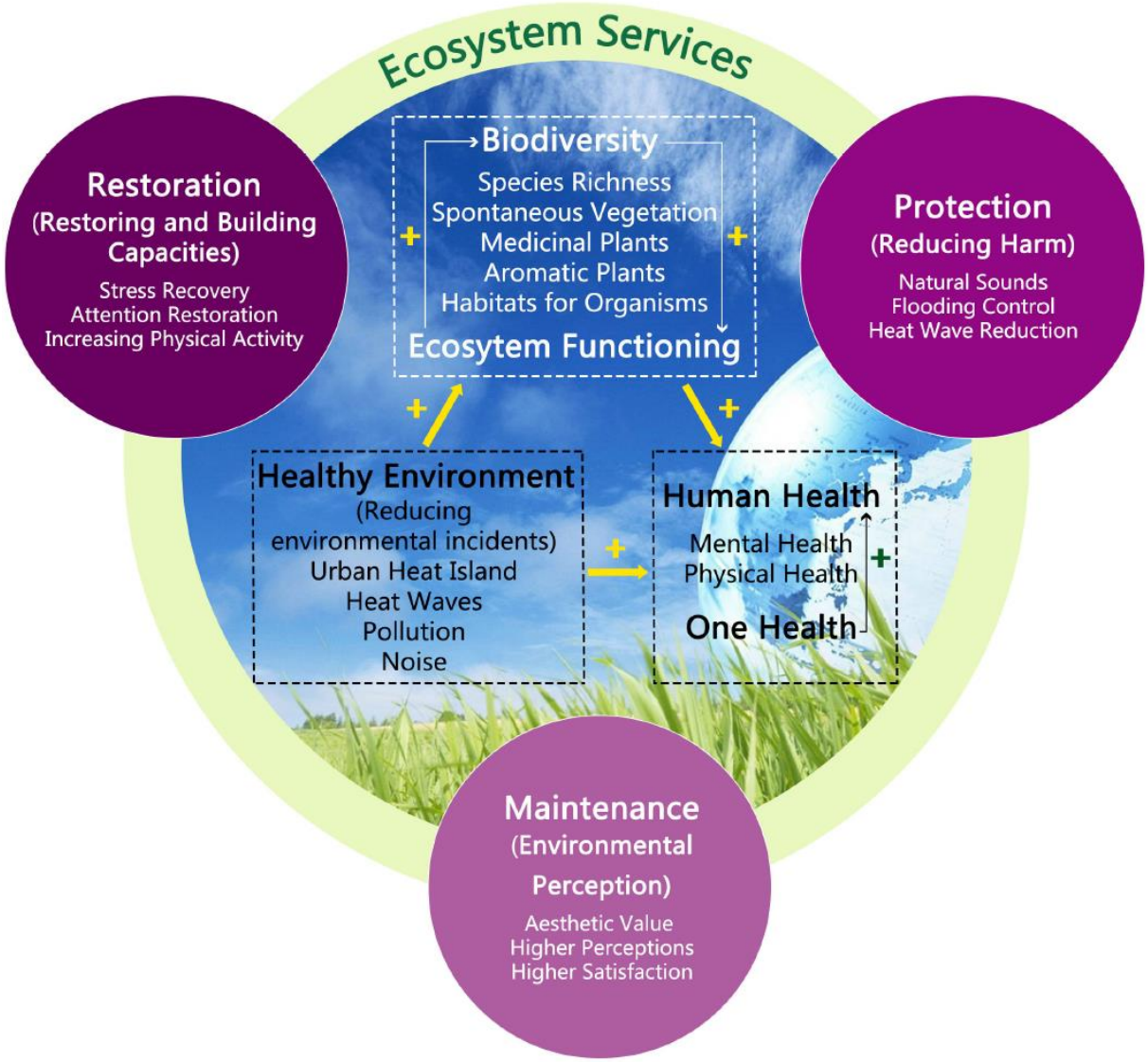


Figure 5: Impacts of ecosystem services, biodiversity and ecosystem functioning on human mental and physical health. Dependencies and interactions between them, with a yellow plus sign indicating a positive influence. Licensed under CC BY 4.0. Credit: © Zhang et al. (2022b) (doi.org/10.3389/fevo.2022.1086408).

In addition to health, basic materials for a good life, and security, other components of human well-being are impacted (though to a lesser extent) by ecosystem services. Social relations such as social cohesion, mutual respect, the ability to help others and ultimately freedom of choice and action are affected (Sarukhán and Whyte 2005). Although most introduced species are relatively harmless to the new biome, some invasive species, such as the chestnut blight – a fungus native to East Asia that is responsible for the death of many chestnut trees in Europe and North America – have not only ecological but also high economic impacts, causing severe damage to the forest ecosystem and high economic losses (Holmes et al. 2009). Many plants are sources of medicines. Loss of biodiversity could also mean that the full potential of new medicines and treatments cannot be realised.

1.1.2.5 Future projections and conservation efforts

If current trends in extinction rates and the loss of nature and habitats continue, hundreds of thousands of species will become extinct in the next few decades. The “sixth mass extinction” is underway, the first ever triggered by a single species – humans (Ceballos et al. 2015; Cowie et al. 2022).

Part of the solution to halting this massive loss of biodiversity is to increase forest cover. To expand and enforce forest cover, especially primary and secondary but near-natural forests that are protected for their recognised natural values, can help protect tree species and the ecosystems they support (WWF 2024). This mitigates climate change by maintaining the CO_2-sink function, contributing to the well-being of the Earth and humanity (Beck-O’Brien et al. 2022). 18 % of the world’s forest area (> 700 Mha) is already protected by legal status such as national parks, protected areas or game reserves (FAO 2020). To conserve forests and prevent deforestation, forest indicators and data need to be monitored comprehensively in as many countries and regions as possible. Monitoring capacity has increased, but there are still information gaps that need to be filled to provide a more complete picture, to make more data available, and to strengthen the capacity of each reporting country to continuously monitor forests in order to identify trends early and act accordingly in the future. Therefore, the reporting burden for individual countries should be reduced and legal requirements should be adapted (FAO 2020).

In addition to the protection of areas, reforestation and afforestation are important tools for maintaining forest cover. A considerable number of initiatives have been launched with the objective of restoring forests and plantations, with the ultimate goal of reinstating lost biodiversity and ecosystem services. One example is the Bonn Challenge, which aims to restore 350 Mha of forest worldwide by 2030. With more than 60 countries participating, the milestone of 150 Mha by 2020 was already exceeded in 2017, according to International Union for Conservation of Nature (IUCN) and German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMUV) (IUCN and BMUV 2011). Another example is the global 30 by 30 initiative (30x30), which aims to protect 30 % of land and marine areas by 2030 and has been signed by 119 countries to date (13 September 2024), according to the United Nations Environment Programme (UNEP) and the High Ambition Coalition for Nature & People (HAC for N&P) (UNEP 2022; HAC for N&P 2024), based on the paper by Dinerstein et al. (2019).

It is important to define sustainable forestry and wood use in the context of the climate and biodiversity crises (Beck-O’Brien et al. 2022) and to not think of them in isolation, for example by trying to achieve net zero emissions by producing biofuels on a large scale, while ignoring their ecological impacts. Therefore, nature positivity is the future, combining human development inequities and planetary boundaries, especially climate change and biodiversity loss, into one complex system (Davis 2022; Locke et al. 2021). The nature positive goal was defined by a group of scientists and CEOs of environmental organisations as “halting and reversing nature loss by 2030 based on 2020” and from there to full nature recovery by 2050, as shown in the timeline (Figure 6). These goals are measurable through metrics of

conservation and restoration, species, ecosystems and natural processes at all scales. According to the nature positive initiative (NPI), they include species diversity and distribution, extinction rates and risks, habitat status and carbon storage amongst others (NPI 2023). It is also important to involve local communities in conservation through education on the sustainable use of ecosystem resources, such as sustainable forestry and logging, but also environmentally friendly agriculture, to ensure minimum habitat destruction and maintain biodiversity.

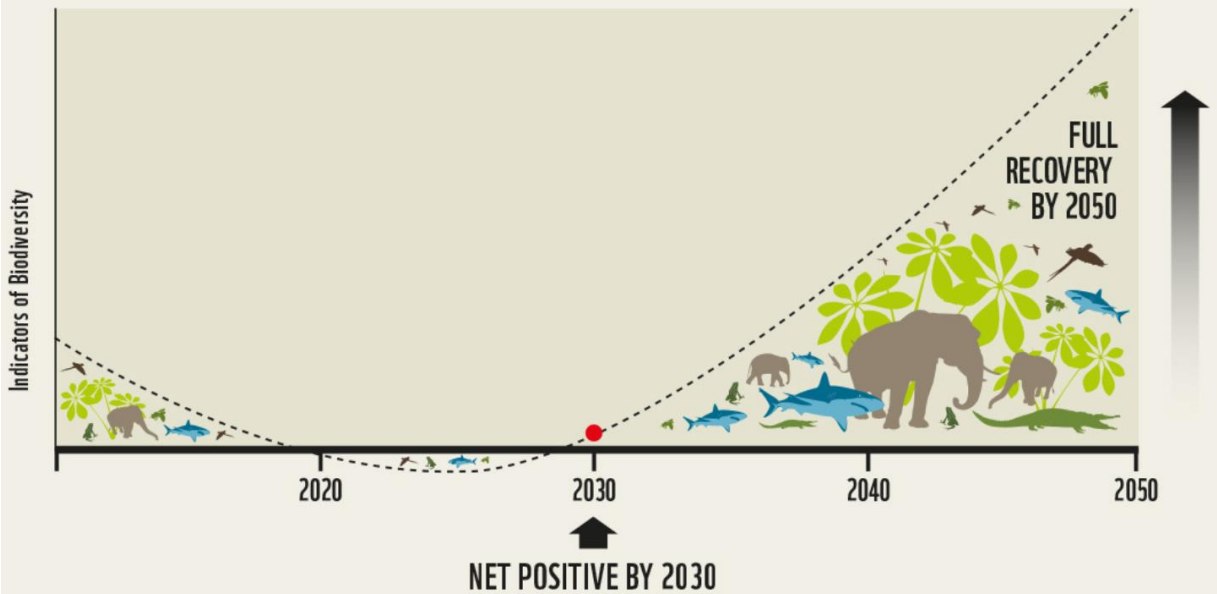


Figure 6: Timeline of the nature positive conservation project based on the reference year 2020. Goal of full nature recovery by 2050. Credit: © Nature Positive Initiative (NPI 2023), original graphic by Locke et al. (2021).

A recent study by Kireta et al. (2024) on cleared pieces of land post-cleared revegetated sites and remnant natural sites examined plant and bee diversity and ecosystem services measured by pollination/fruit set of different plants. They found that revegetated sites with high plant biodiversity, especially in flowering species, contributed almost as much to bee diversity as remnant sites. Revegetated sites with high diversity had a higher diversity of native buzz pollinating bees than cleared or low diversity revegetated sites. Although high diversity sites were richer in bees and also produced more fruit than low diversity and cleared sites, pollination services were only statistically significant for remnant sites. They conclude that conserving remnant nature is most effective and needs to be a top priority, at least for South Australia and for revegetation sites. It needs to be done right – with attention to high diversity! (Kireta et al. 2024)

In order to conserve and sustainably regenerate as much forest as possible, there are a number of possible actions and tools to support these goals. For example, the general consumption of wood products needs to be changed and made more sustainable in order to minimise it in the first place. As Martin and Haggith (2018) write, “The answer to the question ‘paper or plastic’ must more often be ‘neither’.” But reuse and recycling, repair and sharing

are also very important after material use, which should always be the first option and used as long as possible before other cascades and energy production follow. The setting of public agendas by governments, such as the 30x30 target explained above, but also by organisations, companies, communities and individuals, raises awareness and attention to the problems of wasteful and excessive consumption and how wood should be used most wisely. In addition to promoting long-life wood products, policy makers should invest in the circular economy and cascading as an efficient use of wood, for example through education on reuse design and additional uses of waste wood. Unnecessary uses, such as disposable coffee cups, should also be reduced by making them more difficult to use or by banning them. Through these examples and other greater political engagement as well as coordination and collaboration across sectors, forest companies could be encouraged to invest in the bioeconomy (Nikolakis and Innes 2020). In addition, environmental and forest crimes need to be stopped, for example through legislation, as discussed in more detail in section 1.1.4 and 1.1.5.

It is also important for policy makers to track and benchmark wood consumption to ensure that use remains within the planetary boundaries and to guide policies to prevent overconsumption. Additionally governments should lead by example and invest in research to improve data quality, modelling capacity and up-scaling of sustainable practices (Beck-O'Brien et al. 2022).

1.1.3 Political commitments and binding legislation

As explained in section 1.1.2.5, a variety of conservation efforts are key to conserve wildlife species, natural habitats and whole ecosystems. There are two principal methods by which this may be achieved: firstly, the protection and maintenance of intact biomes through protected areas such as national parks and secondly the restoration through the expansion of existing natural areas or the creation of completely new rewilded areas, e.g. through afforestation. Both protection and restoration require legally binding rules and agreements signed by as many governments as possible and supported by as many NGOs, initiatives and companies as possible to ensure compliance (Dinerstein et al. 2019).

1.1.3.1 International frameworks for biodiversity conservation

There are many multilateral environmental agreements and international frameworks, strategies, policies and initiatives that aim to conserve and restore biodiversity, the most important of which are listed below. The aforementioned NPI defined its global target as "Halt and Reverse Nature Loss by 2030 on a 2020 baseline, and achieve full recovery by 2050", as illustrated in Figure 6, to address the global biodiversity loss (NPI 2023; Locke et al. 2021). Although the 2015 international climate change treaty, the Paris Agreement, already included encouragement to conserve carbon sinks such as forests and to reduce emissions from deforestation, its main goal was to limit global warming to a maximum of +1.5 °C compared to pre-industrial levels (UNFCCC 2015). The nature positive goal can thus be seen as an additional sovereign goal with a focus on biodiversity. Many governments have committed to the international nature positive targets and incorporated them into their own national laws

and strategies. For example, the UK government under Prime Minister Boris Johnson committed to a nature positive future in 2021 in response to Professor Dasgupta's international review, according to Her Majesty Treasury (HMT, UK Government 2021).

The Japanese government also updated its "National Biodiversity Strategy and Action Plan 2023–2030" in March 2023 (Government of Japan 2023) in response to the Kunming-Montreal global biodiversity framework (GBF) adopted at the 15th conference of the parties (COP15) at the UN convention on biological diversity (CBD) (UNEP 2022). The Japanese plan includes five basic strategies to achieve the 2050 biodiversity vision of a society "living in harmony with nature by 2050" (UNEP 2022) and nature positive by 2030 (Government of Japan 2023). For example, Basic Strategy 1 "Restoration of the Integrity of Ecosystems" includes status targets such as increasing the area and quality of ecosystems, reducing the risk of extinction of individual species, and conserving genetic diversity. This is supported by action goals such as the 30 by 30 target, reducing pollution and others (Government of Japan 2023). Basic Strategy number 3 promotes the "Transition Strategies toward Nature Positive Economy" as a key priority measure, focusing on the need for businesses to assume a more prominent role in driving the green transformation, including the transition to carbon neutrality, the development of circular economy as well as nature conservation and restoration (Ministry of the Environment, Government of Japan 2024).

Another recent example of progress is the Australian Government's launch of the second stage of its "Nature Positive law reforms" in April 2024 (Australian Government 2024), following its "Nature Positive Plan: better for the environment, better for business" (Australian Government 2022). These last reforms include the establishment of Environment Protection Australia (EPA), a national and independent statutory agency that will be responsible for regulatory activities, to enforce Australia's environmental laws, but also to carry out assessments, to approve projects, to issue permits and to audit businesses to ensure compliance with the law under strict new monitoring authorities (Australian Government 2024).

A G7-Alliance on Nature Positive Economies (G7ANPE) was launched at the Ministers' Meeting on Climate, Energy and Environment of the G7 Summit in Hiroshima/Sapporo in 2023 (G7ANPE 2023). Besides the G7 countries, a number of international companies and the EU are also members of the G7ANPE and have committed themselves to "halting and reversing the loss of biodiversity and becoming nature positive" by 2023 and to efficient and effective implementation of the GBF (G7ANPE 2023). The EU also promotes nature-based solutions policies such as the European Green Deal (European Commission 2019) which must benefit biodiversity and support ecosystem services. It aims to transform the EU, its society and Europe to become the first climate-neutral continent by 2025, with clean energy, circular economy, clean transport, sustainable investment and biodiversity protection to future-proof the Union, supported by research and innovation. Another nature-based solution in line with the Green Deal is the Biodiversity Strategy (European Commission 2020), which commits to "legally protect a minimum of 30 % of the land, including inland waters, and 30 % of the sea

in the Union, of which at least one third should be under strict protection, including all remaining primary and old-growth forests". To achieve this goal, the unsustainable use of land and sea, overexploitation of natural resources, pollution and invasive alien species must be tackled as drivers of biodiversity loss. Therefore, in addition to conservation efforts, other measures such as greening our cities, promoting organic farming and improving the health of Europe's forests should be taken (European Commission 2020).

The Leaders' Pledge for Nature (LP4N 2020) has been signed by 158 world leaders and the President of the European Commission as of September 2024. Together they represent more than 2 billion people and 39 % of the global gross domestic product (GDP). In addition, 10 000 businesses and organisations have committed to bold action. To achieve the goal of "putting nature and biodiversity on a path to recovery by 2030" and the 2050 vision, they are committed to take necessary actions, as in "real accountability and the appropriate legal, economic and financial tools and incentives", combined with a strong political will. At the UN General Assembly in New York in September 2023, many countries announced new contributions to the 2050 goals. Among these were the German government which committed to providing EUR 40 million through its International Climate Initiative (IKI) to the global biodiversity framework fund (GBFF). Colombia announced USD 80 million for the protection of the Amazon and New Caledonia, announced its intention to designate 10% of its marine area as a protected zone (LP4N 2023, 2020).

The term "nature positive" is widely spread across many political agreements, international action plans, strategic programmes and reports. It is no longer used only in the context of conservation strategy, but also in the context of green business development, as mentioned for example by the UN Environment Programme - Finance Initiative (UNEP-FI) in the "Financial Sector Guide for the Convention on Biological Diversity" (UNEP-FI et al. 2021). The financial sector plays a critical role in the process of transforming financial institutions (banks, insurers, investors) and the whole financial system by adapting investment decisions and financial flows to mitigate nature loss and deliver nature benefits (UNEP-FI et al. 2021).

The 30x30 initiative unites 119 supporting countries with the goal of protecting at least 30 % of land and sea areas by 2030 to preserve enough of nature to prevent the collapse of natural systems (HAC for N&P 2024). The EU passed a law in August 2024 requiring all member states to restore/protect at least 20 % of their land and sea areas by 2030. Among other measures, national restoration plans must identify critical species/habitats and ecosystems, as well as timetables and funding plans. These will be assessed for adequacy by the EU and regularly monitored (EU Parliament and EU Council 2024).

There are numerous additional examples, such as the above-mentioned IUCN and Bonn Challenge (IUCN and BMUV 2011), which was signed by 150 heads of government at the 1992 Rio Earth Summit (UNEP 2022). Another initiative is REDD+ (Reducing Emissions from Deforestation and Forest Degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries), launched by

Papua New Guinea and Costa Rica at the United Nations Framework Convention on Climate Change (UNFCCC) negotiations in 2005. The Forest Carbon Partnership Facility (FCPF), comprising 47 developing countries and 17 financiers (governments, NGOs, private sector), provides support to participating developing countries through two funds. The Readiness Fund provides support for strategies and policies, measurement, reporting and verification systems, and environmental safeguards. The international Carbon Fund makes payments for the emissions that have demonstrably been reduced or sequestered by participants in their forestry sector. Both funds have collected USD 1.3 billion in contributions by 2023 (FCPF 2013, 2023).

1.1.3.2 Specific local and global laws

In addition to conservation efforts and international treaties, agreements, initiatives and political commitments, which are often not legally binding, there are also specific laws and binding regulations with specific objectives. The laws listed below generally aim to reduce deforestation and thereby protect forests as habitats for many different species. Examples for this approach include the European Union Deforestation Regulation (EUDR) (EU 2023), the United Kingdom Timber Regulation (UKTR) (UK Government 2013), the US Lacey Act (US Government 2008), Australia's Illegal Logging Prohibition Act (Australian Government 2021) and Japan's Clean Wood Act (Government of Japan 2016). They all regulate deforestation indirectly through imported and traded products and the resources they consume throughout the supply chain. It is a legal obligation for those importing wood-based products to demonstrate, through a process of due diligence, that the products in question have been produced from sustainable goods and resources that do not contribute to deforestation.

The EUDR was originally scheduled to replace the European Union Timber Regulation (EUTR) and come into force on 30 December 2024. However, in response to significant criticism of the unresolved implementing provisions, the European Commission has proposed to postpone the introduction by one year. A guidance document on the EUTR urges to "make sure that the specimens mentioned in the export permit corresponds to the specimen harvested" (European Commission 2018) in case authorities have doubts about the accuracy of documentation and endangered Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES)-listed species. In addition, the geographical origin of harvested timber should be declared by the exporting country. The document also states that expert groups may be consolidated and that the EU Commission may establish general restrictions on timber imports from certain countries or origins, based on the opinion of the scientific review group (European Commission 2018).

The main difference between the EUDR and the EUTR is the shift from targeting exclusively illegal logging to encompassing deforestation-free products and supply chains in general. In addition to wood-based products, the commodities most relevant to the EU's global impact on deforestation and forest degradation are oil palm (34 %), soy (33 %), cocoa (8 %), coffee (7 %), cattle (5 %) and rubber (3 %). Article 18(2)(c) of the EUDR states that checks on suspected operators and traders may include "any technical and scientific means adequate to

determine the species or the exact place where the relevant commodity or relevant product was produced, including anatomical, chemical or DNA analysis" (EU 2023).

Signed into law in 1900, the US Lacey Act (after John F. Lacey) is one of the earliest pieces of environmental legislation, originally designed to protect game and wild birds. It was amended in 2008 to also prohibit the import, export, transportation, sale or receipt of any plant that is taken or traded in violation of any law, including the laws of the United States, any state of the United States, and any foreign country (US Government 2008). The Lacey Act makes it illegal to place any wood-based product, including pulp and paper, containing protected species or wood from illegal logging, on US soil/markets. It can be deduced that a species which is not globally threatened with extinction but endangered in a particular region due to declining numbers of living trees, may still fall under the Lacey Act if its origin is covered by a domestic or foreign law. This applies to species listed in an appendix to CITES and those listed as endangered or threatened under the Endangered Species Act of 1973 (16 U.S.C. 1531). Indigenous species which are threatened with extinction and therefore protected under state law, as provided in section 8 204 (Prevention of illegal logging practices) of the Food, Conservation and Energy Act of 2008, also fall under the Lacey Act (US Government 2008). Importers are required to submit a declaration (including scientific name of species, quantity, value, country of origin, etc.) at the time of import, similar to the UKTR and EUDR.

The CITES is an international treaty that has been signed by 184 countries. The treaty restricts the trade of all species included in its appendices. Countries that have voluntarily joined CITES (parties to the convention) are legally bound to respect the convention and implement it into national law. Appendix I includes species threatened with extinction. The trade of Appendix I species is only allowed in exceptional circumstances. Appendix II lists species that are not yet threatened with extinction whose status may change if overexploitation is not stopped and trade is not regulated. Many timber species, such as ramin (*Gonystylus* spp.) and rosewood (*Dalbergia* spp.), are listed under CITES. This has a direct impact on their legal trade, as participating countries such as the USA – the US Lacey Act mentioned above includes a section on CITES protected species – are required to regulate and restrict the import and export of these endangered species. In addition to species-specific protection, CITES also addresses their origin. CITES Appendix II listings can include stricter regulations on timber from specific regions where the species is at greater risk. For example, ebony (*Diospyros* spp.) is not yet globally threatened, but the population in Madagascar is under extreme pressure. The same accounts for mahogany (*Swietenia* spp.) from the Neotropics and afzelia (*Afzelia* spp.) from African populations (CITES and UNEP 2024).

After leaving the EU, the UK implemented the UKTR (UK Government 2013), which mirrors the EUTR, in order to have their own timber regulation. It prohibits the placing of illegally harvested timber and timber products on the UK market and requires companies to exercise due diligence. It further compels UK companies to ensure that imported timber was legally harvested in its country of origin and sustainably sourced, particularly from endangered

species. It thus addresses wood provenance. It requires due diligence to verify both the legality and sustainability of the source (UK Government 2013).

The Australian Illegal Logging Prohibition Act, originally enacted in 2012 (Australian Government 2021), also prohibits the import of timber and timber products into Australia, which have been manufactured from wood derived from trees that were harvested illegally. Additionally, the processing of illegally harvested raw logs has also been prohibited. The act requires importers to undertake due diligence to ensure that timber imports, particularly from high-risk countries, are legally sourced and meet sustainability standards. This law is closely aligned with CITES and other laws such as the Lacey Act and the EUDR, and focuses on stopping illegal logging practices by prohibiting the import of timber harvested in violation of national or international laws. The penalty for importing illegally logged timber can be 5 years imprisonment (Australian Government 2021).

All of these laws and regulations place a responsibility on companies to exercise due diligence in verifying the legal origin of timber and to demonstrate that their supply chains are deforestation-free. These measures protect endangered species (CITES/IUCN Red List) and forests threatened by illegal deforestation, with enforcement mechanisms ranging from fines to trade bans and imprisonment, contributing to the global conservation efforts discussed in section 1.1.3.1.

1.1.4 Illegal logging and trade

Environmental crime is defined by Nellemann et al. (2016) as illegal activities harming the environment, such as exploitation of, damage to and trade or theft of natural resources. With an interest to benefit from the illegal revenue, actors from all scales are identified – individuals, companies, governments or other groups on a small or large scale up to transnational organised crime. Environmental crimes include illegal logging, accounting for USD 51–152 billion per year – by far the biggest share of all environmental crimes, ahead of illegal wildlife trade (USD 7–23 billion per year), illegal fishing (USD 11–24 billion per year), illegal mining (USD 12–48 billion per year) and illegal waste disposal (USD 10–12 billion per year). Illegal logging alone is very lucrative and is estimated to be the third or fourth largest black-market value after drug trafficking (USD 344 billion) counterfeit crimes (USD 288 billion) and human trafficking (USD 157 billion). It is estimated that 15–30 % of globally traded timber is illegally sourced. Although it is difficult to estimate the exact growth rate based on registered trade statistics, seizures and reported incidents, it is estimated that environmental crime as a whole is growing at a rate of at least 5–7 % per year, which is 2–3 times the growth rate of the global economy (GDP approx. 2.4–2.8 % in reference year 2016). Between 2014 and 2016, the total value of environmental crime increased by 26 % to an estimated USD 91–258 billion annually. It is estimated that governments lose significant proceeds through reduced tax revenues, estimated at USD 9–26 billion, due to the illegal exploitation of the environment. Forestry crimes, including logging and deforestation for agriculture, are estimated to be responsible for 50–90 % of all tropical deforestation and are probably the

greatest threat to life on Earth, as no other human activity currently kills as many species (Nellemann et al. 2016; Nellemann et al. 2020).

Small-scale forest crime by local residents and economically disadvantaged groups, who illegally harvest wood from the forest mainly for heating and cooking, appears to be common even in European countries such as Belgium, Bulgaria, France, Romania, Slovakia and Ukraine (WWF 2021). Once organised criminals identify an opportunity to utilise people's plights and basic needs to disguise their own illicit activities, numbers of illegally harvested wood rise far beyond personal use. So-called "white-collar" environmental crimes through international networks supposedly undermine legal markets and sometimes even influence stock markets. According to research conducted by non-governmental organisations (NGOs), the transnational organised crime even shift their illegal business from drugs and arms to environmental crime – or launder money through it – since natural resources do not attract as much attention (WWF 2021).

1.1.4.1 Deforestation and gold-mining in Latin America

Most of the following paragraph is based on numbers and statements from a report by the independent, international NGO called Global Initiative against Transnational Organised Crime (GI-TOC).

According to GI-TOC (2016), Colombian cocaine money was laundered with apparent ease and considerable profitability through illegal gold mining. And there is more to it: these exact mining operations have resulted in significant environmental degradation, with the clearance of over 1 680km² of forest between 2001 and 2013 in the Amazon rainforest in Colombia. This equates to more than 5 % of Colombia's total forest area cleared for mining alone (Alvarez-Berríos and Mitchell Aide 2015). A similar situation has occurred with regard to much of the rainforest in Guyana, Peru and Brazil, according to the initiative.

GI-TOC further reports how illegal gold miners would discharge about 30 tonnes of mercury a year into the Amazon river systems, poisoning wildlife, especially aquatic animals, drinking water and humans (e.g. permanent brain damage in new-born children). Besides deforestation and pollution, the human cost of illegal gold mining is said to be huge, as it would fuel human trafficking and slavery. In several communities, Yanomami Indians were discovered to work as inexpensive labour in the mines, with a slave number tattooed on their shoulders. Furthermore, child labour and sex trafficking are also often linked to illegal mining. According to GI-TOC, women and girls (some as young as 12) from all over Peru are trafficked to mining towns such as Delta 1 (Madre de Dios region), where police estimate 2 000 sex workers in illegal brothels. In 2012, one in ten individuals in Colombia are said to have been victims of trafficking, with 87 % of these cases occurring in mining regions. These numbers make Colombia the tragic world leader in human trafficking.

The strategy shift by organised crime groups from drugs to illegal mining has been so successful, that the financial value of the world's largest cocaine productions from Colombia and Peru has now been estimated to be surpassed by their illegal gold exports – with 80 % of

Colombia's gold mined illegally. In total, 68 tonnes of gold were extracted illegally from the Amazon and exported mainly to the USA and Switzerland (GI-TOC 2016).

1.1.4.2 Kosso-logging and trade in West-Africa

Illegal logging is easy in some countries, for example by forging permits or exceeding permitted limits, and the potential profits are high. Also the risk of being caught is low, due to extraordinary, sometimes foreign, bribes and the difficulty of detection with low enforcement capacity and corrupt structures (Beck-O'Brien et al. 2022; Nellemann et al. 2020). Illegal deforestation has far-reaching negative environmental impacts, as explained in 1.1.2.2 (Habitat loss).

If endangered species that are already threatened with extinction are systematically overexploited, the threat to and damage to functioning ecosystems intensifies. Unfortunately, in the case of timber, illegal and endangered species are often very valuable and fetch high prices on the black and white markets, which is why these species are often logged and traded. A notable example is the illegal trade of rosewoods – 33 different species of which six are endangered such as the CITES Annex II protected *Dalbergia cochinchinensis* and *Pterocarpus erinaceus* (Nellemann et al. 2020). Together they accounted for 35 % of the aggregated value of all illegal wildlife seizures between 2005 and 2014 according to a report by the United Nations Office on Drugs and Crime. The trade in rosewood consequently has the highest value and volume of any crime involving timber, ahead of all other trees, elephants, reptiles, pangolins, rhinos, big cats and parrots (UNODC 2016).

The species *Pterocarpus erinaceus* for instance, known by the trade name kosso, is favoured by Chinese businessmen (EIA 2017; Nellemann et al. 2020). In the rather new Chinese middle class is a huge demand in fine antique-style furniture as a status symbol. Between January 2015 and December 2016, West Africa (e.g. Nigeria, Gambia, Senegal, Ivory Coast, Ghana, Liberia, Mali) accounted for about 61 % of China's total rosewood imports. On average, China imported 764 000 m^2 of rosewood from the region each year, with an estimated annual value of USD 840 million. It seems that almost all the rosewood traded from Nigeria during this period was illegal: CITES documents for ~1.4 million illegal logs were signed by Nigeria's former Environment Minister (Mrs Mohammed, now UN Deputy Secretary-General) in her last days in office. This suggests that these multi-hundred-million-dollar trades were characterised by high levels of corruption and criminality. Much of the kosso timber was cut in an area controlled by the terrorist organisation Boko Haram. It can be reasonably deduced that the profit increase derived from rosewood logging would indirectly provide financial backing to terrorist activities.

Rosewood was also illegally harvested and traded from West Africa's largest national park, Gashaka Gumti, on the border with Cameroon. Government officials are reported to be collecting fines for illegal timber at checkpoints, but letting it pass (EIA 2017; Nellemann et al. 2020).

1.1.4.3 Charcoal and timber for Europe

A European case reported by Worth (2021) involves the purchase of 300 m^3 of illegally logged teak from valuable natural forests in Myanmar by the German Federal Agency for Agriculture and Food (BLE). The wood was mainly used in the manufacture of premium wooden decks for yachts. According to the WWF, these imports were in breach of EU regulations (Worth 2021).

Charcoal is another commodity often produced from illegally logged timber – and traded in large quantities to the EU and the USA for barbecuing. Charcoal-logging presumably contributes significantly to the deforestation in African countries (e.g. Nigeria, Namibia, South Africa, Somalia, Eritrea and Mali) (Nellemann et al. 2020; Zahnen et al. 2020). These African charcoal black markets are often associated with organised crime, such as militias and terrorist groups, which are indirectly funded by the purchase of illegally sourced charcoal, as they illegally tax up to 30 % of the value of the charcoal (Nellemann et al. 2020).

In addition to African countries of origin, much of the charcoal entering European markets comes from Ukraine, Cuba, Russia and Indonesia (Zahnen et al. 2020; Haag et al. 2020). As charcoal was not regulated by the EUTR at this time, it was imported into and traded within the EU without restrictions, with an estimated illegal volume (wood equivalent) of between 1.6 and 6 million m^3 annually (Haag et al. 2020). The Polish company Dancoal for instance was prosecuted in 2018 for processing huge amounts of illegal tropical timber from Namibia into charcoal. The product lost its FSC certificate and was no longer sold at Aldi and Lidl supermarkets (Zahnen et al. 2020).

1.1.4.4 Tropical rainforest wood in paper

The paper industry has also been accused of breaking forest laws and using illegal wood and endangered species. In 2015, the German BLE investigated allegations made by the WWF which found 20 % tropical rainforest wood in paper products from various stationary retailers and importers (WWF 2015). The accused companies would have neglected to claim the tropical timber content during the importation and subsequent sale of these products on the EU's internal market – a potential violation of the EUTR (WWF 2015).

In Indonesia, the number of such cases has been particularly high in recent years. One is the logging company PT Merbau Pelalawan Lestari, a mixed tropical hardwood supplier to Asia Pacific Resources International Limited (APRIL). PT Merbau Pelalawan Lestari, which is the second largest pulp and paper company in Indonesia, was ordered to pay the highest fine for environmental crimes in Indonesian history of USD 1.7 billion (Shah 2016). Between 2004 and 2006 they cleared 5 500 ha of protected trees in Riau province and another 1 873 ha illegally in other areas. In a separate case in 2016, APRIL's competitor Asia Pulp and Paper (APP) was found guilty of setting fires and fined USD 5 million by a Sumatran High Court (Shah 2016). A 2022 report also suspected pulpwood and palm oil producers of illegally deforesting large areas of the Papua region of Indonesia, including swamps, lowland forests and peatlands (Jong 2022).

These Indonesian cases are particularly sensitive, as deforestation in Indonesia has often been associated with the degradation of swamps and peatlands (Baffoni et al. 2017). Tree plantations for the pulp and paper industry are established on drained peatlands. The oxidation and burning of dry peat have made Indonesia one of the world's largest emitters of greenhouse gases. To turn this development over, it is crucial that the Indonesian paper industry reforests and rewets the drained peatlands and manages these areas sustainably. One method for achieving this is to begin incorporating local flora, such as gelam (*Malaleuca cajuputi*) – in harmony with the nature, history and culture of the indigenous people. Gelam is a fast-growing pioneer species that grows on wet peatlands and produces fibre suitable for pulp production. It could be seen as a high yielding, good quality and more environmentally sustainable alternative to timber from plantations on drained land. Indonesia's nature and enormous wildlife richness are particularly valuable to people, but also to the planet in terms of greenhouse gas emissions. As they are severely threatened by plantations (not only for pulp, but also for palm oil, amongst others, and the pulp industry, species and forest conservation in this region is of particular interest (Baffoni et al. 2017). Therefore, in publication III (section 3.3), teak pulps from Indonesia and other Asian provenances were successfully differentiated by analysing the remaining pulp extractives by GC-MS.

1.1.5 Effectiveness of implemented regulations and policies

Notwithstanding a variety of laws, policies and international regulations designed to combat international organised environmental crimes such as illegal logging, mining of conflict minerals and the trade of these products, significant gaps remain, according to GI-TOC. A number of countries where illegal activities occur have not yet ratified key legal instruments and/or the enforcement of these instruments is weak. Voluntary corporate principles lack sufficient impact, particularly in the artisanal and small-scale sectors, while certification schemes remain vague and implementation is limited. In addition, private sector initiatives often overlap and leave gaps (GI-TOC 2016).

Despite these criticisms, there are successful examples of policies designed to combat illicit activities in forest areas. A seemingly efficacious policy is the EUTR – although it has been the subject of considerable criticism. A fitness check of the EUTR shows that it may have led to 12–29 % fewer imports of illegally harvested timber into the EU and helped raise awareness and transparency in supply chains (European Commission 2021). However, productions in countries with known illegal logging problems, such as Ukraine, Myanmar or Belarus, have actually increased. This means that the EUTR has been more effective in keeping illegal wood-based products out of the EU than in preventing illegal logging in the countries of origin. Operators' biggest problems were validating the information provided by their supply chain. Inadequate due diligences were also difficult to challenge in court, resulting in hesitation to file lawsuits. National effective enforcement of implemented EUTR legislation varied, leading importers to avoid strict countries for importing timber products (European Commission 2021).

Another example is the Brazilian sector-wide Action Plan for the Prevention and Control of Deforestation in the legal Amazon (PPCDAm). Launched in 2004 to reduce deforestation and greenhouse gas emissions caused by the loss of native vegetation in the Amazon region, it already is considered a highly effective strategy as it has led to the highest reductions of deforestation in Amazonia. This success was mainly due to the central coordination of the PPCDAm and the implemented policies (Nikolakis and Innes 2020; West and Fearnside 2021).

1.1.5.1 A closer look at the Brazilian Amazon protection strategies

Some of the most effective strategies leading to the highest reductions in Amazon deforestation include: expansion of protected areas, rural credit restrictions, community blacklisting, improved satellite monitoring, and land tenure registration/regularisation (West and Fearnside 2021). The effectiveness of the policies delineated below is demonstrated through an examination of their implementation in the Brazilian legal Amazon as it contains almost 30 % of the world's rainforest and is therefore of particular importance. The strategies should be transferable to many other forests and countries, thereby serving as a model and orientation for their respective catalogues of measures.

The expansion of protected areas, whether they are indigenous peoples' lands, nature or wildlife reserves, also has a broadly positive effect on the larger picture, reducing deforestation rates, in some cases significantly. It is also worth taking a closer look at connections between the location and the pressure on protected areas. Areas close to roads and cities for instance are under greater pressure and therefore have a higher impact on reducing deforestation if they are protected. Another point to mention is the relocation of production, as people start producing at higher intensities in unprotected areas and exploit them even more when their formal production area is protected. Another negative side effect is rising prices due to scarcity, which further fuels greed, investment and production elsewhere (Pfaff et al. 2015).

The relatively new incentive-based tool of rural credit restrictions in the Amazon region, introduced by the Brazilian Central Bank (Resolution 3545) in 2008, was analysed by Assunção et al. (2020). The resolution required proof of legal land ownership and compliance with environmental laws to qualify for subsidised rural credit. This regulation effectively restricted access to one of Brazil's main agricultural support mechanisms by making credit conditional on meeting legal and environmental criteria. Assunção et al. (2020) found that the estimated deforested area in the restricted credit area was ~60 % lower than it would have been in the absence of Resolution 3545. These figures were based on a difference-in-difference approach using municipalities outside the restricted area as reference. The effect of the rural credit policy was greater in municipalities where livestock farming was the main source of economic income compared to crop-oriented farming (Assunção et al. 2020). As described above, the overall evolution of land use change can be strongly influenced by existing large-scale policy and financial incentives. On a smaller scale, Pailler (2018) showed how re-election years of mayors in Brazil can also negatively influence deforestation. She found an 8–10 % increase in deforestation rates associated with corruption of local politicians, combined with weak

institutions with limited authority over decentralised structures of local governance (Pailler 2018).

From the beginning of 2008, 43 municipalities responsible for 46 % of the total deforestation in the Amazon were given special treatment to reduce deforestation. These blacklisted communities faced stricter environmental monitoring with reduced access to deforestation permits, sanctions on illegally cleared land and restricted market access as part of the enforcement measures (West and Fearnside 2021). As shown by Arima et al. (2014) the deforestation rate fell by 60 % in the three years from 2009 to 2011 in the blacklisted municipalities, while the non-blacklisted comparison group only saw an average reduction of 47 %. The impact of the blacklisting policy was evident in various statistical tests, taking into account the 2008 stock market crash and crop prices, as well as the distance of the treated communities from roads and cities, which also affects the policy impact (Arima et al. 2014).

An improved near-real-time satellite-based monitoring system called DETER-B, developed by the Brazilian National Space Research Institute, has helped detect deforestation early and accurately, as Diniz et al. (2015) note. The system takes pictures almost every day, facilitating constant monitoring and early detection of deforestation and thus allowing authorities to react quickly and prevent further damage. Other benefits of the new system include the detection of deforestation areas smaller than 25 ha, which is important as they have been decreasing in size. Deforestation areas between 25 and 100 ha can be detected with greater precision and information density on the actual type and stage of deforestation. In a test step of the methodology, burnt areas, clear-cuts, conventional selective logging, mining and others are differentiated, while data on the intensity of each is also collected. Through these refinements, the technique allows the detection of deforestation patterns (Diniz et al. 2015).

Another key factor in the success of the Brazilian policy mix, according to Azevedo et al. (2017), is the regulation of land ownership. A comprehensive, nationwide land registry has been established, as deforestation could not be formally attributed to specific landowners – who were often unknown. Between 2002 and 2009, around 69 % of all deforestation took place on land where the ownership and boundaries were not clear. Landowners were therefore asked to map their land, georeference their boundaries and provide information on the state of deforestation on their land based on satellite imagery. This made it possible to monitor whether landowners in the Amazon were complying with the Forest Code's requirement to maintain 80 % of the natural state, to restore it if not, and to identify offending landowners and hold them accountable if they were not. The main incentives to join the registry were financially. The first incentive for compliance was the avoidance of fines for non-registration. However, the risk of being identified and fined was relatively low due to the lack of robust enforcement. Secondly, access to rural credit was contingent upon registration. Thirdly, slaughterhouses were encouraged to refrain from purchasing cattle from farms that violated labour or environmental laws, thereby exerting market pressure. Azevedo et al. (2017) reported clearly positive effects for small properties, resulting in significantly lower deforestation in each of the years observed (2009 to 2011), but the effects for medium and

large properties were mixed. To comply with the requirements, forest owners are supposed to restore deforested land, but 76 % said in a questionnaire that they would only do so if the financial incentive was again high enough, either through government fines that are higher than the income they could earn from the deforested land (e.g. through cattle or through market incentives). Overall, land tenure regularisation works and is an important part of the policy mix, but it needs the right amount of financial pressure and benefits from public and private actors, and greater prosecutorial power by the government, to have a higher positive impact on deforestation than in the Brazilian example (Azevedo et al. 2017).

There were two important moratoria in Brazil. The first was the soy moratorium in 2006, in which vegetable oil producers and grain exporters agreed not to buy or sell soybeans from land that's been cleared from the Amazon rainforest. The moratorium had a number of complex effects. Since the price of pasture land suitable for soya cultivation in regions not restricted by the moratorium rose, cattle ranchers made profits by selling their land and moving their cattle to other Amazonian land – only to deforest it (West and Fearnside 2021; Macedo et al. 2012). Kastens et al. (2017) found that pre-soy moratorium deforestation rates based on satellite imagery were 2–5 times higher than post-soy moratorium rates, which contrasts with Svahn and Brunner (2018) who found no discontinuity in regions affected by the soy moratorium around the Amazon biome boundary before and after implementation, meaning that the moratorium had no visible effect on deforestation reduction. As market price declines occurred in parallel with the introduction of the policy by the government and industry, Macedo et al. (2012) are also unsure of the actual impact of the soy policy on deforestation reduction. The second moratorium on actors in the beef supply chain from slaughterhouses to meatpacking companies was to only purchase beef from deforestation-free and environmental and labour-friendly sources. It had no relevant impact on deforestation in the vicinity of slaughterhouses in the Amazon biome in the states of Mato Grosso and Pará between 2009 and 2014 (Alix-Garcia and Gibbs 2017).

GI-TOC recommends that countries involved in the export and import of illegally sourced materials, including timber, improve compliance with international laws to ensure accountability. Exporting countries could improve the situation by increasing funding and simplifying the formalisation process. Importing countries should strengthen the enforcement of environmental standards on companies and their subcontractors by requiring them to comply with both international and national regulations. Companies themselves should improve their structures to map supply chains and ensure that materials are sourced in a legal, sustainable and human rights-respecting manner, in essence in line with environmental, social, and corporate governance (ESG) principles (GI-TOC 2016).

1.2 Identification of wood

In order to achieve the above-mentioned global future goals and political commitments (1.1.3), it is essential to provide comprehensive support for the selected strategies, which entail the utilisation of international agreements and cooperation, followed by the

implementation of national laws and regulations. Illegal logging and international criminal organisations need to be tackled through strong enforcement. Hence, capacity and power, as well as expertise and coordination, need to be strengthened and pooled. Increased funding for enforcement, as well as harmonisation of international laws with stronger penalties for violations, could help to facilitate prosecution and cooperation. Cooperation and mutual support through the sharing of information, knowledge and law enforcement is essential, as criminal organisations operate internationally. Identification techniques play an important role in verifying the due diligence of market participants and law enforcement. In order to implement these laws effectively and to facilitate law enforcement, as well as to gain knowledge and monitor market developments (e.g. trends in species used to anticipate which species may be threatened in the near future), science must support and provide identification methods.

1.2.1 Analytical methods

The following section provides an overview of the analytical methods available and used by scientists for the identification of solid wood and wood-based products such as particle board, and pulp and paper.

Firstly, methods for solid wood taxon identification (taxonID) and origin identification (originID) are reviewed. Secondly, wood-based products are looked at, again divided into methods suitable for taxonID and originID. Figure 7 illustrates this overview of suitable methods depending on the type of wood product and the question of taxonID or originID.

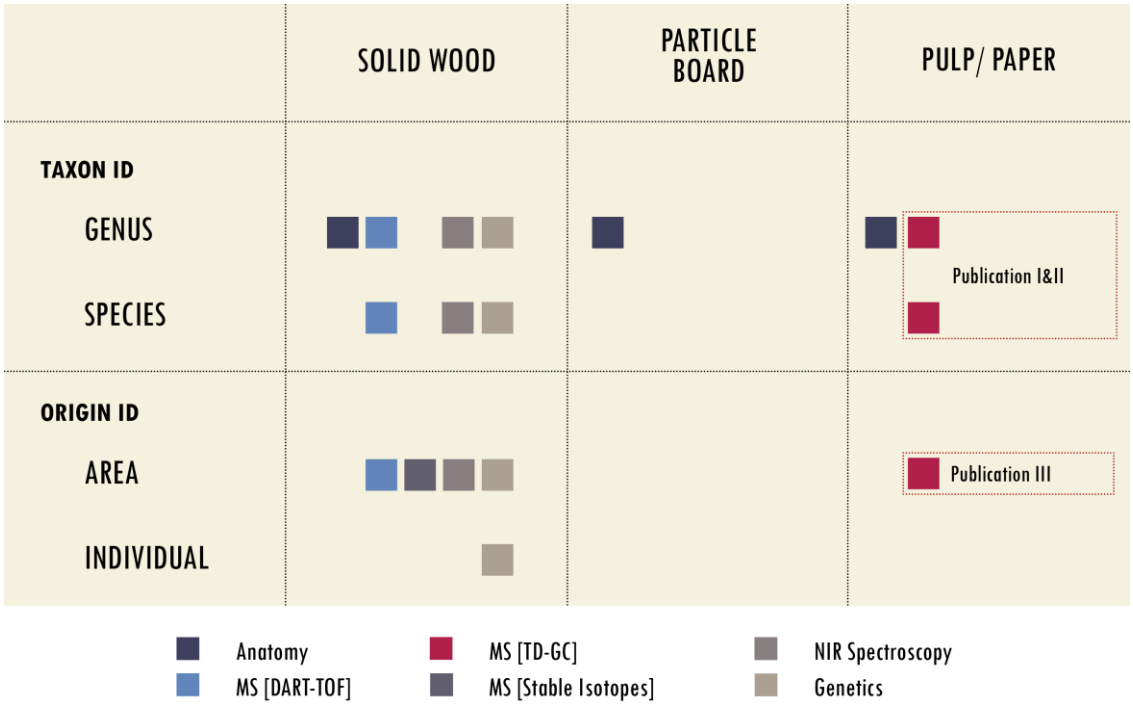


Figure 7: Overview of taxonID and originID methods suitable for different wood products. For pulp and paper, the chemotaxonomic TD-GC-MS method was developed, tested and validated in Pub I & Pub II and applied to the originID question in Pub III. Figure adapted from © Flaig et al. (2023) and Flaig et al. (2025), published by De Gruyter, licensed under CC BY 4.0.

1.2.1.1 Solid wood taxonID

Solid wood is the easiest of all wood products to identify. Almost all available methods are suitable for solid wood identification in general. Techniques particularly addressing the taxonID question of solid wood can be sorted into three groups: Anatomy, chemistry and genetics.

1.2.1.1.1 Genus level

Anatomical methods use macroscopic or microscopic examination of characteristic anatomical wood features in transverse, tangential and radial histological sections to identify wood genera in solid wood (Silva et al. 2021). Various databases with vast library reference collections are available. Two examples of databases are InsideWood (2004 onwards) with over 10 000 descriptions of modern hardwoods and softwoods, and over 70 000 images by Wheeler (2011) and the DELTA (Description Language for Taxonomy) database "Commercial timbers" by Richter and Dallwitz (2000 onwards).

The chemical approaches established for solid wood taxonID are spectrometry and spectroscopy. These chemical methods will also be referred to as chemotaxonomic in the course of this work, as chemotaxonomy generally means taxonomic differentiation or grouping based on chemical characteristics – finding similarities or dissimilarities within different organisms based on shared/non-shared chemical composition. These analysed specific chemical compositions of components present in solid wood are mainly secondary metabolites. Since 2012, direct analysis in real time - time of flight - mass spectrometry (DART-TOF-MS), a type of chemical fingerprinting, has been developed and is becoming the most present method, often used in law enforcement (Cody et al. 2012; Espinoza et al. 2015; Lancaster and Espinoza 2012). A new approach of creating an inter-laboratory database with shared references, accessible to colleagues and laboratories in different locations, has been successful (Lancaster et al. 2024) This is very valuable as joined forces can act faster and stronger against illegal logging. The advantages of DART-TOF-MS are rapid results and smaller/portable instruments. Another suitable method is near-infrared spectroscopy (NIRS), which measures the absorption rates of functional groups. There are some disadvantages to this method, including dependence on fibre orientation and surface roughness (which can be sanded and cleaned) and it is also limited by the minimum amounts required for detection (Schwanninger et al. 2011). However, there are many advantages such as speed and non-destructive analysis with minimal sample preparation. Examples of successful developments using near-infrared spectroscopy (NIRS) for CITES-relevant taxa are the analysis of *Dalbergia* spp. by Snel et al. (2018) and *Swietenia* spp. by Braga et al. (2011).

Genetic analysis of deoxyribonucleic acid (DNA) is another area where various successful technical approaches have been used to identify taxa in solid wood (Jiao et al. 2020). In terms of law enforcement and CITES protection, it would be practical to have a simple, easy-to-use and comparatively inexpensive technology that's affordable and accessible, especially to countries that do not have a lot of financial and technological resources, but have valuable

tropical timber species that are in urgent need of protection. A recent example of a suitable technology, although long known, that still ticks all these boxes for modern needs, is a DNA macroarray as presented by Bogun et al. (2024) for relevant tropical species and their look-alikes. In general, DNA extraction and analysis are easier and more successful when the wood is fresh and untreated. As demonstrated by Rachmayanti et al. (2009), the identification of old wood that has been stored for a long time and even kiln-dried wood using DNA barcoding is more challenging, yet still feasible.

1.2.1.1.2 Species level

Of the solid wood identification techniques listed above that are suitable at the genus level, the chemical and genetic methods are also applicable at the species level.

Chemical species identification of solid wood also works well with DART-TOF-MS. Cody et al. (2012) for instance, managed to find significant differences between the two *Quercus* species *Q. alba* and *Q. rubra* in 11 peaks of the negative mass spectra. Deklerck et al. (2019) successfully established an automated DART-TOF-MS based protocol for species identification, which was tested on different tropical wood species, using standardised pre-processing parameters and a standardised classification via the random forest algorithm. The chemotaxonomic differentiation of many different *Dalbergia* species including the CITES Annex I protected *D. nigra* was achieved by Espinoza et al. (2015). Classical NIRS technology using light absorption has been successful in many species identification (speciesID) cases, for example even with a portable handheld device by Snel et al. (2018), which also differentiated *Dalbergia nigra* from other high-value *Dalbergia* species. Lazarescu et al. (2017) used NIRS to identify *Tsuga heterophylla* and *Abies amabilis* with 94 % and 86 % accuracy, respectively. Light scattering, on the other hand, is dominant over absorption in solid wood surfaces and has great potential for identification as well (Ma et al. 2019). This has been used to identify different soft and hardwood species with 94 % accuracy by combining the NIR-SRS (spatially resolved spectroscopy) with hyperspectral imaging (HSI) (Ma et al. 2019).

Genetic identification of solid wood species has been achieved many times. It is very accurate and the most widely used and successful method for this purpose. For example, Yu et al. (2016) found a DNA barcode region (trnH-psbA) suitable for speciesID of *Dalbergia odorifera* and *Dalbergia tonkinensis*. This DNA sequence is relatively short, making it easier to amplify even in the case of highly degraded DNA from treated wood or wood products (Yu et al. 2016). The genus *Dalbergia* is of particular importance as one species, *D. nigra*, is protected under CITES Appendix I and all others are protected under Appendix II (CITES and UNEP 2024). Asif and Channon (2005) were also able to amplify large PCR fragments from two chloroplast and one mitochondrial region in all samples tested and successfully identified the CITES protected species *Gonystylus bancanus* in both fresh and dry wood samples.

1.2.1.2 Solid wood originID

The originID of solid wood is important, given that in numerous instances, populations from specific regions are more endangered than others. This has led to disparate levels of protection and trade capabilities, as outlined above.

1.2.1.2.1 Area level

Geographical provenance adds stable isotope analysis to the available chemical methods for solid wood products. Stable carbon isotopes ($\delta^{13}C$) have been successfully used to identify different provenances of two Pinus species from the southwestern USA with a precision of 114–302 km (Kagawa and Leavitt 2010). Analysis of $\delta^{18}O$, $\delta^{2}H$, $\delta^{13}C$, $\delta^{15}N$, and $\delta^{34}S$ by elemental analysis isotope ratio mass spectrometry (EA-IRMS) of freshly cut *Aucoumea klaineana* samples from two sites 240 km apart in Gabon showed significant differences in the resulting stable isotope ratios (Watkinson et al. 2022). Another recent study of the two economically important tropical timbers azobé and tali from Gabon, Republic of Congo and Cameroon showed good country determination but little potential for sub-country regional application due to high local isotope variations. They also found species and site interferences, so that stable isotope analysis may be improved by species-specific isotope databases, and even species identification may be possible (Boeschoten et al. 2023).

DART-TOF analysis works not only for taxon, but also for origin. Deklerck et al. (2020) were able to discriminate between solid wood samples of *Terminalia superba* originating from the Democratic Republic of Congo and the Ivory Coast using DART-TOF-MS spectra along a pith to bark gradient combined with the random forest classification algorithm.

Infrared spectroscopy methods can also help to differentiate wood provenance. Even wood samples of unknown species can be separated from planted to natural origin with some limitations. In an attempt by Ramalho et al. (2018), growing environments such as native forests and plantations were differentiated using NIRS. Traoré et al. (2018) successfully used attenuated total reflectance-Fourier transform infrared (FTIR-ATR) to analyse *Pinus* sylvestris and *Pinus nigra* samples from two regions in Spain. A discriminant analysis test of polysaccharide bands in the absorbance range of wave numbers from 400 to 4 000 cm^{-1} was found to be useful not only for species identification, but also for provenance determination (Traoré et al. 2018).

Schroeder et al. (2016) genetically analysed over 1 000 white oak individuals of 13 different species from three continents: Asia, Europe and North America. Five indels and one single nucleotide polymorphism (SNP) – both genetic mutations – were found to be specific to the continents, meaning that each of the white oak species tested can be assigned to a continent of origin based on these genetic characteristics (Schroeder et al. 2016).

1.2.1.2.2 Individual level

At the individual level, only genetic methods are applicable. To prevent illegal logs from entering the timber supply chain, the critical point is between the logging concession and the

timber mill, where many unsustainably harvested logs enter the supply chain with falsified documentation (Lowe et al. 2010). Lowe et al. (2010) developed a genetic fingerprinting method for individual logs of merbau (*Intsia palembanica*) to verify certification and detect illegal substitution of individual logs along the supply chain of this high-value Indonesian timber.

1.2.1.3 Particleboard and other products

A study conducted by Sieburg-Rockel and Koch (2020) shows that traditional wood anatomy can be used for genus-level identification of wood in particleboard, although sample preparation is laborious due to small particle sizes and the manual, time-consuming cutting of microscopic sections of three anatomical planes per sample. The authors analysed 57 particleboards and identified 68 different taxa from Central Europe, tropical and subtropical Asia, temperate Asia, tropical Africa, tropical South America and some plantation-grown taxa worldwide. Two taxa, *Diospyros* and *Swietenia*, were protected under CITES Appendix II (Sieburg-Rockel and Koch 2020). Genetic methods cannot be applied as a standard to the analysis of particleboard due to the product composition of many different species in some cases combined with small particle sizes and heat-induced DNA degradation during the manufacturing process.

Charcoal is another very environmentally critical wood product, produced in large quantities and often from illegal sources, as described in section 1.1.4.3. There are two methods for identifying wood species used for charcoal production. The anatomical technique uses either 3D reflectance light microscopy (3D-RLM) or field emission scanning electron microscopy (FESEM) to examine the anatomical characteristics of wood. 3D-RLM can be used on fresh fracture planes without any further sample preparation. High contrast images are produced using polarised light, which facilitates analysis (Zemke et al. 2020).

NIRS is also suitable for charcoal identification. In a study by Nisgoski et al. (2021), charcoal from different eucalyptus species was distinguished from charcoal from different cerrado species. The best results, in some cases even to the species level, were obtained in transverse section at wave numbers between 4 000 and 6 000 cm^{-1} (Nisgoski et al. 2021).

Suitable identification methods for other wood products such as medium density fibreboard (MDF) (Koch et al. 2017), oriented strand board (OSB) and plywood mainly depend on the particle size that can be sampled and the production process (Schmitz et al. 2020). In general, the larger the particles and the milder the treatment during production, the more likely it is that all solid wood practices such as anatomy, genetics and NIRS will come into play. The smaller the particles, and the more intense the treatments, the more likely it is that only anatomy can be applied (Schmitz et al. 2020).

1.2.1.4 Pulp and paper – taxon and originID

For pulp and paper products the anatomical method of genus identification is long established (Ilvessalo-Pfäffli 1995) and is currently the only method in use. It uses characteristic

anatomical features of wood vessels that are visually examined under a light microscope (Helmling et al. 2016). The technique is practical, relatively quick, inexpensive and in most cases very accurate to genus level. It has a growing number of reference samples, such as those in the Atlas of Vessel Elements, against which unknown samples can be compared (Helmling et al. 2018). Microscopic images of fibrous material have recently been stored in a digitised database on which the wood taxonID has been successfully tested for automation by image recognition and deep learning (Nieradzik et al. 2024).

To date, the only other technical possibility for wood taxonID in pulp and paper is chemotaxonomic fingerprinting based on residues of wood extractives remaining in pulp and paper after the pulping process. The extractives are analysed by thermal desorption - gas chromatography - mass spectrometry (TD-GC-MS) and compared with reference chromatograms in a database approach searching for marker peaks and other similarity features, as developed and presented in Pub I, section 3.1 of this work (Flaig et al. 2023). The method was developed to potentially support the anatomical method as an independent alternative in difficult cases such as heavily beaten pulp samples with unrecognisable anatomical features or taxa with particularly high anatomical similarity. The method was tested on 38 selected genera and species in a comprehensive blind test against the traditional anatomical method in Pub II of this dissertation (Flaig et al. 2024). The test results showed that the chemotaxonomic method achieved a similarly good overall average hit rate of 86 %. Although some specific species were included in the blind test and successfully identified, they were always the only ones of their genus, requiring a further test with different species of the same genus to verify the identification at species level and to draw final conclusions on the possible taxonID depth. It is likely that some species are distinct while others are chemotoaxonomically indistinguishable. In Pub III, the method was also applied to pulp samples of ramin (*Gonystylus* spp.), red meranti (*Rubroshorea* spp.) and teak (T*ectona grandis*) originating from different geographical areas (Flaig et al. 2025). The results of the provenance analysis by principal component analysis (PCA) were very positive with clear differentiations in many cases.

2 Structure of the dissertation

This thesis presents the development of a chemotaxonomic method based on chemical fingerprinting using pulp extracts for the identification of mainly tropical hardwood species in pulp. It further describes the creation of a database, the method validation and its application to a second problem. The method-development included the selection of a suitable mill, an appropriate extraction process and solvent to obtain the characteristic wood extractives remaining in the paper, and the development of a gas chromatographic measurement method. A suitable database was then set up to allow comparison of unknown pure or mixed samples with the reference fingerprint extracts stored in the database.

Figure 7 shows the chronological sequence of the publications and thus the structure of the main part of this dissertation. The method-development, including sample preparation, data pre-treatment, database storage and querying, is described in detail in Pub I. Subsequently, the new chemotaxonomic method was applied in a blind test. 38 pulps were externally mixed in 15 test sheets of unknown composition. The blind samples were analysed anatomically by three independent institutes (TI, TUDa, ISEGA) and chemotaxonomically by the author of this work at the University of Hamburg who employed the novel method. The chemotaxonomic results were directly compared with the results of the well-established anatomical identification method. The approach was then validated for instrument independence by determining the repeatability of the results on different GC-MS systems. In the third publication, it was applied to the problem of tracing the origin of wood in paper.

In the following main part of this thesis (section 3), the three publications (Pub I–III) by Flaig et al. (2023), Flaig et al. (2024) and Flaig et al. (2025) are reprinted with the permission of the journal Holzforschung and the publisher De Gruyter. The contents of the original Pub I and Pub II have been updated with the latest knowledge on the classification of the wood genera *Anthoshorea*, *Richetia*, *Rubroshorea*, and *Shorea*. The style and format of the original papers have been adapted to align with the format of this document, ensuring consistency and aesthetic coherence. Furthermore, the figures and tables have been renumbered, and the references have been merged and relocated to the end of the dissertation. The results are discussed and evaluated in section 4, after which a conclusion is presented in section 5 and a future outlook is given in section 6.

3 Publications

3.1 Publication I: Identification of tropical wood species in paper: a new chemotaxonomic method based on extractives

Original Article

Max L. Flaig, Jens Berger, Philip Wenig, Andrea Olbrich and Bodo Saake*

Identification of tropical wood species in paper – a new chemotaxonomic method based on extractives

http://doi.org/10.1515/hf-2023-0048
Received May 17, 2023; accepted Oct 05, 2023;
published online November 7, 2023

*__Corresponding author: Bodo Saake,__ Institute of Wood Science, Chemical Wood Technology, University of Hamburg, Haidkrugsweg 1, 22885 Barsbüttel-Willinghusen, Germany, E-mail: bodo.saake@uni-hamburg.de
Max L. Flaig and Jens Berger, Institute of Wood Science, Chemical Wood Technology, University of Hamburg, Haidkrugsweg 1, 22885 Barsbüttel-Willinghusen, Germany https://orcid.org/0009-0000-0585-1734 (Flaig, M.L.)
Philip Wenig, Lablicate GmbH, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany
Andrea Olbrich, Johann Heinrich von Thünen Institute (TI), Institute of Wood Research, Leuschnerstraße 91, 21031 Hamburg, Germany

Abstract: The European Deforestation Regulation 2023/1115 (EUDR) prohibits trading of wood and wood products obtained from illegal logging on the EU market (EU 2023). While the identification of solid wood via anatomy, chemistry and genetics has already been established, there is a lack of identification methods for pulp and paper that complement anatomy. This publication presents a newly developed chemotaxonomic method for identifying mixed tropical hardwood (MTH) species in pulp and paper products based on their extractives analysed with thermal desorption - gas chromatography - mass spectrometry (TD-GC-MS). The measured data was processed and compared to identify marker substances and was then merged into a fingerprint database for identifying MTH species in paper of unknown composition. As database references, fully bleached kraft pulps were produced from 38 anatomically identified wood samples and then cryo-ball milled and extracted successively with n-hexane and acetone. The results show that the remaining wood extractives generated from bleached pulps are specific enough to find chemical relevant marker substances to detect MTH species. As chemical composition and anatomy are independent characteristics of wood, this paper makes a completely independent method available, which potentially improves the screening for CITES protected species.

Keywords: CITES; database; extractives; GC-MS; mixed tropical hardwood; pulp and paper.

1 Introduction

Most of the time, pulp and paper are products obtained from the renewable raw material wood. Since quality requirements for pulp are simply focused on the individual fibre and not on the wood structure itself, inferior quality and mixed wood can be involved. Thus, this can be connected to the use of wood from clear-cutting of natural forests, such as those in Indonesia for example, for pulp production from so called mixed tropical hardwood (MTH) (Uryu et al. 2008). There is a suspicion for the whole of South East Asia regarding tropical woods being illegally cleared e.g. to make room for palm oil plantations while their wood is used for pulp production (Hirschberger et al. 2010). Over a period of time the annual pulp and paper production in Southeast Asia and China was steadily increasing, reaching a record high in 2018. Since then, production volume has stagnated at a level close to the all-time high (FAO 2021). Even in industrialized countries such as Canada or Australia (e.g. Tasmania) virgin forests are still being clear-cut for pulp and paper production (Hirschberger et al. 2010).

In order to preserve biodiversity and to protect the rainforest as an effective ecosystem it is necessary to substantiate the suspicion named above by identifying tropical timber species in pulp and paper within global trade specified by the CITES as proof. In relation to that, the EUTR from 03 March 2013 prohibited the imports of illegally harvested timber and timber products on EU internal markets (EU 2010) which also applied to pulp and paper. This regulation was updated by the EUDR on 31 of May 2023 (EU 2023) placing due diligence onto market participants to take appropriate measures and comply with the standards required. Considering the fact that manufactured paper production represents a large consumption of the raw material wood (Windhagen et al. 2019), determining its composition is of high importance. Nevertheless, up to now there has been a lack of available reference samples of relevant species and analysis options for checking the manufacturer's data, especially for the detection of tropical woods used in the production of paper. In reaction to that, the goal is to methodically support the enforcement of the EUDR (EU 2023).

In challenging circumstances, such as those involving highly beaten pulp or cases where the established anatomical method may yield inconclusive results, the chemotaxonomic approach can serve as a valuable alternative. This method provides additional information and can serve as a second independent method, strengthening the accuracy and reliability of wood species identification. However, the authors acknowledge that the current state of the chemotaxonomic method is more time-consuming and requires extensive laboratory work compared with the frequently and traditionally used anatomical method. Additionally, the database of reference samples of the chemotaxonomic approach consists of only 38 entries at present, which limits its performance. Nevertheless, further enlarging this database can provide a promising outlook.

To forensically identify **solid** wood various methods can apply (Figure 8). These can be grouped into anatomical, chemical (compound analysing mass spectrometry (MS), composition of isotopes of elements analysing MS (stable isotopes), near-infrared spectroscopy (NIRS)) and

Flaig, M.L. et al. (2023). *Holzforschung* 77: 860–878,

genetic techniques (DNA barcoding, DNA profiling, population genetics and phylogeography) (Low et al. 2022; Lowe and Cross 2011; Schmitz et al. 2020). Depending on the focus of the research question being asked, either one of these can represent the most suitable technique for identification (Braga et al. 2020; Schmitz et al. 2019). For genus identification anatomy, MS, NIRS and genetics can be used. MS, NIRS and DNA barcoding are suitable techniques to differentiate species. Additionally, the differentiation of geographic origin can be achieved through a range of analytical techniques, including MS (Deklerck et al. 2020; Espinoza et al. 2014), stable isotopes analysis, NIRS (Silva et al. 2018) and genetic methods. These methods besides stable isotopes are currently only useful for identifying geographic origin when fixed options are available for comparison, such as distinguishing between specific known locations (e.g., Location A vs. Location B vs. Location C). However, they may encounter challenges in cases where samples are completely blind or when attempting to differentiate over a large spatial area without prior reference points. Lastly, Individual tree identification is only possible using genetic methods (Degen et al. 2017).

Wood anatomy works very well down to the family and genus level (Braga et al. 2020) and includes a high number of available database references (Wheeler 2011; InsideWood 2004 onwards; Richter and Dallwitz 2000 onwards). In contrast, stable isotopes cannot be used for species identification. They can only be used for origin determination of mainly solid woods (also semi-processed wood products and wood-based panels), successively shown by Watkinson et al. (2020) for different *Quercus* spp. origins throughout the United States. NIRS is suitable for taxonomic species identification of solid wood (Tsuchikawa and Kobori 2015). Snel et al. (2018) even identified seven CITES listed *Dalbergia* species with an accuracy of 90 % while the available NIRS reference data is still rather limited (Low et al. 2022). DNA barcoding, as another example, is well suited for genus and species and even individual tree identification as well as for distinguishing between origins (Ng et al. 2017). There are many available solid wood DNA reference samples collected in databases (Low et al. 2022).

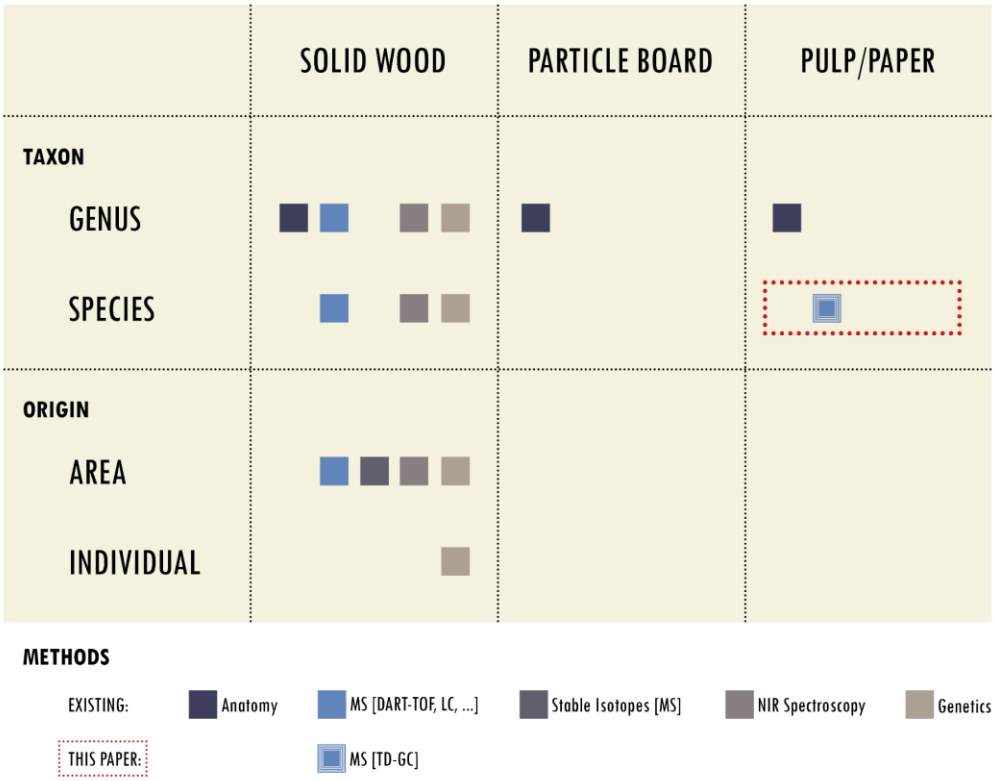


Figure 8: Comparison of different established identification methods related to specific wood products, taxon and origin.

Compared to **solid** wood, the identification of wood products such as particle board and paper entails more challenges. Depending on the type, Figure 8 shows that a range of methods can be suitable. Regardless, anatomy is the only technique useful for particle board but requests many objects of study per board which make sample preparation time consuming and laborious (Sieburg-Rockel and Koch 2020). To the authors' knowledge it is not possible to extract high quality DNA from particle boards but from processed wood such as dried and oven heated sawn wood and glued wood like window frames and other products (Rachmayanti et al. 2009; Asif and Channon 2005). In this process, the minimal test specimen size for DNA extraction needs to be 1 cm^3 (Schmitz et al. 2019). To analyse charcoal, specific anatomy and NIRS can be used (Zemke et al. 2020; Nisgoski et al. 2021). In contrast to genus/species identification of solid wood, bleached pulps and papers present an even bigger challenge. Throughout the entire manufacturing process – from pulping to the various bleaching stages – many wood constituents are removed and the original DNA is destroyed by the chemicals. NIRS is also influenced by surface roughness, fibre orientation, sample thickness and, most importantly, the number of minor compounds for analysis is limited to a minimum mass fraction of 0.1–0.5 % (Schwanninger et al. 2011). This explains why anatomy was the singular method possible for pulp, paper and fibreboard up to now. Challenging this view, this publication shows how thermal desorption - gas chromatography - mass spectrometry (TD-GC-MS) can provide an even more detailed method for species

identification while also providing solutions for some of the difficulties of the anatomical method.

Although anatomical identification via morphological characteristics of wood vessels has already been established (Helmling et al. 2018; Helmling et al. 2016; Ilvessalo-Pfäffli 1995), this method reaches a limit when fibres are modified by refining, a common process in papermaking, which destroys the remaining vessel elements (Helmling et al. 2018). In addition to that, there are wood genera that have only a few distinctive anatomical structural features and can be confused with other genera – even with non-closely related woods (Gasson 2011). A particularly problematic example is the genus Gonystylus (ramin), which is strictly protected under CITES Appendix II (CITES and UNEP 2024). In order to have another independent method complementary to the anatomical method for pulp and paper, the chemotaxonomic method was developed. A major advantage is that this method is not impaired by the mechanical transformation of anatomical features, such as the refining of fibres.

Overall, chemotaxonomy or fingerprint analysis can be used to classify and identify plants based on their secondary plant compounds (Alaerts et al. 2014). The potential of chemotaxonomy in CITES enforcement using different MS techniques was demonstrated many times before. For example, Kite et al. (2010), analysed solid heartwood of various *Dalbergia* species successfully using LC-MS. Further, Lancaster and Espinoza (2012) showed how DART-TOF-MS coupled with the multivariate data analysis methods PCA and LDL can chemotaxonomically distinguish different solid wood *Dalbergia* species. Cody et al. (2012) classified White Oak and Northern Red Oak with a success rate of 100 % while this rate can vary when used on closely related species (Deklerck et al. 2019). Moreover, Chemotaxonomy using DART-TOF-MS can be used to distinguish wild from cultivated solid *Aquilaria* spp. (Espinoza et al. 2014). This shows that the DART-TOF-MS method is highly beneficial for solid wood analysis (Dormontt et al. 2015; Low et al. 2022). However, in relation to genetics and anatomy it still has a relatively small albeit growing body of collected reference spectra. However, what must be considered is that while previous studies concentrated on solid wood, the chemotaxonomic studies of this paper were performed on the **pulp and paper** deriving from tropical wood species. A high degree of care is applied during the industrial pulping and bleaching process in order to remove wood extractives as far as possible, without risking poorer paper quality, still ensuring a faultless production. Because this results in low extractive content, it was not clear whether wood species in pulp could be identified by chemotaxonomic methods at the start of this study. Also, the composition of the pulp extractives differs from that of the original solid wood extractives. For this reason, information on chemotaxonomically relevant substances in solid woods, as described by Hegnauer (1986), cannot simply be relied upon for pulp. Nevertheless, this investigation showed that extractives suitable for chemotaxonomic identification can be extracted from pulp in sufficient quantities by n-hexane. In the end, TD-GC-MS analysis of these reference extracts was performed, specific marker substances were found, and a fingerprint database was built up using specialized software.

2 Materials and methods

For the method development, a purchased industrial MTH pulp from Sumatra served as the initial material for the development of suitable grinding and extraction parameters for MTH pulps. Subsequently, 38 single variety solid wood samples of approximately 2 kg each (Table 2) were obtained from different, mostly commercial sources (no documented origin) and generated the reference database of the chemotaxonomic study. Genera and species of these solid wood samples were morphologically identified by wood anatomists with the assistance of the Xylothek at the Thünen Institute, Hamburg, Germany. Then, each individual reference wood was pulped and bleached to produce respective kraft pulps for further analysis within the database created.

2.1 Production of the reference pulps

To start with, the solid wood samples were manually cut into 3 x 3 cm wood chips with a self-built semi-automatic chipping machine. Afterwards, they were steamed with saturated steam at 1 bar prior to pulping for 30 min. The kraft pulping was carried out in a program-controlled 7-liter M/K digester (M/K Systems INC., Williamstown, USA) with liquor circulation. Depending on the quantity of the raw material, between 550–1 000 g dry matter wood chips were used per cooking. The liquor to wood ratio was 4:1 (v/w). Total NaOH chemical usage was 22–25 % with a sulfidity (Na_2S) of 35 %. The heating time (t) was 90 min followed by 120 min at a maximum temperature (T_{max}) of 165 °C. The obtained kraft pulps were slot screened to a maximum of 0.15 mm and subsequently bleached in five bleaching stages, adapted from the industrial bleaching practice of the Mercer Stendal pulp mill. The target brightness was 90 % ISO. The oxygen stage (O) was carried out with 12 % consistency (c), 2.5 to 4.0 % NaOH content and 0.4 % Mg SO_4 content for 120 min at 98 °C. The complexing agent stage (Q) was carried out with 3 % consistency and 0.2 % DTPA content for 30 min at 60 °C. The oxygen-enhanced peroxide stage (OP) was carried out with 12 % consistency, 2 % NaOH content, 0.1 % $MgSO_4$ content and 0.05 % DTPA content for 120 min at 95 °C. The chlorine dioxide stage (D) was carried out with 10 % consistency, 2–2.5 % ClO_2 content, 0.1 % DTPA for 180 min at 70 °C. Lastly, the peroxide stage (P) was carried out with 10 % consistency, 1.25–1.75 % NaOH, 0.1 % $MgSO_4$ content and 1–2 % H_2O_2 content for 120 min at 80 °C.

The described kraft pulping process was based on the process steps and parameters commonly used in the industry. The fibres of the pulps produced are therefore similar to the fibres of industrially produced ECF pulps. A short discussion on the chosen pulping and bleaching parameters is included.

2.2 Grinding

The pulp was ground in a programmed and liquid nitrogen cooled ball mill (CryoMill, Retsch GmbH, Haan, Germany) with a frequency of 25 Hz using a 50 ml stainless steel grinding jar and one grinding ball of 25 mm diameter. Each pulp sample was ground five times for 1 min. Prior to the first grinding cycle, the grinding jar including the pulp was automatically cooled for

2 min. Cooling was applied for 0.5 min between all grinding cycles. The knife mill (M20 universal mill, IKA-Werke GmbH & CO. kg, Staufen, Germany) with a rotating vane blade was used with a speed of 20 000 1/min. The samples were ground three times for 30 sec.

2.3 Extraction

The solvents were freshly distilled by rectification from both technical quality n-hexane and acetone (VWR International, LLC., Radnor, USA). The quality was analysed via GC-MS. It is identical to GC quality n-hexane and PESTINORM acetone from VWR (VWR International, LLC., Radnor, USA). The extraction was performed by a Soxtherm SOX 6 from the company C. Gerhardt GmbH (Königswinter, Germany), run with the automatic MULTISTAT controlling system. The cellulose extraction thimbles and the viscose wadding used to plug the thimbles were pre-washed by a short Soxhlet extraction of 5–6 cycles for 60 min using a self-made azeotrope of n-hexane and acetone. The ground pulps of 5 g per extraction thimble were extracted with 140 ml of solvent with the automated self-programmed n-hexane extraction program: T-Class 200 °C; hotplate temperature 180 °C; setback interval 3.5 min; setback pulse 2 sec; boiling phase 45 min; distilling off A, 4 intervals; extraction time 3 h; distilling off B, 1 interval; general distilling off interval time 5 min; total time 4 h 7 min. For acetone the extraction program was equal apart from the setback interval, which was 4.5 min. After extraction, the extracts were adjusted to a volume of 50 ml in volumetric flasks.

2.4 Sample preparation

The extracts were characterized by TD-GC-MS. Approximately 90 ± 5 µg of extract was applied into each sample cup by adding the extract solutions. Therefore, a defined volume of 50 µl of the solvents was gently evaporated by being placed under the fume cupboard at room temperature covered with a paper sheet to avoid dust pollution. Afterwards, the concentrations in the extracts were measured and sample cups filled up multiple times until the target mass was reached. After that, the sample cups were introduced into the pyrolysis system for TD-GC-MS analysis.

2.5 TD-GC-MS parameters

The TD-GC-MS analysis was performed using a micro furnace Double-Shot Pyrolyzer (Py-2020iD) equipped with an Auto-Shot Sampler (AS-1020E) both produced by Frontier Laboratories Ltd. (Koriyama, Japan). The pyrolysis system was interfaced to a GC-MS (6890N/5973N, Agilent Technologies Inc., Santa Clara, USA). The TD temperature for the dried and solid extractives was 325 °C and the interface was set to 330 °C. For pyrolysis of low-density polyethylene (LD-PE) retention index standards the pyrolizer was set at 500 °C. The GC inlet and the GC-MS interface temperature were kept at 320 °C. A low polarity column (ZB-5HT, Phenomenex Inc., Torrance, USA) of 30 m × 0.25 mm i.d. and 0.25 µm film thickness was used with helium as carrier gas. The split ratio was set to 20:1. A flow rate of 1 ml/min (constant flow) was set for gas chromatographic separation. The signal data rate was 20 Hz, scan frequency was 2.22 scans/second and scan speed (u/s) was 1 562 [N = 2].

The oven temperature of the GC started at 45 °C, which was held for 2 min and subsequently increased up to 340 °C with decreasing heating rates. This temperature was kept constant for 30 min. The exact configuration of the GC oven temperatures is listed in Table 1 and visualized in Figure 10. The total run time was 134.17 min. For mass spectral detection an 5973N MSD (Agilent Technologies Inc., Santa Clara, USA) was used with electron impact ionization energy of 70 eV. The scanning range for measurement in total ion current (TIC) mode was 29–700 m/z with a threshold of 100.

Table 1: GC oven temperature ramps.

Ramp	Rate (°C/min)	Final temperature (°C)	Hold time (min)
Initial		45	2.0
1	10.0	100	
2	5.0	180	
3	2.0	280	
4	1.5	320	
5	5.0	340	30.0

2.6 Data evaluation

Every dried pulp extract was analysed at least twice. The multiple chromatograms per extract were visually compared in an overlay before checking its individual peaks. In case of unusually missing or additional peaks, these were screened and its reasons, such as a small air infiltration during the GC measurement for instance, were analysed. Those samples were prepared again and measured twice to make sure the GC-MS measurement worked without any disturbances and that the chromatograms were equal. One of the representative chromatograms was further used. The data sets were pre-treated and put together in a database of single-variety pulp extract reference chromatograms with the continuously developing database software OpenChrom® (version: 1.4.0.2021 04 21 10 10, Lablicate GmbH, Hamburg, Germany) (Wenig and Odermatt 2010b; Wenig and Odermatt 2010a) and its ChromIdent® database tool. Compound identification of single peaks was performed using the NIST Mass Spectral Library (NIST20) (National Institute of Standards and Technology, USA).

Based on the fact that developing a new methodology is significant to this paper and therefore, its result, all key aspects of the methodology development are placed in 3 Results and discussion. For instance, the outcome of comparing different mills for pulp milling and the discussion of suitable solvents for ground pulp extraction are part of the technique development process. Subsequently, the following chapter presents these alongside achieved bleached pulp brightnesses for different genera, their extract amount yield, the pulp extract chromatography data pre-treatment and setting up/functioning of the reference chromatogram database.

3 Results and discussion

Supporting the development of the presented general sample preparation method, a broadly MTH kraft pulp of Asian origin seemed most appropriate. The industrial MTH pulp, was described in Section 2 (Materials and methods). MTH is a raw material segment with a composition that is subject to wide fluctuations. It can also be defined as a more or less natural mixture of tropical hardwood types. For example, in some countries MTH consists of 15–20 species, but in others MTH can contain up to 100 species (Grützmann 2013). As a result, the densities of mixed tropical hardwoods can vary from 700 kg/m^3–800 kg/m^3. Considering individual species within a forest area, the variation increases even further and can vary from 150 kg/m^3–1 300 kg/m^3 (Grützmann 2013). As a kind of by-product, MTH is largely produced during large-scale conversions of natural forests in the tropics into palm oil and timber plantations (Broich et al. 2011).

To establish the pulp extract database, the anatomically determined woods underwent kraft pulping and bleaching processes. According to the current FAO pulp and paper capacity report (FAO 2022b) the global capacity of chemical paper grade pulp is 2.2 MT of sulphite pulp compared to 125 MT of kraft pulp. With the kraft pulp procedure, the dominant share of the market is covered at this point of the study. A combination involving oxygen, peroxide, chlorine dioxide, and complexing agents for bleaching kraft pulps appears most relevant. While on the one hand, the specific reaction conditions vary due to factors like different kinds of raw material, mill design and philosophy, on the other hand, the overall exposure to pH and chemicals of extractives, remains relatively consistent. Although bleaching conditions vary, the authors believe that they won't alter the fingerprint of extractives. Therefore, sensible choices for the study's starting point were made.

3.1 Choice of wood species

The tropical forests of Southeast Asia are characterized by high species diversity among the tree species growing there. In cooperation with the NGOs Greenpeace and World Wide Fund For Nature (WWF), a list of 38 particularly relevant and interesting wood species was compiled (Table 2).

The pulp industry likes to use cheap and locally available woods. Therefore, the selection of relevant woods for the reference database was also based on the distribution and availability for the industry and correspondingly good procurement opportunities for the pulp industry. Some of the woods originate from plantations, for example *Hevea brasiliensis* or *Paulownia tomentosa*. They also find use as pulp as they are often co-products.

The list also includes *Acacia mangium* and *Eucalyptus globulis*, which are often legally used for pulp. These serve as a proof that the methods developed in this study can be applied to industrially produced pulps. In addition, these pulps are often found in papers and must be distinguished from tropical woods. The palms *Cocos nucifera* and oil palm *Elaeis guineensis* were included in this study because they can also be used for pulp production and it is

anticipated that these palms, which are currently grown in plantations on former rainforest areas, will be increasingly cleared and utilized for pulp production (Onuorah E. O. et al. 2015; Welling and Liese 2019; Wan Daud and Law 2011). Table 2 also shows the achieved ISO brightness and extractive content of every samples bleached pulp.

According to the authors information, whole debarked logs are used in MTH pulp production and not sawmill residues. The authors therefore assume that heartwood always feeds into pulp production in large proportions and thus the use of heartwood as reference material was decided. Origins are not given in Table 2 since 2 kg of solid wood material were required for pulp production, which was not possible to obtain with a documented origin for most of the woods.

Table 2: Wood taxa/species and pulp properties. The extractive content means only the n-hexane extractives.

			Bleached pulp	
No.	**Wood taxon**	**Trade name**	**Brightness (% ISO)**	**Extractive content (%)**
1	*Acacia mangium* Willd.	Acacia	90.9	0.52
2	*Alniphyllum pterospermum* Matsum.	Mee Dong	90.1	0.20
3	*Avicennia marina* (Forssk.) Vierh.	Api Api	86.5	0.52
4	*Calophyllum* spp.	Bintangor	88.9	0.23
5	*Canarium* spp.	Kedondong	90.2	0.14
6	*Castanopsis argentea* (Blume) A. DC.	Berangan	83.1	0.58
7	*Cocos nucifera* L.	Coconut Palm	89.3	0.24
8	*Cunninghamia lancelota* (Lamb.) Hook.	Chinese Fir	82.9	0.63
9	*Dendrocalamus latiflorus* Munro	Bamboo	86.9	0.24
10	*Dipterocarpus* spp.	Keruing	85.5	1.19
11	*Durio* spp.	Durian	89.1	0.17
12	*Elaeis guineensis* Jacq.	Oil Palm	89.6	0.06
13	*Eucalyptus globulis* Labill.	Eucalyptus	90.1	0.73
14	*Fagus sylvatica* L.	Beech	89.5	0.04
15	*Gonystulus* spp.	Ramin	89.0	0.16
16	*Heritiera* spp.	Mengkulang	87.8	0.07
17	*Hevea brasiliensis* (Willd. ex A. Juss.) Müll. Arg.	Rubberwood	88.7	0.16
18	*Ilex triflora* var. *kanehirai* (Yamamoto) S. Y. Hu	Holly, Kecemang	89.2	0.20
19	*Intsia* spp.	Merbau	86.8	0.34
20	*Koompassia malaccensis* Maingay ex Benth.	Kempas	84.5	0.29
21	*Lophopetalum* spp.	Perupok	89.3	0.11
22	*Mangifera* spp.	Ambacang, Mango	89.7	0.45
23	*Nyssa javanica* (Blume) Wangerin	Tupelo, Nyssa	88.8	0.09
24	*Palaquium* sp.	Niato/Suntai	87.9	0.30
25	*Parashorea* spp.	Gerutu	82.1	0.34
26	*Paulownia tomentosa* (Thunb.) Steud.	Paulownia	87.3	0.48
27	*Phellodendron* sp.	Amur Cork Tree	88.4	0.48
28	*Pterygota* sp.	Koto	91.1	0.04
29	*Rhizophora* spp.	Red Mangrove	85.8	0.23

30	*Schima superba* Gardn. & Champ.	Samak, Puspa	86.4	0.13
31	*Anthoshorea* spp.	White Meranti	89.4	0.20
32	*Richetia* spp.	Yellow Meranti	89.1	0.19
33	*Rubroshorea* spp.	Dark/Light Red M.	85.0	0.32
34	*Shorea* spp.	Bangkirai, Balau	88.6	0.28
35	*Swintonia* spp.	Merpauh	90.8	0.25
36	*Tectona grandis* L.F.	Teak	82.8	0.10
37	*Terminalia tomentosa* Willd.	Limba	90.4	0.14
38	*Tetramerista glabra* Miq.	Punah	83.6	1.01

3.2 Grinding

For the selection of a suitable mill, the knife mill with rotating vane blade and the cryo ball mill were compared for pre-treatment of MTH pulp before the extraction. Differences between the results may indicate that changes occur in the pulp during grinding. This could be caused not only by the different grinding techniques, but also by the temperature differences within the mills. The ball mill is cooled with liquid nitrogen (-196 °C) and produces a very fine powder, which provides a large surface area and thus good accessibility for the solvent (Figure 9c). It allows the solvent to flow uniformly through the sample and dissolve repeatable amounts of extractives, resulting in reproducible peak areas (Table 3). The knife mill, on the other hand, produces an inhomogeneous fluffy grind that contains both fibrous large particles (Figure 9b), but also very small powdery particles that are not visible in Figure 9b.

The differently ground MTH pulps were extracted with petroleum ether (six samples per mill of 5 g each) and analysed both quantitively, by gravimetric determination of the amount of extract, and qualitatively, by the detected number of peaks and peak areas in the TD-GC-MS chromatograms (Table 3). The quantitative results showed slightly higher extract amounts of 0.026 % (0.2 % CV) after the first extraction step for knife mill ground pulp in comparison to 0.020 % (2.3 % CV) for ball mill ground pulp. Furthermore, the number of peaks were higher within the knife mill treated pulp. These chromatograms had a mean number of 125.8 peaks (variance 17.1) whereas the ball mill ground pulps had a mean of 104.8 peaks (variance 6.2). The ball mill provided at least twice as good reproducibility of the peak areas from the extract chromatograms: The coefficient of variation (CV) of the areas of the 15 biggest peaks of the ball mill extracts chromatograms is 0.128 in contrast to 0.301 for the knife mill. The higher extract quantity and peak numbers of the knife mill spoke at first sight for this mill. But for the following analytical evaluation and the purpose of this paper the GC-MS measurement precision and the reproducibility of the results are crucial. The function of the fingerprint database relies on reproducible peak areas and proportions. Therefore, the extracts chromatograms of the pulps ground with the ball mill are more suitable for this study. The clearly better reproducibility of the peak areas when using the cryo ball mill could be due to the finer grinding under cold conditions.

Flaig, M.L. et al. (2023). *Holzforschung* 77: 860–878,

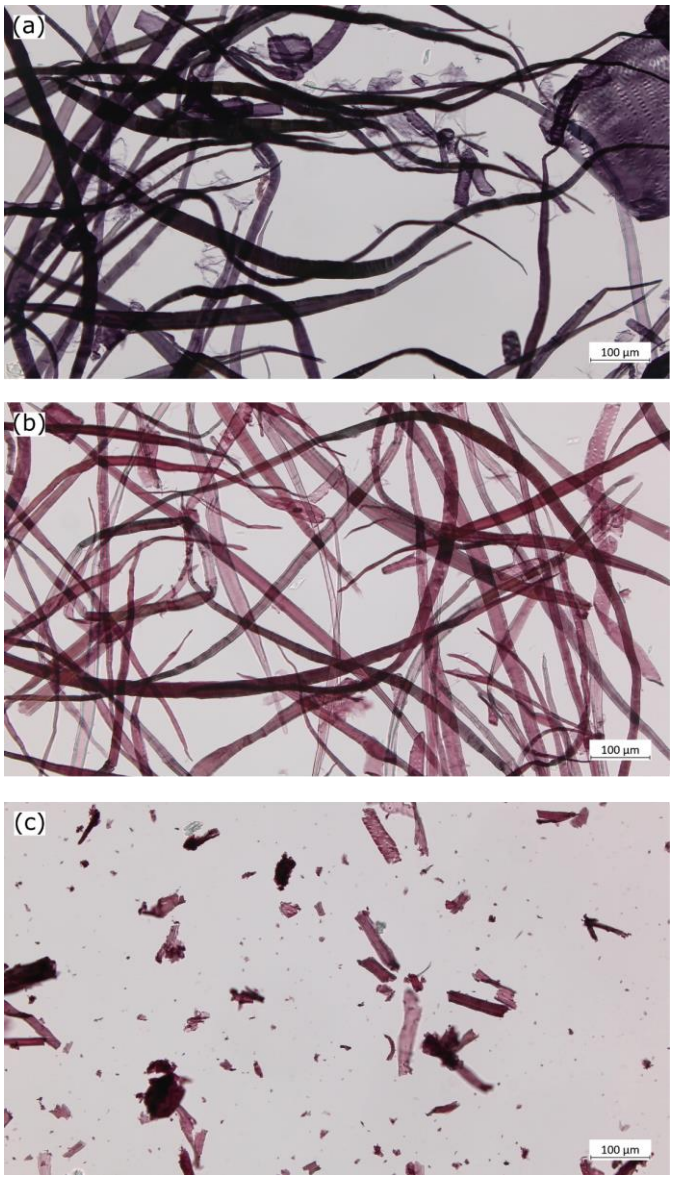


Figure 9: Microscopic images of MTH pulp: (a) raw (b) knife mill ground (c) cryo ball mill ground. Staining was done with Alexander Herzberg stain.

Table 3: Comparison of cryo ball mill and knife mill ground petroleum ether extracted MTH pulp. Peak area of 15 biggest peaks per chromatogram, determined by one chosen SIC area for every peak.

	Ball mill		Knife mill	
	Mean	**CV (%)**	**Mean**	**CV (%)**
Number of peaks	104.8	2.4	125.8	3.3
Peak area	2 891 103	12.8	2 571 478	30.1

3.3 Extraction

The successive extraction was also optimized on basis of the MTH industrial pulp milled with the cryo ball mill. The extract quantity was maximized through the selection of suitable solvents, extraction time and number of extraction cycles. The extract quantities of the individual extraction steps were determined. The extraction process used a combination of nonpolar and polar solvents, including three times petroleum ether (6 samples) followed by two times acetone and stand-alone n-hexane (5 samples) to obtain both fractions of the extracts as exhaustively as possible, as shown by Ponnuchamy et al. (2021) for kraft lignin and Krogell et al. (2012) for Norway spruce bark. Because the third extraction step obtained hardly any measurable extracts (Table 4), this method effectively extracts ground pulps exhaustively with two steps per solvent. Water was not used because it was assumed that water-soluble components were already discharged with the process water during pulping and bleaching. The results showed that most of the absolute extractive amounts were obtained in the first extraction step with either solvent. Table 4 also shows the relative extractives amount in percent and extract yields per solvent and extraction step.

Table 4: Comparison of the successive extraction with petroleum ether (3 times) followed by 2 times acetone and a sole n-hexane extraction. Total and relative extract amounts are determined based on extraction of 20 g of ball milled MTH pulp.

	Petroleum ether				Acetone (suc. after pet. ether)				N-Hexane			
Extraction step	Amount (mg)	Amount (%)	SD (%)	CV (%)	Amount (mg)	Amount (%)	SD (%)	CV (%)	Amount (mg)	Amount (%)	SD (%)	CV (%)
1	4.05	0.020	0.0005	2	10.24	0.051	0.0046	9	4.12	0.021	0.0003	2
2	1.46	0.007	0.0006	8	3.06	0.015	0.0049	32	1.43	0.007	0.0005	7
3	0.77	0.004	0.0014	36								
Sum:	6.28	0.031			13.29	0.066			5.55	0.028		

Acetone was found to be quantitively more effective in extracting pulps than hexane, but high yield does not imply specificity of the extractives. Additionally, the acetone extracts were difficult to measure by GC-MS without derivatization due to their polarity. These polar components can be found in paper making effluents (Willför et al. 2006; Örsa and Holmbom 1994; Björklund Jansson 2005; Valto et al. 2012; Holmberg 1999), meaning they are affected at least in their quantity by pulping and bleaching. It was assumed that acetone mainly extracts polar extractives and hexane the more non-polar, more interesting extractives. The **non**-polar extractives are **least affected** by pulping and bleaching, as the process takes place in polar media, as water is used to dissolve the applied chemicals. Therefore, the hexane soluble extractives are more valuable for the database. The focus was on developing a practical method. That is why it was decided to leave the acetone extracts aside and concentrate on the hydrophobic fraction of the extractives. Another reason for the pre-fractionation is reduction of the complexity of the chromatograms. However, one problem with using the solvent petroleum ether is that it is not clearly defined in terms of composition, leading to a

risk that extraction conditions may not be identical when changing batches or suppliers. Moreover, the ground pulps were not only cool extracted but boiled in the solvent in the first step. Monitoring the petroleum ethers (boiling range 40–60 °C) boiling temperature during the extraction process showed a temperature rise over time. This supported the assumption that the cool extraction was not carried out with the entire solvent mixture, but only with the less volatile part, because the composition of the petroleum ether presumably changed over time in favour of the higher-boiling fractions. So, petroleum ether is not perfectly suitable for extraction systems without pressure like Soxtherm or Soxhlet. In contrast, n-hexane is a well-defined solvent. The results of using n-hexane were found to be similar in terms of efficiency and reproducibility when compared to using petroleum ether (Table 4). Therefore, n-hexane was deemed the best choice for this purpose.

During the optimization process, sources of impurities for example in the extraction wadding, the extraction thimbles, the solvent or on glass surfaces and others were identified. They were largely eliminated by establishing a laboratory routine. These measures led to an improved quality of the pulp extracts with a significant reduction of impurities. A chromatogram of a blank extraction of empty and unwashed extraction thimbles and wadding with petroleum ether showing the impurities peaks (Supplementary Figure S5) and a list of identified compounds (Supplementary Table S40) were added to the supplementary material.

3.4 Optimization of TD-GC-MS chromatogram quality

Prior to development of data pre-processing systems, an optimization of the GC-MS analysis was a necessity. In this way, the quality of the chromatograms was increased by defining suitable GC-MS parameters such as the TD temperature of 325 °C and developing a fitting GC oven temperature program for the n-hexane pulp extracts. A better chromatographic separation of overlapping or sometimes called co-eluting peaks, which occasionally occur in one-dimensional gas chromatography (Blumberg 2012), was achieved especially of the mainly phytosterol peaks eluting around 280 °C in the range of RT 65–85 min (Figure 10). Their overlap was satisfactorily separated by the slow heating rates of 2.0 and 1.5 °C per minute. This was achieved by using polyethylene analyses as a neutral measure of the effect of the heating ramps. The heating ramps were adjusted so that the distances of the alkane and alkene peaks of the polyethylene were approximately equidistant over the relevant chromatogram range instead of increasingly narrow. Thus, a better capture of the information hidden behind coelution in this RT range of the chromatograms was achieved. In order to prevent artefacts of the high-boiling extract components in the following measurements, the final bake-out time needs to be 30 min long. Still the overall measurement time per extract of less than 140 min wasn’t unnecessarily extended. The optimized oven temperature program is shown in Figure 10.

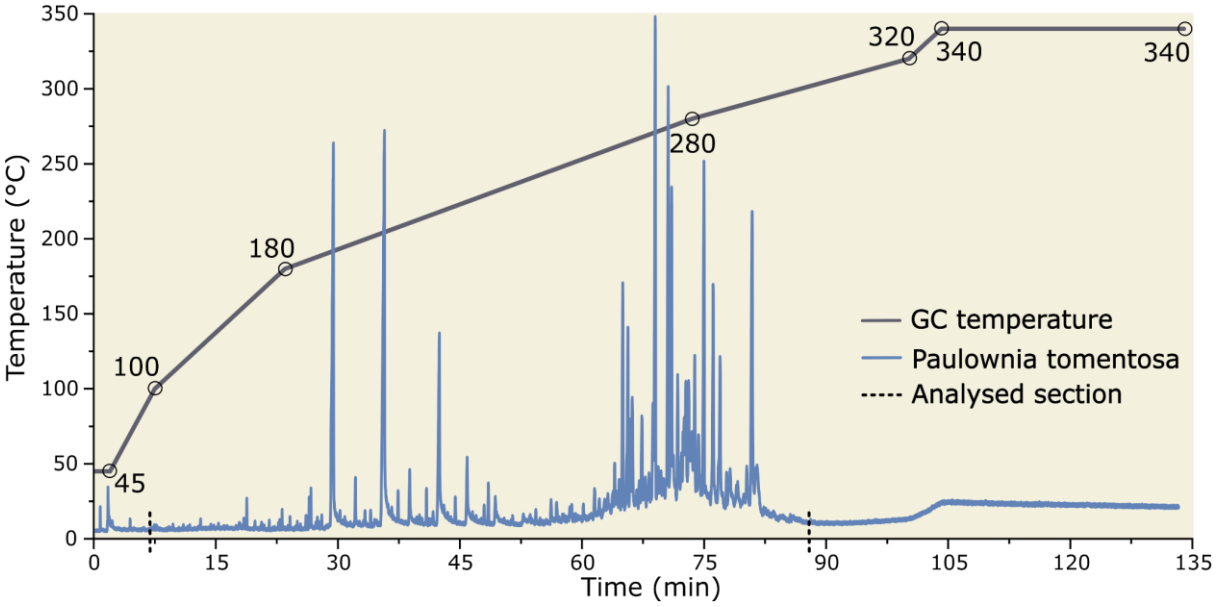


Figure 10: GC oven temperature program and chromatogram of *Paulownia tomentosa* (Thunb.) Steud. pulp extract before preprocessing. For the database references the chromatogram section of RT 7–88 min was analysed.

3.5 Pre-processing of GC-MS data

The extractable constituents of the pulps were largely captured by the described conscientious sample preparation. For the following analysis they are available in the form of GC-MS chromatogram data sets. With the help of data pre-processing steps, it was possible to record the chromatographed analytes, i.e. their chemotaxonomic characteristics, and store them in a database. Appropriate pre-treatment of the data eliminated or compensated artefacts. The quality of the chromatograms was thus improved and for each peak in the chromatogram, which in turn represents a component of the complex extractive mixture, characteristic properties such as percental peak area and retention index can be determined more reliably. The following nine most important pre-processing steps were found to be crucial for the pulp extract data matrix and the purpose of a subsequent fingerprint database. They need to be performed on the raw GC-MS data files for both **references** for the database and **unknown** pulp/paper samples for comparing against the database.

3.5.1 Rounding of detected mass traces

Each mass trace (m/z) is rounded to nominal masses (Khrisanfov and Samokhin 2022). Because the mass axis in the used GC-MS system tends to shift upwards, rounding of decimal places was performed from inclusive -0.3 to exclusive +0.7. Whole numbers are important for later trouble-free smoothing, which in turn is important for the peak detection algorithm.

3.5.2 Deleting empty scans

Empty MS scans are those that don't contain any data. They are removed to shrink the dataset size and to speed up the data processing time.

3.5.3 Selecting chromatogram range

The beginning and the end of the recorded raw chromatograms are cut. They are not of interest for this analysis because the front part consists of irrelevant peaks, which don't belong to the pulp extracts but for example arise from solvent residues or the atmosphere. The back part of the chromatogram is also discarded due to mainly noise and column bleeding. Also, to reduce the dataset size and speed up the processing the chromatogram section used for the database was selected from 7–88 min.

3.5.4 Removing unwanted ions

Mass traces, that do not (only) belong to the pulp extracts but (also) to the atmosphere or the GC column should be avoided. The ions 18, 28, 32, 44 result from the atmospheres gases (water, nitrogen, oxygen, carbon dioxide). 207 and 281 are attributed to column bleed, the loss/decomposition of the stationary phase, at elevated temperatures. Therefore, they are removed from the whole dataset. Also, the mass trace 84 m/z is removed as it occurs regularly and often and is not specific to pulp extractives at all.

3.5.5 Smoothing chromatograms

For a well working peak detection and clean peak integration the whole chromatograms were smoothed. For this purpose, the chromatogram filter Savitzky-Golay smoothing with the settings order = 2 and width = 5 was applied. This filter algorithm was developed by Savitzky and Golay (1964). It is used to smoothen single ion signals or TIC signals and remove electric noise (Wenig 2011). The smoothing was done per ion, meaning every single m/z was smoothed separately, which is important for the following peak detection via deconvolution. Figure 11 shows how smoothing the raw m/z data also affects the peak shape in the TIC chromatogram. One can see how the blue raw chromatogram in Figure 11a looks unsmoothed in comparison to the smoothed brown TIC signal underneath (Figure 11b). For instance, the three mid-sized peaks in the centre (RT 66.5, 67.7, 69.1 min) show rough detector signals and therefore tiny double peaks, which are removed in the smoothed data set below. Choppy signals are risky because they can be falsely detected as multiple peaks on both the ion level and the TIC level.

Flaig, M.L. et al. (2023). *Holzforschung* 77: 860–878,

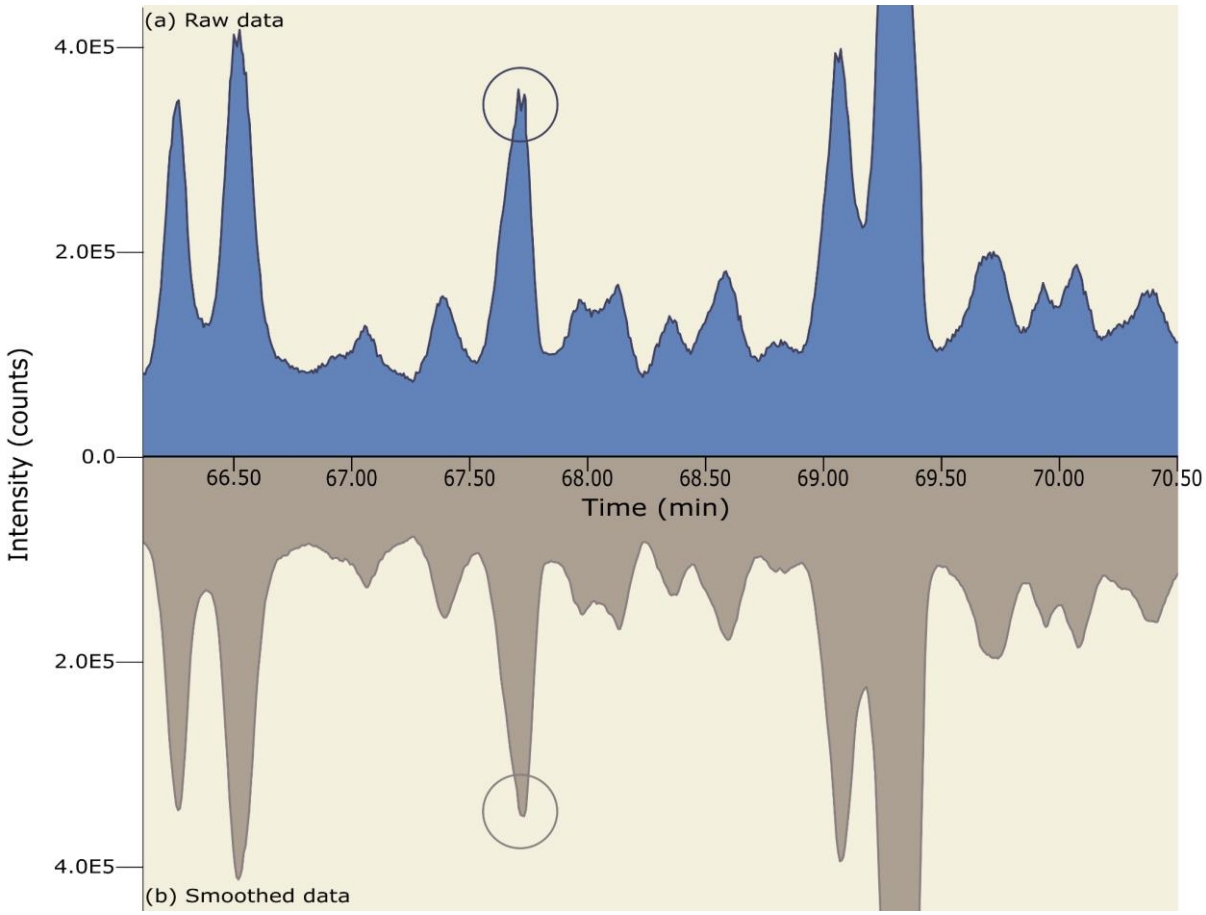


Figure 11: Savitzky-Golay smoothing: TIC chromatogram sections from database reference sample *Paulownia tomentosa* (Thunb.) Steud., RT 66.10–70.50 min: (a) raw data versus (b) smoothed data.

3.5.6 Deconvoluting/Detecting peaks

Multivariate curve resolution by alternate regression (MCR-AR) according to Gerber et al. (2012) was used to decompose the multivariate GC-MS data into individual pure component spectra and to detect the peaks. MCR-AR enables the detection of hidden peaks underneath the baseline noise or other peaks via their single ion traces. Thus, the underlying and overlapping peaks can be detected individually as shown in Figure 12. MCR involves a segmentation of chromatograms into non-overlapping minimum 50-scan windows using local minima/maxima. Mass channels are baseline-corrected within these windows by linear interpolation and area subtraction. Each window forms a data cube. This data cube is then unfolded into a matrix (N × K) × L, where N represents the number of samples, K represents scans, and L represents mass channels. The decomposition obtained from the initial MCR step is used as a starting point for AR. It alternates between deconvoluting chromatographic and mass spectral profiles until convergence. It starts by assuming a single distinct compound (rank = 1) and applies constraints like non-negativity to the solution. The pure component spectra are used as constraints in the regression process. The data is iteratively reprocessed („MCR-AR max iteration“: 50.0), by incrementally increasing the rank by one and calculating solutions with each iteration refining the decomposition. It continues until a predefined

stopping criterion is met, indicating that the decomposition has reached a stable state. This means that the algorithm has found a solution that best represents the underlying pure components in the mixture (Gerber et al. 2012). This peak detection is very useful for the creation of the fingerprint database, as the complex wood extract samples often contain many similar substances that elute very close to each other in time. The overlaps cannot be completely separated chromatographically and sometimes appear as “mountains” in the chromatogram (Figure 10, RT 65–85 min). In the following some MCR-AR settings, that were used for best results, are given and explained. The “Local Maxima/Minima” classifier uses an algorithm that follows the signal from left to right to find local maxima/minima based on the first derivative/slope of the signal. Although this classifier does not search for chromatographic peaks with start/maximum/stop, some kind of start point must still be found internally, to be able to evaluate the parameter “Signal difference threshold between start and extrema”, which was set to 1000.0. To set a starting point it is required that the slope exceeds a certain threshold (“Local Maxima Scan Slope Threshold”: 0.001). Then the “Local Maxima/Minima Consecutive Scans” parameter comes into play: This defines how many scans must remain above this threshold after it has been exceeded for a start point to be set. This helps to filter out the smallest local maxima in the signal noise range. It was set to 1.0. All detailed MCR-AR parameter adjustments including explanations can be found in Supplementary Table S39.

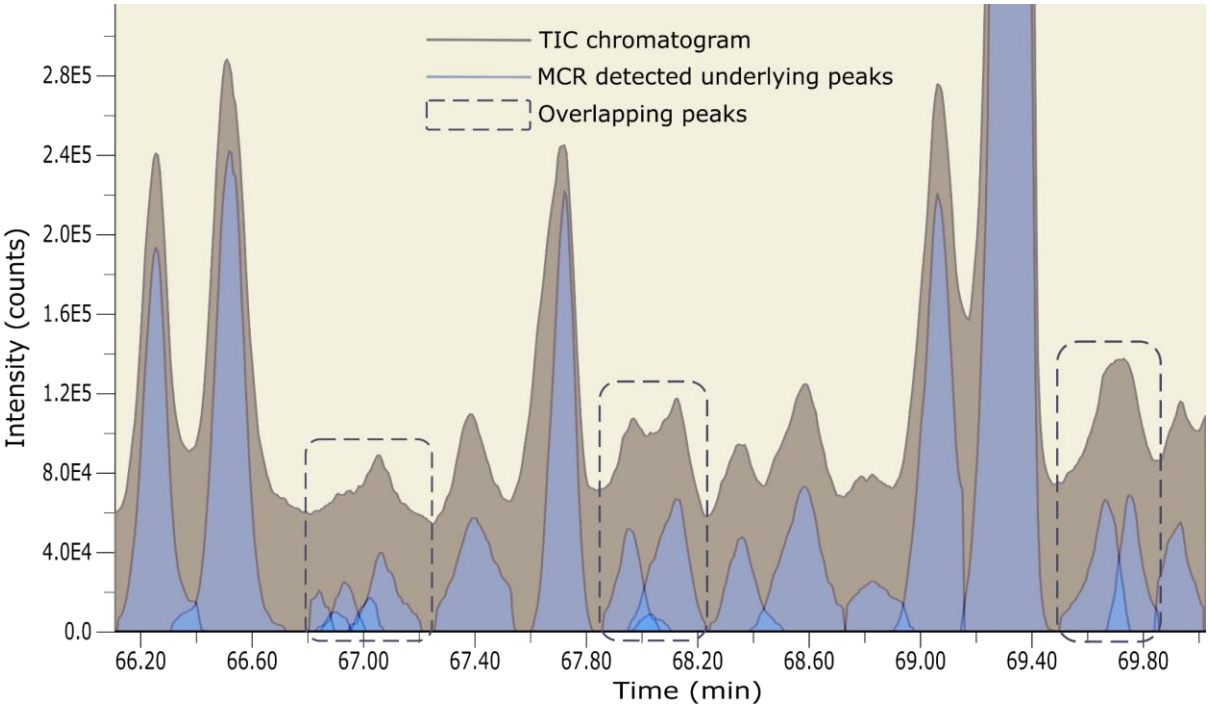


Figure 12: Peak detection using deconvolution (MCR-AR): section from smoothed database reference sample *Paulownia tomentosa* (Thunb.) Steud., RT 66.11–70.03 min.

3.5.7 Integrating peaks

To determine their peak areas the MCR-AR detected peaks were integrated with the trapezoid peak integrator. The deconvoluted TIC signal was integrated excluding the background. The

peak areas were chosen to be put into the database as one of the comparison features rather than the intensity because the intensity and peak shape can vary whereas the peak areas and proportions to each other, which are used for fingerprinting, stay the same. The peak areas are normalized so that only a percentage value is used. Since this is not a quantitative but a qualitative identification method, the absolute content is not decisive here, thus only the ratios need to be repeatable.

3.5.8 High pass ions and peaks

The high pass filters keep all ions or peaks with the biggest intensity or area. For this database the 400 biggest peaks by area per reference chromatogram worked the best and were kept for further analysis. Of every of those peaks, the 150 ions with highest signal intensities were kept. Keeping more than 150 ion traces didn't approve the database performance on random pulp mixtures because in the mixture the amount of every pulp is lower. Therefore, the extracts peaks are smaller and consequently due to the threshold settings the smallest single ions aren't detected by the MS anyway.

3.5.9 Retention Index (RI) calculation

The retention index (RI) is a measure used to characterize the retention time of a compound relative to the elution times of reference compounds. It is dimensionless and represents the relative affinity of a compound for the stationary phase compared to the reference compounds, meaning it converts retention times into system independent constants. An internal RI value is calculated for each peak maximum with LD-PE as standard reference, which is separated under the same conditions as the sample of interest. The RI are determined by relative distances of the compounds of interest to the LD-PE alkene peaks in the chromatogram: Retention Index $RI = 100 \times [n + (N - n) * (RT(unknown) - RT(n)) / (RT(N) - RT(n))]$ where n is the carbon number of the n-alkene, $N = n + 1$ and RT = retention time. In every GC-MS sequences LD-PE RI standards are measured minimum after every tenth measurement. This is how a system-independent index value without time unit could be given, which is not influenced by any kind of RT shift due to column abrasion, cutbacks of the column or a changed column/measurement program. It is very important to calculate the RI for every peak because the RT shift can even on the same GC-MS system be substantial with longer times between measurements and therefore influences the matching accuracy of the database massively if no RI are used.

3.6 Functioning of the database

The pre-treated chromatograms contain several pieces of information important for the functioning of the database: deconvoluted peaks consisting of the individual mass traces, their peak areas and RI. These are associated with the peaks. For each reference extract of a pulp produced from a single wood species, the above information results in an individual fingerprint of contained compounds. Although most of the substances are not wood species-specific marker substances, the complex combination of substances with their proportional

Flaig, M.L. et al. (2023). *Holzforschung* 77: 860–878,

ratios to each other is unique. This fingerprint information is fed into the database (Wenig and Odermatt 2010b; Wenig and Odermatt 2010a), as shown in Figure 13.

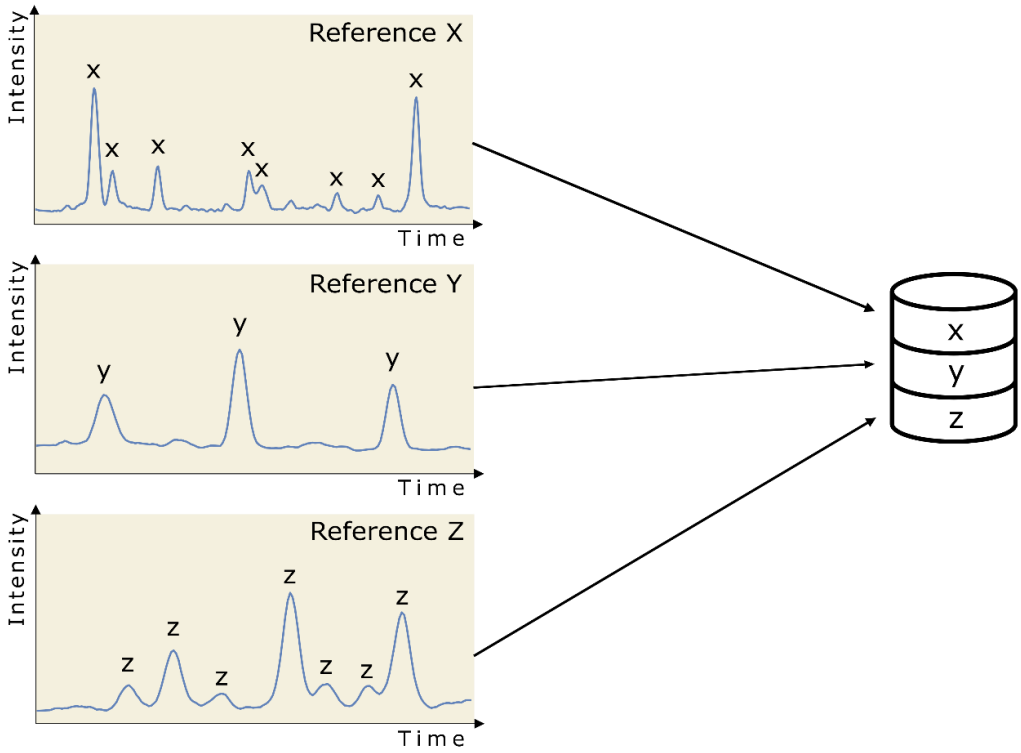


Figure 13: Schematic illustration of the database approach.

When **adding** a **reference** chromatogram to the database, each peak within a fixed RI window of ± 10 is compared with all other reference peaks in the same frame. In contrast to the NIST library peak comparison for compound identification by Stein and Scott (1994) for the fingerprint matching the cosine algorithm by Alfassi (2004) is used. The peaks are basically compared via their mass spectral match quality. Using retention indices is essential, otherwise the rate of false positive matches rises due to the similarity of wood extractive peaks, especially of alkanes due to their similar mass fragmentation pattern. Only if the peaks from different reference extracts match 80 % or better (Match Factor 80+), they are rated as the same compound and merged and added collectively to the database as **one** library peak. This combined library peak links to all the corresponding references where it originates from. If peaks at the same RT/RI match with a smaller percentage than 80, they are rated to be different peaks and put into the database separately as individual library peaks. The matching parameter was set to a high value of greater than MF 80 because the wood samples contain confusingly similar components.

When **matching** an **unknown** mixed species pulp extracts chromatogram for identification against the database, the first step again is to align the unknown chromatogram with the reference chromatograms in the database using a RI corridor of ± 10. After that searching for matches in that RI corridor using the cosine algorithm starts. Also, peak areas and ratios (fingerprinting) are compared to find the best matches. The results, based on the similarity between the unknown and reference mass spectra, are ranked/sorted according to the scores,

with higher scores indicating stronger matches. These MF scores are penalized by RI distance: the higher the RI (within the determined RI window) are lying apart from each other, the more the MF values are reduced. A minimum MF of 75 for every peak comparison of the unknown against the database reference peaks was chosen somewhat lower, since impure mass traces are expected due to overlaps within the mixed pulp extract sample. In addition, particular mass traces of wood species contained in small amounts in the mixed pulps may be below the mass spectrometers threshold intensity of 100. As soon as a peak of an unknown extract is identified against the database subset (RI ± 10, MF > 75), the information is available if this peak occurs only in one or more references. A peak, that occurs only in one reference sample or group is a direct hit/marker peak (Figure 16). Therefore, a relatively high number of marker peaks, compared to the number of total matched peaks assigned to a specific database reference, is an important indicator in the identification of unknown samples. All other peaks are ambiguous hits. The mathematics behind it is quite simple: principally it's just counting the references. A direct hit counts 1. An ambiguous hit, which occurs in 2 references, counts with 0.5 and so on. The quantity of peaks and number of counting's provide a hint of the unknown extract. When evaluating the comparison results of an unknown sample against the database the reverse similarity index (RSI) and reverse match factor are most important. The unknown pulps are rarely pure. They are mostly mixed samples from different wood species. Therefore, the match against the database is tested in **reverse** to see how well the fingerprint peaks of the single pure database references match the peaks of the unknown sample mixture, i.e., whether every pure reference can be found in the unknown mixed sample. Since the mixed sample consists of several species, matching the whole unknown mixture sample **forwards** against each reference of the database results in lower MF, because the entire sample mixture can never be found in one pure reference sample. As an example: a mixture of X, Y and Z is identified by a database consisting of the references X, Y and Z. Reverse matching results: X is 100 % included in the mix, Y is 100 % included in the mix and so is Z. Forward matching results: the whole mixture matches 33 % with each X, Y and Z. The results of a database matching are mean values of the matching of all individual peaks against each other. The individual peak matching can be represented as in the following Figure 14. Figure 14a shows the mass traces of an unknown deconvoluted peak of an unknown pulp extract mixture in the RT range of 67.50–68.00 min with its maximum at the RI of 3 127 in comparison to the mass traces of the *Paulownia tomentosa* database reference peak at the same RI in an inverted overlay format with the legend of all mass traces (m/z) on the bottom right. Figure 14b shows the comparison of the same *Paulownia tomentosa* database reference peak in detail with all its mass traces not over a RT range but a single scan against the NIST20 library. The unknown extract substance is identified as (+)- Sesamin with a MF of 80.3 and a reverse MF of 87.2.

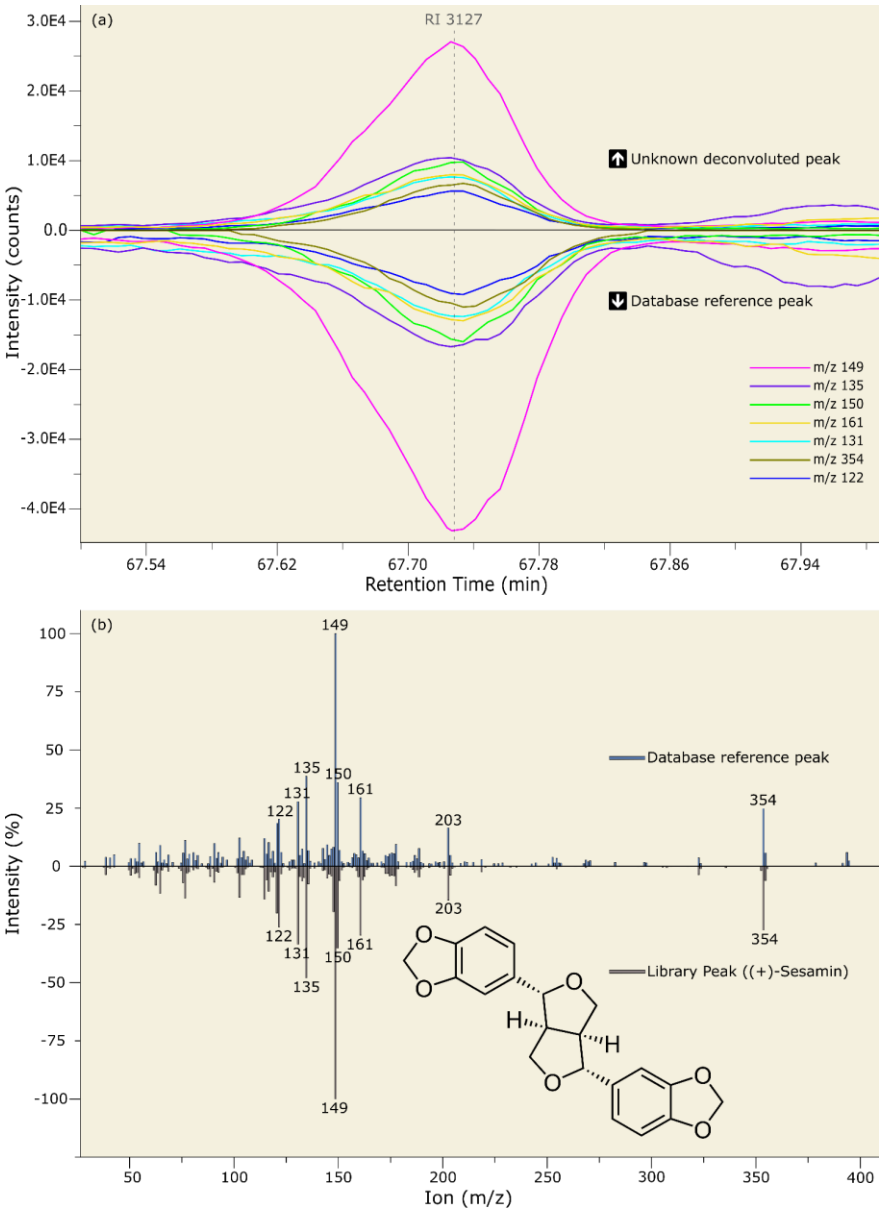


Figure 14: Single peak database comparison: (a) comparison of an unknown deconvoluted peak of an unknown pulp mixture extract against the database library peak from the *Paulownia tomentosa* (Thunb.) Steud. reference pulp extract at RI 3 127 (b) comparison of the database reference peak against the NIST20 library peak of the substance (+)- Sesamin.

In order to represent the database graphically and to make the dimension of the differences of the individual references visible, a principal component analysis (PCA) was calculated from the database entries with the above-mentioned information. In the score plot of the PCA (Figure 15) the PC1 (principle component 1) with 10.98 % on the x-axis and the PC2 on the y-axis with 7.96 % represent only about 19 % of the total database information. This is due to the fact that the constituents of the wood species are often similar and a huge data cloud of more than 6 000 peaks times 38 species is difficult to reduce in its complexity to two dimensions. In the 3-dimensional interactive score plot (PC1, PC2, PC3) (Supplementary Figure S3), a total of about 26 % of the database information is represented and the distances are clearer than in 2-D.

Flaig, M.L. et al. (2023). *Holzforschung* 77: 860–878,

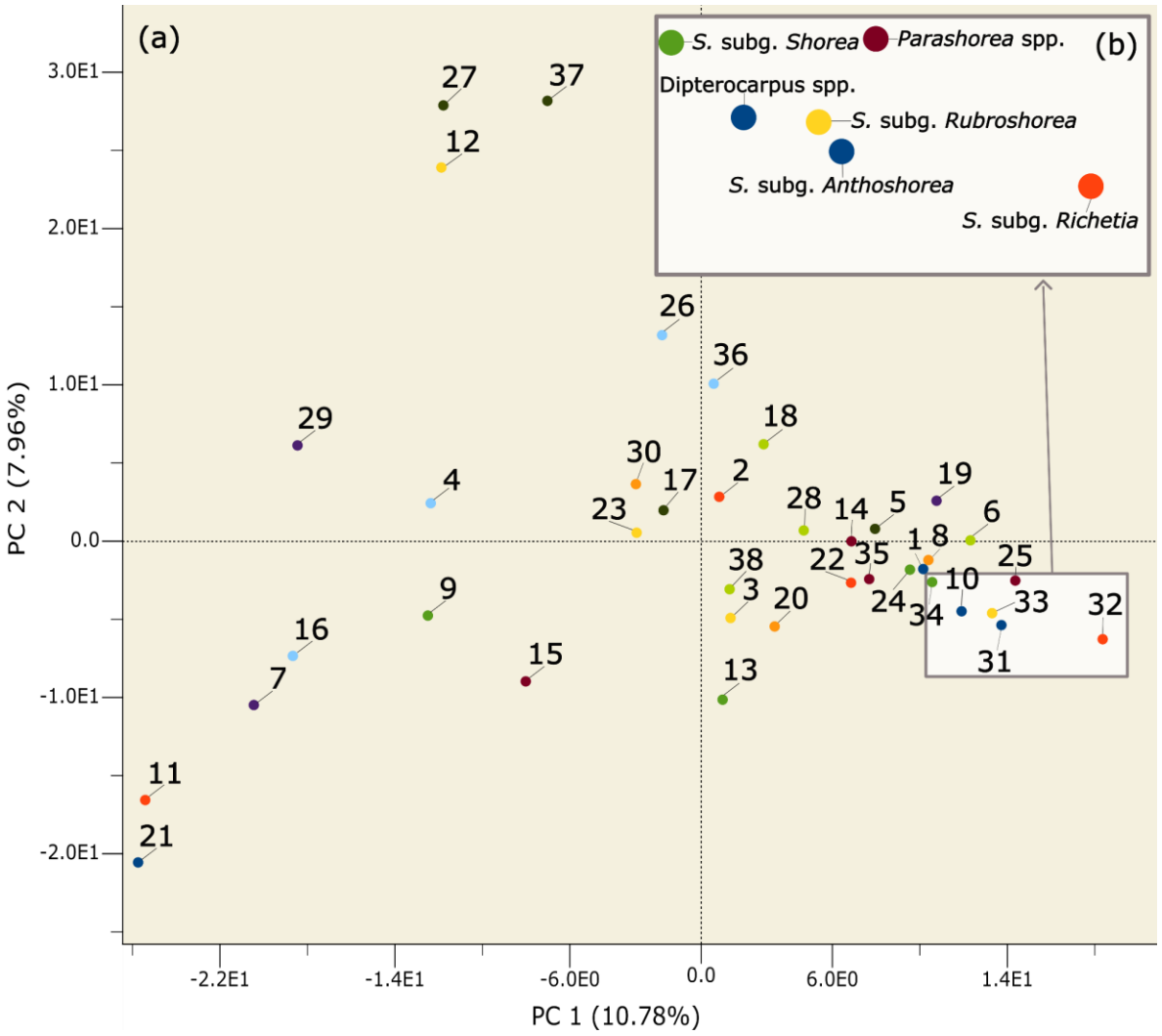


Figure 15: (a) PCA-score plot of the pulp extract database, numbered after Table 2 (b) enlarged section of the *Dipterocarpaceae* family.

Figure 15b shows an enlarged section of the score plot exposing the *Dipterocarpaceae* family. As anticipated, the close relationship between the *Dipterocarpaceae* family members, in particular the former subgenera of *Shorea*, which have recently been reclassified as distinct genera (Ashton and Heckenhauer 2022), is evident in the chromatograms – they cluster together in the PCA representation (Figure 15b). Thus, when identifying one of these species in an unknown sample, a closer look must be taken to ensure that there is no confusion. In order to refine the identification of *Diptocarpaceae* members, a separate database containing only these very similar species could be built in the future. Possibly, the matching parameters can then be chosen even more narrowly, and thus the distinctiveness of the matching results can be increased.

For each database reference chromatogram, there is a table of recorded library peaks including their RT, RI, peak area, NIST20 identified substances (MF > 80) and their 10 biggest ion traces as well as the other database references in which these peaks are also contained (Supplementary Tables S1–38). The most decisive peaks are the specific marker peaks

originating from only **one** reference. The whole database, including all original GC-MS data, can be found under Supplementary Figure S4. The reference chromatograms of *Acacia mangium* and *Paulownia tomentosa* pulp extracts are also given as examples in Supplementary Figures S1–2.

The comparison of unknown paper extracts chromatograms against the database is supported by a visual comparison tool which shows the unknown chromatogram with unmatched, matched but ambiguous, meaning they are contained in more than one reference, and marker peaks (Figure 16).

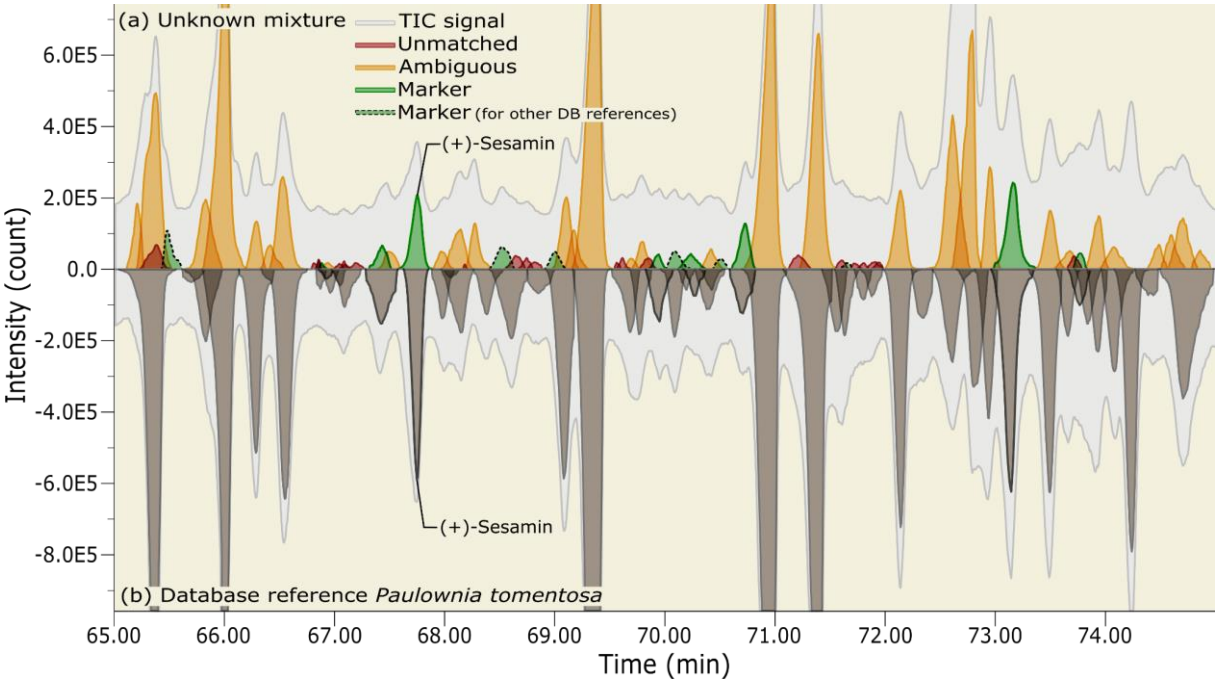


Figure 16: Comparison of an (a) unknown pulp mixture extract chromatogram with (b) the database reference chromatogram of Paulownia tomentosa (Thunb.) Steud.

When matching an unknown pulp extract chromatogram with the database, several comparison factors and statistics are crucial: the total RSI value, the number of marker peaks/ambiguous peaks and the total matching area which is calculated for every database reference compared to the unknown chromatogram. Ultimately, the observer must incorporate all of these comparison results/statistics into the decision-making process. An example of an actual species identification in an unknown pulp mixture extract is shown in Figure 17, a screenshot of the software interface. It shows the database query results from a chromatogram of a mixed pulp extract against the reference database. For the database reference of *Paulownia tomentosa* an RSI of 88.7 % is given, 58 marker and 141 ambiguous peaks are matched. Additionally, 10.4 % of the peak area of the unknown chromatogram was identified as *Paulownia tomentosa*. Analysing these clear results, the authors correctly concluded that *Paulownia tomentosa* is contained in the mixture. *Tectona grandis* (15 matched marker peaks) and Fagus sylvatica (22 matched marker peaks) are also contained one-third each in the mixed sample and correctly identified. As the sample was randomly

mixed by a colleague, the composition was unknown to the authors until the analysis and identification decisions were made.

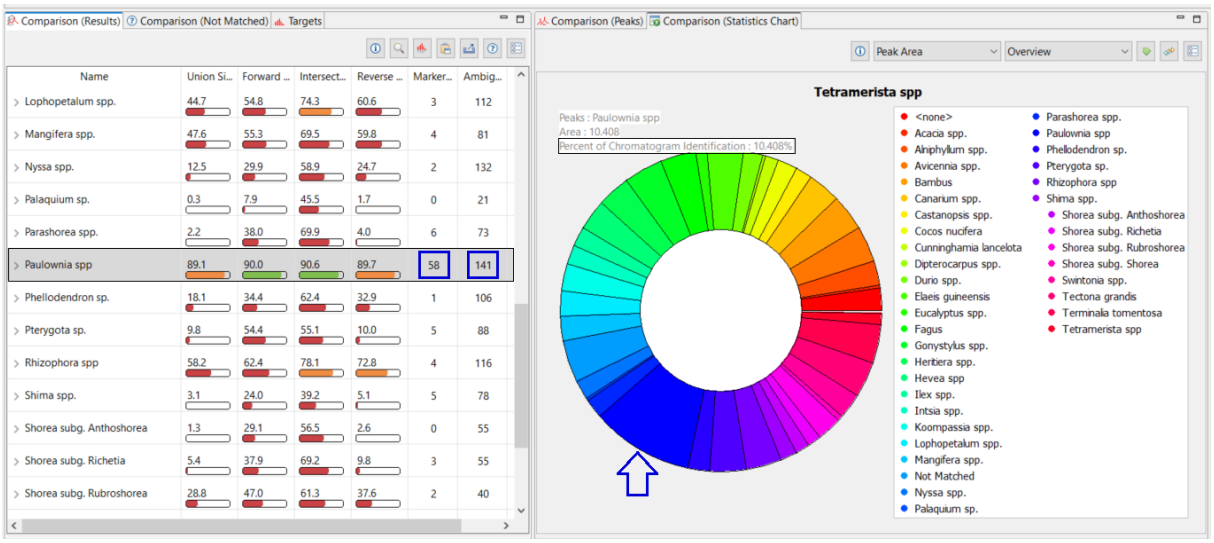


Figure 17: Database query results from a chromatogram of a mixed pulp extract (containing *Paulownia tomentosa*) against the reference database: Screenshot of the software interface.

3.7 Limitations

The authors acknowledge that limitations exist in the studies ability to comprehensively represent all possible variations within tree species. The researchers have taken these limitations into account.

Due to the requirement of 2 kg of material for pulp production, it was not possible to obtain samples with documented origin for most of the woods. As a result, there could be variations within tree species/genera that the current database is not capable of identifying, particularly in cases where trees are grown under extreme conditions.

In conclusion, the inability to obtain samples with a documented origin for most of the wood genera/species, along with the practical constraints on the number of samples analysed, constitute limitations of the study. They highlight the future need for expanding the database to account for variations in tree species grown in diverse environments.

4 Conclusion

The objective of this research was to develop a chemotaxonomic technique for identifying tropical wood species in pulp and paper using a chemical fingerprint database. As a result, it enabled the use of chemotaxonomy based on extractives and thus, a new method for extracting and analysing the chemical components of wood pulps. The grinding and extraction methodology used was selected on the basis of two important criteria. Firstly, it should ensure a good extraction of pulps. This means that a large number of chemotaxonomically relevant extractives is obtained ensuring sufficiency for the next steps of the process. Secondly, the methodology should be simple and reproducible so that a large number of prospective samples can be processed quickly and reliably. With these criteria in mind, the combination of grinding with the cryo-ball mill and the extraction with the Soxtherm extractor using n-hexane was chosen. Although the extractable content did not give the highest quantity in all tested options, many extracts were captured and the reproducibility was of high standard. Moreover, such a combination is an advantageous application since its grinding and extraction processes are partly automated.

Furthermore, all extractives within this method are measured in a standardized manner with the same temperature program by TD-GC-MS. This ensures pre-processing the obtained data continuously in the same way using a standardized batch process. A database for pulp extract chromatograms has been built up from 38 reference chromatograms. When matching an unknown pulp extract chromatogram with the database, several factors such as the RSI value and the number of marker peaks as well as ambiguous peaks are crucial. It shall be added that the ChromIdent® software tool for matching unknown samples with the database will be further developed and improved in the future making this method even more reliable.

New wood species are constantly coming into focus and gain relevance due to increased use in paper production and in connection to new classifications of CITES protection statuses. Thus, the current chemotaxonomic n-hexane extractives database needs to be extended on a regular basis. The larger the reference database, the more powerful the method will become. In perspective, its performance will be tested in a blind test of pulp samples with unknown composition and on other GC-MS systems to validate the method. Additionally, the potential of the acetone extracts will be included for increased resolution of critical species. Future investigations will also focus on chemotaxonomic variations among different provenances of wood. This is essential for the protection of natural forests, as no other method can distinguish between pulp from plantations and pulp from natural forests yet. In general, the database could even be extended to include wood products such as particle board, which is laborious for anatomists to work with, thus widening the usage of this method. All in all, although more work can and should develop this method even further, it already contributes to supporting the EUDR and sustainable forestry by adding new possibilities for its protection.

Flaig, M.L. et al. (2023). *Holzforschung* 77: 860–878,

Acknowledgments: PD Dr. habil. Jürgen Odermatt, who initiated the research for this paper and passed away suddenly in 2019, deserves special thanks. The authors would also like to thank Birte Buske, who supported with laboratory work, as well as Othar Kordsachia, Nils Grützmann and Alina Wassink for their exploratory work on pulping.
Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: This work was funded by the German Environmental Foundation (DBU) (Grant/Award Number: AZ 34295/01) in connection with the project "Detection of Tropical Hardwood in Paper – Chemotaxonomy and Anatomy for the Identification of Mixed Tropical Hardwood".
Conflict of interest statement: The authors declare that they have no conflicts of interest regarding this article.

References

All **References** are listed at the end of the dissertation on page 134.

3.2 Publication II: Chemotaxonomic and anatomic wood species identification in bleached pulp: blind test and method validation

Original Article

Max L. Flaig, Jens Berger, Stephanie Helmling, Andrea Olbrich, Heinz Joachim Schaffrath, Daniel Zahn and Bodo Saake*

Chemotaxonomic and anatomic wood species identification in bleached pulp – Blind test and method validation

http://doi.org/10.1515/hf-2024-0025
Received April 04, 2024; accepted July 3, 2024;
published online August 1, 2024

*__Corresponding author: Bodo Saake,__ Institute of Wood Science, Chemical Wood Technology, University of Hamburg, Haidkrugsweg 1, 22885 Barsbüttel-Willinghusen, Germany, E-mail: bodo.saake@uni-hamburg.de
Max L. Flaig and Jens Berger, Institute of Wood Science, Chemical Wood Technology, University of Hamburg, Haidkrugsweg 1, 22885 Barsbüttel-Willinghusen, Germany https://orcid.org/0009-0000-0585-1734 (Flaig, M.L.)
Stephanie Helmling and Andrea Olbrich, Johann Heinrich von Thünen Institute (TI), Institute of Wood Research, Leuschnerstraße 91, 21031 Hamburg, Germany
Heinz Joachim Schaffrath, Papierfabrikation und Mechanische Verfahrenstechnik, TU Darmstadt, Alexanderstr. 8, 64283 Darmstadt, Germany
Daniel Zahn, Physikalische Analytik und Materialprüfung, Sterilisierverpackung, ISEGA Forschungs- und Untersuchungsgesellschaft mbH, Zeppelinstraße 3, 63741 Aschaffenburg, Germany

Abstract: This paper presents a comparative analysis of the blind test outcomes of two independent methods for the identification of tropical wood species in pulp and paper products. Both, the established anatomical and the relatively new chemotaxonomic method support the European Deforestation Regulation 2023/1115 (EUDR), which aims to ensure that only legally harvested timber that has not contributed to deforestation is traded in the EU. The blind test involved 570 decisions on 15 test sheets of 37 self-manufactured mixed tropical hardwood pulps and an industrial beech pulp, used as a matrix. Both detection techniques demonstrated robust performance with over 80 % hit rates. The results show that the synergies and combination of the strengths of both methods can be utilized and lead to even better combined performance. In order to establish the chemotaxonomic identification method as a complement to the conventional anatomy-based method, statistical analyses were performed to assess its intermediate precision between three different GC-MS systems. In most cases, the method gave consistent results independent of the instrument used.

Keywords: CITES; database; GC-MS; mixed tropical hardwood; pulp and paper, wood anatomy.

1 Introduction

Pulp and paper are primarily derived from renewable wood resources. The raw material plays a pivotal role in sustainable paper production. To address concerns related to biodiversity conservation and the preservation of crucial rainforest ecosystems, it becomes imperative to substantiate the identification of tropical wood species in pulp and paper products, along with the recognition of CITES-listed timber species in global trade. In alignment with the European Union's Deforestation Regulation 2023/1115 (EUDR), the placement of unlawfully harvested timber on EU internal markets is explicitly prohibited (EU 2023). This regulation should replace the EUTR that has been in effect since 2013 and is essentially an additional extension to deforestation-free agricultural supply chains (Köthke et al. 2023). But still the regulation's purview encompasses timber, timber products, as well as pulp and paper, beginning at the point of their initial introduction to the EU markets.

In light of the substantial consumption of wood resources in paper manufacturing (Windhagen et al. 2019), the ability to ascertain the composition of paper products assumes paramount significance. An essential objective shared by both anatomical and chemotaxonomic methodologies is to systematically bolster the enforcement of the EUDR. Adherence to this regulation necessitates market participants to exercise due diligence.

To date, the lack of accessible reference samples of relevant species and analytical methodologies designed for the verification of manufacturer-supplied data has been constantly shrinking. Forensic identification techniques for wood, can broadly be categorized as anatomical, chemical, and genetic methodologies (Flaig et al. 2023; Low et al. 2022; Lowe and Cross 2011; Schmitz et al. 2020). Solid wood is generally easier to identify using different methods with similarly good results (Ravindran and Wiedenhoeft 2020). Even some endangered tree species can already be classified using deep learning methods (Zheng et al. 2024). Wood-based products such as pulp, paper and particle board tend to be more difficult to identify than solid wood (Sieburg-Rockel and Koch 2020). The selection of the most appropriate technique depends on the specific identification inquiry in question (Braga et al. 2020; Schmitz et al. 2019; Schmitz et al. 2020). Notably, for the genus-level identification of pulp and paper products, exclusively anatomical and chemotaxonomic approaches prove applicable. Wood anatomy provides robust performance, delivering accurate results when applied on commercial paper samples at the genus taxonomic level. It also benefits from a well-established and extensive database of reference materials (Wheeler 2011; InsideWood 2004 onwards; Richter and Dallwitz 2000 onwards), which can be supported by deep learning techniques (Nieradzik et al. 2024). Chemotaxonomy based on pulp extractives/ chromatographic fingerprint data represents a relatively nascent field, with its database currently containing a modest 38 reference extracts from 37 tropical plant pulps and 1 beech pulp (Flaig et al. 2023). For the validation of the chemotaxonomic identification method of wood taxa in pulp and paper, it is crucial that the method is independent of the GC-MS measurement system used. In general, the variability between repeated measurements is referred to as precision. The intermediate precision lies between the reproducibility with

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

maximum variability, defined as “the closeness of agreement between independent results obtained with the same method on identical test material but under different conditions (different operators, different apparatus, different laboratories and/or after different intervals of time)”, and the repeatability with minimum variability under constant conditions (Gold 2019; ISO 1994).

Across the entirety of the paper manufacturing process, which encompasses stages ranging from pulping to various bleaching procedures, the chemicals employed in these processes effectively obliterate the DNA originally present in solid wood. Although anatomical identification through the examination of morphological attributes of wood vessels has been well-established (Helmling et al. 2016; Helmling et al. 2018; Ilvessalo-Pfäffli 1995), complications can arise, especially when fibres undergo modification through refining – a common occurrence in paper production. This refining process typically results in the destruction of residual vessel elements (Helmling et al. 2018). Furthermore, some wood genera exhibit limited distinctive anatomical structural features, making them susceptible to confusion with other genera, including some that are taxonomically distant (Gasson 2011). A significant advantage of the chemotaxonomic approach lies in its immunity to the mechanical alterations of anatomical features, as can occur during fibre refining. In return identifying wood genera in pulp through chemotaxonomic methods is challenging due to the minimal extractive content. In industrial pulping and bleaching processes, extensive efforts are made to eliminate wood extractives to enhance paper quality and streamline production. Moreover, the composition of the extractives in pulps differs from that in solid wood.

Typically, chemotaxonomy, or fingerprint analysis, has been employed to classify and distinguish plants based on their secondary compounds (Alaerts et al. 2014). While prior chemotaxonomic research primarily focused on solid wood, this study conducts chemotaxonomic and anatomical analysis on the pulp and paper derived from tropical and temperate wood genera (including the three monocot plants Bamboo, Coconut and Oil palm).

2 Materials and methods

Obtained from various mainly commercial sources (lacking documented origin), a collection of 37 distinct solid mixed tropical hardwood (MTH) samples (Table 7, first column), each weighing around 2 kg, was amassed. The relevant genera and species were chosen in cooperation with the NGOs Greenpeace and World Wide Fund for Nature (WWF). The focus of the wood selection was on the presence of woods in South East Asia, as this region and its wood species were found to be of particular relevance in the research carried out as part of the "Detection of Tropical Hardwood in Paper – Chemotaxonomy and Anatomy for the Identification of Mixed Tropical Hardwood" project (Berger et al. 2021). Wood anatomists, aided by the resources of the Xylothek at the Thünen Institute in Hamburg, Germany, carried out the morphological identification of the genera of these wood samples.

2.1 Blind test

Subsequently, each specific reference wood belonging to a different genus/species underwent a pulping and bleaching process. The wood samples were cut into 3 x 3 cm chips and kraft pulped using a 7-liter M/K digester (M/K Systems INC., Williamstown, USA). Pulping involved 550–1 000 g wood chips, 4:1 liquor-to-wood ratio, and 22–25 % NaOH with 35 % sulfidity. The process included 90 minutes at 165 °C. The produced pulps underwent slot screening and a 5-stage bleaching sequence, targeting 90 % ISO brightness. Bleaching stages included oxygen (O), complexing agent (Q), oxygen-enhanced peroxide (OP), chlorine dioxide (D), and peroxide stages (P). This laboratory-scale kraft pulping mirrors industrial standards, yielding modern ECF pulps. A detailed overview of pulping and bleaching methods as well as a discussion on the chosen parameters is given by Flaig et al. (2023).

The fully bleached kraft pulps were then utilized for the production of the blind pulp mixture samples. Therefore, the reference pulps (Table 7, first column) were entrusted to an external paper research institute (Hochschule München). There, the blind samples were created with specific compositions without revealing these to the participating working groups. A total of 15 blind samples (Figure 18a) were mixed according to the guidelines for the production of the blind samples by the external institute. From each blind sample round test sheets (Figure 18b) were produced.

2.2 Guidelines for the production of the blind samples

- The beech pulp served as the matrix pulp into which the other pulps were mixed.
- All pulps were included at least once in a blind sample with a minimum of 5 %, but other proportions such as 10 % and 20 % were also used.
- As *Shorea* and the 3 former subgenera (*Anthoshorea*, *Rubroshorea*, and *Richetia*) are difficult to distinguish, they were included in the samples also as mixtures.
- Due to its CITES II protection status, at least 4 samples containing *Gonystylus* spp. were prepared.

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

- *Gonystylus* spp. can easily be confused with *Durio* spp. and *Lophopetalum* spp. and were therefore included in the samples separately and in mixtures.
- *Swintonia* spp. was used when *Durio* spp. was used because the *Durio* spp. pulp was contaminated with traces of *Swintonia* spp.
- Until the evaluation of the blind test, the information on the composition was kept secret.

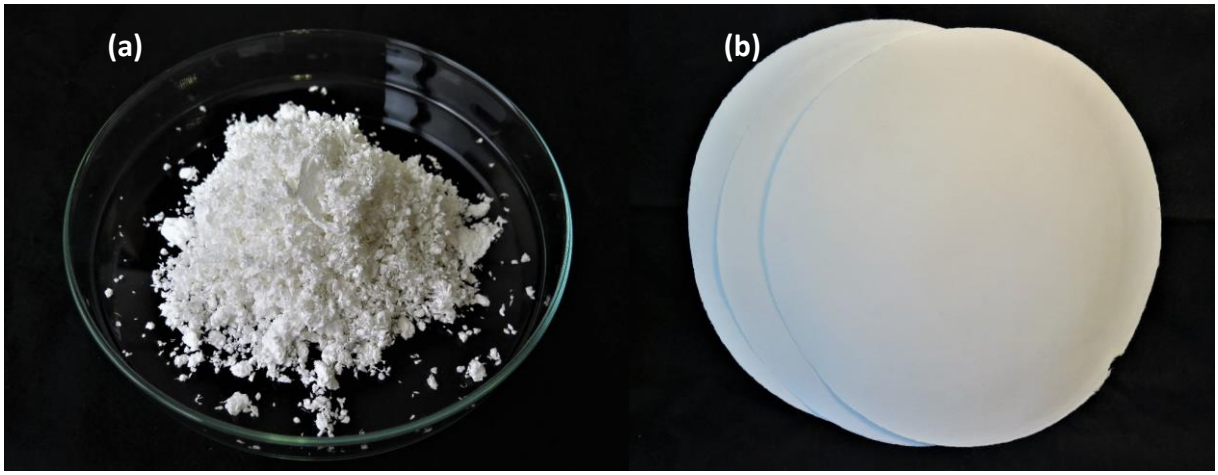


Figure 18: Blind test: (a) exemplary blind pulp mixture sample (b) test sheets.

All participating entities, including Thünen Institute (TI), Technical University of Darmstadt (TUDa), ISEGA and University of Hamburg (UHH), were informed about these guidelines. The test sheets were divided and made available to the participating institutes. Additionally, the TUDa and ISEGA received the "Atlas of vessel elements" in digital format (Helmling et al. 2018), along with descriptions of unpublished references for six pulps.

2.3 Chemotaxonomy

2.3.1 Grinding

The blind test sheets were milled in a CryoMill (Retsch GmbH, Haan, Germany), a liquid nitrogen-cooled ball mill, at a frequency of 25 Hz. Each sample was milled five times for 1 minute, with cooling applied for 0.5 minutes between cycles according to Flaig et al. (2023).

2.3.2 Extraction

Freshly distilled n-hexane (VWR International, LLC., Radnor, USA), equivalent to GC-grade, was used. Extraction was performed using a Soxtherm SOX 6 apparatus (C. Gerhardt GmbH, Königswinter, Germany) with an automated MULTISTAT control system. The ground blind test sheets (5 g per cellulose extraction thimble) were extracted with 140 ml of solvent. The extraction parameters described in detail by Flaig et al. (2023) resulted in a total extraction time of 4 h 7 min. Extracts were then adjusted to a final volume of 50 ml.

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

2.3.3 Sample preparation

The extracts were applied to the pyrolysis crucibles for TD-GC-MS analysis by evaporating the solvents of the applied extract solutions. This process was repeated with adjusted volumes until the desired dry mass of 90 ± 5 µg per crucible was obtained.

For the analysis of the system independence of the chemotaxonomic method (section 3.2), 6 randomly selected mixed blind samples and 7 randomly selected pure samples were uniformly prepared and analysed as described above and below. The samples were prepared and analysed by the same operator in the same laboratory, with more than one year elapsing between measurements on GC-MS system A and systems B and C. From each sample extract, 8 pyrolysis crucibles were prepared and measured.

2.3.4 TD-GC-MS systems and parameters

The TD-GC-MS analysis systems **A** and **B** consisted of a Frontier Laboratories Ltd. micro furnace Double-Shot Pyrolyzer (Py-2020iD) with an Auto-Shot Sampler (AS-1020E). They were coupled to Agilent Technologies Inc. GC-MS: system **A** (6890N/5973N), system **B** (6890N/5973 inert). System **C** consisted of a Frontier Laboratories Ltd. Multi-Shot Pyrolyzer (EGA/Py-3030D) with an Auto-Shot Sampler (AS-1020E) connected to an Agilent Technologies Inc. GC-MS (8890/5977B).

The TD temperature was 325 °C, while the interface was set to 330 °C. For polyethylene (PE) pyrolysis, the pyrolizer operated at 500/600 °C. The GC inlet and the MS interface temperature were maintained at 320 °C. A 30 m × 0.25 mm i.d. and 0.25 µm film thickness low polarity column (ZB-5HT 7HG-G015-11, Phenomenex Inc., Torrance, USA) was used with helium as carrier gas, split ratio set to 20:1, and a constant flow rate of 1 ml/min.

The GC oven temperature started at 45 °C for 2 min, then increased to 340 °C with decreasing heating rates and was maintained at this temperature for 30 min. The total analysis time was 134.17 min. The GC oven temperature settings followed Flaig et al. (2023). An electron impact ionization energy of 70 eV was used for mass spectral detection. The measurement scanned a range of 29–700 m/z in total ion current (TIC) mode with a threshold of 100. Data was recorded with a scan frequency of 2.22 scans/second and sample rate [N = 2].

2.3.5 Data processing

Each blind test pulp extract underwent a minimum of two analyses. One representative chromatogram from each set was selected for subsequent analysis. The chromatogram was subjected to pre-treatment procedures such as rounding and smoothing of mass traces, peak deconvolution, integration and retention index (RI) calculation and finally matched against the reference database of single-variety wood pulp extracts (DB) following Flaig et al. (2023).

For the analyses of the system independence of the chemotaxonomic method (Section 3.2), the pre-processing method for the raw GC-MS data and the DB settings were slightly adjusted compared to the parameters of Flaig et al. (2023) used for the blind test (Section 3.1). Table 5

shows the updated DB settings and Table 6 a short version of the updated pre-processing steps. The full version of the pre-processing operations in tabular form is provided as Supplementary Table S1 and in OpenChrom Method (.OCM) file format for use with the software as Supplementary File S1.

The process of identifying a pre-treated unknown mixed or pure sample extract chromatogram containing information on the deconvoluted peaks with individual mass traces, RI and peak areas involved aligning the chromatogram using a RI corridor of ± 40, followed by searching for peak matches using the NIST algorithm (Stein and Scott 1994) based on mass spectral match quality. A minimum match factor (MF) of 70 % was used for peak comparisons due to potential impurities and low intensity mass traces of small peaks in the mixtures. The cosine algorithm (Alfassi 2004) was used for whole chromatogram comparison and the OpenChrom® software (version: 1.5.0.20231117-1619, Lablicate GmbH, Hamburg, Germany) was utilized for all pre-processing and DB operations.

Table 5: Updated pulp extract reference database settings.

Name	Description	Setting
Min Match Factor (MF)	Matches below the given value are ignored.	80.0 %
Min Reverse Match Factor	Matches below the given value are ignored.	80.0 %
Mass Spectrum Comparator	Comparison algorithm used.	NIST (Identity Normal)
Delta Window	RI corridor used for peak comparison.	34.0
Calculate Penalties		false
Retention Index (RI) Limit	The penalty is applied on all values outside of this RI window.	0.0
Penalty Level Factor	Calculates a MF penalty, depending on RI delta between peaks.	0.0
Max Penalty	Maximum penalty applied on the MF.	0.0 %

Table 6: Updated raw data pre-processing operations (short version). Underlined parts were adjusted.

Name	Description	Settings
Scan Cleaner (Remove Empty)	This filter removes empty scans.	
High Pass Ions	This filter preserves the n highest ions per scan.	{„Number Highest“:125}
Ion Round Method	This filter sets the system ion round method settings.	{„Ion Round Method“:“MINUS_0.38 (incl.) to PLUS_0.62 (excl.)“}
Nominalize (Unit Mass)	This filter condenses the scan to nominal mass.	
Ion Remover Filter	This filter removes all specified ions from a mass spectrum.	{„Mode“:“INCLUDE“, “Ions“:“18 28 32 44 84 207 281“}
SNIP	This baseline detector uses the SNIP algorithm.	{„Window Size“:13, “Number of Iterations“:120}
Chromatogram Baseline Subtract Filter	This filter enables to remove the baseline from the chromatogram.	
Chromatogram Selection (Select Range)	This filter selects the chromatogram range.	{…, “Start RT (Minutes)“:0.0,…, “Stop RT (Minutes)“:115.0}
Scan Duplicator	This filter duplicates series of scans.	{„Duplicated Scan (Merge Traces)“:true}
Savitzky-Golay Smoothing (4x)	This filter applies the Savitzky-Golay filter.	{„Filter Ions“:true,“Order“:2,“Width“:13}
First Derivative	Implementation of a first derivative peak detector.	
High Pass Peaks	Keep the n-highest peaks	{…, “Window Size“:13,…, “Optimize Baseline (W)“:true, “Min S/N Ratio“:100.0,…, “Threshold“:“HIGH“, “Filter Mode“:“EXCLUDE“}
MCR-AR [Targeted Chained]	This Peak Detector provides targeted peak deconvolution where targets are generated by first derivate peak detector	{„Filter Option“:“HEIGHT“, “Number Highest“:6} {„Settings“:{„0.DefaultProfile“:“{\“Local Minima Scan Slope Threshold\“:-0.001,\…:1,\…:0,\…:3}“, “1.DefaultProfile“:“{\“Minimum Analysis Segment Width\“:50}“, “2.DefaultProfile“:“{\“Local Maxima Scan Moving Average Window Size\“:3,\…:2,\…:1300,\…:0.001}“, “3.DefaultProfile“:“{\“Percentage Height after Peak\“:1.0,…:1.0,\…:50,\… “:11,\…:5,\… “:1.0,\…:1,\…:45}“}…}
Peak Integrator Trapezoid	This extension point tries to implement a ChemStation similar peak integrator.	{„Include Background“:false,“Area Constraint“:true, “Traces to Integrate“:“0“}
High Pass Peaks	Keep the n-highest peaks	{„Filter Option“:“AREA“, “Number Highest“:206}
Retention Index Calculator (embedded)	Calculates the retention indices for scans and peaks in the chromatogram.	Example: {…, “Extrapolate (Right)“:true,…, “Retention Index Marker“:“4.3260 | 800.0 | C8 (Octene);5.9690 | 900.0 | C9 (Nonene);…;125.5020 | 5100.0 | C51 (Henpentacontene)“, “Extrapolate (Left)“:true}
NIST (extern)	This plugin uses the NIST library to identify peaks.	{„NIST Folder (MSSEARCH)“…, “Min. Match Factor“:50.0, “Min. Reverse Match Factor“:50.0,…, “Number of Targets“:15}
OpenChrom Chromatogram (*.ocb)	Reads and writes OpenChrom chromatograms	{„File Name“:“{chromatogram_name}“, “Export Folder“:“…}

2.4 Anatomy

2.4.1 Maceration

For the purpose of comparing unknown paper samples and analysing the blind test samples in this paper, anatomical references were established. To accomplish this, samples of all reference woods were macerated according to Franklin (1945). The maceration made the cell structure, especially the wood vessel elements and their individual features, more visible under a microscope.

2.4.2 Sample preparation

The samples were prepared analogous to the description of the preparation of fibrous materials in the “Atlas of vessel elements” (Helmling et al. 2018). For the hardwoods, vessel elements were systematically separated from the samples, prepared on microscope slides, and embedded. Since softwood species do not contain vessel elements, the reference permanent slides for *Cunninghamia lanceolata* were prepared from its tracheids. Subsequent visual documentation was performed using a light microscope. Initially, microscopic images of at least 36 vessel elements from the hardwoods were taken in various focal planes. The images of the vessel elements were examined for their qualitative characteristics and systematically evaluated. The characteristics of the vessel elements that were analysed and documented are: vessel element description, tails, perforation plates, intervessel pits, vessel-ray pits, pits to axial parenchyma cells, areas without any pits, tyloses, vessel length and width plus their l/w ratio, intervessel pit borders in vertical diameter, pit apertures, fibber length, fibre width and fibre wall thickness.

2.4.3 Analysis

Different numbers of slides were made from each test sheet; TUDa made 5, TI 10 and ISEGA 10. The analysis involved comparing the structural characteristics of the vessel elements found in the blind test samples with the descriptions and images in the “Atlas of vessel elements” (Helmling et al. 2018). Morphological key features included the shape of the vessel elements and their tails, the type of vessel perforation plates, the presence of tyloses or helical thickenings, as well as the arrangement and shape of intervessel pits and vessel-ray pits. Using histometric measurements, parameters such as vessel length and width, as well as the size of the intervessel pit apertures and borders were determined from the images and compared with the descriptions in the atlas.

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

3 Results and discussion

3.1 Blind test

From the 37 self-manufactured single variety bleached kraft pulps and the industrial beech matrix pulp, 15 blind samples of defined composition were externally produced as described in section 2.1 without informing the authors of the composition. The aim for the participating working groups was to correctly identify the composition of the test sheets. In the blind test, 15 x 38 = 570 decisions had to be made by every participating testing laboratory. The results of three institutes using the anatomical method and one institute using the chemotaxonomic method are shown in Table 7 for every wood taxon.

Table 7: Blind test results for every institute and taxon (genus or species).

Taxon	Chemotaxonomy	Anatomy			
	UHH	∑	TUDa	ISEGA	TI
	% right	% right	% right	% right	% right
Acacia spp.	100	100	100	100	100
Alniphyllum spp.	73	98	93	100	100
Anthoshorea spp.	73	89	73	100	93
Avicennia spp.	87	98	93	100	100
Calophyllum spp.	80	98	93	100	100
Canarium spp.	87	93	100	100	80
Castanopsis spp.	87	89	67	100	100
Cocos nucifera	80	96	87	100	100
Cunninghamia lancelota	100	98	93	100	100
Dendrocalamus spp.	87	93	87	100	93
Dipterocarpus spp.	100	91	73	100	100
Durio spp.	87	87	73	100	87
Elaeis guineensis	100	96	87	100	100
Eucalyptus spp.	87	91	73	100	100
Fagus sylvatica	100	100	100	100	100
Gonystylus spp.	87	96	87	100	100
Heritiera spp.	67	87	87	93	80
Hevea spp.	80	89	87	93	87
Ilex spp.	87	98	93	100	100
Intsia spp.	100	89	80	100	87
Koompassia spp.	87	87	87	100	73
Lophopetalum spp.	80	98	93	100	100
Mangifera spp.	87	82	73	87	87
Nyssa spp.	60	98	93	100	100
Palaquium sp.	93	96	87	100	100
Parashorea spp.	100	91	93	87	93

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

Paulownia tomentosa	80	91	87	100	87
Phellodendron sp.	87	91	73	100	100
Pterygota sp.	73	82	73	87	87
Rhizophora spp.	73	100	100	100	100
Richetia spp.	100	82	73	93	80
Rubroshorea spp.	87	80	67	80	93
Schima spp.	80	98	93	100	100
Shorea spp.	87	91	73	100	100
Swintonia spp.	87	84	60	100	93
Tectona grandis	87	89	73	93	100
Terminalia tomentosa	87	91	73	100	100
Tetramerista spp.	93	96	87	100	100
Total hit rate	**86**	**92**	**84**	**98**	**95**

In addition to the results for each taxon, Table 8 shows the results for all members of the family Dipterocarpaceae (*Anthoshorea* spp., *Dipterocarpus* spp., *Parashorea* spp., *Richetia* spp., *Rubroshorea* spp., *Shorea* spp.) and separately for the former subgenera of *Shorea* alone (*Anthoshorea* spp., *Richetia* spp., *Rubroshorea* spp., *Shorea* spp.), which also belong to the family. They are particularly important because of their dominance in the tropical rainforests of Southeast Asia, with over 200 species (Bansal et al. 2019).

Table 8: Comparison of the percentage of correct choices by group: all taxa, the family Dipterocarpaceae and the former subgenera of the genus *Shorea*.

	Chemotaxonomy	**Anatomy**		
Group	**UHH (%)**	**TUDa (%)**	**ISEGA (%)**	**TI (%)**
All taxa	86	84	98	95
Dipterocarpaceae family	91	76	93	93
Former *Shorea* subgenera	87	72	93	92

3.1.1 Chemotaxonomy

The chemotaxonomic analysis of the test sheets was performed on the GC-MS system A – the same system on which the DB references were measured. The identification process relied on the marker peak score (MPS), peaks unique to a single reference, and ambiguous peaks occurring in multiple references. The reverse similarity index (RSI) was also crucial when comparing an unknown extract to the DB, as the RSI assesses the extent to which the library peaks of a single DB reference match the peaks in the mixed sample (Figure 22). The matched chromatogram area (MCA), i.e. the percentage of the total unknown peak area that was identified as a specific reference (Figure 19c), and the comparison factors mentioned above, required extensive analysis to make an informed decision.

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

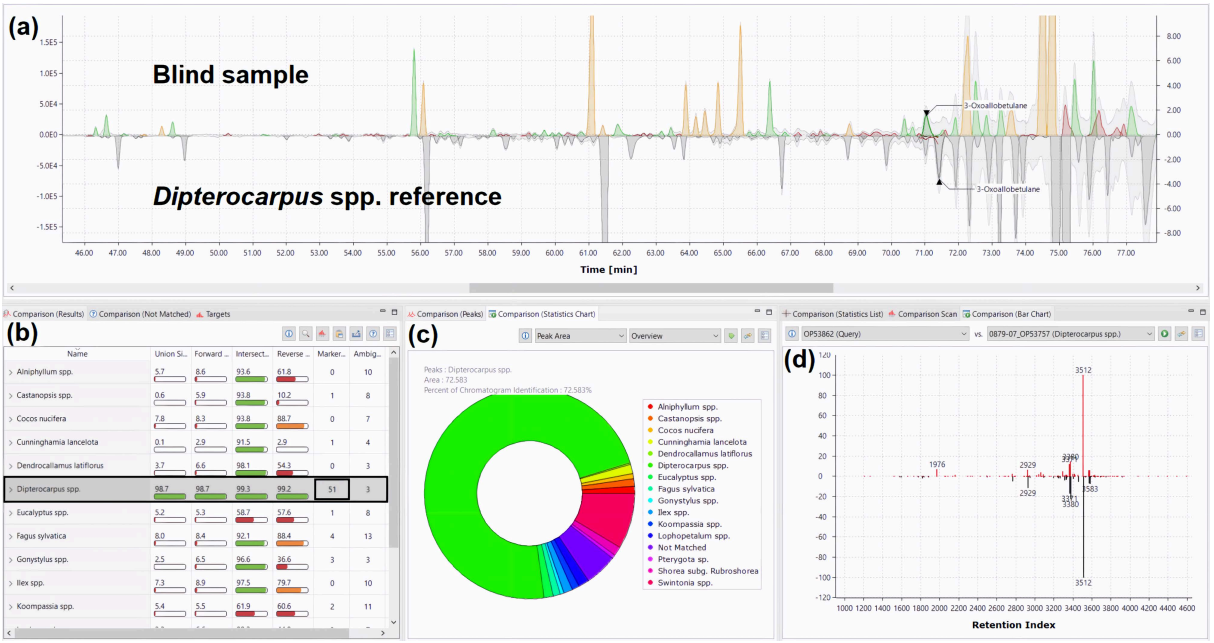


Figure 19: Database query results from the blind sample 07 (containing *Dipterocarpus* spp.) against the reference DB: screenshot of the software interface: (a) chromatogram comparison (b) statistic query results (c) graphical area distribution (MCA) (d) normalized peak area comparison aligned by RI.

The screenshot of the software interface (Figure 19), displays the outcomes of a query performed on blind sample 07, compared to the reference DB. A visual comparison tool (Figure 19a) shows the unknown chromatogram at the top with red unmatched, yellow ambiguous and green marker peaks compared to the grey DB reference chromatogram of *Dipterocarpus* at the bottom. The retention times of both chromatograms are shifted in the figure, but this was corrected for analysis by RI calculation and alignment. For *Dipterocarpus* an RSI of 99.2 % is reported, along with 51 matched marker peaks and 3 ambiguous peaks (Figure 19b). Furthermore, a substantial portion, specifically 72.6 %, of the peak area within the chromatogram of the unknown pulp extract was attributed to *Dipterocarpus* (Figure 19c), which the authors correctly concluded to be present in the mixture after analysing these distinct findings.

In the end eight woods were identified 100 % correctly as present or absent in the 15 blind samples and an overall blind test evaluation rate of 86 % correct decisions was achieved (Table 7). Nevertheless, there were certain wood species/genera that were more difficult to identify. These included *Avicennia*, *Calophyllum*, *Castanopsis*, *Eucalyptus*, *Hevea*, *Ilex*, *Koompassia*, *Lophopetalum*, *Nyssa*, *Pterygota*, *Rhizophora* and *Anthoshorea*. These genera were never clearly identifiable with the current DB version. There are several possible reasons for this: One is that the chemical fingerprints were not always specific enough to be separated from all other references. Although isolated markers pointing to these genera were sometimes detected in some blind samples, the expressions of these markers were too weak to make clear decisions. Another reason could be that even if a specific marker substance was detected in a pure reference extract, its amount/peak size may be too low in a mixture containing only

a small proportion of that pulp, so that the marker signal falls below the detection threshold of the MS. In the evaluation of the blind test, the Dipterocarpaceae family as a whole was distinguished from representatives of other families by clear marker substances (Table 8). However, the close relationship between members of the Dipterocarpaceae family, especially the former *Shorea* subgenera, was evident in the chromatograms. Nevertheless, *Richetia* was 100 % correctly identified in the blind test. To refine the identification of the other Diptocarpaceae members, a separate database containing only these very similar species could be created in the future. It may then be possible to narrow down the matching parameters even further, thereby increasing the accuracy of the matching results.

The authors acknowledge that the present chemotaxonomic approach requires additional time and involves a more extensive laboratory process when compared to the effort required by the anatomical method.

3.1.2 Anatomy

The blind test participants using the anatomical method prepared coloured microscopic specimens from the received 15 test sheets and examined them under a light microscope. The effectiveness of identifying individual taxa is remarkable (Table 7). For instance, in the case of the genus *Rhizophora*, known for its distinctive and less common features (scalariform perforation plates and scalariform intervessel pits), all participants correctly identified the genus as present or absent in the 15 blind samples with 100 % accuracy. This was the case for only 3 taxa in total. However, 20 additional taxa were identified with 100 % accuracy by two of the three participants, and the third participant made only one or two incorrect identifications for most of these taxa. These taxa can be reliably identified by anatomical analysis, considering the anatomical references and the characteristics of each taxon. In contrast, the former subgenera of the genus *Shorea* and the genera *Dipterocarpus* and *Parashorea* were described by all participants as challenging to differentiate because they appear very similar in terms of the dimensions of the vessel elements and the arrangement of the pits. In particular, the reclassified genera *Richetia* and *Rubroshorea* could not be correctly identified with 100 % accuracy by any of the participants. When evaluating these genera of the family Dipterocarpaceae separately, the success rates were lower than the average for all taxa to be determined (Table 8). Therefore, for the genus *Shorea*, as described by Helmling et al. (2018), one should be aware of the high risk of confusion within the family and should refrain from making determinations down to the former subgenus level.

Other genera with challenging anatomical differentiation included *Mangifera* and *Swintonia*, both belonging to the Anacardiaceae family. There were also groups of genera, that were not even closely related, but shared highly similar anatomy of vessel elements and therefore had a high risk of being confused. Of particular importance was the group containing the genera *Durio* (Figure 20a), *Lophopetalum* (Figure 20b) and CITES Annex II protected *Gonystylus* (Figure 20c). It's worth noting that the rate of correct identifications for the genus *Gonystylus* in the 15 blind samples was 100 % by two institutes. However, when all three genera were

considered together, the rates were 96 %, 87 % and 100 % for only one participant. The genera *Koompassia* and *Heritiera* also show a striking similarity in vessel anatomy. In a blind sample containing *Koompassia*, only one participant correctly identified this genus, while all participants initially misidentified it as *Heritiera*. It's important to note that real samples may contain other wood species for which references are not yet available, and these species may also exhibit high similarity to *Gonystylus*. The detailed descriptions and figures of the microscopic images of all these challenging genera are described in the Atlas of vessel elements (Helmling et al. 2018). Their similarity can be seen there. If only pure samples were to be analysed, identification could be made more clearly by the total number of vessel elements belonging to only one genus. Identification in mixed samples is therefore more difficult because each vessel element has to be evaluated individually, if possible. If a vessel element with a meaningful pit arrangement is turned sideways on the slide or shows no meaningful features at all, it cannot be assigned to a single genus.

As TI and ISEGA microscoped 10 slides per sample and achieved better results than TUDa who only prepared 5 slides, this suggests that a larger number of slides can increase the accuracy of the results.

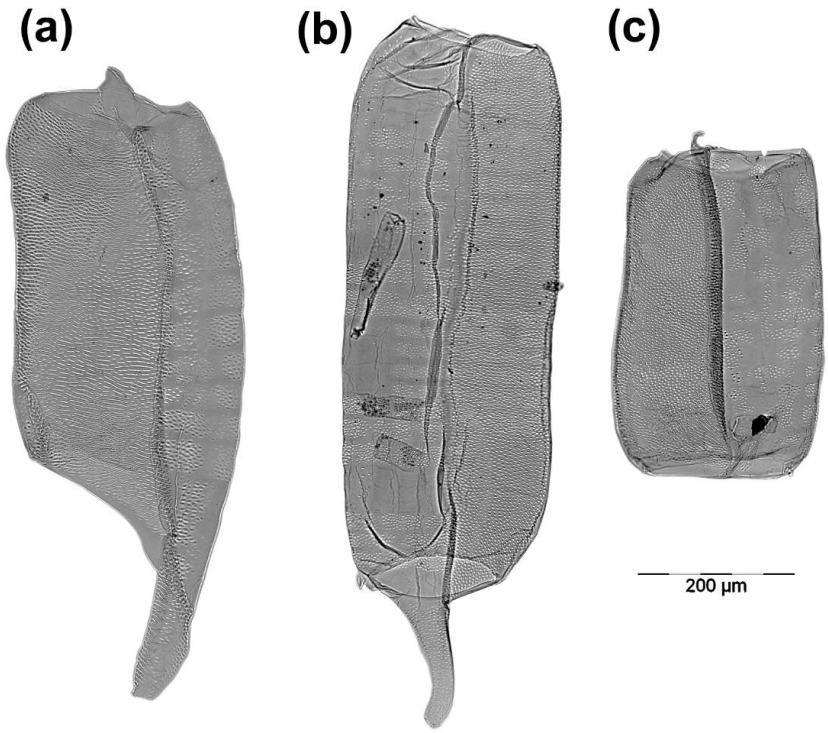


Figure 20: Microscopic images of vessel elements: (a) *Durio* spp. (b) *Lophopetalum* spp. (c) *Gonystylus* spp. (Helmling et al. 2018).

3.2 System independence of the chemotaxonomic method

The independence from the analytical instrument used is crucial for the practical application of the chemotaxonomic method. Therefore, three parameters (MPS, RSI and MCA) were analysed in two different types of samples (mixed and pure) on three different instruments

(TD-GC-MS systems A, B and C) as schematically illustrated in Figure 21. As in this investigation the laboratory and operator remained constant, but the GC-MS systems used were changed, neither repeatability nor reproducibility was determined, but what is known as intermediate precision.

From each mixed sample extract two pyrolysis crucibles were measured on GC-MS system A and three on each of the systems C and B. All 8 chromatograms per sample underwent identical pre-processing as described in 2.3.5 Data processing and were matched against the DB. Each DB query gave scores for MPS, RSI and MCA for each of the 38 reference extracts included in the DB. MPS, RSI and MCA are most important for the identification of wood taxa and therefore need to be of high precision across different instruments. The most promising candidates predicted in the blind test identification were also selected for the statistical analysis of the system independence. Therefore, a selection of 6 blind samples and 1–5 MPS of potentially containing reference taxa per blind sample chromatogram were analysed. A total of 16 MPS times 8 chromatograms per sample equals 128 MPS, divided into the systems A: 32, B: 48 and C: 48 (Figure 21). The same applies to the RSI and MCA values, i.e. a total of 384 data points of the mixed samples were analysed. The results are shown in Table 9–11.

For each GS-MS system (A, B and C) the mean values of the three parameters studied are given. As suggested by (ISO 1994; Zwanziger and Sorkau 2020; Gold 2019) the intermediate precision variance (Var), intermediate precision standard deviation (SD) and intermediate precision relative standard deviation or coefficient of variance (CV) were calculated to analyse the system independence of the method. The overall CV is given at the end of each table as an average of all calculated CV values. In addition, Analysis of Variance or ANOVA was conducted to assess whether the means of the three groups (GC-MS systems) were significantly different from each other. It accomplished this by partitioning the total variance observed in a dataset into different components: variance **between** groups and variance **within** groups. If the variance between groups was significantly greater than the variance within groups, ANOVA suggested that there were differences among the means of the groups. The null hypothesis (H0) stated that there was no significant difference among the group means.

If ANOVA yielded **non**-significant differences, this indicated that there was insufficient evidence to reject the H0. In other words, the observed differences between the groups were within the range of what could be expected due to random chance alone. This meant that any observed differences were not statistically significant based on the chosen significance level (p) of 0.05 and H0 was accepted.

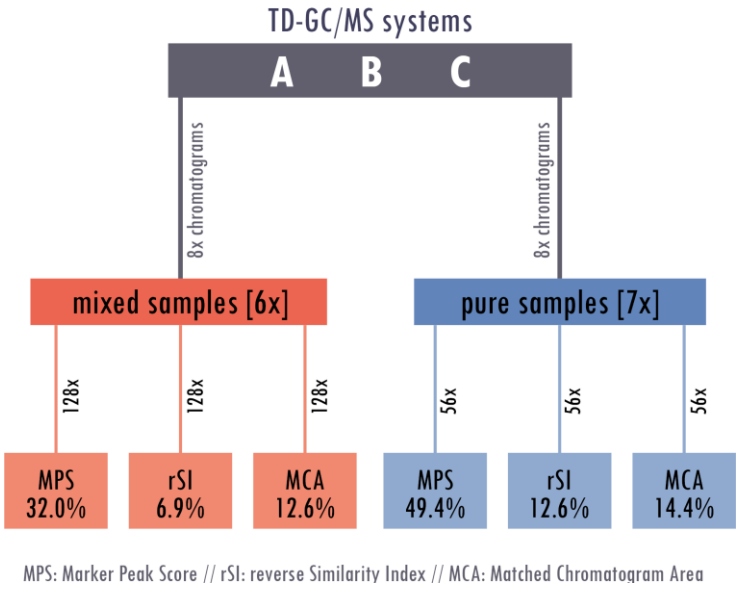


Figure 21: Schematic illustration of the variance analysis of the system independence of the chemotaxonomic method.

Table 9: Statistics results for marker peak scores (MPS) of **mixed** samples (MS).

No.	Sample	Analyzed taxon	GC-MS system			Total marker peak scores				
			A	B	C	∑ABC				
			Mean	Mean	Mean	Mean	Var	SD	CV	ANOVA
1	MS-07	*Dipterocarpus*	61.0	38.3	45.7	46.75	94.50	9.72	21 %	$F(2, 5) = 39.54$, $p < 0.05$
2	MS-07	*Fagus*	3.0	2.3	2.3	2.50	0.57	0.76	30 %	$F(2, 5) = 0.65$, $p = 0.562$
3	MS-07	*Lophopetalum*	6.5	8.0	11.7	9.00	6.86	2.62	29 %	$F(2, 5) = 8.25$, $p < 0.05$
4	MS-05	*Acacia*	12.5	7.3	8.3	9.00	8.00	2.83	31 %	$F(2, 5) = 3.91$, $p = 0.095$
5	MS-05	*Canarium*	4.0	6.7	4.3	5.13	2.70	1.64	32 %	$F(2, 5) = 3.93$, $p = 0.094$
6	MS-02	*Castanopsis*	4.0	5.3	4.3	4.63	1.13	1.06	23 %	$F(2, 5) = 1.19$, $p = 0.377$
7	MS-02	*Palaquium*	4.0	4.7	6.7	5.25	2.21	1.49	28 %	$F(2, 5) = 4.77$, $p = 0.069$
8	MS-02	*Schima*	2.5	3.3	2.3	2.75	1.07	1.04	38 %	$F(2, 5) = 0.71$, $p = 0.534$
9	MS-02	*Richetia*	6.5	10.0	6.7	7.88	6.98	2.64	34 %	$F(2, 5) = 1.99$, $p = 0.230$
10	MS-02	*Shorea*	3.0	3.3	4.0	3.50	0.57	0.76	22 %	$F(2, 5) = 1.25$, $p = 0.363$
11	MS-12	*Fagus*	32.5	7.7	12.3	15.63	114.55	10.70	68 %	$F(2, 5) = 201.4$, $p < 0.05$
12	MS-13	*Acacia*	5.0	3.0	5.0	4.25	2.79	1.67	39 %	$F(2, 5) = 1.56$, $p = 0.297$
13	MS-13	*Elaeis*	1.5	2.3	1.7	1.88	0.41	0.64	34 %	$F(2, 5) = 1.42$, $p = 0.325$
14	MS-13	*Terminalia*	1.5	1.7	1.7	1.63	0.27	0.52	32 %	$F(2, 5) = 0.06$, $p = 0.945$
15	MS-15	*Cunninghamia*	61.0	41.0	57.3	52.13	152.13	12.33	24 %	$F(2, 5) = 3.36$, $p = 0.119$
16	MS-15	*Tectona*	5.0	6.0	4.0	5.00	1.71	1.31	26 %	$F(2, 5) = 2.50$, $p = 0.187$
	Mean								32.0 %	

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

Table 9 shows the mean marker peak scores of the systems A, B and C and the overall statistics (∑ABC) per sample and taxon in the mixed samples. The overall CV of 32 % indicates that the variation of the MPS between all the chromatograms tested was quite high, but interestingly the F-tests of the ANOVAs were not significant in 13 of the 16 cases tested (*p*-value > 0.05). This means that the differences between most of the group means were smaller than the differences within the groups. The H0 – groups are equal – was accepted in most cases. Most of the mean values of the MPS for the three GC-MS systems are therefore not statistically significantly different, i.e. GC systems A, B and C are statistically equal, although the total variance is quite high! For example, the mean MPS of system C of 4.3 for *Castanopsis* in the mixed sample 02 (Table 9) is composed of three single values 4, 3 and 6, giving a CV of 35 % within this group. The total CV of all individual values of all groups A, B and C for MS-02 *Castanopsis* is much smaller – only 23 % (Table 9). In this case, looking only at the group means; 4.0, 5.3 and 4.3, they have an even smaller CV of 15 %. Therefore, it makes sense that the ANOVA also concludes that the variation within groups is higher than between groups. Raw data and calculations can be found in the Supplementary Tables S2.

As explained in detail by Flaig et al. (2023), the RSI is the most important similarity index for the identification of mixed samples in particular. Figure 22 shows that only the peaks that are present in the DB reference chromatogram are compared with the unknown peaks from the mixed sample to calculate the RSI value.

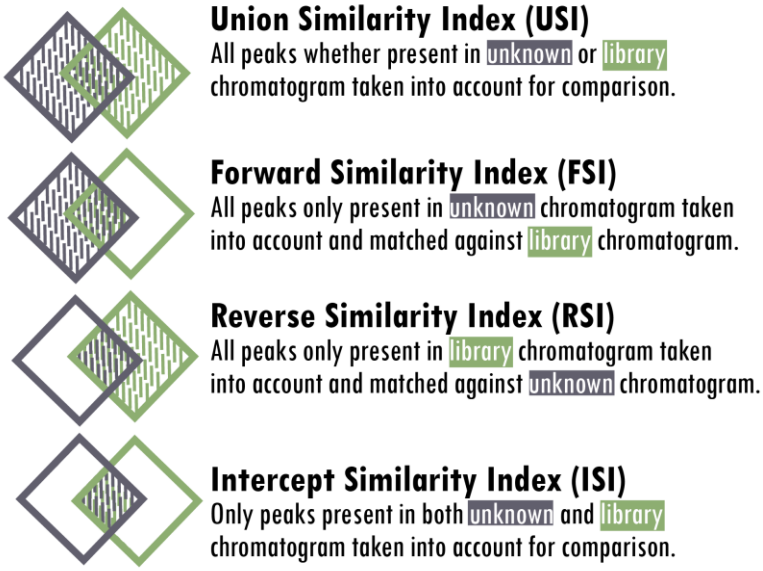


Figure 22: Graphical explanation of the similarity indices.

The results in Table 10 clearly show that the precision of the RSI values in the mixed samples was very accurate with an overall CV of 6.9 %. There were also no statistically significant differences in the RSI values for the three different GC-MS systems, except for MS-12 Fagus. For this sample, the *p*-value for the systematic measurement error of the different GC-MS systems was 0.0005, well below the 0.05 significance level, which means that H0 had to be rejected and the alternative hypothesis (H1) – groups differ – was accepted in this case.

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

Table 10: Statistics results for reverse similarity index (RSI) of **mixed** samples.

Sample	Analyzed taxon	GC-MS system			Total reverse similarity index				
		A	B	C	∑ABC				
		Mean	Mean	Mean	Mean	Var	SD	CV	ANOVA
MS-07	*Dipterocarpus*	0.99	0.99	0.99	0.99	0.00	0.01	1 %	F(2, 5) = 0.01, p = 0.989
MS-07	*Fagus*	0.49	0.45	0.45	0.46	0.00	0.02	5 %	F(2, 5) = 2.84, p = 0.150
MS-07	*Lophopetalum*	0.51	0.49	0.49	0.49	0.00	0.01	3 %	F(2, 5) = 1.55, p = 0.230
MS-05	*Acacia*	0.98	0.99	0.98	0.99	0.00	0.01	1 %	F(2, 5) = 2.81, p = 0.152
MS-05	*Canarium*	0.93	0.94	0.94	0.94	0.00	0.00	0 %	F(2, 5) = 2.00, p = 0.223
MS-02	*Castanopsis*	0.47	0.46	0.46	0.46	0.00	0.00	1 %	F(2, 5) = 0.23, p = 0.804
MS-02	*Palaquium*	0.44	0.44	0.44	0.44	0.00	0.00	1 %	F(2, 5) = 1.96, p = 0.235
MS-02	*Schima*	0.48	0.37	0.36	0.39	0.00	0.07	17 %	F(2, 5) = 4.12, p = 0.088
MS-02	*Richetia*	0.99	0.99	0.99	0.99	0.00	0.00	0 %	F(2, 5) = 4.83, p = 0.068
MS-02	*Shorea*	0.95	0.95	0.95	0.95	0.00	0.00	0 %	F(2, 5) = 0.78, p = 0.507
MS-12	*Fagus*	0.86	0.76	0.82	0.81	0.00	0.04	5 %	F(2, 5) = 48.20, p < 0.05
MS-13	*Acacia*	0.87	0.95	0.89	0.91	0.00	0.04	5 %	F(2, 5) = 4.39, p = 0.079
MS-13	*Elaeis*	0.95	0.80	0.94	0.89	0.03	0.16	18 %	F(2, 5) = 0.65, p = 0.560
MS-13	*Terminalia*	0.74	0.64	0.62	0.66	0.03	0.16	24 %	F(2, 5) = 0.28, p = 0.767
MS-15	*Cunninghamia*	0.50	0.55	0.49	0.52	0.02	0.16	30 %	F(2, 5) = 0.10, p = 0.906
MS-15	*Tectona*	0.84	0.84	0.84	0.84	0.00	0.00	0 %	F(2, 5) = 0.71, p = 0.534
Mean								6.9 %	

Table 11 shows the results of the MCA analyses for the mixed samples, i.e. the percentage of the peak area of the mixed sample that matched a particular reference. According to the recommended acceptance criteria for intermediate precision by Little (2016) the overall CV of 12.6 % is excellent and all but one of the ANOVAs are not significant, confirming the equivalence of the results from the three instruments. The MCA values given are not absolute but partial. For instance, for the mixed sample 07 measured on system A, a mean MCA of 28.9 % is given for *Dipterocarpus* (Table 11). The total sum of the areas of all peaks detected in MS-07 and assigned to the reference *Dipterocarpus* is actually 82.1 %. This is much higher than the 28.9 % given in Table 11. The difference is due to the many ambiguous peaks that are matched by more than one reference peak from the DB. In this example of MS-07, the largest peak (RI 3509, Dammaran-3-one, 20,24-epoxy-25-hydroxy) represents approximately 50 % of the total area of all peaks in the sample. This ambiguous peak was matched with *Dipterocarpus* and 6 other references. Therefore, the peak area of 50 % is divided by 7 and then only this 7.1 % portion is included in the sum of the matched peak areas for *Dipterocarpus* and each of the other 6 DB references, resulting in much smaller MCA values listed in the results tables than the actual total values.

Table 11: Statistics results for matched chromatogram areas (MCA) of **mixed** samples.

Sample	Analyzed taxon	GC-MS system			Total matched chromatogram area (%)				
		A	B	C	∑ABC				
		Mean	Mean	Mean	Mean	Var	SD	CV	ANOVA
MS-07	*Dipterocarpus*	28.9	22.0	24.3	24.59	12.24	3.50	14 %	F(2, 5) = 5.02, p = 0.064
MS-07	*Fagus*	0.3	0.3	0.3	0.27	0.00	0.04	15 %	F(2, 5) = 0.51, p = 0.626
MS-07	*Lophopetalum*	10.7	12.2	13.7	12.39	2.82	1.68	14 %	F(2, 5) = 3.13, p = 0.131
MS-05	*Acacia*	15.1	14.4	15.7	15.08	0.59	0.77	5 %	F(2, 5) = 4.34, p = 0.081
MS-05	*Canarium*	11.3	12.2	12.7	12.17	0.90	0.95	8 %	F(2, 5) = 1.78, p = 0.261
MS-02	*Castanopsis*	7.3	8.2	7.5	7.72	0.36	0.60	8 %	F(2, 5) = 2.80, p = 0.153
MS-02	*Palaquium*	5.6	5.8	6.8	6.13	0.51	0.71	12 %	F(2, 5) = 3.42, p = 0.116
MS-02	*Schima*	11.3	13.2	12.6	12.51	0.87	0.93	7 %	F(2, 5) = 5.71, p = 0.051
MS-02	*Richetia*	7.1	8.5	8.6	8.18	0.88	0.94	11 %	F(2, 5) = 2.96, p = 0.142
MS-02	*Shorea*	4.7	5.1	6.5	5.51	0.98	0.99	18 %	F(2, 5) = 4.87, p = 0.067
MS-12	*Fagus*	6.8	8.9	9.4	8.58	1.54	1.24	14 %	F(2, 5) = 8.89, p < 0.05
MS-13	*Acacia*	8.2	9.0	8.2	8.51	0.30	0.55	6 %	F(2, 5) = 4.50, p = 0.076
MS-13	*Elaeis*	1.0	1.2	1.0	1.09	0.05	0.22	20 %	F(2, 5) = 0.99, p = 0.434
MS-13	*Terminalia*	1.5	1.8	1.6	1.66	0.09	0.30	18 %	F(2, 5) = 0.99, p = 0.434
MS-15	*Cunninghamia*	11.4	12.5	11.3	11.75	0.57	0.75	6 %	F(2, 5) = 4.49, p = 0.077
MS-15	*Tectona*	1.1	1.5	1.4	1.37	0.11	0.33	24 %	F(2, 5) = 1.36, p = 0.338
Mean								12.6 %	

For the pure samples the statistical analyses were performed in the same way as for the mixed samples, but for MPS, RSI and MCA only one genus per sample was analysed. In Table 12, the statistical results for the MPS of the pure samples reveal high Var, SD and an overall CV of 49.4 %. The ANOVA analyses results – in all cases $p < 0.05$ – further underscore the significant differences between these three GC-MS systems for MPS in pure samples. System B and C perform similarly. System A achieves more than twice as many marker peaks in most cases which can be explained by the fact that it is the original system on which the database was created. The results of B and C allow a better judgement about how universally the database can be used. Therefore, the intermediate precision of MPS for pure samples may not be as bad as it appears from Table 12. Furthermore, wood identification is not only about the MPS number itself. Even in the by far worst example with a CV of 73 % for the reference *Hevea* in pure sample 17, an average of 10.5 MPS was identified on system B (Table 12). This score still allows a very clear identification decision in favour of *Hevea*, as the query results for almost all other reference woods in the DB are 0 or in two cases maximum 3 MPS (Supplementary Figure S1). Since the decision making process is always a combination of considering not only the MPS but also the RSI (0.92, Table 13) and MCA (4.6 %, Table 14, 79.4 % in total), it is obvious that even in this drastic example with the highest variation between the systems the correct decision for *Hevea* is easily made on all systems. In the end, this is what counts in practice, as demonstrated by the excellent blind test results (3.1 Blind test).

Table 12: Statistics results for marker peak scores (MPS) of **pure** samples (PS).

Sample/taxon	GC-MS system			Total marker peak scores				
	A	B	C	∑ABC				
	Mean	Mean	Mean	Mean	Var	SD	CV	ANOVA
PS-06 *Castanopsis*	29.0	20.3	10.7	17.43	56.62	7.52	43 %	F(2, 4) = 13.68, p < 0.05
PS-19 *Intsia*	36.0	14.7	16.7	18.57	67.95	8.24	44 %	F(2, 4) = 15.23, p < 0.05
PS-09 *Dendrocalamus*	72.0	26.3	30.3	34.57	288.62	16.99	49 %	F(2, 4) = 45.23, p < 0.05
PS-17 *Hevea*	39.0	10.5	13.0	17.20	155.70	12.48	73 %	F(2, 2) = 26.68, p < 0.05
PS-13 *Eucalyptus*	87.0	46.3	29.0	44.71	432.24	20.79	46 %	F(2, 4) = 89.53, p < 0.05
PS-29 *Rhizophora*	41.0	16.7	19.7	21.43	77.29	8.79	41 %	F(2, 4) = 276.23, p < 0.05
PS-21 *Lophopetalum*	53.0	20.7	20.3	25.14	151.48	12.31	49 %	F(2, 4) = 543.31, p < 0.05
Mean							49.4 %	

The overall intermediate precision CV for RSI values in pure samples stands at 12.6 % (Table 13). Although this is less favourable than the outcomes of 6.9 % for mixed samples (Table 10), it still reflects a satisfactory level of reliability.

Table 13: Statistics results for reverse similarity index (RSI) of **pure** samples.

Sample/taxon	GC-MS system			Total reverse similarity index				
	A	B	C	∑ABC				
	Mean	Mean	Mean	Mean	Var	SD	CV	ANOVA
PS-06 *Castanopsis*	0.99	0.97	0.98	0.98	0.00	0.01	1 %	F(2, 4) = 3.08, p = 0.155
PS-19 *Intsia*	0.99	0.97	0.98	0.97	0.00	0.01	1 %	F(2, 4) = 1.90, p = 0.263
PS-09 *Dendrocalamus*	0.74	0.69	0.71	0.71	0.00	0.03	4 %	F(2, 4) = 0.88, p = 0.484
PS-17 *Hevea*	1.00	0.92	0.89	0.93	0.00	0.04	5 %	F(2, 2) = 608.38, p < 0.05
PS-13 *Eucalyptus*	1.00	0.97	0.97	0.98	0.00	0.01	1 %	F(2, 4) = 6.54, p = 0.055
PS-29 *Rhizophora*	0.99	0.15	0.86	0.57	0.16	0.40	70 %	F(2, 4) = 40634.96, p < 0.05
PS-21 *Lophopetalum*	0.61	0.54	0.58	0.57	0.00	0.04	6 %	F(2, 4) = 2.62, p = 0.188
Mean							12.6 %	

Table 14 shows the results of the MCA analysis for the pure samples. Only two of the seven ANOVAs are not significant but the overall CV of 14.4 % is acceptable showing that MCA values for pure samples are quite reproducible on different instruments.

Table 14: Statistics results for matched chromatogram areas (MCA) of **pure** samples.

Sample/taxon	GC-MS system			Total matched chromatogram area (%)				
	A	B	C	∑ABC				
	Mean	Mean	Mean	Mean	Var	SD	CV	ANOVA
PS-06 *Castanopsis*	32.1	32.0	31.8	31.93	1.17	1.08	3 %	$F(2, 4) = 0.04$, $p = 0.962$
PS-19 *Intsia*	16.2	11.3	12.1	12.32	4.57	2.14	17 %	$F(2, 4) = 4.33$, $p = 0.100$
PS-09 *Dendrocalamus*	35.4	29.8	24.9	28.54	15.24	3.90	14 %	$F(2, 4) = 712.27$, $p < 0.05$
PS-17 *Hevea*	5.3	4.6	6.0	5.34	0.51	0.71	13 %	$F(2, 4) = 58.27$, $p < 0.05$
PS-13 *Eucalyptus*	17.8	14.1	9.0	12.47	12.50	3.54	28 %	$F(2, 4) = 41.20$, $p < 0.05$
PS-29 *Rhizophora*	24.0	20.8	18.8	20.37	4.09	2.02	10 %	$F(2, 4) = 11.64$, $p < 0.05$
PS-21 *Lophopetalum*	14.3	11.8	10.1	11.43	2.91	1.71	15 %	$F(2, 4) = 10.09$, $p < 0.05$
Mean							14.4 %	

To improve the intermediate precision, some of the pre-processing parameters of the original method (Flaig et al. 2023) were adjusted prior to this investigation. Rounding was set to the range of m/z -0.38 to m/z +0.62 according to Khrisanfov and Samokhin (2022). In addition, the 206 largest peaks by area were analysed per reference chromatogram and the 125 highest mass traces per scan were retained. The mass traces were smoothed using the Savitzky-Golay filter (Savitzky and Golay 1964) 4 times in a row for better smoothing results, signal-to-noise ratio and peak detection. The mass spectra were compared using the "NIST (Identity Normal)" algorithm, which goes a step beyond the cosine comparison used in the original DB. This approach, outlined by Stein and Scott (1994), involves additional comparison of the highest intensity m/z values. The originally used cosine method exhibits slight weaknesses in distinguishing isomers and alkanes. Detailed pre-processing operations and DB setup parameters are given in section 2.3.5 Data processing. The latest version of the pulp extract DB is available from the authors on request.

Even with the updated and improved DB, reproducing MPS for pure samples proved challenging, with unfavourable results from ANOVA analyses and an overall CV of 49.4 %. However, given that real-life samples primarily consist of mixtures, the focus shifts to the crucial precision of MPS, RSI and MCA values in mixed samples. For mixed samples, the overall CV results of MPS (32.0 %), RSI (6.9 %) and MCA (12.6 %) scores were satisfactory, or RSI and MCA even excellent according to the definition by Little (2016). Encouragingly, the ANOVA results for mixed samples were also highly positive, providing assurance that the three GC-MS instruments did not exhibit statistically significant differences.

To reduce the system dependency and increase the overall robustness of the chemotaxonomic method across different laboratories, standardized procedures for the analysis are proposed. These should include protocols for data processing, analysis, and sample preparation, including guidelines for extraction methods as described in this paper and by Flaig et al. (2023). Calibration procedures should consistently use standard spectra tune (s-tune) and the same amount of polyethylene for retention index calibration. Quality control

samples with known marker peak signatures could be included in each batch of analysis to monitor system variability in real time. To ensure consistency between laboratories, maintaining and expanding reference libraries for known genera and including metadata such as analytical equipment, calibration details, and sample preparation methods, can facilitate cross-referencing and verification of results.

3.3 Limitations

It shall be noted that the same pulps used for the references were used to create the blind samples of unknown composition. In addition, the knowledge of the participants that each of the pulps must be present at least once in one of the blind samples changes the attention to it. The search for each of the 38 taxa probably leads to more species being named by the testing laboratories in the blind test than in real samples.

For both methodologies, a critical factor revolves around the extent of reference samples present in their databases. Genera can only be identified if there are references for them. The expansion of the reference database is therefore very important for reliable identification. For the anatomical method, if an unknown sample contains cells for which there are no references, but there is a declaration of the genus being processed, a new reference can be created relatively quickly using the Thünen wood collection. The declaration can then be checked for plausibility.

Notably, the chemotaxonomic method faces limitations regarding the determination of MPS, which is currently dependent on the GC-MS system used. Other factors that can hamper this method and make the accurate identification of certain taxa difficult are the low extractive content of the pulp and the limited number of marker substances available.

Also, its capacity to comprehensively encompass the full spectrum of potential variations within tree species is limited. This limitation arises from the well-known phenomenon of extractives compositions undergoing alterations in diverse environmental settings (Deklerck et al. 2020; Silva et al. 2018), whereas anatomical features tend to remain relatively stable (IAWA Committee 1989).

Flaig, M.L. et al. (2024). *Holzforschung* 78: 487–502,

4 Conclusion

A comprehensive blind test was conducted, allowing for a comparative evaluation of the outcomes produced by two independent identification methods: chemotaxonomy and anatomy. All testing laboratories demonstrated a high level of accuracy in identifying the wood taxa present in the blind test pulp samples. The anatomical method proved highly effective in identifying 23 wood taxa 100 % correctly by at least two laboratories, with particularly notable accuracy in the case of the genera *Acacia* and *Rhizophora*. These results underscore the reliability of anatomical analysis for specific taxa, even in cases where differentiation can be challenging. Using the chemotaxonomic method, the authors identified 8 wood genera with 100 % accuracy. The method and software proved effective in matching blind samples against the DB. Key factors for identification included the total RSI value, the count of marker/ambiguous peaks and the total matching area. Some challenging genera, such as *Eucalyptus*, *Avicennia* and *Koompassia* displayed weak marker expressions. The study aimed to compare the individual strengths of each method and optimize their collective capabilities, especially in complicated cases such as highly beaten pulp, to improve overall identification performance. An examination of Table 7 revealed interesting patterns in the reliability of identification by anatomy and chemotaxonomy. For example, the genus *Nyssa* was correctly identified 98 % of the time by the anatomically trained participants, but was unrecognizable by chemotaxonomy. Conversely, taxa such as *Intsia*, *Parashorea* and *Richetia* were confidently identified by chemotaxonomy, while anatomical identification was more uncertain. This clear contrast highlighted the complementary nature of the two methods. By integrating both methods, synergies could be used to achieve even more robust and accurate results. This study emphasizes the importance of a comprehensive and multi-faceted approach to wood species identification.

The systematic validation, in which three parameters were analysed in two different sample types on three different instruments, affirmed the system-independence and robustness of the chemotaxonomic technique under certain conditions. Uniform preparation and analysis of mixed and pure tropical hardwood sample extracts on the GC-MS systems A, B and C demonstrated the consistent performance of the method for RSI and MCA values. MPS values were less reproducible. The poor statistical results were due to much higher numbers of marker peaks achieved by system A, while the average results of systems B and C were quite similar. The outperformance effect of system A was more or less observable in all investigations, but it was worst for MPS statistics in pure samples. It was expected that the original system A, on which all the DB reference samples were measured, might have an advantage over the other instruments. However, systems B and C, representing all other GC-MS instruments, performed well and reproducibly even for MPS in most pure samples. As real-life samples are mainly mixtures, the noteworthy intermediate precision results for RSI (6.9 % CV) and MCA (12.6 % CV) in mixed samples underline the reliability of our approach and reinforce its usefulness as a versatile tool for wood taxa identification in different research settings and laboratories. To further reduce the inter-instrument variance, it would be useful

in the future to enrich the pulp extract DB not only with additional wood references from only one instrument, but also with chromatography data references from other instruments.

Whilst the first attempt to create a chemotaxonomic database proved successful with equivalent blind test results to the established anatomical method, the results of the validation showed promising results with potential improvements to the method. For general application, it would be useful to identify and refine the parameters associated with its partial system dependency and use standardized protocols for sample preparation, calibration, data processing and analysis. Further extension with a diverse range of samples to account for variation between different wood provenances is anticipated in future chemotaxonomic studies. Although more work needs to be done, the combination of both approaches is particularly valuable when dealing with taxa that present challenges to either method alone and contributes to the support of EUDR and sustainable forestry.

Acknowledgments: PD Dr. habil. Jürgen Odermatt, who initiated the research for the chemotaxonomic method and passed away suddenly in 2019, deserves special thanks. The authors would also like to thank Prof. Dr. Helga Zollner-Croll for preparing the blind test samples and Doris Helm, Sergej Kaschuro, Anne Wettich and Birte Buske for their assistance with microscopic and laboratory work. For the preparatory work the authors also thank Othar Kordsachia, Nils Grützmann and Alina Wassink.

Research ethics: As the research does not involve the use of humans or animals, the Declaration of Helsinki does not apply. Therefore, the research ethics statement is not applicable. The manuscript has not been published previously and is not under consideration for publication elsewhere.

Author contributions: MLF, JB, AO and SH conceived and planned the experiments. MLF and JB carried out the chemotaxonomic analysis. AO, SH, HJS and DZ carried out the anatomic analysis. MLF conceived, planned and carried out the method validation and wrote the original manuscript. BS revised and edited the manuscript and supervised the project. All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This work was funded by the German Federal Foundation for Environment (DBU) (Grant/Award Number: AZ 34295/01) in connection with the project “Detection of Tropical Hardwood in Paper – Chemotaxonomy and Anatomy for the Identification of Mixed Tropical Hardwood”. The funding organization played no role in study design; analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Conflict of interest statement: The authors declare that they have no conflicts of interest regarding this article.

Data availability: The raw data and the latest version of the pulp extract database can be obtained on request from the corresponding author.

References

All **References** are listed at the end of the dissertation on page 134.

3.3 Publication III: Tracing the geographic origin of wood in pulp and paper by GC-MS analysis of extractives

Original Article

Max L. Flaig*, Jens Berger and Bodo Saake

Tracing the geographic origin of wood in pulp and paper by GC-MS analysis of extractives

http://doi.org/10.1515/hf-2024-0080
Received August 12, 2024; accepted November 4, 2024;
published online November 28, 2024

***Corresponding author: Max L. Flaig,** Institute of Wood Science, Chemical Wood Technology, University of Hamburg, Haidkrugsweg 1, 22885 Barsbüttel-Willinghusen, Germany, E-mail: max.lauritz.flaig@uni-hamburg.de
https://orcid.org/0009-0000-0585-1734
Jens Berger and Bodo Saake, Institute of Wood Science, Chemical Wood Technology, University of Hamburg, Haidkrugsweg 1, 22885 Barsbüttel-Willinghusen, Germany.

Abstract: This study explores the application of chemotaxonomic methods to trace the geographic origin of wood used in paper production. Utilizing thermal desorption - gas chromatography - mass spectrometry (TD-GC-MS) and principal component analysis (PCA), distinct chemical profiles were identified in both hydrophobic and hydrophilic extracts. Residual wood extractives, that are capable of origin differentiation, were found in acetone extracts of the CITES-protected genus ramin (*Gonystylus* Teijsm. & Binn. spp.), petroleum ether extracts of red meranti (*Rubroshorea* (Meijer) P.S.Ashton & J.Heck. spp.) and n-hexane extracts of teak (*Tectona grandis* L.f.) pulp. The PCA results demonstrate the potential of these chemical fingerprints to serve as reliable indicators of wood provenance, which could potentially be crucial for the practical execution of the European Deforestation Regulation (EUDR) and the US Lacey Act in the paper industry. Despite the promising findings for ramin and teak, the effectiveness of the method to distinguish the origin of red meranti pulps was limited in some cases.

Keywords: CITES; extractives; GC-MS; mixed tropical hardwood; provenance; pulp and paper.

1 Introduction

1.1 Problem

Biodiversity conservation enhances ecosystem resilience and protects both human and natural systems. Adaptation and mitigation strategies that incorporate biodiversity needs are essential for preventing ecosystem damage (Rathore and Jasrai 2013). Illegal practices in the forest sector, driven by poor governance, corruption, unclear legal frameworks, weak law enforcement and the demand for cheap timber, result in lost revenues and hinder economic growth for local and national governments (FAO and European Union 2022). These practices cause forest degradation, deforestation and biodiversity loss, and negatively impact forest-dependent communities, increasing conflicts over land and resources and disempowering indigenous peoples. Such activities undermine sustainable forest management and broader sustainable development goals like poverty reduction, food security and climate change mitigation (FAO and European Union 2022). Illegal logging is the most profitable natural resource crime, surpassing wildlife trafficking, illegal fishing, illegal mining, and crude oil theft. It accounts for 15–30 % of the global timber trade, valued at USD 51–152 billion annually, causing severe economic, environmental, and social consequences (INTERPOL 2021; May 2017). Some tropical countries in Southeast Asia, Central Africa and South America see 50–90 % of their timber obtained illegally (Deklerck 2023). The global forestry products trade was worth USD 522 billion in 2014, making illegal logging a substantial portion of this market (May 2017).

1.2 Regulations

Regulations to curb illegal logging and trade, explained by Deklerck (2023) operate in two main spheres: supply-side and demand-side. Supply-side regulations involve in-country laws that create a legal framework for logging, commerce, and shipping. Demand-side regulations, such as the European Union Deforestation Regulation (EUDR), UK Timber Regulation (UKTR), and US Lacey Act, control timber and forest products at the border or market entry. These regulations require importers to perform due diligence (EUDR and UKTR) or due care (US Lacey Act) to ensure the legality of the timber. Species and origins listed by the CITES are of particular interest due to the threat of extinction and protection urgency. Due diligence or due care, according to Deklerck (2023), often involve answering two key questions using scientific methods:

Can the genus/species be determined? (taxonID)

Can the geographical harvest origin be determined? (originID)

While scientific proof isn't mandatory for court cases under UKTR, EUDR, or the US Lacey Act, it can support legal proceedings. However, the inherent uncertainty in science can conflict with the legal system's need for certainty, particularly in verification (checking claims against reference data) and determination (identifying origins without prior information), which differ in their reliance on prior information about species and harvest origin. Disproving a claimed

Flaig, M.L. et al. (2025). *Holzforschung* 79: 15–29,

origin by finding mismatches is generally easier than proving it. Since trees do not adhere to country boundaries, is was assumed that local growing environments would be more appropriate for provenance clusters than artificial country borders. The construction of chemotaxonomic fingerprints based on local environments would therefore be more logical and easier to determine (Deklerck 2023).

1.3 Natural variability – OriginID of wood

Generally, there are significant differences between softwood and hardwood extractives, and also between wood species (Holmbom 1999). The biosynthesis and accumulation of secondary metabolites are thought to be influenced by various genetic, ontogenic, morphogenetic, and environmental factors. Environmental conditions such as light, temperature, soil water, soil fertility, and salinity are considered particularly important, as changes in any of these factors can reportedly affect metabolite content even when other conditions remain unchanged (Yang et al. 2018). It is suggested that the variation and abundance of metabolites may be influenced by genotype-environment interactions (Deklerck 2023). Metabolite abundance is assessed to have higher heritability than morphological and physiological properties, with for example the heritability of monoterpenes in scots pine (*Pinus sylvestris*) estimated at 70–98 % (Iason et al. 2005).

Natural variability depending on the origin, growing conditions, accessibility of nutrients, tree age and other factors were observed many times for different tree species and parts of the trees. For example, studies on Norway spruce (*P. abies*) needles suggest that northern provenances tend to have higher levels of stilbene glucosides compared to southern ones, with this increase reportedly occurring primarily during late summer and autumn (Solhaug 1990; Kaufmann et al. 1974). It is generally assumed that older trees tend to have higher extractive content in their heartwood compared to younger trees (Hillis 1962). There were variations found in polyphenol content between tree provenances of sitka spruce (*Picea sitchensis*) origins (Forrest 1975) and also great chemical composition differences were observed in Norway spruce by northern and southern locations in Finland (Kaakinen et al. 2007). In addition Piispanen et al. (2008) found, the total concentration of knotwood lignans in Norway spruce wood from northern Finland to be significantly higher with 14 % of the dry mass than in wood from a southern Finland origin with approximately 5.4 % of the dry mass.

Not only trees but also other plants are chemotaxonomically dividable. Fingerprint analysis combined with correlative chemometric methods of the Chinese medical herb capillary wormwood (*Artesmisia capillaris*) from four different locations showed slight differences in their compositions (Guo et al. 2004). Also, not only extractive compositions are influenced by origin and growing conditions, but also physical properties. For big leaf mahogany (*Swietenia macrophylla*) for example the density, total tangential and radial shrinkage, modulus of elasticity, static bending strength and crushing strength in compression parallel to grain were studied by Langbour et al. (2011), who found significant differences for plantation and natural forest wood. Whereas another study shows no significant effect of wood origin on modulus

of elasticity and modulus of rupture in yellow birch (*Betula alleghaniensis*) and sugar maple (*Acer saccharum*) from two regions in Canada (Duchesne et al. 2015).

1.4 Analytical solutions

As the morphology of wood vessel elements is not affected by provenance, the anatomical method is not applicable for originID (Helmling et al. 2018). A comparison of the taxonID and originID techniques suitable for solid wood and wood products is shown in Figure 23.

1.4.1 Genetics for solid wood

Genetic techniques are not feasible for pulp and paper due to the extensive processing and bleaching that degrades DNA beyond usability. For solid wood Cronn et al. (2021) found that genetic differentiation, geographic structure and geolocalization accuracy for big leaf maple (*Acer macrophyllum*) showed strong spatial genetic structures. This resulted in higher heterozygosity and better individual discrimination for Pacific Northwest populations, while Californian populations had lower heterozygosity and discrimination power. However, the geographic origin of Southern California trees was predicted with less error than Pacific Northwest trees. This study is a good example of how genetic techniques work well for originID.

Generally speaking, there are two main challenges in the genetic analysis of solid wood from different origins. Firstly, extracting high-quality DNA from processed wood had been difficult, limiting the practicality of timber tracing. Recent advancements improved this by replicating short DNA fragments from low amounts of DNA. Secondly, DNA provenance worked best in natural forests with clear genetic structures, while artificial regeneration with seeds of unknown origin posed greater difficulties (Deklerck 2023).

1.4.2 Mass spectrometry

1.4.2.1 For single-origin samples: Stable Isotopes Ratio Analysis (SIRA)

Stable heavy isotope ratios of hydrogen (δ2H), carbon (δ13C), oxygen (δ18O) nitrogen (δ15N) and sulphur (δ34S) are found in plant materials and substrates and can be used to verify the origin of solid wood or theoretically for single-origin pulps. Their abundances, modulated by environmental influences and physiological processes, can be measured with high precision. Tree ring series, for instance, can provide time-integrated records of stable isotope abundances (Augusti and Schleucher 2007; Deklerck 2023). Because stable isotope variation is a continuous spatial variable in nature, isoscapes (isotope landscapes) have been developed to map the variation. For US oak (*Quercus* spp.), for example, model isoscapes such as January precipitation for oxygen isotopes, July precipitation for hydrogen isotopes, April water vapor for carbon isotopes and a combination of March shortwave radiation and May sulphate concentration for sulphur isotopes have been constructed and used as predictors. These isoscapes helped determine the regions from which an unknown solid oak wood sample has originated, with resolution depending on the rarity of the stable isotope signature (Watkinson et al. 2020).

Flaig, M.L. et al. (2025). *Holzforschung* 79: 15–29,

Other examples of stable isotope ratios used to verify geographic origin are found in food products. Grape products, orange juice, olive oil, cheese, butter, and caviar are subject to legal applications and can be reliably measured in routine work and successfully compared across different laboratories (Camin et al. 2017). Since pulp and paper products are mostly mixtures of different origins and species (implying that the specific stable isotope ratios for each origin are also mixed), it is assumed that current stable isotope analysis techniques are not yet suitable for these types of products.

1.4.2.2 For single-origin solid wood AND mixed-origin pulp & paper samples: TOF-MS & GC-MS

DART-TOF-MS has emerged as a valuable tool for identifying the origin of solid wood, complementing genetic methods, as reported by Deklerck (2023). Before that it has already proven effective in taxonID, which has been piloted by the US Fish and Wildlife Forensic Service. The method involves placing wood slivers in a heated stream of helium gas, causing thermal desorption and ionization of molecules, resulting in chemical fingerprints of mostly low molecular weight metabolites, some extractives and some lignin degradation products, which are sometimes called metabolite profiles or chemotypes. Successful application of these classification models is facilitated by selecting from only limited options of origin, such as specific countries or regions, or plantation versus natural forest (Deklerck 2023).

DART-TOF-MS has not only shown success in identifying at species level and but also in distinguishing geographic origin. For example, even the high-heated (120 °C) solid padauk (*Pterocarpus* Jacq.) wood species *P. santalinus* and *P. tinctorius*, both CITES Appendix II protected, were 100 % successfully differentiated using the machine learning method similarity network fusion (SNF) to combine headspace GC-MS and DART-MS data for optimal taxonID results (Zhang et al. 2022a). The same two species were also successfully differentiated using DART coupled to Fourier transform ion cyclotron resonance MS (DART-FTICR-MS) and supervised orthogonal partial least squares discriminant analysis (OPLS-DA) models on differently treated solid wood samples (Zhang et al. 2019). The analysis of 211 samples of 12 padauk species from different regions has achieved 98 % accuracy in identifying broad geographic origins, demonstrating the method's ability to detect phytochemical trends between wood species from different regions (Price et al. 2021). In another study, DART-TOF-MS differentiated douglas-fir (*Pseudotsuga menziesii*) samples from western Oregon, USA into two geographic classes with 70–76 % accuracy by detecting region-specific molecular differences (Finch et al. 2017). The technique has also been used to distinguish wild from cultivated agarwood (*Aquilaria* spp.) and to identify the country of origin for limba (*Terminalia superba*) by analysing growth rings along a pith-to-bark gradient combined with random forest classification (Espinoza et al. 2014; Deklerck et al. 2020). However, differences in DART-TOF mass spectra between earlywood and latewood samples in ring-porous trees can be challenging (Deklerck 2023).

While DART-TOF-MS has shown promise for solid wood, its application to pulp and paper has not been tested and would be very challenging due to the scarcity of metabolites, extractives and lignin in bleached pulp. They have to be extracted and concentrated before measurement. Therefore, the method of Flaig et al. (2023) using thermal desorption gas chromatography and mass spectrometry (TD-GC-MS) of concentrated extractives was considered a good option for pulp products.

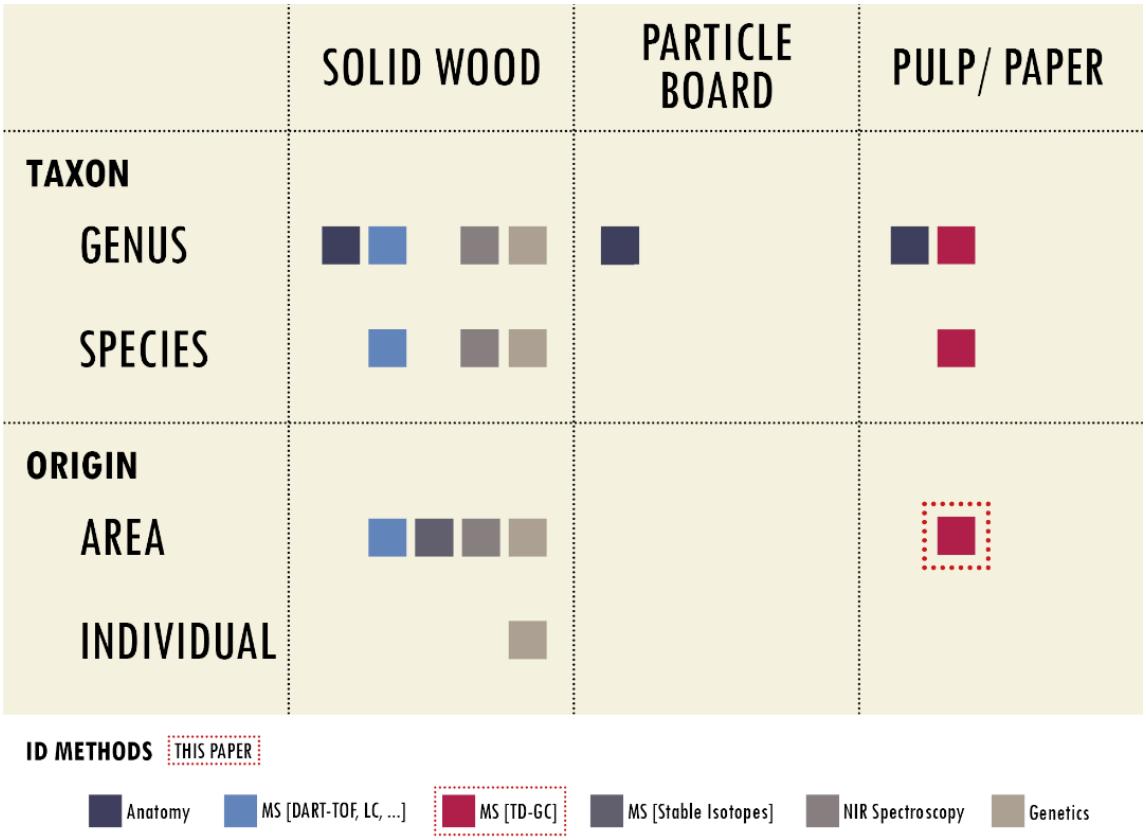


Figure 23: Solid wood and wood products with their corresponding taxonID and originID techniques; diagram adapted from Flaig et al. (2023), published by De Gruyter, licensed under CC BY 4.0.

1.5 PCA and its applications in wood originID

Generally, the principal component analysis (PCA) is a multivariate data analysis technique, that is very well suitable for finding hidden differences and similarities/groups in complex data matrices of many variables by focusing on the most important characteristics. It is used to reduce the dimensionality of a dataset while preserving as much variability as possible. It transforms the data into a new set of orthogonal axes, called principal components, where each successive component captures the maximum remaining variance and can be plotted in a two-dimensional graph, to simplify and visualize the data (Backhaus et al. 2016).

PCA has been applied in various originID contexts, such as in NIRS based differentiation of plantation and natural forest solid wood, where spectral information revealed the potential to distinguish between wood from planted and native forests (Ramalho et al. 2018). In another

example, cedrela (*Cedrela odorata*) samples from South America were analysed using UHPLC-ESI-QTOF-MS to determine geographic origin. By examining wood extractives with PCA, numerous origin-dependent differences were identified. The marker substances indicated that several metabolic pathways were influenced by geographical factors, some of which suggested pest infections (Creydt et al. 2021). Successful PCA discrimination of Norway spruce from two German provenances 250 km apart was also achieved by Creydt et al. (2022) based on secondary metabolites.

2 Materials and methods

2.1 Samples

The provenances of the ramin (*Gonystylus* Teijsm. & Binn. spp.) solid wood samples are unknown, but they definitely came from six different sources (producers and at least different packages, purchase and storage times – most likely not from the exact same origin). They were obtained from three different old stocks of North German carpenters: Olbrich (source 1), Steffen (source 2) and Ratheiser (source 3). In the warehouse of the Ratheiser carpentry there were 4 different packages that were sampled for this study (source 3 A–D).

For the solid wood samples of red meranti (*Rubroshorea* (Meijer) P.S.Ashton & J.Heck. spp.) the geographical origin of harvest is known. They originate from seven different regions: Central Kalimantan, East Kalimantan, West Kalimantan, Mentawai Islands, Moluccas, West Malaysia and Sabah. Two different wood samples were obtained from each origin.

For the teak (*Tectona grandis* L.f.) samples, the provenance, growing conditions and importing companies are known. One sample came from a natural forest in Myanmar, purchased from GARPA (Hamburg, Germany), another sample came from an Asian (not closer defined) plantation, purchased from Ploß (Stelle, Germany), another sample came from a plantation in Brazil (GARPA, Hamburg, Germany), and the last sample came from the “Perhutani” plantation on the island of Java, Indonesia, also purchased from GARPA (Hamburg, Germany).

The declared biological taxa (ramin, red meranti and teak) of all solid wood samples collected were verified by wood anatomy specialists from the Thünen Institute (Hamburg, Germany).

2.1.1 Pulping

All wood samples were processed into 3 × 3 cm chips and subjected to kraft pulping using a 7-liter M/K digester (M/K Systems INC., Williamstown, USA). The pulping process utilized a liquor-to-wood ratio of 4:1 and 22–25 % NaOH with 35 % sulfidity, conducted for 90 minutes at 165 °C. The resulting pulps were then slot-screened and bleached through a five-stage sequence to achieve 90 % ISO brightness. The bleaching stages comprised oxygen (O), complexing agent (Q), oxygen-enhanced peroxide (OP), chlorine dioxide (D), and peroxide (P). This laboratory-scale kraft pulping process is designed to replicate industrial standards, producing modern ECF pulps. Flaig et al. (2023) provide a comprehensive overview of the pulping and bleaching methods used.

2.1.2 Grinding

Pulps were ground using a CryoMill (Retsch GmbH, Haan, Germany), a liquid nitrogen-cooled ball mill, at 25 Hz. Each sample underwent five milling cycles of 1 minute each, with a 0.5-minute cooling period between cycles.

Flaig, M.L. et al. (2025). *Holzforschung* 79: 15–29,

2.1.3 Extraction

A Soxtherm SOX 6 apparatus (C. Gerhardt GmbH, Königswinter, Germany) with an automated MULTISTAT control system performed the extraction. Ground pulps (5 g each) were extracted with 140 ml of solvent over a total time of 4 h and 7 min per extraction cycle. The extraction was performed successively in four cycles per sample: two times with petroleum ether (40–60 °C) or n-hexane and afterwards two times with acetone. The extracts of the same solvent were combined and then concentrated to a final volume of 50 ml according to Flaig et al. (2023).

2.1.4 Sample Preparation

Extract solutions were applied to pyrolysis crucibles for TD-GC-MS analysis. The desired dry mass of 90 ± 5 µg per crucible was achieved by evaporating the solvent with concentration adjusted solution volumes.

2.2 TD GC-MS analysis

Two TD-GC-MS analytical instruments were employed, each with its own set of parameters, to measure the samples. Ramin and red meranti samples were analysed using system 1 (S1), while teak samples were analysed using system 2 (S2). Both systems have been used and described by Flaig et al. (2024).

System 1 (S1) utilized a Frontier Laboratories Ltd. (Koriyama, Japan) micro furnace Double-Shot Pyrolyzer (Py-2020iD) with an Auto-Shot Sampler (AS-1020E), coupled to an Agilent Technologies Inc. GC-MS (6890N/5973N). The thermal desorption (TD) temperature for S1 was set at 300 °C, with the interface temperature at 360 °C and the inlet temperature at 320 °C. A 30 m × 0.25 mm i.d., 0.25 µm film thickness low polarity ZB-5 column (7GH-G002-11, Phenomenex Inc., Torrance, USA) was used with helium as the carrier gas at a constant flow rate of 1 ml/min and a split ratio of 30:1. The total analysis time was 170 minutes. The mass spectral detection scanned a range of 15–550 m/z in total ion current (TIC) mode with a threshold of 150.

System 2 (S2), on the other hand, consisted of a Frontier Laboratories Ltd. Multi-Shot Pyrolyzer (EGA/Py-3030D) with an Auto-Shot Sampler (AS-1020E), connected to an Agilent Technologies Inc. GC-MS (8890/5977B). The TD temperature for S2 was set at 325 °C, with the interface temperature at 330 °C and the inlet also set at 320 °C. Similar to S1, S2 used a 30 m × 0.25 mm i.d., 0.25 µm film thickness low polarity ZB-5HT column (7HG-G015-11, Phenomenex Inc., Torrance, USA) with helium as the carrier gas, but with a split ratio of 20:1. The total analysis time was 134 min. The mass spectral detection scanned a range of 29–700 m/z in TIC mode with a threshold of 100.

Either three or four replicate measurements were made per sample. For consistency, all samples compared in the same PCA were measured on the same instrument under identical parameters. Retention indices (RIs) were calculated using pyrolysis products of polyethylene standards (Flaig et al. 2023).

2.2.1 Data processing

Based on Flaig et al. (2024), raw data pre-processing was performed using the OpenChrom® software (version: 1.5.0. 2024 07 05-08 11, Lablicate GmbH, Hamburg, Germany). Ions (m/z) were rounded from -0.38 (inclusive) to +0.62 (exclusive) according to Khrisanfov and Samokhin (2022). Ions originating from either the atmosphere or the stationary phase of the separation column (m/z 18, 28, 32, 207, 281) were removed from the total ion count (TIC) signal. The Savitzky-Golay filter (settings: order = 2, width = 13) was applied per m/z to smoothen the signals according to Savitzky and Golay (1964). The retention time (RT) range of the chromatograms, that is of interest for the analysis was selected. For the ramin acetone extracts and the red meranti petroleum ether extracts measured on system 1, the retention time (RT) range from 5 to 150 min was used and for teak hexane chromatograms, the RT range from 5 to 101 min was analysed. Afterwards multivariate curve resolution – alternating regression (MCR-AR) according to Gerber et al. (2012) was used to deconvolute and detect the peaks. The peaks were integrated to determine their areas using the "Peak Integrator Trapezoid". For each chromatogram the 220 largest peaks by area were kept and automatically (which is not always precise for small peaks) identified using the NIST 2020 library. A few smaller peaks arising from plasticizer impurities (Flaig et al. 2023) were deleted by the targets "phthalic", "phthalate" and "hexanedioic acid, bis(2-ethylhexyl) ester" contained in their NIST identified names.

2.3 PCA parameters

All PCAs were fed with the peak area values for feature (variable) comparison grouped by retention index (RI) in a RI window of ± 8 or 10. The nonlinear iterative partial least squares (NIPALS) algorithm was used to compute the PCAs according to Wold et al. (1987). Because not all samples contain all peaks, peak area values for some variables were missing from the entire PCA data table. These missing values had to be replaced for the NIPALS algorithm to function. Generally, one approach is to substitute the invalid data with the mean value, preventing it from affecting feature discrimination. Additionally, centring the input variables, for example by subtracting their mean, ensures that the PCA focuses on the variance rather than the mean. Scaling the data is also crucial, as it removes bias and ensures that all dimensions are equally important (Wenig and Mailänder 2024). Slightly different combinations of PCA pre-processing parameters were used with the OpenChrom® software for each set of samples.

2.3.1 Ramin

Missing values were replaced by the mean. Data were mean centred, range scaled and power transformed.

2.3.2 Red meranti

Data were normalized using the "Normalization Inf-Norm". Missing data were replaced by the small random value. Data were mean centred, automatically scaled and power transformed.

2.3.3 Teak

Data were normalized with “Normalization 2-Norm”. Missing values were replaced by the mean value and the data were mean centred and level scaled.

3 Results and discussion

Until now, it hast not been possible to determine the origin of wood used in pulp and paper. The following results were obtained by applying a chemotaxonomic method developed for species identification (Flaig et al. 2023; Flaig et al. 2024) to this problem. In this study, it was applied to extracts of self-produced and fully bleached kraft pulps from three different wood taxa of different origin, growing conditions and possibly age. The method is based on GC-MS measurements of pulp extracts.

3.1 Differentiation of ramin wood sources after pulping

In order to differentiate between different wood sources of bleached pulps of the genus ramin, six differently sourced solid wood samples with little information on their origin, but very unlikely to have grown in the same forest and been harvested in the same year, were analysed.

After pulping and bleaching, the pulps were milled and extracted first with petroleum ether and then with acetone as described in section 2.1.2 and 2.1.3. As the extractive content in ramin pulp and especially in the petroleum ether fraction was very low the acetone extracts were selected for the PCA.

Figure 24 shows the PCA score plot of acetone extracts from all six ramin sources. Sources 3-A and 3-C were very clearly separated – each in a quarter of the score plot plotted by the principal component 1 (PC1) (16.57 %) and PC3 (9.61 % of the variance). The other four sources on the right side of the score plot had more similarities and are located closer together. They were harder to distinguish, but the repeated measurements of the individual extracts from each source also clearly formed groups that are distinct.

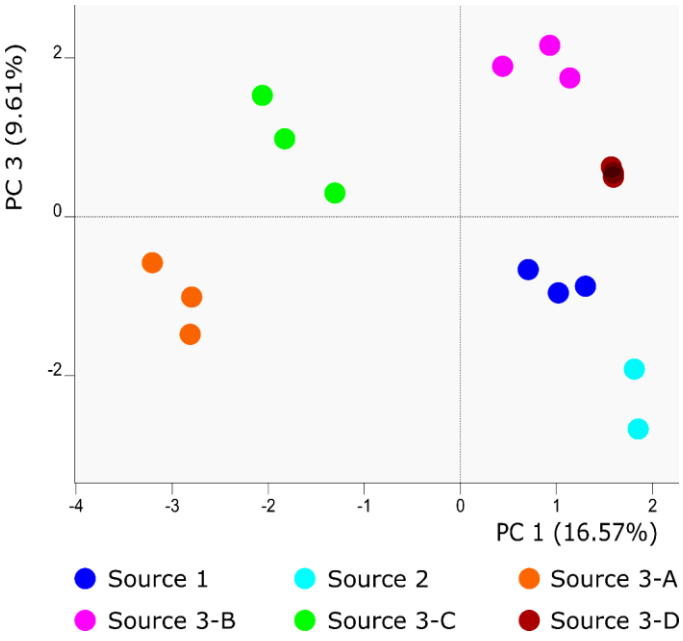


Figure 24: PCA score plot of acetone extracts of bleached ramin pulps from 6 different wood sources, 3 replicate measurements per sample.

As there is no further information on the ramin woods used, it is not possible to discuss whether the clear differentiation of the sources 3-A and C and the similarities between the other sources are related to the provenance, the age or the growing conditions of the trees. However, the 14 most influential extractive compounds on the PCA model were identified from a total of 359 variables analysed. The loadings plot (Figure 25) shows that most of the variables with less influence, coloured in blue, are close to the centre, but those that were most important in separating the different sources on the score plot are further away from the cluster centre and are coloured in red.

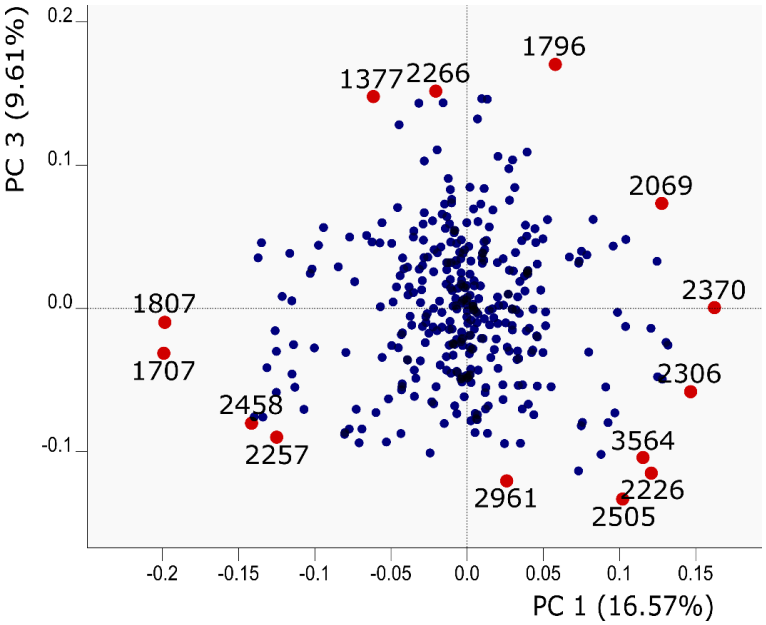


Figure 25: PCA loadings plot of the variables of the ramin pulps' acetone extracts. The numbers are also the retention indices of the extractives listed in Table 15. The 14 red ones are the most important for the differentiation of the wood sources.

These key chemical compounds of the ramin pulps extractives were identified using the NIST database and grouped into chemical clusters. Table 15 shows the variables/RIs of the individual compounds identified and the main clusters: alcohols, alkanes, esters and fatty acids.

Although the specific provenances of the ramin woods were not known, differentiation of the sample sources was achieved using the acetone extractives of their fully bleached kraft pulps. Two of the six sources were obviously different and easily distinguishable, while the others were more similar but still separable in the PCA score plot. These results are particularly valuable in the light of the CITES Appendix II protection status and the listing of some ramin species as critically endangered according to the IUCN Red List (Barstow 2018; CITES and UNEP 2024).

Table 15: Clustered acetone extractive compounds of ramin pulp that are most important for the differentiation of wood sources.

Cluster	Variable (RI)	Chemical compound (autom. identif. by NIST)
Alcohols	1 796	4-Octanol, 2-methyl
		3,5-Dimethyl-2-octanol
	1 807	1-Octanol, 2-butyl
Alkanes	1 707	Decane, 6-ethyl-2-methyl
		Hexadecane
		Heptadecane
	1 807	Octadecane
		Nonadecane
	2 266	Heptadecane, 4-propyl
	2 306	Docosane, 11-butyl
		Tetracosane
		Eicosane, 7-hexyl
	2 505	Pentacosane
		Hexacosane
		Heptacosane
Esters	1 807	Carbonic acid, decyl prop-1-en-2-yl ester
	2 257	Palmitic acid vinyl ester
		Hexadecanoic acid, 2-hydroxyethyl ester
	2 458	Octadecanoic acid, 2-hydroxyethyl ester
		Octadecanoic acid, phenyl ester
	2 961	Octadecanoic acid, 2-oxo-, methyl ester
	2 961	Oxalic acid, allyl pentadecyl ester
Fatty acids	1 377	8-Methylnonanoic acid
		n-Decanoic acid
	2 069	Heptadecanoic acid
	2 226	Octadecanoic acid
	2 370	Eicosanoic acid
		8-Methylnonanoic acid
Others	1 792, 1 796	Benzenesulfonamide, N-butyl
	2 961	Citronellol epoxide (R or S)
	3 564	Acetic acid, 1-methyl-3-(1,3,3-trimethyl-bicyclo[4.1.0]hept-2-yl)-propenyl ester

3.2 Determination of red meranti wood provenances after pulping

Another objective of this study was to determine the provenances of red meranti wood after pulping using PCA of petroleum ether extracts measured by TD-GC-MS. The analysis focused on distinguishing the chemical profiles of the extracts from seven different Indonesian and Malaysian wood provenances: Moluccas, Mentawai Islands, Sabah, East Kalimantan, West Kalimantan, Central Kalimantan and West Malaysia, as shown on the map (Figure 27). The sampling sites are relatively close geographically, so climate and light conditions are expected to be similar. Figure 26 shows the PCA score plot of the petroleum ether extracts of the bleached pulps from all seven regions. For each origin, two solid wood samples (light and dark shade of the same colour in Figure 26) were collected and pulped and extracted. Each extract was measured four times. The replicate measurements per extract, except for the two from West Malaysia (light and dark grey), cluster together in tight groups, indicating that the sample preparation and GC-MS measurement were consistent.

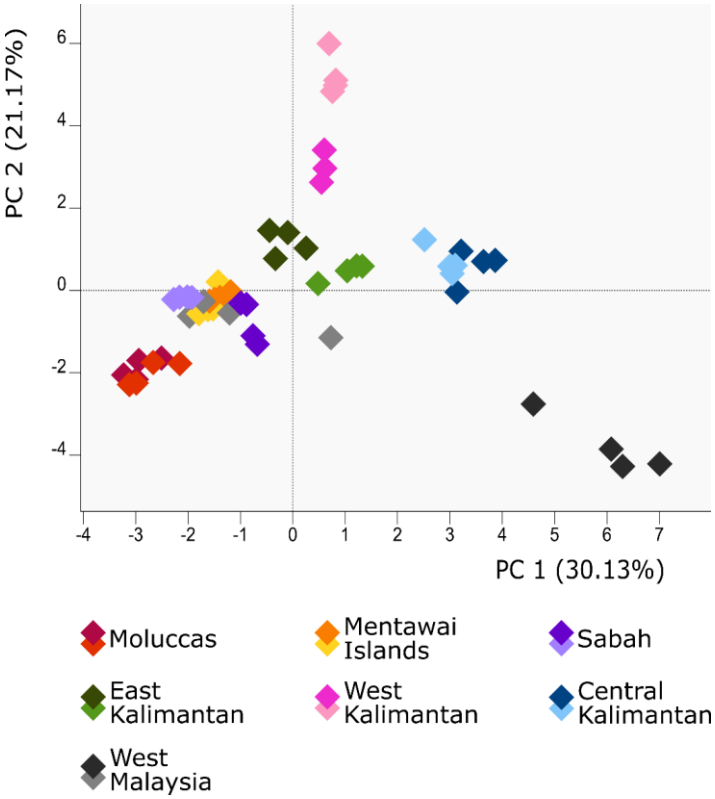


Figure 26: PCA score plot of 14 petroleum ether extracts of bleached red meranti pulps from 7 different Asian wood provenances; 2 pulps from 2 different wood samples per provenance (light and dark shade of the same colour); 4 replicate measurements per sample.

The Malaysian samples each have one conspicuous outlier, indicating that something went wrong in either preparation, measurement or both. Unlike all other provenances where the two related samples form a coherent group, the two Malaysian samples were also very far apart from each other. Since both contained unexplainable outliers and the double determination of this origin did not form a connected group, the West Malaysian samples were excluded for the following analysis. The corresponding loadings plot to the score plot of

all red meranti samples (Figure 26) is not pictured, because it contains the same variables as the one excluding West Malaysia (Figure 28b), only the configuration looks slightly different. Also, the list of important components (Table 16) for the separation along the PCs is valid for both score plots.

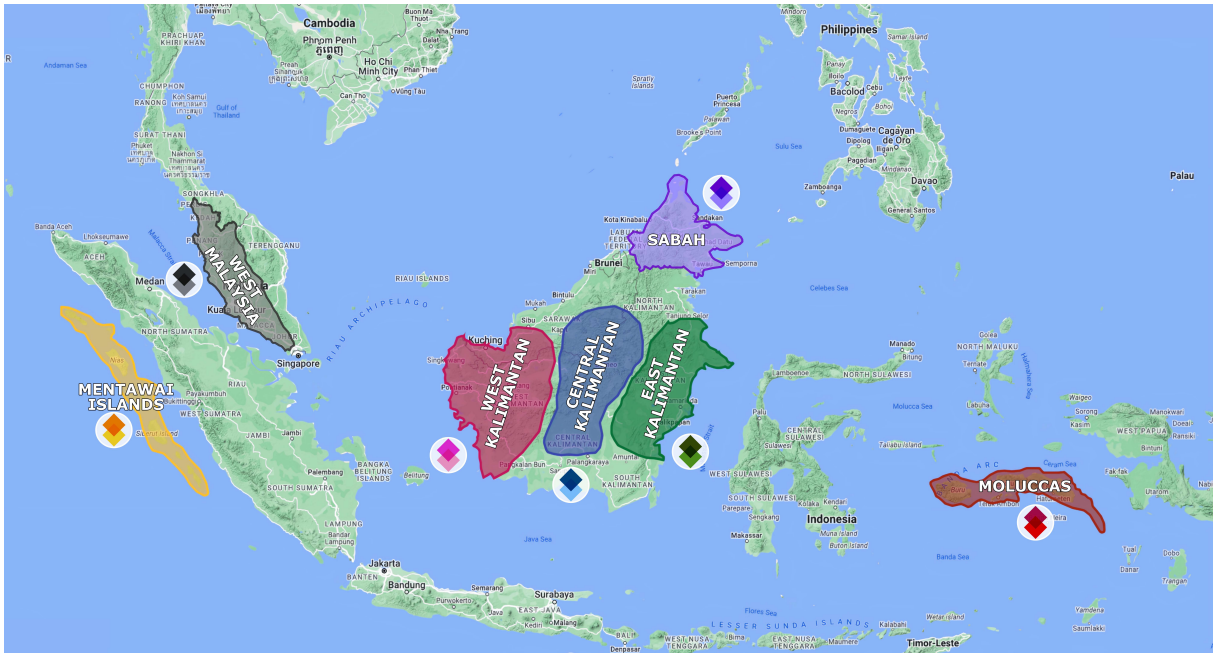


Figure 27: Map of the different regions of origin of the red meranti solid wood samples (regions of origin do not correspond to country borders), created with Google My Maps (2024), based on map data ©2024 Google, TMap Mobility.

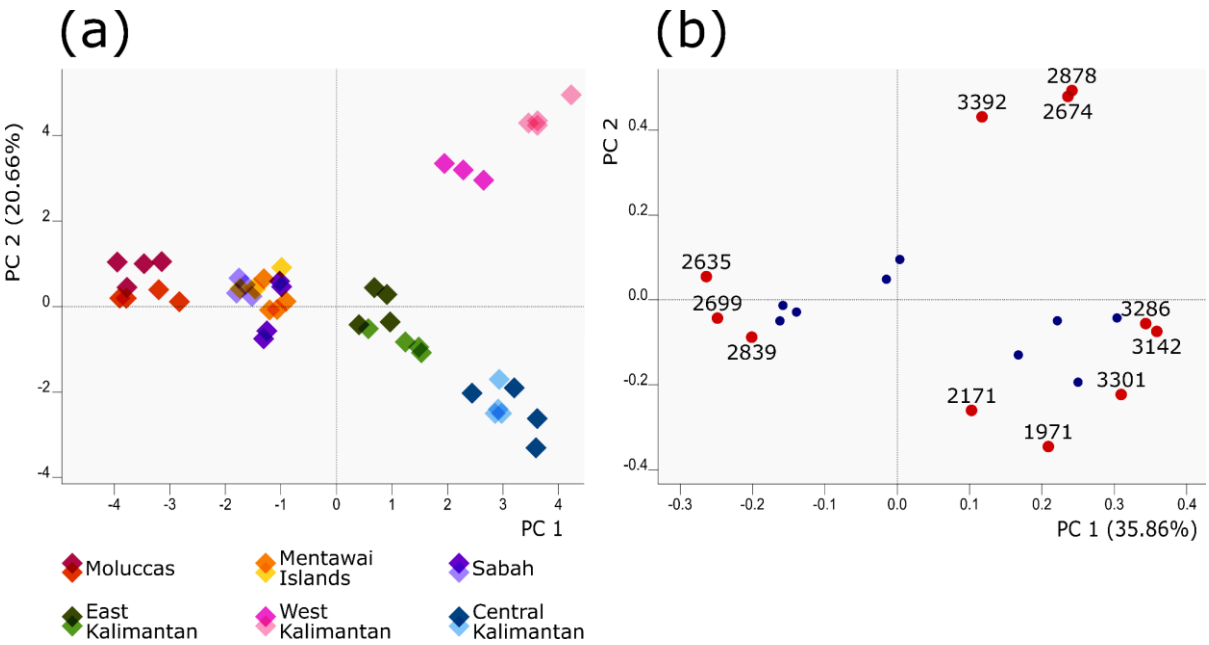


Figure 28: PCA plots of petroleum ether extracts of bleached red meranti pulps from 6 different Asian wood provenances, with 2 samples per origin: (a) score plot (b) loadings plot of the variables/peaks. The variable numbers are also the retention indices of the extractives.

Figure 28a presents the PCA score plot showing the distribution of all origins except West Malaysia. The two principal components, PC1 and PC2, explain 35.86 % and 20.66 % of the variance, respectively. The clustering of the two samples from the same provenance indicates distinct chemical profiles for each geographic origin. Moluccan samples (red diamonds) are grouped on the left side of the plot and show a clear separation from other regions. Samples from the Mentawai Islands (orange/yellow diamonds) and Sabah samples (purple diamonds) form a compact cluster close to the Moluccas, but with a slight shift along PC2. Samples from East (green diamonds), West (pink diamonds) and Central Kalimantan (blue diamonds) are a bit more scattered, but still form identifiable clusters. This clear separation in the PCA score plot, except for Mentawai and Sabah, suggests that the chemical composition of these petroleum ether extracts can effectively distinguish four of the six wood provenances of red meranti. Using the current variables, the Mentawai and Sabah samples could not even be divided in a separate PCA. But the results at least support the conclusion that in a practical scenario, the origin of an unknown single variety red meranti pulp extract declared to be from the Mentawais, located within the orange and purple group in Figure 28a, could be safely excluded as being from either the Moluccas, East, Central or West Kalimantan. This is already a positive achievement in the light of the difference between originID determination and verification, as explained in section 1 (Introduction). The given declaration could be verified as plausible.

Figure 28b provides the PCA loadings plot, showing the contribution of the 20 variables to the principal components of Figure 28a. The RIs of these compounds are labelled on the plot, corresponding to the extractives identified in Table 16. These compounds are pivotal in distinguishing the provenances due to their significant outer loading values, indicating strong influence on the separation along the principal components. The main groups of chemical compounds contributing to the differentiation are: alcohols and aldehydes, alkenes and cycloalkanes, esters, fatty acids, sterols and terpenoids. These compound clusters are common in the *Dipterocarpaceae* family, to which red meranti belongs (Ashton and Heckenhauer 2022), and were already found by Geevandana et al. in 1980.

Table 16: Clustered petroleum ether extractives of the pulp of red meranti that are most important for distinguishing different Asian wood origins.

Cluster	Variable (RI)	Chemical compound (autom. identif. by NIST)
Alcohols & aldehydes	2 839	Docosanal
		Hexacosanal
		Tetracosanal
	2 699	1-Heneicosanol
		Behenic alcohol
		n-Nonadecanol-1
Alkenes & cycloalkanes	2 635	1,21-Docosadiene
	2 699	Cyclohexadecane
		1-Hexacosene
	2 674	1,19-Eicosadiene
		1,21-Docosadiene
		9-Octadecen-1-ol, (Z)
Esters	2 878	(Z)-14-Tricosenyl formate
Fatty acids	1 971	n-Hexadecanoic acid
	2 171	Octadecanoic acid
Sterols & terpenes	3 142	9,19-Cycloergost-24(28)-en-3-ol, 4,14-dimethyl-, acetate, (3.beta.,4.alpha.,5.alpha.)
		9,19-Cyclolanostan-3-ol, 24-methylene-, (3.beta.)
	3 286	.gamma.-Sitosterol
		14,17-Nor-3,21-dioxo-.beta.-amyrin, 17,18-didehydro-3-dehydroxy
		Betulin
		Epilupeol
		Olean-18-ene
		Stigmastanol
	3 301	.beta.-Amyrin
	3 392	24-Methylenecycloartan-3-one

Previously, the differentiation of red meranti provenances was only possible for solid wood products, e.g. using stable isotope ratios as described by Fujii et al. (2007) for veneers from Borneo and the Philippines. The results of this study confirm that the chemical profiles of petroleum ether extracts of red meranti pulps can also be used to effectively differentiate at least some wood provenances. The distinct clustering observed in the PCA score plot underlines the influence of geographical origin on the extractive composition, which is supported by the loadings plot, highlighting specific compounds responsible for these differences.

3.3 Determination of teak origin

In addition to the differentiation of ramin sources and red meranti provenances, four solid wood samples of the species teak were also pulped, milled, extracted and analysed. From these samples not only the country of origin was known, but also the growing conditions and the approximate age of the trees, which were also very interesting for the differentiation of the chemical fingerprints. For these last samples, n-hexane was used for extraction because previously gained knowledge (Flaig et al. 2023) suggests that it is better suited for use with the Soxtherm extraction unit than the formally used petroleum ether. This is due to its clear definition as opposed to petroleum ether, which is a mixture of varying alkanes and isoalkanes in the boiling range of 40 to 60 °C.

Figure 29 displays the PCA score plot of the n-hexane extracts of fully bleached teak pulps from a natural forest in Myanmar and from three plantations: one in Brazil, one in Java, Indonesia and another Asian plantation not further defined. At least five replicate measurements were performed for each sample. PC1 explains 28.34 % and PC2 16.67 % of the variance. The samples were clearly clustered according to their origin, one per quadrant. This separation suggests significant chemical differences in the n-hexane extracts due to geographic origin and growing conditions. The Myanmar natural forest samples and the Indonesian plantation samples are both to the right of the score plot and a little closer together in the score plot (PC1 vs. PC2) than the other samples.

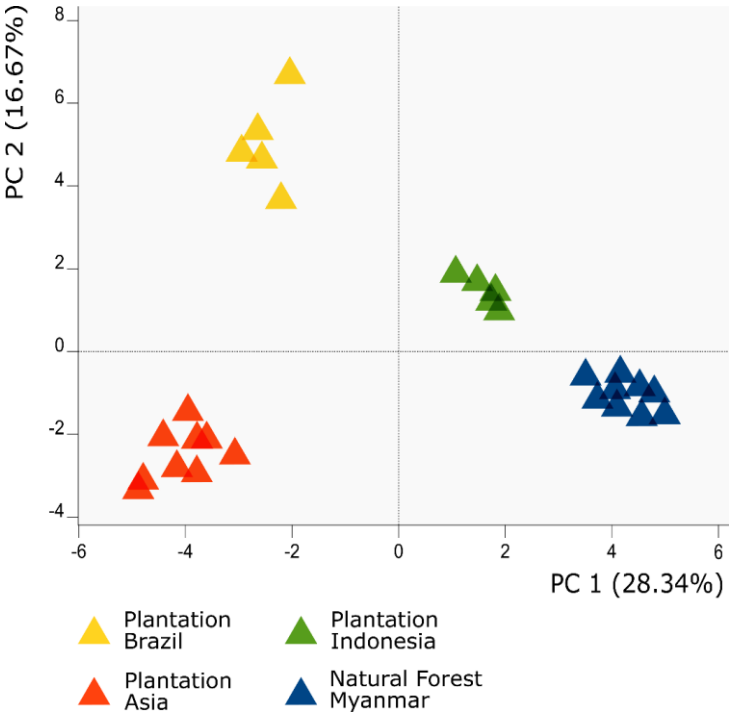


Figure 29: PCA score plot of 4 n-hexane extracts of bleached teak pulps from 4 different wood provenances and growing conditions.

This is largely due to the extractive substance 9,10-anthracenedione, 2-methyl, found at RI 2 114. It belongs to the chemical class of quinones. In particular, the samples from Myanmar

and Indonesia contain proportionally a lot of it, as it is their largest peak, whereas in the other samples this substance is not as prominent, as can be seen in Figure 31, an overlay of exemplary sample chromatograms for each group. The figure also highlights other key peaks/extractives that are important for the origin differentiation: the aromatic structure naphthalene, 1,6-dimethyl-4-(1-methylethyl) at RI 1 694, which is also the most abundant in the natural forest sample and the pentacyclic triterpenoid A-neooleana-3(5),12-diene at RI 2 877. The eicosanoic acid eluting at RI 2 370 belongs to the fatty acids and is very important for the placement of the Brazilian and Asian samples (red & yellow) on the left side of the PCA score plot (Figure 29), which is clearly due to the variable 2 370 also being on the far left of the loadings plot (Figure 30).

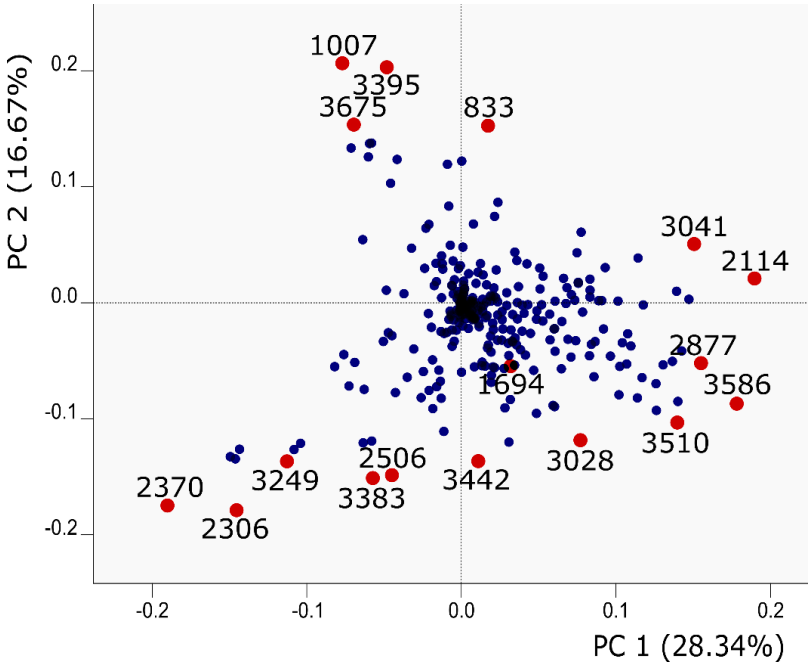


Figure 30: PCA loadings plot of the variables of the hexane extracts of teak pulps from different wood origins. The variable numbers are also the retention indices of the extractive compounds.

Again, in the loadings plot, the red dots represent the variables/compounds with their RIs that indicate a strong influence on the differentiation of the samples because they are far from the cluster centre. Although RI 1 694 is closer to the centre of the PC1 vs. PC2 loadings plot (Figure 30), it was found to be important in combination with others for the differentiation of the Myanmar natural forest samples, also visible in Figure 31. These compounds, which are probably responsible for the distinct chemical profiles observed in the score plot, are listed in Table 17. They belong to the chemical clusters of alkanes, carboxylic acids, cholestanes, fatty acids, phenols & other aromatics, quinones, steroids & derivatives and tri- & pentacyclic triterpenoids, which were also mostly found in solid teak wood (Yang et al. 2020).

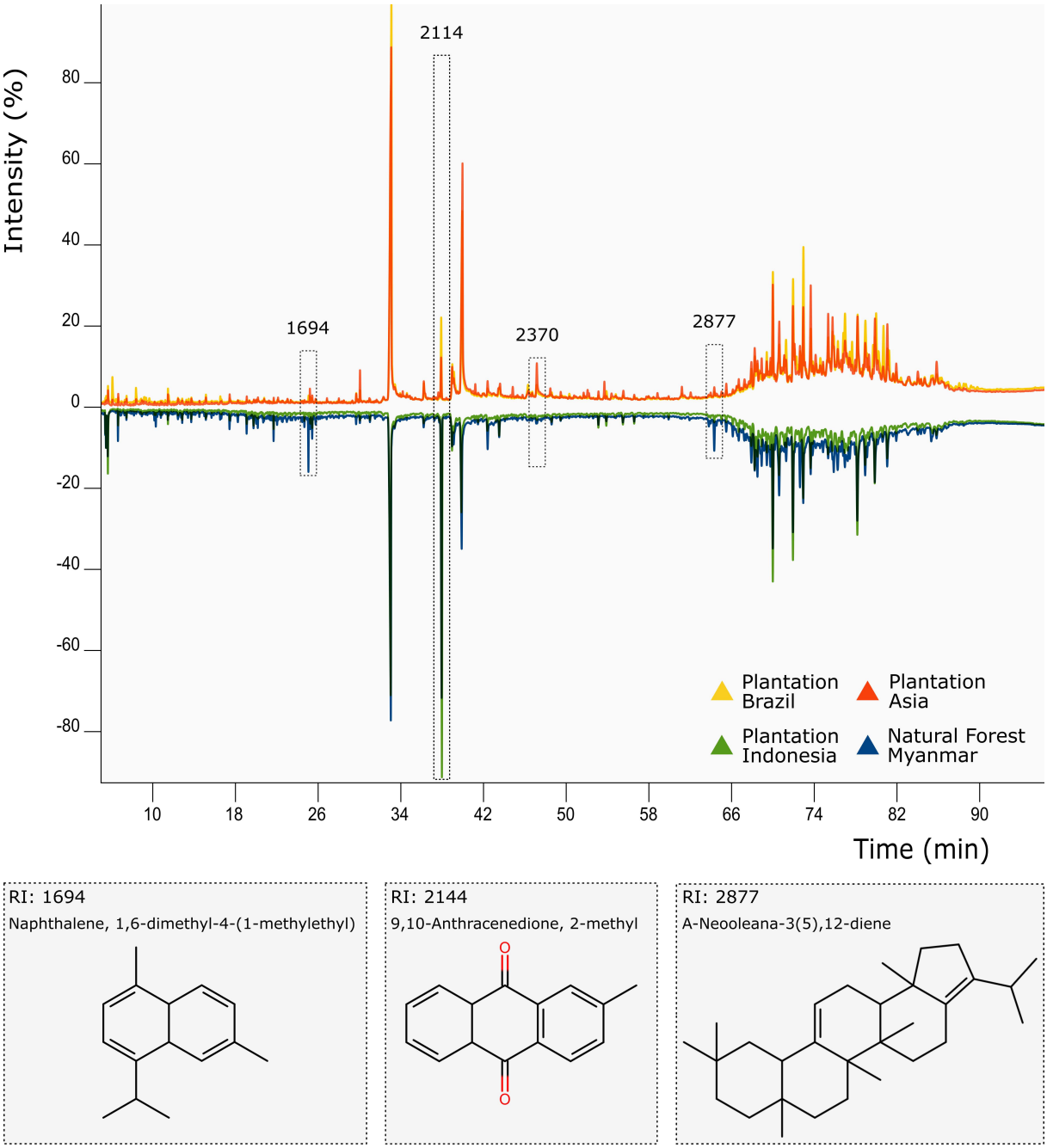


Figure 31: Overlay of GC-MS chromatograms of n-hexane extracts of bleached *Tectona grandis* pulps from different wood provenances and growing conditions. Important peaks/compounds for the origin differentiation are highlighted including their retention indices and structural formula.

In general, solid teak wood contains anthraquinone, naphthoquinone and various natural quinone derivatives that are responsible for its durability and termite resistance (Niamké et al. 2021; Sari et al. 2024; Chávez-Salgado et al. 2022). Lukmandaru (2012) conducted a chemotaxonomic differentiation of solid plantation teak wood from different growing sites in Java, Indonesia. He found the quinone profiles and in particular the concentration of tectoquinone (RI 2114; 9,10-anthracenedione, 2-methyl) to be distinctive. As described above, it was also found to be very distinctive for the Javanese Indonesian plantation in this study,

confirming the studies of Lukmandaru (2012). The presence of the same compound in solid Myanmar teak wood is also confirmed by (Anita et al. 2021).

Table 17: Clustered n-hexane extractives of the pulp of teak that are most important for distinguishing different wood origins and growing conditions.

Cluster	Variable (RI)	Chemical compound (autom. identif. by NIST)
Alkanes	2 306	Docosane
		Tricosane
	2 506	Pentacosane
		Hexacosane
Carboxylic acids	833	3,4-Dimethoxycinnamic acid
		3,5-Dimethoxycinnamic acid
Cholestanes	3 041	(3.beta.,5.alpha.)-4,4-Dimethylcholesta-8,14,24-trien-3-ol
	3 249	Pregnane-11,20-dione, 17-[[(1,1-dimethylethyl)borylene]bis(oxy)]-3-hydroxy-, (3.alpha.,5.beta.)
	3 510	Cholest-7-en-3-ol-15-one, 14-methyl
		Cholestan-26-oic acid, 3,7,12-trihydroxy, (3.alpha.,5.beta.,7.alpha.,12.alpha.)
		Cholestan-3-ol, 4,4-dimethyl
		Cholestan-3-ol, 5-chloro-6-nitro-, (3.beta.,5.alpha.,6.beta.)
Fatty acids	**2 370**	Eicosanoic acid
		Octadecanoic acid
Phenols & other aromatics	1 007	2,3-Dimethylhydroquinone
		2,5-Dihydroxybenzaldehyde
	1 694	Naphthalene, 1,6-dimethyl-4-(1-methylethyl)
Quinones	**2 114**	9,10-Anthracenedione, 2-methyl
		Anthraquinone, 1,2,4-trimethyl
Steroids & derivatives	3 028	Stigmasta-5,22-dien-3-ol, acetate (3.beta.,22Z)-, Stigmastan-3,5,22-trien
	3 395	Stigmasterol
	3 041	4,22-Stigmastadiene-3-one
		Stigmasterol
	3 442	Stigmast-4-en-3-one
		.gamma.-Sitostenone
	3 675	Ergosterol peroxide Me derivative
		Stigmasta-7,22-dien-3-ol, acetate, (3.beta.,5.alpha.,22E)
Tri- & pentacyclic triterpenoids	**2 877**	A-Neooleana-3(5),12-diene
	3 028	(1S,6R,9S)-5,5,9,10-Tetramethyltricyclo[7.3.0.0(1,6)]dodec-10(11)-ene
	3 383	.alpha.-Amyrin
	3 395	Betulin
	3 586	.beta.-Amyrin
		11-Oxo-.beta.-amyrin

The PCA score plot (Figure 29) demonstrated a clear differentiation of the teak samples based on their geographical origin and growing conditions, suggesting distinct chemical profiles. The loadings plot (Figure 30) and the corresponding Table 17 identified specific compounds (with

their RIs) that contribute to this differentiation. These compounds serve as chemical markers to distinguish the origin of the wood samples. These findings were corroborated by the GC-MS chromatograms (Figure 31), which also show distinct patterns of chemical compounds across different samples. Taken together, these analyses indicate that n-hexane extracts of teak pulps contain specific chemical profiles that vary by origin, which can be effectively analysed using GC-MS and PCA techniques. This information can be valuable for quality control, authentication, and provenance determination of teak pulp products.

Hopefully, there are very few pulps made from ramin, red meranti and teak today. Nevertheless, it is important to have the technical means to identify the origin of wood even after the industrial processes of pulping and bleaching. Therefore the three exemplary tested wood taxa of different origin and growing conditions were helpful to apply the method of chemotaxonomic taxa identification (Flaig et al. 2023) to the problem of origin identification.

3.4 Limitations

As ramin is CITES protected and currently very difficult to import, the solid ramin wood samples used for pulping and analysis were obtained from joineries from a previous project with limited information on their provenance, which cannot be improved. This lack of information makes it difficult to fully interpret the results and draw comprehensive conclusions about geographical origins. Nevertheless, the authors decided to analyse these samples as they represent the only opportunity to gain valuable knowledge about the extractant profiles of ramin pulp and to contribute to its urgent conservation. Although it is not yet possible to name originating countries to which specific extractives, responsible for the successful chemotaxonomic differentiation, can be attributed to, the crucial substances for ramin pulp are now known. This is also an important proof of concept and applicability of the method, as ramin pulp contains very low amounts of extractives.

Differentiation between red meranti wood provenances was possible in most cases, showing the potential, but not yet possible for the Mentawai Islands and Sabah, based on petroleum ether extracts and using the current variables for PCA. However, as explained in section 3.3, n-hexane is better suited for consistent sample extraction with the setup used, as confirmed by excellent PCA results for n-hexane teak extracts. The authors consider that the use of n-hexane instead of petroleum ether may also have been better for distinguishing between the Mentawai Islands and Sabah. Acetone also gave satisfactory results with the ramin samples. Simply using acetone as an all-rounder to get all the extractives in one solvent, rather than fractionating the components into lipophilic and slightly less hydrophobic fractions by successive extractions, could be a way of getting even more information per chromatogram and increasing the chances of successful PCA discrimination. On the other hand, the increased complexity of these full acetone chromatograms could make peak separation on the GC column and subsequent peak detection more difficult.

Although the study shows that the variability of wood pulps from different origins can be analysed and differentiated by GC-MS and PCA, this has been tested on single-variety lab-

made samples. The method and the positive results are not easily transferable to mixed real-life industrial samples.

The natural and anthropogenic variability, as studied by Kaakinen et al. (2007) who measured a small effect of nitrogen fertilization on wood chemistry and heartwood content, are crucial to be better understood for the evaluation of chemotaxonomic results in general. As noted by Deklerck (2023) for mass spectrometry techniques, an investigation of what drives the differences in metabolomic/extractive fingerprints between different locations and how much of this difference can be attributed to environmental or genetic effects is needed to improve these methods for use in timber and pulp originID questions.

4 Conclusion

The successful application of a chemotaxonomic method to identify the origin of wood used in pulp was demonstrated. By employing GC-MS and PCA analyses, the research highlights the distinct chemical profiles associated with different geographical sources of wood. Specifically, the results show that it is possible to differentiate between wood provenances based on the chemical composition of acetone extracts of the CITES-protected genus ramin, petroleum ether extracts of red meranti and n-hexane extracts of teak. This differentiation shows the practical possibility of using pulp extracts for chemotaxonomic originID on these three exemplary wood taxa – proof of concept was provided. This could be crucial for quality control, authentication and provenance determination of pulp and paper products to support local legislation such as EUDR, UKTR and the US Lacey Act, especially since anatomical and genetic methods are not applicable to this problem.

However, the study also acknowledges limitations. Firstly, distinguishing between certain provenances, such as red meranti pulp from the Mentawai Islands and Sabah, remains challenging with the current setup. Therefore, the potential of the successive acetone extracts, as shown for ramin, or even combining all compounds by extracting only with acetone rather than separating them, could be tested to increase the information density in the chromatograms and enhance the chances of successful PCA discrimination. It could also simplify the extraction by reducing the number of steps in the process and by making solvent handling easier, as acetone is less toxic and less environmentally harmful than petroleum ether and n-hexane. Secondly, the current method has been tested on single-variety pulps prepared in the laboratory, so it may be challenging to transfer the method and results to mixed samples. As demonstrated by Flaig et al. (2023) and Flaig et al. (2024), taxonID was possible in pulp mixtures, which gives confidence that originID will also be achieved in mixtures in the future.

Future research should focus on understanding the drivers of chemical differences in wood/pulp and improving the method to account for the natural variability. By doing so, the applicability of these techniques to real-world industrial samples could be enhanced, aiding

efforts to combat illegal logging and promote sustainable forest management contributing to the broader goals of biodiversity conservation.

Acknowledgments: PD Dr. habil. Jürgen Odermatt, who initiated the research for the chemotaxonomic method and passed away suddenly in 2019, deserves special thanks. The authors would also like to thank Birte Buske for her assistance with laboratory work. For the preparatory work we also thank Othar Kordsachia, Nils Grützmann and Alina Wassink.
Research ethics: As the research does not involve the use of humans or animals, the Declaration of Helsinki does not apply. The manuscript has not been published previously and is not under consideration for publication elsewhere.
Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: This work was funded by the German Environmental Foundation (DBU) (Grant/Award Number: AZ 31 759-31 and AZ 34 295/01) in connection with the project “Detection of Tropical Hardwood in Paper – Chemotaxonomy and Anatomy for the Identification of Mixed Tropical Hardwood”. The funding organization played no role in study design, analysis and interpretation of data, in the writing of the report or in the decision to submit the report for publication.
Conflict of interest statement: The authors declare that they have no conflicts of interest regarding this article.
Data availability: The raw data can be obtained on request from the corresponding author.

References

All **References** are listed at the end of the dissertation on page 134.

4 Discussion

The following discussion is divided into three parts. The analytical protocol (4.1) discusses sample preparation and process improvements as well as preliminary study results on these improvement ideas. Section 4.2 discusses the database approach, data pre-processing improvements, and limitations such as the partial system dependency of the approach. The third part (4.3) discusses blind test results, taxonID, and originID in comparison with the anatomical method.

4.1 Analytical protocol

The analytical protocol developed for the identification of wood taxa and origin in pulp, as described in Pub I, comprises the following steps: pulping and bleaching of reference pulps, milling of the fibrous materials (pulp/paper) with a liquid nitrogen-cooled cryo ball mill, and finally sequential extraction (2 x hexane + 2 x acetone) and preparation of the extracts for TD-GC-MS analysis. To reduce the volume of the extract solutions, the dissolved extracts are pre-concentrated, but halted before precipitation occurs. To apply 90 ± 5 µg of dry extracts to the pyrolysis sample cups (crucibles) for TD-GC-MS, the pre-concentrated extract solutions are transferred to the crucibles and evaporated. After subsequent data pre-processing, the resulting chromatograms are either stored in the DB as a reference sample or, if unknown, identified through automated matching against the existing set of DB reference. For the first time in history, this protocol has led to promising chemotaxonomic identification and differentiation results for bleached pulp and paper, though it could still be further optimised. Some process improvements that have led to this success are discussed below, as well as some further ideas.

4.1.1 Improvement of extract quality and impurities

During the process of method development, a number of sources of error were identified in the analysis of small amounts of extractives in pulps. These included impurities that were partly contained in the solvents used, particularly in the case of petroleum ether, as well as additional contaminations from the extraction thimbles and wadding. One challenge was the need to concentrate the extracts as the solvent volumes are large compared to the small amounts of extractables. The concentration resulted in the accumulation of impurities, which derived from the cellulose extraction thimbles that were used to hold the milled sample during the Soxtherm extraction and the wadding used to cover the samples in the thimble. The issue of impurities from thimbles and wadding was dealt with by pre-washing. Unwashed thimbles contained an average of 0.27 mg extract each and the amount of unwashed wadding required for one thimble contained approximately 0.38 mg of impurity extract. A standard MTH pulp sample of 6 g with an extractives concentration of ~0.1 % yields a total extract of approximately 6 mg. The summed impurities of ~0.55 mg, extracted from both the thimble and wadding make up a significant proportion of the total extract obtained, particularly in pulp samples with low extract concentrations. Pre-washing the thimbles and wadding with a prior

Soxhlet extraction reduced contamination by 88 % to a total of only 0.08 mg from the extraction thimble and wadding together.

Another source of error was the large surface area of glass that came into contact with the extracts during the process. Therefore, all glassware washed in the laboratory dishwasher with a special program and inorganic cleaner was rinsed by hand (with the same solvent used for sample extraction) immediately before use. Care was also taken to ensure that all materials were handled only with clean gloves.

To control for possible sources of error in the process, blind extractions were performed without samples but with prewashed extraction thimbles and absorbent wadding. These blind extracts were analysed to check for impurities in the chromatograms. The laboratory routine was improved by laboratory protocols for the main sample preparation steps, such as grinding, extraction, treatment of extracts and preparation of analytical samples, which defined precise workflows. This ensured comparability of samples regardless of operator.

Another important point is the purity of the solvents used. This also determines the quality of the chromatograms, which should contain only the peaks of the pulp extractives. Investigations of earlier samples extracted with petroleum ether showed a certain degree of contamination in the chromatograms of some extracts. There was a correlation between the degree of contamination and the time of extraction. Pulps extracted later were more contaminated overall. The proportion of impurities in the chromatograms also depends on the extractives content of the extracted pulps due to the preparation of the analytical samples per weight. The lower the extractant content of the pulp, the higher the level of impurities. These were mainly substances used as plasticisers. Different types of phthalic acid and adipic acid esters were identified.

The interfering substances found prompted the search for an alternative solvent to petroleum ether. In addition to the interfering substances described in Pub I, the use of petroleum ether as a solvent mixture in the Soxtherm apparatus can be expected to produce variable extraction results. Unlike n-hexane, petroleum ether is a mixture of different alkanes with a boiling range of 40–60 °C. During the extraction process it is partially distilled off to reduce the solvent volume. It is hypothesised that this has resulted in a modification to the composition of the petroleum ether in favour of the higher boiling solvent fraction. This issue can be increased by the use of different charges of petroleum ether. N-hexane and petroleum ether with high purity were compared in terms of its solvent potential, both quantitatively and qualitatively. As already described in Pub I, both the quantity and reproducibility of the first two extraction steps of petroleum ether and n-hexane are comparable (Table 4) and are therefore likely to be equally suitable for the extraction of non-polar constituents. There are also no significant qualitative differences between the two extracts. The following Figure 32 shows the chromatograms of two representative extracts of an industrial MTH pulp: petroleum ether (purple) and n-hexane (red) shows that the chromatograms of the two extracts are in high qualitative agreement, making hexane a suitable alternative.

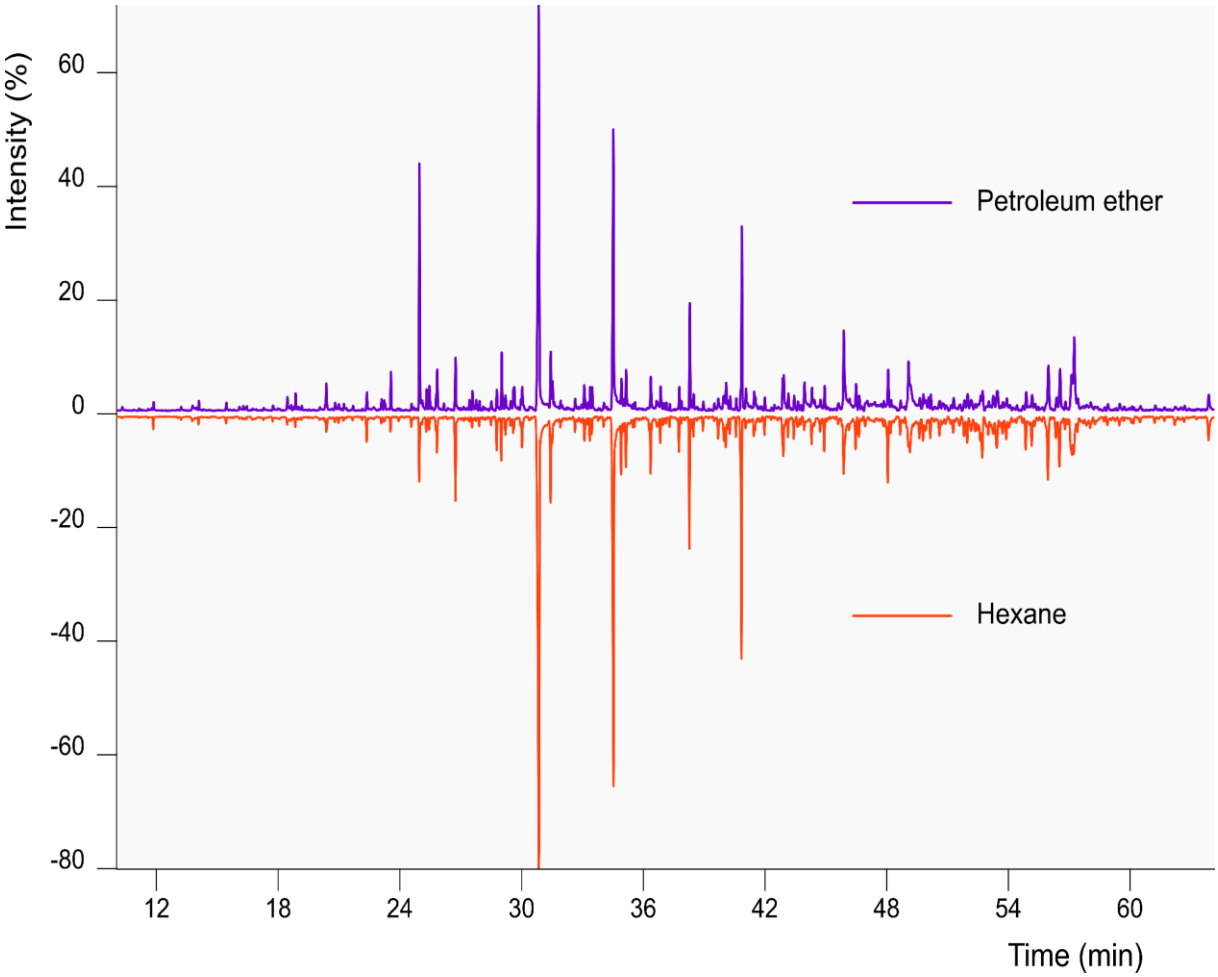


Figure 32: Qualitative comparison between TD-GC-MS chromatograms of MTH pulps extracted with the solvents petroleum ether and n-hexane.

Pre-distillation studies have shown that the purity of even high-quality solvent grades, such as high-performance liquid chromatography (HPLC) grade solvents, can be significantly improved in the laboratory. For acetone with a guaranteed evaporation residue of less than $0.5e^{-3}$ %, Soxtherm blind extractions without extraction thimble, absorbent wadding and sample resulted in an evaporation residue of $0.54e^{-3}$ %, which is very close to the manufacturer's specification. After in-house distillation and subsequent blank extraction, the acetone evaporation residue was only $0.06e^{-3}$ %. This corresponded to an increase in purity of almost 90 %. For n-hexane, an increase in purity of more than 50 % was achieved.

4.1.2 Extraction procedure & derivatisation

The idea of pre-fractionating the pulp extracts into non-polar and semi-polar fractions prior to the actual analysis was ultimately not precise and may not have been a necessary inclusion. The theory and reasoning behind the decision for successive extraction was first to extract exhaustively to obtain as many extractives as possible from the limited amount of far less than 0.01 % extractive content in bleached pulp in many cases (Table 2). Both the first two non-polar extractions of the pulp and the acetone extractions recovered almost all the soluble extractives of the material, as preliminary tests showed that further triple or quadruple

extractions with the same solvent did not yield any more weight. In this sense, the exhaustive extraction was successful. Nevertheless, the acetone fraction after the hexane extraction of the same sample still contained many of the non-polar substances, in a much lower percentage – indicated by much smaller peaks – but still clearly detectable. As an example, Figure 33 shows the hexane extract in red (combined extracts of the first and second extractions) and the acetone extract in blue (combined third and fourth Soxtherm extractions) of a successively extracted Brazilian plantation teak pulp. The peaks of the fatty acids hexadecanoic acid (C16) and octadecanoic acid (C18) are marked in the chromatogram. Their intensities in acetone are about 25 % of those in hexane. Shorter carboxylic acids are generally expected to be more soluble in more polar solvents due to the polarity of their carboxyl group, but the longer the aliphatic alkyl residue of the fatty acids, the more soluble they become in non-polar solvents such as hexane. The C16 and C18 fatty acids from pulp samples are obviously soluble in both hexane and acetone. Later in the chromatogram of the acetone extract, the sterols, which were also present in a large mountain in the hexane extract in the RT range between 44 and 54 min, were also found in acetone. This raises the question of why the non-polar fraction did not contain all of the aforementioned components. As explained above, more extraction cycles did not produce more extract quantity overall. It is assumed that acetone, in combination with a longer total extraction time, manages to extract the rest of these particular fatty acids as well as many other more polar components such as phenolic compounds, aldehydes, shorter carboxylic acids and shorter fatty alcohols earlier in the chromatogram.

The more important question to discuss is this: if the selectivity of separate solvent pre-fractionation is not as clear and good as expected, why go through the trouble of complicating the process unnecessarily and using more resources? Why not just use acetone for extraction, saving time and money and using the less toxic and environmentally damaging solvent? This would also reduce the amount of glassware used and the risk of contamination.

Having two independent extracts from one sample, separated by polarity, was considered practical for two reasons. Firstly, it was theoretically positive to be able to analyse both fractions separately and draw conclusions only on these constituents. This is only true for the hexane extract, which contains less total extract and the more hydrophobic peaks in the chromatogram. As previously outlined and illustrated Figure 33, the acetone chromatogram comprises a combination of additional, more hydrophilic peaks. However, numerous peaks present in the hexane chromatogram are also present in the acetone extract with reduced intensity. The use of acetone alone for pulp extraction may result in the incomplete extraction of strongly lipophilic substances present in the hexane extracts. This implies that a small degree of information may be lost as a consequence of this approach.

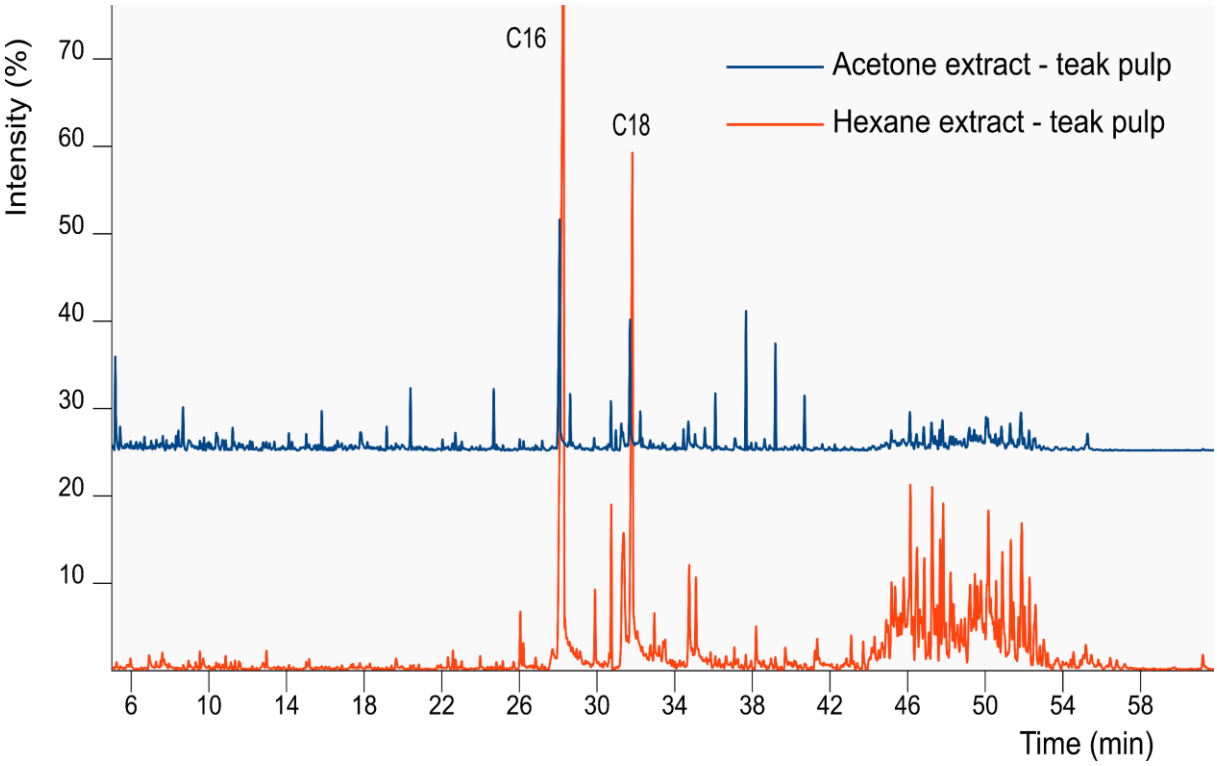


Figure 33: Chromatograms of exemplary hexane and acetone fractions of successive extractions of teak pulp from a plantation in Brazil; C16: hexadecanoic acid, C18: octadecanoic acid.

Secondly, the separated chromatograms should be cleaner and less overloaded. A reduction in the overall number of peaks per sample also results in larger peaks – and can be more effectively separated through offset elution from the separation column at different RT. The more peaks there are, the greater the chance that they will overlap, making later peak detection more difficult.

In fact, the RTs of compounds primarily extracted by hexane and those unique to the acetone extracts elute in different RT ranges anyway. As visible in the overlay in Figure 33 the first part (RT 5–26 min) of a combined chromatogram would be dominated by the acetone extractives, the second part (RT 26–36 min) would be dominated by the hexane extractives, the third part (RT 36–42 min) again by peaks from the acetone fraction and the last part primarily by the more hydrophobic hexane extractives. This indicates that an extract made either by a combination of separate hexane and acetone extracts or an acetone only total extract would certainly have more peaks in total. However, they would probably be well distributed over the whole RT time axis trough peak separation by the interactions with the internal surface of the column in combination with its length of 30 m and the temperature program run in the GC oven. Surely, there are always some overlapping peaks in such complex chromatograms of hundreds of compounds. With advanced peak detection technologies such as the MCR-AR deconvolution algorithm described in Pub I, peaks that elute at the same time can be mathematically separated by analysing their individual mass traces, even if they are not visible

in TIC mode. To not only rely on the deconvolution technology, in addition, the temperature programme of the GC oven could be adjusted to stagger the peaks over a longer total RT.

However, a longer measurement time means that the sample cups sit even longer in the autosampler before their turn. This means that the solid extractives in the cups are exposed to room temperature, air and possibly sunlight, which can alter the samples. It is therefore preferable to limit the time per measurement in order to ensure the quality of the GC-MS analysis and to facilitate more efficient laboratory work.

Furthermore, it is anticipated that a comprehensive acetone extract, comprising both lipophilic and slightly more hydrophilic compounds, will result in diminished peak intensities and sizes. This is due to the presence of a greater number of peaks in the total mixture, which consequently reduces the proportion of both the lipophilic and the more hydrophilic fractions. The hexane-soluble fraction will be discriminated to a greater extent than the acetone fraction, as the amount of hexane extract is less than acetone in the total pulp extract. This issue of reduced peak intensities can be addressed through two measures. Firstly, the quantity of analyte applied to the sample cups can be increased, for example to 200–300 µg. Secondly, the split ratio can be reduced, which also results in a greater amount of analytes in the column and their detection in the MS.

The objective of developing the chemotaxonomic method was to create a practical and straightforward procedure that would minimise the risk of error and facilitate implementation in other laboratories. The reason for starting the development process with the hexane extracts was that, in the event of these extracts failing to provide an adequate quantity and specificity of extractives for identification purposes, the acetone extracts could be employed to provide support. Retrospectively, it could be argued that this option was not pursued because it would have required considerably more effort. Consequently, in terms of practicability and simplification it may be more prudent to combine both factions from the outset.

The approach to use acetone only if absolutely necessary was also planned in light of two assumptions regarding the acetone extracts. The first assumption was that the higher polarity of acetone compared to hexane would result in lower amounts of extract from the bleached pulps. This was based on the premise that during kraft pulping and the five bleaching stages, the pulping and bleaching chemical dissolved in water would effectively remove all polar and most semi-polar extractives through a series of washing steps with changing pH values from basic to acidic. However, it turned out that acetone, which is well in the middle of the eluotropic series and has much greater eluotropic power than hexane, still had the ability to extract about twice as much extractives by weight as the previous non-polar extraction (hexane/acetone ratio of ~1:2). The second assumption was that it would be more difficult to measure acetone extractives on the low-polarity ZB-5 column currently installed in the laboratory's TG-GC-MS analytical setup. The aim was to keep the process as lean as possible and to avoid extra sample preparation step of derivatising the more polar components to

make them better detectable with the current analytical setup. Therefore, the hexane extracts were used first. When the results with hexane extractives were positive, there was no need to go back to the acetone extracts. In the end, acetone extracts of ramin pulps without derivatisation were successfully used for originID analysis in Pub III. The acetone extracts of ramin were used, because these samples originate from a previous project in which petroleum ether was initially used for the first extraction stages before acetone. The acetone extracts were used rather than the petroleum ether ones because the petroleum ether extracts contained some impurities, as previously discussed, and more importantly they had extremely low extractive amounts.

Taking all these arguments into consideration, it is plausible that acetone alone could be tested as a universal solvent for the extraction of pulps for taxonID, given its capacity to dissolve a diverse range of polarity components and its proven efficacy in ramin originID. It is therefore essential to address the aforementioned limitations, namely the more complex chromatograms and the lower intensities caused by the smaller proportions of both fractions in a combined extract.

Although the acetone extractives were measurable with the analytical set-up used, the resulting chromatograms appeared relatively empty in comparison to the hexane fraction (Figure 33), particularly given the ~90 µg of dry extractives applied per sample. One potential explanation is that high molecular weight extractives, such as waxes and other esters, are present in the acetone extracts. These extractives are not volatile and therefore not detectable under the current thermodesorption temperature conditions used for fractionation. The outcomes of size exclusion chromatography (SEC) of acetone extracts of solid woods from another project support this hypothesis. Another explanation for few and small peaks could be the polarity of the compounds. As mentioned above, a column coated with a non-polar stationary phase (95 % dimethylpolysiloxane and 5 % phenyl-dimethylpolysiloxane) does not match the more polar compounds. The better the interactions between the inner surface of the column and the analyte, the better the separation. Both must be of similar polarity to achieve the best results. If polar compounds pass through a non-polar column, there will be minimal interaction, resulting in quicker elution and smaller signals, poor retention, resolution, separation, and peak shapes, making them harder to detect. There are two solutions: One possible solution could be a polar column, which on the other hand cannot be heated as high and as the extractives are mixtures, the opposite problem would occur with the hydrophobic analytes in the sample. The other solution is to derivatise polar functional groups, such as those present in carbocyclic acids, phenols or fatty alcohols, to reduce their polarity. One method is methylation of hydroxy and carboxy groups (Challinor 2001). A thermally assisted hydrolysis and methylation (THM) derivatisation with tetramethylammonium hydroxide (TMAH) was tested in preliminary studies on acetone extracts of teak pulp (Brazilian plantation). They showed promising results, yielding even more information from the extract.

Figure 34 illustrates an untreated acetone extract of Brazilian plantation teak pulp measured regularly by TD-GC-MS compared to the same extract after THM through the application of TMAH. The comparison clearly demonstrates an increase in the number of peaks, particularly in the early part of the chromatogram, as well as a notable expansion in the total peak area. The additional peaks occur for three reasons: 1. There are by-products and residues of TMAH which are also detected. 2. The hydrolysis reaction breaks down esters, e.g. from fats, oils and waxes present in wood/pulp extracts, into their constituent alcohols (glycerol) and fatty acids, which are methylated and detected. A former wax component, for instance, represented by a single peak or, in some cases, not even detected, undergoes breakdown by the hydrolysis reaction into two distinct peaks. 3. Some compounds were already visible in the untreated extract, but with smaller peak intensities due to the polarity incompatibility between the analyte and the inner column surface. After the methylation a new peak can occur, albeit at a different RT, as illustrated in Figure 35.

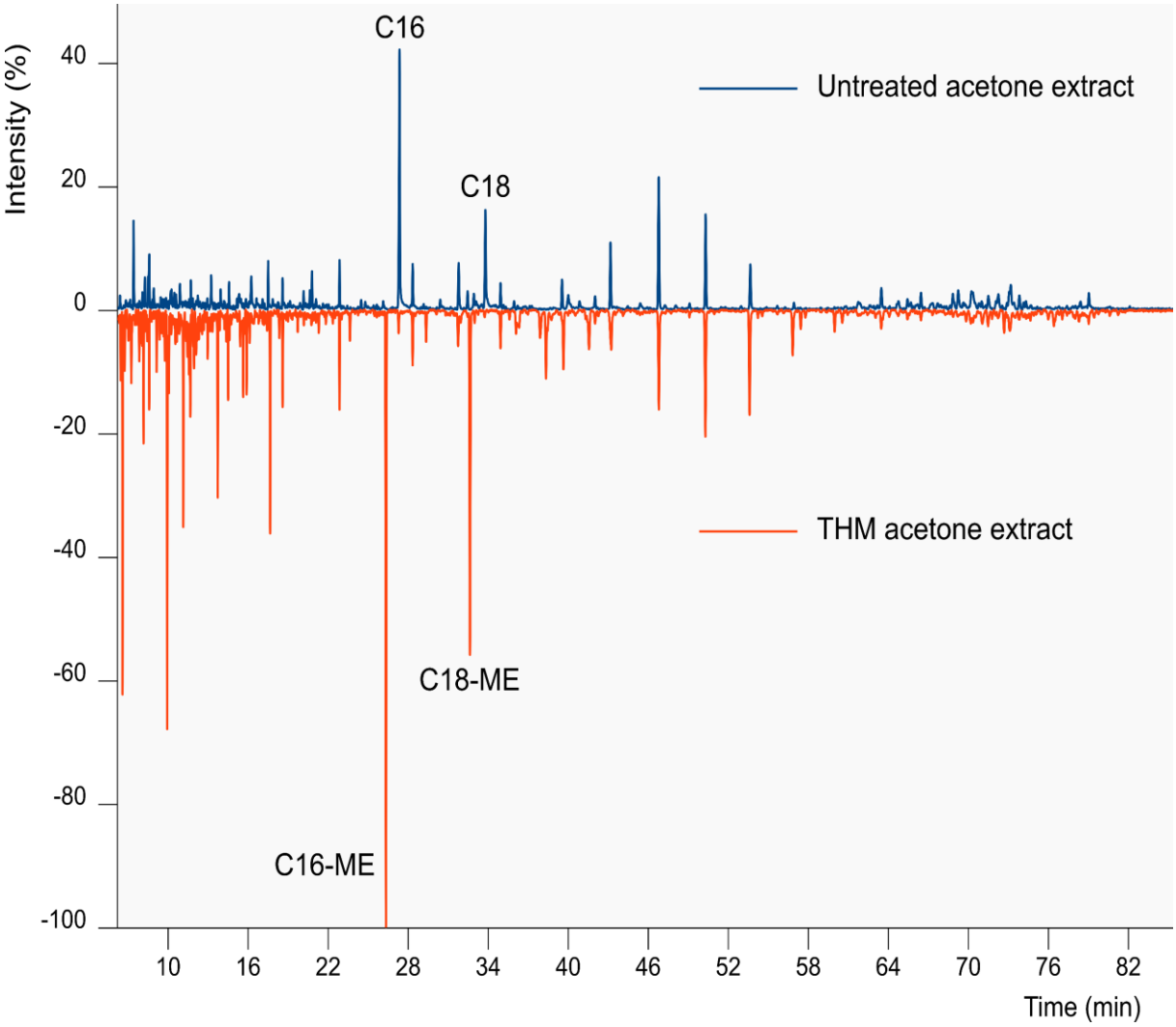


Figure 34: Comparison of chromatograms of Brazilian plantation teak acetone extracts: untreated (TD-GC-MS) vs. TMAH treated (THM-GC-MS).

The example of two fatty acids shows how the sample preparation and treatment method also affects the efficiency of the derivatisation (Figure 35). The untreated blue extract shows both C16 and C18 fatty acids in the chromatogram with a relatively small peak intensity and area. The partially methylated yellow chromatogram also shows the fatty acid methyl esters (FAMEs) formed. The red chromatogram shows the almost completely methylated extract with large FAMEs and hardly any original acids.

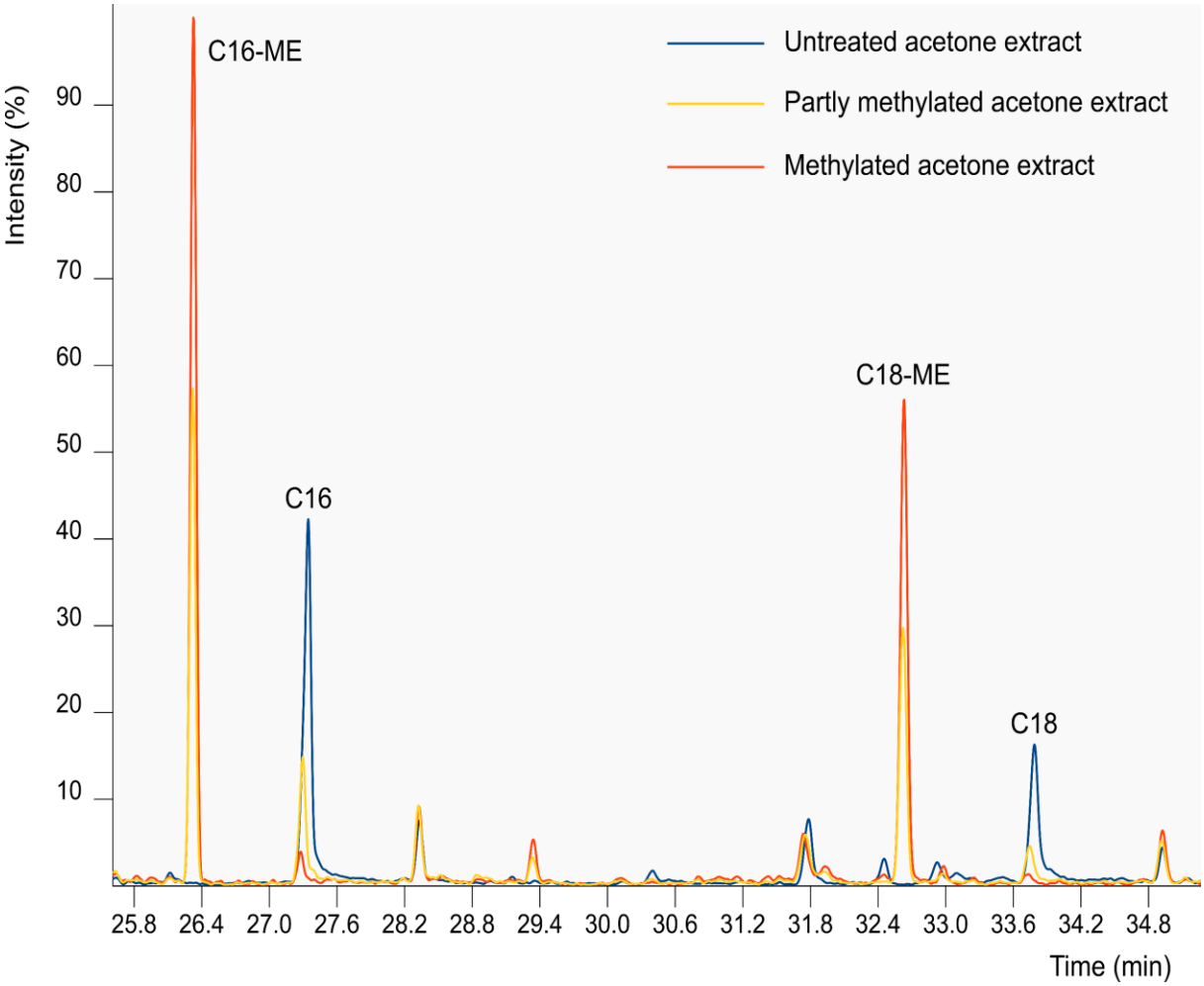


Figure 35: Comparison of chromatograms of acetone extracts; untreated (TD-GC-MS), partially methylated (-ME), and almost completely THM with TMAH.

Finally, the comprehensive acetone extracts or just the acetone fraction of a sequential extraction can be derivatised. The process of methylation has the effect of producing a greater number of peaks in the data set. Some of these peaks undoubtedly contribute to the overall value of the data, as they were not previously detectable. This enriches the information and has the potential to provide more specific marker peaks for taxon and originID. It is crucial to ensure that the derivatisation method is fully efficient to prevent the generation of false marker peaks due to partial methylation of some samples or peaks while others get fully methylated. Consequently, further development and optimization of the derivatisation method would be required.

4.1.3 TD-GC-MS: Tuning opportunities

The mass spectrometers of GC-MS instruments have to be calibrated (tuned) regularly to optimise MS performance. Tuning ensures accurate and consistent data by adjusting instrument parameters such as ion optics and detector settings. An important goal of tuning is to achieve reproducible results and a high accuracy of each mass trace (m/z). The m/z axis can shift and report incorrect masses (m/z values) if not tuned regularly. This does not mean that correct m/z values are always in whole numbers – due to the presence of isotopes, accurate mass trace values are often expressed as decimal numbers. These are reasons why MS data is rounded throughout the pre-processing steps, as explained in section *3.5.1* of Pub I based on the findings by Khrisanfov and Samokhin (2022). In addition, the sensitivity is optimised to detect low intensity signals. This results in better peak shape, narrower and higher intensities and therefore better resolution. Tuning is done by analysing perfluorotributylamine (PFTBA) as calibration substance. The tuning algorithm adjusts the MS settings until the known calibration ions are detected with the correct m/z values and the expected peak shape and intensity (signal-to-noise ratio).

There are two basic tuning algorithms on board of every GC-MS system – standard spectra tune (s-tune) and autotune (a-tune), elucidated by Agilent Technologies (2010). The s-tune has specific ratio targets for the specific calibration substance ions it is looking for, while the a-tune calibrates based on the maximum response of the calibration mass trace over a specific range of the tune. The a-tune only has a minimum target but an open-ended maximum depending on the current state of the system, while the s-tune has both minimum and maximum targets to be achieved for the standard compound (PFTBA) ions (m/z 69, 219, and 502). M/z 69 is set at 100 % and the other two (m/z 219 and 502) have specific limits at either end of the tuning process (s-tune) or only a minimum (a-tune) as explained above. The s-tune provides a standardisation suitable for a wide range of samples and typical applications. It provides a balanced performance between sensitivity, resolution and mass accuracy. This usually results in a better match quality of measured compounds with spectra databases. When used as the default in different systems, data sets measured in s-tune mode are more consistent within one system and across different systems. Because a-tune customises the instrument settings to be optimised at its individual level, typically in favour of higher mass traces and therefore larger molecules, the configurations do not necessarily match other systems (Agilent Technologies 2010).

The majority of the samples in this work have been measured in a-tune mode. This approach was selected at the outset of the method development process because the a-tune is highly effective when used with a single GC-MS system, as it does not impose calibration constraints on the maximum intensities. Additionally, it was assumed that the a-tune, which modestly elevates the abundances of larger m/z could be advantageous for taxonID. This was based on the assumption that larger non-polar components could be distinctive and, therefore, valuable marker compounds for individual species. It seems reasonable to posit that s-tuning and the creation of a database based on s-tuned chromatograms could enhance the precision

and reproducibility of both taxonID and originID, as indicated by method validation and system dependence results in section 3.2 of Pub II. This is because the a-tune parameters vary even within the same system, especially if there are extended periods of time between tunings. As s-tuning improves analytical consistency, such an approach would certainly improve interlaboratory repeatability and the shared use of an s-tuned reference database.

4.2 Database approach

In addition to the improved analytical protocol, laboratory routine and purity of solvents, the chosen software, which provides all sorts of functional filters for individual data processing, and its database and PCA tools, contributed greatly to the promising results of this work.

4.2.1 Pre-processing of GC-MS data

This chapter presents a discussion of data processing steps and potential improvements. The pre-processing steps and settings, described in section *3.5* of Pub I, have been subjected to continuous improvement and adaptation until Pub III in order to achieve optimal results. As explained above, the rounding of the mass traces (m/z) is important for the reproducibility of the generated data sets, since slight shifts in the mass-axis and isotopes lead to differing decimal places of the m/z values, which, for example, hinder the subsequent smoothing of the mass traces. Therefore, it needs to be compensated by efficient tuning (section 4.1.3) and normalisation, which was performed by rounding from -0.38 (incl.) to +0.62 (excl.) according to Khrisanfov and Samokhin (2022) (Pub I, section *3.5.1*).

The aforementioned shift in the mass-axis (m/z) is independent from the shift in the RT-axis. To compensate the RT shift, independent RIs are calculated for each peaks individual RT using the alkenes of the polyethylene standard's pyrolysis products as references (Pub I, section *3.5.9*). This is very important for comparison and matching against the reference peaks in the DB, especially when comparing across different GC-MS systems or after a change of the separation column.

The scan duplicator function has proved to be useful prior to smoothing, especially for low resolution data sets, as it takes two scans and creates a synthetic new scan in the middle, merged from the mass traces of the actually measured neighbouring scans. The new scans between the old ones give more data points throughout the chromatogram, which makes smoothing (Savitzky and Golay 1964) easier for small and low resolution peaks with poor and rough shapes. This filter is not necessary for good quality MS data at high scan rates and, if used, the operator must always be aware that the data will be significantly altered. This step also increases the size of the data, almost doubling it. The ion remover filter is also useful to quickly remove atmospheric and system inherent ions, but it must be used carefully as experience has shown that it is not always as easy as it seems. In the course of this work, the use of the ion remover option has steadily decreased. Experiments showed that some ions are more important for peak identification than they are detrimental to the data set. In some cases this was the case for m/z 207, which is normally associated with column bleed. This ion was deleted by default, but in some cases this particular mass is also present in peaks,

affecting the comparison with the NIST DB for compound identification, leading to poorer match qualities or no match results at all. All in all, removing air, water, nitrogen and some column bleed ions is generally a good idea, but the current standpoint for pulp extracts with database use would be do not remove too many or uncertain ones. Less is more, especially as the MCR-AR peak deconvolution recognises the mass traces belonging to a particular peak anyway, so there is no need to remove them beforehand.

The statistics-sensitive non-linear iterative peak-clipping (SNIP) baseline removal filter also lowers the baseline, making smaller peaks visible (Ryan et al. 1988; Wenig and Mailänder 2024). The SNIP baseline remover filter is not described in Pub I as it was later found to be useful and implemented into the set of pre-processing steps. The SNIP algorithm works by making an initial assumption about the relative height of the baseline (consisting of noise and background signals) in an area of the chromatogram where no peaks are detected. From this approximate starting point, it iteratively starts clipping. Data points lower than the surrounding points are progressively reduced. Sensitivity to true analyte peaks is achieved through statistical properties. That is the advantage over linear baseline corrections. SNIP can adapt to non-linear baseline changes. The process continues until a statistically stable baseline is found that will not be changed much by further iterative corrections (Ryan et al. 1988). This algorithm does not delete the noise signals completely, it just merely minimises their intensity while leaving the true peak signals intact. This is beneficial for visual peak recognition by the operator, who will be able to see peaks that stand out from the lowered baseline better and it supports the peak detection algorithms. However, it can also affect the area of the analyte peaks. Peak areas and intensities tend to become smaller, which is logical due to the removal of background signals. In theory, the post-SNIP peak size is close to the true and pure peak size. Depending on the actual baseline level, which in turn depends on the system used, vacuum quality, column wear, tuning and other things, it may be difficult to find the best SNIP settings to always remove the same amount and thus give reproducible peak area results. In the light of this concern, the PCA analysis for provenance discrimination in Pub III, which is highly dependent on comparable peak sizes, was based on data without baseline correction. In addition, MCA outcomes also depend on peak sizes and can therefore also be slightly influenced by varying SNIP settings.

The MCR-AR peak deconvolution and detection, according to Gerber et al. (2012), in combination with the RI calculation, are the most important tools and defiantly contribute the most to the positive result of this work. However, experience shows that even the MCR algorithm does not always work perfectly, or rather, quite regularly a few obvious peaks in the chromatogram (estimated 0–3) are not correctly detected. In most cases these are medium sized peaks. For this reason, a manual targeted peak deconvolution has been added to the pre-processing method. This so called MCR-AR UI allows the operator to manually set targets on the peaks or scans that were not detected before, thus initiating a second search for peaks at the defined location. This manual step in automated data processing is useful to check for

missed peaks, improve overall peak detection and obtain complete data sets as a basis for further analysis.

The high pass ions filter was used in the early days of method development in two different ways, to retain either the 150 highest intensity ions per **scan** (Figure 36) or per **peak**, as described in section *3.5.8* of Pub I. The idea behind this was data reduction: by getting rid of unnecessary data from very low abundance mass traces, memory was saved and data processing accelerated. This option is no longer used in the current database setup as it discriminates the smaller peaks in overlapping peaks (Figure 36).

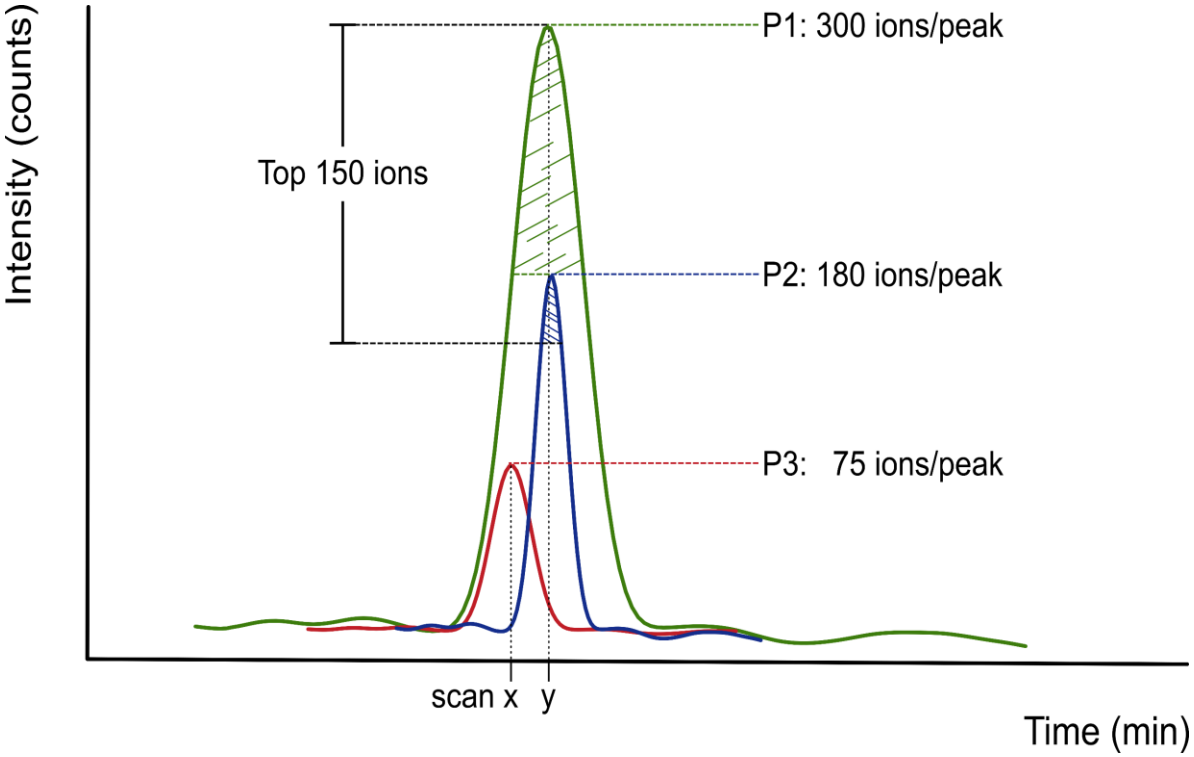


Figure 36: Schematic diagram of the high pass ions filter applied per scan, with three exemplary peaks 1–3 (P1–3).

There is a risk that the peak deconvolution will be disturbed, which could result in suboptimal matching of overlapping peaks. This can be seen in Figure 36: If only the 150 highest intensity m/z are retained from scan y, this means that 130 m/z of the green peak 1 (P1) and only the 30 highest of the smaller blue peak 2 (P2), consisting of 180 ions in total, are retained due to the same RT of both. In another wood pulp, P1 may be absent, resulting in the retention of the 150 highest ions of P2 in the reference DB. This would result in a poor MF between these two P2 references and both would be stored separately as markers for their respective species, even though they represent the same extractive substance. The red peak 3 (P3), the smallest of the three, reaches its maximum at scan x. It is less affected by the high pass ions filter, which demonstrates the extent to which it relies on effective peak separation through GC-MS. Keeping the highest ions per **peak** rather than per **scan** might seem better. However, it also carries risks with the MS detection threshold and different peak sizes in pulp blends of different species. At this stage, for pulp with many similar substances and many small and

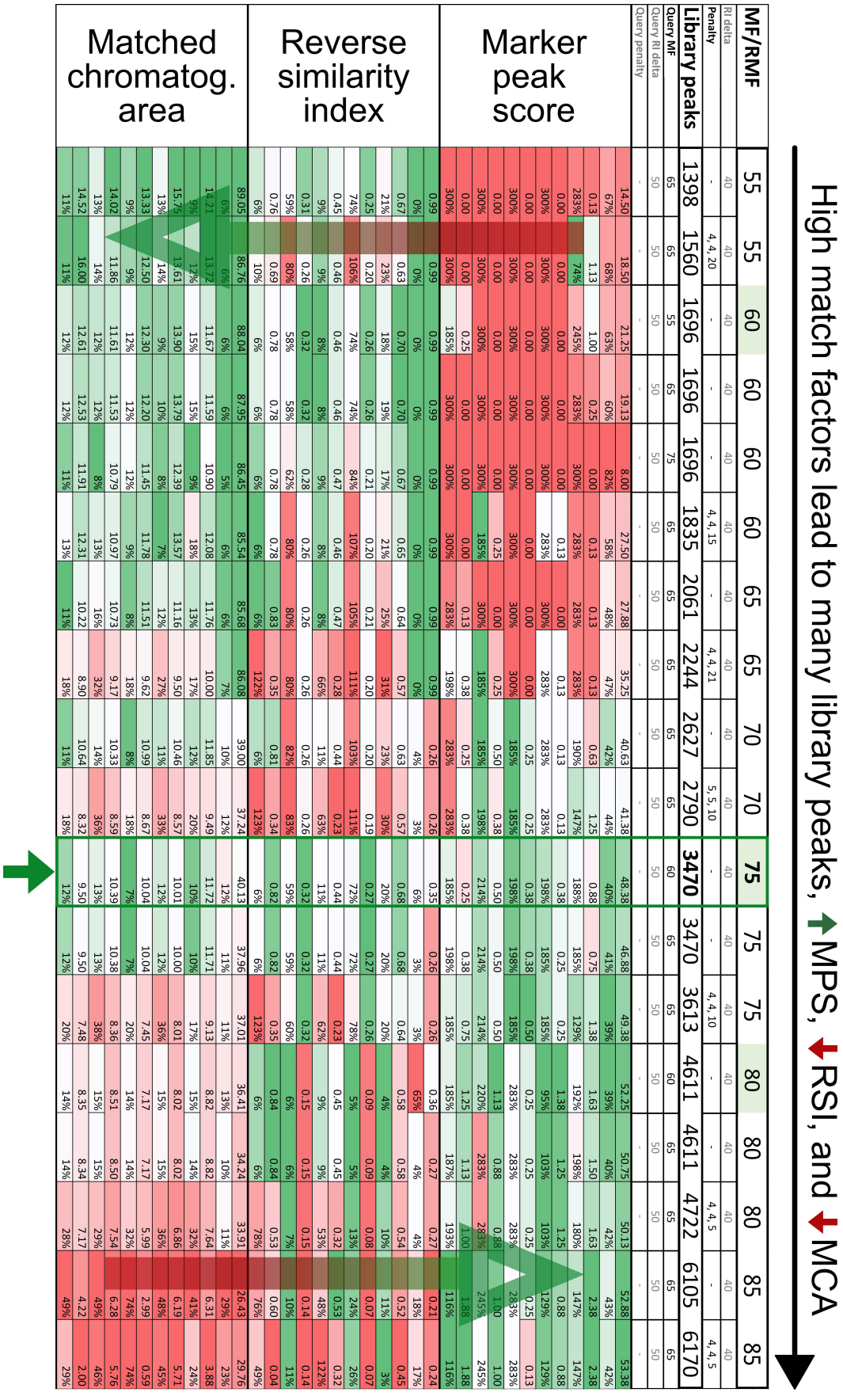
overlapping peaks, it appears that the most optimal approach would be to retain all ions. More information is usually beneficial and the potential memory saving is less important than the proper functioning of the marker peak detection.

4.2.2 Configuration of database settings

The choice of appropriate parameters for creating a database is fundamental to its subsequent performance. Many settings were tried and compared. The latest findings and best practices are discussed in this chapter. There are hundreds of extractives per reference species, many of which are similar. Many species have a high degree of overlap of the same extractives with others species. This means that few marker compounds (Figure 16) are expected, which is an important finding when building the database. The presence of many similar compounds/molecules means that their mass spectra are also very similar. To discriminate between these similar but not identical extractives, the correct MF must be chosen. In general, the less similar the DB reference samples are, the lower the MF can be chosen as there is no risk of mass spectra being falsely matched. In the case of setting up the DB of the wood pulp extractives in this work, the comparison of the mass spectra of the substance works best with a relatively high MF of 75. This indicates that all peaks from all species in a defined RI delta window of ± 40 are matched with the peaks recorded so far. Only if two or more peaks match with a value higher than 75 are they accepted as equal and from the same compound. These matched peaks are stored in the DB as ambiguous peaks for multiple references. If their MF is below the threshold of 75 they are accepted as individual marker peaks and stored in the DB for **one** reference only. Table 18 shows the relationship between MF and library peaks.

The higher the MF threshold for peaks to overcome, the fewer ambiguous peaks are found and the more marker peaks are stored in the DB. The data set has a total of 7 810 reference peaks from 38 wood species with ~200 peaks each. The lower the MF, the more ambiguous peaks were found, leading to a lower total number of library entries (ambiguous and marker) (MF 55; 1 398 entries). The higher the MF, as shown in Table 18 (MF 55–85), the more marker and therefore total library peaks were stored in the DB (MF 85; 6 170 library entries). As a result of the large number of marker peaks in the DB, a query of an unknown extract with this (high MF; high library peaks) DB will produce a high number of matched marker peaks (MPS). This observation is shown in Table 18 as mainly green highlighted MPS on the right-hand side of the MSP table rows. The individual numbers in the table are averages of eight measurements on the three different GC-MS systems described in Pub II. The percentages below each number are the corresponding CVs. The objective of this analysis is not to examine the performance of individual numbers, but rather to gain an overview of the relative efficacy of different database settings. The colour grading demonstrates favourable values in green and unfavourable values in red in relation to all values in the same row. To illustrate, the first row displays 53.4 marker peaks on the far right (dark green = optimal value in this row) and 14.5 on the far left (light red = relatively poor, but not the least favourable value in this row). The same is true for RSI and MCA: higher values are preferable and appear in green.

Table 18: Test of different database settings. The green arrow points to the most effective all-round parameters (MF 75, RI delta 40, no penalty) for balanced MPS, RSI, and MCA results.



	MF/RMF	55	55	60	60	60	65	65	70	75	75	80	80	80	85	85
	RI delta	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
	Penalty	-	4, 4, 20	-	-	4, 4, 15	-	4, 4, 21	-	5, 5, 10	-	4, 4, 10	-	4, 4, 5	-	4, 4, 5
	Library peaks	1398	1560	1696	1696	1835	2061	2244	2627	2790	**3470**	3470	3613	4611	4611	4722	6105	6170
	Query MF	65	65	55	75	65	65	65	65	65	60	65	65	65	65	65	65	65
	Query RI delta	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
	Query penalty	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Marker peak score		14.50	18.50	21.25	19.13	8.00	27.50	27.88	35.25	40.63	41.38	48.38	46.88	49.38	52.25	50.75	50.13	52.88	53.38
		67%	68%	63%	60%	82%	58%	48%	47%	42%	44%	40%	41%	39%	39%	40%	42%	43%	42%
		0.13	1.13	1.00	0.25	0.00	0.13	0.13	0.13	0.63	1.25	0.88	0.75	1.38	1.63	1.50	1.63	2.38	2.38
		283%	74%	245%	283%	300%	283%	283%	283%	190%	147%	188%	185%	129%	192%	198%	180%	147%	147%
		0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.13	0.13	0.13	0.38	0.25	0.25	1.38	1.25	1.25	0.88	0.88
		300%	300%	300%	300%	300%	283%	283%	283%	283%	283%	198%	185%	185%	95%	103%	103%	129%	129%
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.25	0.38	0.38	0.50	0.25	0.25	0.25	0.25	0.13
		300%	300%	300%	300%	300%	300%	300%	300%	185%	185%	198%	198%	185%	283%	283%	283%	283%	283%
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.38	0.50	0.50	0.50	1.13	0.88	0.88	1.00	1.00
		300%	300%	300%	300%	300%	300%	185%	185%	198%	214%	214%	214%	220%	283%	283%	283%	245%	245%
		0.00	0.00	0.25	0.00	0.00	0.13	0.38	0.25	0.38	0.25	0.38	0.75	1.25	1.13	1.13	1.00	1.88	1.88
		300%	300%	185%	300%	300%	283%	198%	283%	283%	185%	198%	185%	185%	187%	187%	193%	116%	116%
		0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.26	0.26	0.35	0.26	0.26	0.36	0.27	0.27	0.27	0.21	0.24
		0%	0%	0%	0%	0%	0%	0%	4%	3%	6%	3%	3%	65%	4%	4%	4%	18%	17%
		0.67	0.63	0.70	0.70	0.67	0.65	0.64	0.63	0.57	0.68	0.68	0.64	0.58	0.58	0.58	0.54	0.52	0.45
		21%	23%	18%	19%	17%	21%	25%	23%	30%	20%	20%	20%	4%	4%	4%	10%	11%	3%
		0.25	0.20	0.26	0.26	0.21	0.20	0.21	0.20	0.19	0.27	0.27	0.26	0.09	0.09	0.09	0.08	0.07	0.07
		74%	106%	74%	74%	84%	107%	105%	111%	103%	111%	72%	78%	5%	5%	5%	13%	24%	26%
		0.45	0.46	0.46	0.46	0.47	0.46	0.47	0.28	0.44	0.23	0.44	0.23	0.45	0.45	0.45	0.32	0.53	0.32
		9%	9%	8%	8%	9%	8%	8%	66%	11%	63%	11%	62%	9%	9%	9%	53%	48%	122%
		0.31	0.26	0.32	0.32	0.28	0.26	0.26	0.26	0.26	0.26	0.32	0.32	0.15	0.15	0.15	0.15	0.14	0.14
		59%	80%	58%	58%	62%	80%	80%	82%	83%	83%	59%	59%	6%	6%	6%	7%	10%	11%
		0.76	0.69	0.78	0.78	0.78	0.78	0.83	0.81	0.34	0.34	0.82	0.82	0.84	0.84	0.84	0.53	0.60	0.04
		6%	10%	6%	6%	6%	6%	6%	6%	123%	123%	6%	123%	6%	6%	6%	78%	76%	49%
Reverse similarity index		89.05	86.76	88.04	87.95	86.45	85.54	85.68	86.08	39.00	37.24	40.13	37.96	37.01	36.41	34.24	33.91	26.43	29.76
		6%	6%	6%	6%	5%	6%	6%	7%	10%	12%	12%	11%	11%	13%	10%	11%	29%	23%
		14.21	13.72	11.67	11.59	10.90	12.08	11.76	10.00	11.85	9.49	11.72	11.71	9.13	8.82	8.82	7.64	6.31	3.88
		9%	12%	15%	15%	9%	18%	13%	17%	12%	20%	10%	10%	17%	15%	14%	32%	41%	24%
		15.76	13.61	13.90	13.79	12.39	13.57	11.16	9.50	10.46	8.57	10.01	10.00	8.01	8.02	8.02	6.86	6.19	5.71
		13%	14%	9%	10%	8%	7%	12%	27%	11%	33%	12%	12%	36%	15%	15%	36%	48%	45%
		13.33	12.50	12.30	12.20	11.45	11.78	11.51	9.62	10.99	8.67	10.04	10.04	7.45	7.17	7.17	5.99	2.99	0.59
		9%	9%	12%	12%	12%	9%	8%	18%	8%	18%	7%	7%	20%	14%	14%	32%	74%	74%
		14.02	11.86	11.61	11.53	10.79	10.97	10.73	9.17	10.33	8.59	10.39	10.38	8.36	8.51	8.50	7.54	6.28	5.76
		13%	14%	12%	12%	8%	13%	16%	32%	14%	36%	13%	13%	38%	15%	15%	29%	49%	46%
		14.52	16.00	12.61	12.53	11.91	12.31	10.22	8.90	10.64	8.32	9.50	9.50	7.48	8.35	8.34	7.17	4.22	2.00
		11%	11%	12%	12%	11%	13%	11%	18%	11%	18%	12%	12%	20%	14%	14%	28%	49%	29%
Matched chromatog. area																			

Table 18 clearly shows that lower MF settings in the DB setup result in higher RSI values and higher MCA, but significantly lower MPS. In consideration of the current data set comprising 38 reference pulp extractives, a medium high MF of 75 (green arrow) was identified as the optimal setting to achieve balanced query scores for all three parameters, thereby establishing the most comprehensive information base for the identification of unknown species.

Other factors that effectively influence the DB performance are the peak comparison algorithm used to match mass traces, the RI window in which this algorithm searches for matches, the penalty settings and the query settings used to match the unknown against the DB: The NIST matching algorithm, based on Stein and Scott (1994) has recently been used as the best performing mass spectrum comparator. The RI window of ± 40 has recently been found to be the optimal choice. This RI delta is considerably larger than the ± 10 window used in Pub I, which remains effective for the majority of peaks due the precision of RI calculations during data pre-processing. However, in certain instances, a larger window was necessary for two reasons: peak shape and size, as illustrated in Figure 37.

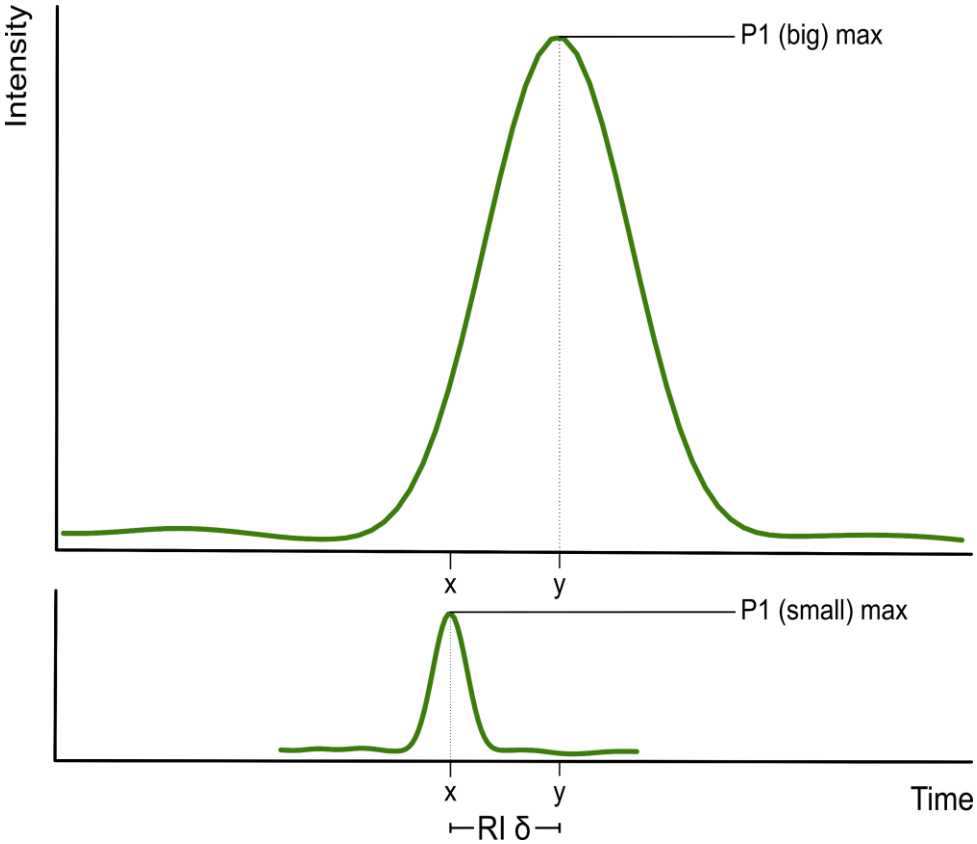


Figure 37: Peak size affecting the position of its maximum on the RT axis.

When the peaks in question are large, such as P1 (big) in Figure 37, the peak maximum is observed to be situated at a later point on the time axis in comparison to the peak maximum of a peak originating from the same compound but of a smaller magnitude, such as P1 (small). The RT at the peak maximum is used for the calculation of its RI. The discrepancy in RT/RI was observed in samples containing the same substance at varying concentrations, resulting in

significant size variations. This necessitated a larger RI delta in the DB configuration. Furthermore, the peak shape also influences the position of its peak maximum, requiring RI adjustments in some cases. The relatively wide RI delta of ± 40 enabled the correct matching of these examples without any adverse effects. The application of penalty settings can reduce the MF if the RI delta is too large. The following case provides an illustration: Two peaks are matched with a MF of 80 and an RI difference of δ20 between them. If the penalty threshold is set to a value of -1 MF per 1 RI above the threshold of 10 RI, this would result in a reduction of 10, thereby yielding a MF of 70 for this particular match of mass spectra. Consequently, if the MF threshold was set to 75, these two peaks would be stored separately as marker peaks in the database, as they are too distant from each other in terms of RI in the chromatograms. This penalty option was systematically tested with different settings, but no superior results were obtained. For this type of pulp extractive data, the results were inconclusive, and therefore the penalty approach was excluded from the current best practice settings.

In order to query the database, it was determined that the optimal threshold for the MF was 60 with a RI delta of 50, with no penalties applied. The MF and RI delta settings were both slightly lower and wider than the settings used for DB creation. This is recommended in the software manual and has been confirmed by experimentation. It follows that the MF should be smaller for querying the DB, given that pulp mixtures may contain species in small amounts, resulting in peaks in the mixture that are much smaller. As a consequence of reduced peak heights, some m/z intensities fall below the GC-MS detection limit, as previously outlined in the section on the high pass ion filter (4.2.1). The MF was therefore less effective for mixtures than for pure samples, thereby justifying the lower MF of 60 for querying. Furthermore, peak shape and size issues are exacerbated in mixed samples containing minute quantities of each species. This reinforces the rationale for selecting a wider RI delta of 50 as an optimal option.

4.2.3 Strengths and limitations

A software advantage of this chemotaxonomic database approach is that it is capable of storing and comparing thousands of peaks in defined RI windows with defined matching settings in very short processing times. Data pre-processing can be partially automated by batch processing – running on many GC-MS data sets at once. The process steps applied to all samples are identical, not only in sample preparation but also in data handling. This ensures that the entire process is highly reproducible. Only in some instances, as described before, the manual addition of a few peaks, which are occasionally not automatically detected, is required. The output of the DB query is not yet sufficiently unequivocal in all cases to base correct identification decisions on. However, the analysis of the marker peaks in combination with the similarity indices and the matched chromatogram area provides valuable indications in most cases. Nevertheless, in some cases the DB query results remain challenging to interpret and may occasionally be misleading, leading to erroneous identification decisions, as demonstrated in section *3.2* of Pub II.

4.2.3.1 System (in)dependence

In Pub II, the method was tested on three different but similar GC-MS systems within the same laboratory setting, primarily by the same operator with the same samples. The results, demonstrated superior intermediate precision values across instruments for mixed samples in comparison to pure samples and the results for RSI and MCA were much better than for MPS. For MPS, the ANOVA tests showed statistically significant differences between instruments, which was already apparent from the variance of the results compared between the different systems. The variance was high, both between and within groups.

The observed, partial system dependency was found to be largely attributable to the original system producing significantly higher MPS, which is both beneficial and detrimental. This indicates that the other systems demonstrated comparable stability and performance. However, it also suggests that the library peaks in the database may contain some instrument-specific ion patterns, which could result in improved MF and increased MPS when samples measured on the original system are queried. This needs to be avoided in further research and use of the newly created DB.

The biggest levers available for optimising the database are the number and selection of peaks and ions, as well as the improved comparability of MS-spectra by using s-tune. Recently, around 200 peaks per chromatogram have been analysed and filtered through the high pass peaks filter. They always retained the same number of the largest peaks (by area) per chromatogram. This approach may be too static and not necessarily the optimal number of peaks per sample. The number of peaks found to be effective is approximately 200, which serves as a useful general guideline. However, in instances where a greater number of peaks is present in a chromatogram, 300 or 400 peaks may still be a suitable option to obtain additional marker peaks that are sufficiently sized for analysis, even in mixtures with low amounts. There are two methods that may be employed to achieve this objective individually adapted peak numbers per sample. Firstly, peak filtering could be implemented by actual peak size ranges, as opposed to high pass filtering by fixed numbers (e.g. all peaks with areas ≥ x should be retained). This approach would be more dynamic than the static high pass filter and might permit the passage of only 50 peaks in some instances, while allowing for the inclusion of more than 400 peaks in others. This would be dependent on the peak sizes present in the chromatogram, rather than a fixed number of peaks. Secondly, the alternative approach would be to configure the MCR-AR peak detection parameters in such a way that peaks with a signal-to-noise ratio below a specified threshold are excluded. This would allow for a dynamic number of peaks, adapted to the individual sample, to be realised.

The best strategy for improving the DB performance and resolving the issue of system dependency is to conduct continuous s-tuning of all instruments. As outlined in section 4.1.3, the s-tune mode strives to achieve a balance between sensitivity and resolution by adjusting the instrument parameters within a uniform predefined range of minimum and maximum values. This results in a quality improvement of the raw data in terms of both comparability and reproducibility.

4.3 Comparison with anatomy

For the identification of wood taxa in pulp and paper, the long-established classical anatomical method is undoubtedly the best and first choice.

4.3.1 Blind test results and synergies

The results of the blind test conducted in Pub II were promising, with partly complementary strengths that can be exploited to improve the combined efficacy of both methods. In practice, the realistic future role of chemotaxonomy is that of a supporting method alongside the anatomical method. The anatomical method is a well-established and widely accepted approach. A significant number of experts and research laboratories around the world are already engaged in this field, having accumulated a substantial amount of experience and expertise, as exemplified by the TI, who participated in the blind test in Pub II (section *3.1.2*). Many reference samples are currently available and continuously increasing. The method is routinely used and can be readily adopted by other laboratories, which is a great advantage. As demonstrated in the blind test, the two other participating laboratories, in addition to the TI, only required the provided microscopic vessel element images and the corresponding descriptions as reference material to conduct the experiment. Maceration is a relatively rapid routine work, whereas microscopy necessitates a high level of expertise and knowledge to identify the relevant features. However, it is less laborious, faster and cheaper than chemotaxonomy. The future of chemotaxonomy will be characterised by its coexistence with anatomy, with a particular focus on assisting in special cases where highly beaten pulp or other issues impede anatomists from finding intact anatomical features.

The other future role of the chemotaxonomic method could be in special cases of uncertainty, where there are mixtures of wood species that are too difficult to distinguish anatomically. The blind test demonstrated difficulties in distinguishing between the Anacardiaceae family members *Mangifera* spp. and *Swintonia* spp. for instance. Additionally, *Durio* spp., *Lophopetalum* spp. and the CITES Appendix II protected *Gonystylus* spp. (Figure 20) showed a notable degree of similarity in vessel anatomy. In the event that these taxa are of high relevance, it may be justifiable to expend additional resources (cost and effort) to create bespoke, small databases with higher discriminatory power than a large collection. It may be more cost-effective to create high-utility, specialised databases before attempting to build a comprehensive, all-encompassing database that includes all traded taxa.

4.3.2 OriginID

To date, chemotaxonomy is the only possible method that has demonstrated to be effective for originID in pulp. Although Pub III was only a preliminary study on this topic, the proof of concept was provided for different taxa at the genus and species levels in various regions and across plantations and natural forests. This could constitute a substantial contribution to the enforcement of CITES, and would undoubtedly be beneficial for the entire field of pulp identification analysis. The chemotaxonomic originID in cooperation with rapid and cost-effective taxonID, provided in advance by anatomy, could be a promising partnership,

particularly in light of the growing importance of provenance information of pulp samples in view of the impending implementation of the EUDR on 30 December 2025 after the current postponement. As detailed in section 1.1.2.2 (Habitat loss), 7 % of the world's forests are planted forests and 3 % are plantations (FAO 2020). The number of plantations is increasing rapidly, especially in China and Indonesia (FAO 2020), which is likely to make this issue even more pressing in the future.

5 Conclusion

The principal objectives of this dissertation were to develop a chemotaxonomic methodology for the identification of wood taxa and determination of wood origin in pulp and paper products. To this end, chemical fingerprinting, based on pulp extractant profiles, was employed to construct a reference database for the comparison of unknown samples.

The chemotaxonomic approach offers a standardised and partly automated method, including cryo-ball milling, Soxtherm extraction and uniform batch pre-processing of TD-GC-MS data for analysis. To ensure the continued effectiveness of the DB and support ecological goals, it is essential to maintain and regularly expand it. Furthermore, references of emerging wood taxa that gain relevance in the industry or in conservation through CITES protection listings have to be included.

The method is largely system independent, generating reproducible results across different instruments for mixed samples. However, for pure samples, the parent system, on which the DB was set up, produced higher scores of marker peaks. This issue is assumed to have partly arisen from individual tunings of the compared instruments. One potential solution is the implementation of consistent s-tuning of all instruments used.

A comprehensive blind test of the established anatomical and new chemotaxonomic method demonstrated positive identification hit rates for both methods. Furthermore, the combination of both methods appears to improve overall performance, particularly for challenging samples. This is due to the complementary strengths of the two methods, which can be used to support each other and enhance identification reliability. As a combined approach, it is therefore applicable to a wider range of taxa and sample conditions, such as highly beaten pulps.

The proof of concept for the chemotaxonomic provenance identification was successfully demonstrated for ramin, red meranti and teak pulp. The discriminating PCA revealed different distinct extractives, associated with certain origins. These findings prove the method's applicability to geographic originID and the potential for distinguishing between plantation and natural forest sources. This could be of particular importance in the context of combating illegal deforestation, as there is currently no other methodological option available for this purpose in pulp and paper products.

As proposed by Dormontt et al. (2015) and Silva et al. (2021), to reinforce global crime prosecution and environmental regulations, robust scientific support based on multiple, combined and collaborative technologies is necessary to achieve the ultimate goal of accurately answering specific identification questions. The chemotaxonomic method has the potential to become an independent and complementary tool to the anatomical method for taxonID, and especially for originID, in pulp and paper products, contributing to the conservation of forests and biodiversity, one of the major challenges of our time.

6 Outlook

There are three outlook topics of interest for the future of the chemotaxonomic method. Firstly, there is a need to validate the method further for the utilisation on a diverse range of industrial samples. Secondly, it is valuable to gain insight into the underlying reasons for the observed differences in extractive fingerprints. Thirdly, there is a potential for improvement and automation of the DB analysis through the application of machine learning techniques.

Further validation steps are required to ensure the applicability of this method to real-market industrial pulp and paper samples across different laboratories. Firstly, the functionality of the methodology must be validated on the same reference taxa as previously, but with the use of different, novel pulps for each taxon. This is a crucial consideration, as the samples employed in the blind test conducted in Pub II are the same ones on which the current DB is based. Secondly, it is essential to ensure the consistency of the method's performance across a variety of sample treatments by testing diverse pulping and bleaching approaches on the same taxa. In the event that differing pulping and bleaching conditions affect the chemical fingerprints of certain species, these pulp extracts should also be incorporated into the DB references. In order to gain the most comprehensive market overview, it would be optimal to conduct data collection directly at the pulp mills. Prior to the chemotaxonomic analysis of the extractives in the pulp, an anatomical analysis of the solid wood used would be required. The data obtained regarding the wood taxa employed, the industrial pulping and bleaching processes conducted, and the resulting extractives profiles would be of significant value as references for the DB. This approach would be complex and could only be realised in cooperation with the manufacturers, preferably directly in the tropical countries. However, it would still be necessary to produce reference pulps of protected taxa in the laboratory. Thirdly, the DB should be validated in addition to the intermediate precision across different instruments within the same laboratory (Pub II), across different laboratories by other operators and over longer periods of time as well. In a noteworthy demonstration, Price et al. (2022) illustrated an effective example on how to test the reproducibility of a database approach based on DART-TOF-MS data of solid heartwood tissue across diverse laboratories. Following the completion of these three future validation stages, the methodology can be applied to real-life industrial samples for a broader utilisation of the method.

More research is needed on the question raised by Deklerck (2023) of what drives extractive fingerprint differences – environment or genetics? The technology capable of identifying wood taxa and wood provenance by chemical fingerprinting is provided and very valuable. However, its further development could be facilitated by a more comprehensive understanding of the reasons and mechanisms underlying the synthesis of different extractives by specific species in specific locations and what influences this the most. This would be particularly beneficial, especially in the context of originID.

The question of chemotaxonomy and AI is an intriguing one, yet it is challenging to provide a definitive answer. The GC-MS data of pulp extractives are highly complex and to some extent

dependent on the detection system used and its current condition. Variance is introduced even in repeated measurements due to factors such as MS sensitivity, vacuum quality and column wear. Consequently, the utilisation of AI would be challenging and require a substantial quantity of training data. It is still promising because deep learning is particularly adept at analysing complex data sets and identifying patterns within them. Presently, the operator is required to make determinations based on their experience, manually examining specific peaks and fine-tuning the data processing if the results are not clear. It is inevitable that there will be discrepancies in measurement conditions, as well as interferences between different samples and within new mixtures of extractives that form new compounds during the process. This will result in varying GC-MS data, as demonstrated by present results of Pub II, illustrating a considerable degree of variance in query outcomes. There is currently no linear pattern or statistical certainty that a reference hit with specific outcomes (e.g. > x MPS, > y RSI, > z MCA) can be definitively identified in a mixture. Consequently, the involvement of an experienced operator is essential at this stage of development. Even with consistent sample preparation and automated data processing, a DB query will always depend on the specific sample, the instrument's condition and other variables. It might be possible that a machine learning approach could compensate analytical variations and recognise underlying patterns which are not yet visible in the current approach. It would be desirable to receive probability value outputs from the AI, which could then be subjected to plausibility verification by an experienced human.

References

Agilent Technologies (2010). How autotune vs. standard spectra tune can affect analyte ion ratios using Agilent mass spectrometers, Available at: https://www.agilent.com/cs/library/support/documents/f05036.pdf (Accessed 26 June 2024).

Alaerts, G., Pieters, S., Logie, H., van Erps, J., Merino-Arevalo, M., Dejaegher, B., Smeyers-Verbeke, J. and Vander Heyden, Y. (2014). Exploration and classification of chromatographic fingerprints as additional tool for identification and quality control of several Artemisia species. *Journal of Pharmaceutical and Biomedical Analysis* 95: 34–46, https://doi.org/10.1016/j.jpba.2014.02.006.

Alfassi, Z.B. (2004). On the normalization of a mass spectrum for comparison of two spectra. *Journal of the American Society for Mass Spectrometry* 15: 385–387, https://doi.org/10.1016/j.jasms.2003.11.008.

Alix-Garcia, J. and Gibbs, H.K. (2017). Forest conservation effects of Brazil's zero deforestation cattle agreements undermined by leakage. *Global Environmental Change* 47: 201–217, https://doi.org/10.1016/j.gloenvcha.2017.08.009.

Alvarez-Berríos, N.L. and Mitchell Aide, T. (2015). Global demand for gold is another threat for tropical forests. *Environ. Res. Lett.* 10: 14006, https://doi.org/10.1088/1748-9326/10/1/014006.

Anita, Y., Tanifuji, K., Ohi, H. and Evelyn, E. (2021). Kraft cooking with teak wood extract and determining residual 2-methylanthraquinone in eucalyptus pulp. *Japan TAPPI Journal* 75: 153–163.

Arima, E.Y., Barreto, P., Araújo, E. and Soares-Filho, B. (2014). Public policies can reduce tropical deforestation: lessons and challenges from Brazil. *Land Use Policy* 41: 465–473, https://doi.org/10.1016/j.landusepol.2014.06.026.

Ashton, P.S. and Heckenhauer, J. (2022). Tribe Shoreae (Dipterocarpaceae subfamily Dipterocarpoideae) finally dissected. *Kew Bull* 77: 885–903, https://doi.org/10.1007/s12225-022-10057-w.

Asif, M.J. and Channon, C.H. (2005). DNA extraction from processed wood: a case study for the identification of an endangered timber species (Gonystylus bancanus). *Plant Molecular Biology Reporter* 23: 185–192,

Assunção, J., Gandour, C., Rocha, R. and Rocha, R. (2020). The effect of rural credit on deforestation: evidence from the Brazilian Amazon. *The Economic Journal* 130: 290–330, https://doi.org/10.1093/ej/uez060.

Augusti, A. and Schleucher, J. (2007). The ins and outs of stable isotopes in plants. *The New Phytologist* 174: 473–475.

Australian Government (2021). *Illegal logging prohibition act 2012 - No. 166, 2012: compilation No. 4*, Canberra, Australia, Available at: https://faolex.fao.org/docs/pdf/aus158000.pdf (Accessed 1 October 2024).

Australian Government (2022). *Nature positive plan: better for the environment, better for business*, Department of Climate Change, Energy, the Environment and Water (DCCEEW), Canberra, Australia, Available at: https://www.dcceew.gov.au/environment/epbc/publications/nature-positive-plan (Accessed 28 September 2024).

Australian Government (2024). *Progressing our nature positive law reforms*, Department of Climate Change, the Environment, Energy and Water (DCCEEW), Canberra, Australia, Available at: https://www.dcceew.gov.au/about/news/progressing-nature-positive-law-reforms (Accessed 29 September 2024).

Azevedo, A.A., Rajão, R., Costa, M.A., Stabile, M.C.C., Macedo, M.N., Dos Reis, T.N.P., Alencar, A., Soares-Filho, B.S. and Pacheco, R. (2017). Limits of Brazil's forest code as a means to end illegal deforestation. *Proceedings of the national academy of sciences of the United States of America* 114: 7653–7658, https://doi.org/10.1073/pnas.1604768114.

Azote for Stockholm Resilience Centre. *Planetary boundries*, Stockholm University, Stockholm, Sweden, Available at: www.stockholmresilience.org/research/planetary-boundaries.html (Accessed 4 September 2024).

Backhaus, K., Erichson, B., Plinke, W. and Weiber, R. (2016). *Multivariate Analysemethoden: eine anwendungsorientierte Einführung,* 14., überarbeitete und aktualisierte Auflage, Springer Gabler, Berlin, Heidelberg, Germany.

Baffoni, S., Miettinen, O., Tinhout, B., Supartinah, W. and Haggith, M. (2017). *Too much hot air: the failure of the Indonesian pulp and paper industry to reform its management of peatlands*, Environmental Paper Network (EPN), Available at: https://environmentalpaper.org/wp-content/uploads/2019/10/Too-much-hot-air-20170426.pdf (Accessed 24 September 2024).

Bansal, M., Nagaraju, S.K. and Prasad, V. (2019). The origin and diversification of plant family Dipterocarpaceae. *OJG* 09: 593–596, https://doi.org/10.4236/ojg.2019.910050.

Barstow, M. (2018). *Gonystylus bancanus. The IUCN red list of threatened species 2018: e.T32941A68084993*, Available at: https://dx.doi.org/10.2305/IUCN.UK.2018-1.RLTS.T32941A68084993.en (Accessed 29 July 2024).

Beck-O'Brien, M., Egenolf, V., Winter, S., Zahnen, J. and Griesshammer, N. (2022). *Everything from wood: the resource of the future or the next crisis? How footprints, benchmarks and targets can support a balanced bioeconomy transition*, Universität Kassel, Center for Environmental System Research, WWF Germany, Berlin, Deutschland.

Bennett, E.M., Carpenter, S.R. and CARACO, N.F. (2001). Human impact on erodable phosphorus and eutrophication: a global perspective. *BioScience* 51: 227–234, https://doi.org/10.1641/0006-3568(2001)051[0227:HIOEPA]2.0.CO;2.

Berger, J., Flaig, M.L., Olbrich, A., Helmling, S. and Saake, B. (2021). *MTH III - Abschlussbericht DBU-Projekt AZ 34295/01: Nachweis von Tropenholz in Papier - Chemotaxonomie und Anatomie zur Identifizierung von Mixed Tropical Hardwood*, Bericht, Hamburg, Available at: https://www.dbu.de/projektdatenbank/34295-01/ (Accessed 20 February 2024).

BGCI (2021). *State of the World's trees*, Botanic Gardens Conservation International (BGCI), Richmond, UK.

Björklund Jansson, M. (2005). *Birch extractives in kraft pulp washing: STFI-Packforsk report no 141*, Stockholm, Sweden.

Blumberg, L.M. (2012). Metrics of separation performance in chromatography: Part 2. Separation performance of a heating ramp in temperature-programmed gas chromatography. *Journal of Chromatography A* 1244: 148–160, https://doi.org/10.1016/j.chroma.2012.04.053.

Boeschoten, L.E., Vlam, M., Sass-Klaassen, U., Meyer-Sand, B.R.V., Boom, A., Bouka, G.U., Ciliane-Madikou, J.C., Obiang, N.L.E., Guieshon-Engongoro, M. and Loumeto, J.J., et al. (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.

Bogun, A.C., Paredes-Villanueva, K., Mascarello, M. and Magel, E.A. (2024). Development of a DNA macroarray for the molecular biological identification of trade-relevant tropical CITES timber species and their look-alikes. *Holzforschung* 78: 471–486, https://doi.org/10.1515/hf-2023-0114.

Braga, J., Deklerck, V., Espinoza, E., Groening, M., Koch, G., Monteiro Pastore, T.C., Ramananantoandro, T., Schroeder, H., Watkinson, C. and Wiedenhoeft, A.C. et al. (2020). *Scientific methods for taxonomic and origin identification of timber: Global Timber Tracking Network (GTTN)*, Available at: https://www.researchgate.net/publication/342003654 (Accessed 15 May 2023).

Braga, J.W.B., Pastore, T.C.M., Coradin, V.T.R., Camargos, J.A.A. and da Silva, A.R. (2011). The use of near infrared spectroscopy to identify solid wood specimens of Swietenia macrophylla (Cites Appendix II). *IAWA J.* 32: 285–296, https://doi.org/10.1163/22941932-90000058.

Broich, M., Hansen, M.C., Potapov, P., Adusei, B., Lindquist, E. and Stehman, S.V. (2011). Time-series analysis of multi-resolution optical imagery for quantifying forest cover loss in Sumatra and Kalimantan, Indonesia. *International Journal of Applied Earth Observation and Geoinformation* 13: 277–291, https://doi.org/10.1016/j.jag.2010.11.004.

Bullock, E.L., Woodcock, C.E., Souza, C. and Olofsson, P. (2020). Satellite-based estimates reveal widespread forest degradation in the Amazon. *Global change biology* 26: 2956–2969, https://doi.org/10.1111/gcb.15029.

Camin, F., Boner, M., Bontempo, L., Fauhl-Hassek, C., Kelly, S.D., Riedl, J. and Rossmann, A. (2017). Stable isotope techniques for verifying the declared geographical origin of food in legal cases. *Trends in Food Science & Technology* 61: 176–187, https://doi.org/10.1016/j.tifs.2016.12.007.

Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M. and Palmer, T.M. (2015). Accelerated modern human-induced species losses: entering the sixth mass extinction. *Science advances* 1: e1400253, https://doi.org/10.1126/sciadv.1400253.

Challinor, J.M. (2001). Review: the development and applications of thermally assisted hydrolysis and methylation reactions. *Journal of Analytical and Applied Pyrolysis* 61: 3–34, https://doi.org/10.1016/S0165-2370(01)00146-2.

Chapman, A.D. (2009). *Number of living species in Australia and the world: second edition*, Australian Biological Resources Study, Canberra, Australia.

Chávez-Salgado, L.P., Vandenbossche, V. and Vilarem, G. (2022). Tectona grandis Linn. f. secondary metabolites and their bioactive potential: a review. *iForest* 15: 112–120, https://doi.org/10.3832/ifor3714-015.

CITES and UNEP (2024). *Convention on international trade in endangered species of wild fauna and flora (CITES): Appendix I, II & III - valid from 23 February 2023*, Available at: https://cites.org/eng/app/appendices.php (Accessed 29 July 2024).

Cody, R.B., Dane, A.J., Dawson-Andoh, B., Adedipe, E.O. and Nkansah, K. (2012). Rapid classification of white oak (Quercus alba) and northern red oak (Quercus rubra) by using pyrolysis direct analysis in real time (DART™) and time-of-flight mass spectrometry. *Journal of Analytical and Applied Pyrolysis* 95: 134–137, https://doi.org/10.1016/j.jaap.2012.01.018.

Costello, M.J., May, R.M. and Stork, N.E. (2013). Can we name Earth's species before they go extinct? *Science* 339: 413–416, https://doi.org/10.1126/science.1230318.

Cowie, R.H., Bouchet, P. and Fontaine, B. (2022). The sixth mass extinction: fact, fiction or speculation? *Biological Reviews* 97: 640–663, https://doi.org/10.1111/brv.12816.

Creydt, M., Lautner, S., Fromm, J. and Fischer, M. (2022). Wood profiling by non-targeted liquid chromatography high-resolution mass spectrometry: Part 2, Detection of the geographical origin of spruce wood (Picea abies) by determination of metabolite pattern. *Journal of chromatography. A* 1663: 462737, https://doi.org/10.1016/j.chroma.2021.462737.

Creydt, M., Ludwig, L., Köhl, M., Fromm, J. and Fischer, M. (2021). Wood profiling by non-targeted high-resolution mass spectrometry: Part 1, Metabolite profiling in cedrela wood for the determination of the geographical origin. *Journal of Chromatography A* 1641: 461993, https://doi.org/10.1016/j.chroma.2021.461993.

Cronn, R.C., Finch, K.N., Hauck, L.L., Parker-Forney, M., Milligan, B.G., Dowling, J. and Scientists, A. (2021). Range-wide assessment of a SNP panel for individualization and geolocalization of bigleaf maple (Acer macrophyllum Pursh). *Forensic Science International: Animals and Environments* 1: 100033, https://doi.org/10.1016/j.fsiae.2021.100033.

Davis, J. (2022). Destruction of forests and grasslands is biggest cause of biodiversity loss: Science news. *Natural History Museum,* 9 November, Available at: https://www.nhm.ac.uk/discover/news/2022/november/destruction-forests-and-grasslands-biggest-cause-of-biodiversity-loss.html (Accessed 12 September 2024).

Degen, B., Blanc-Jolivet, C., Stierand, K. and Gillet, E. (2017). A nearest neighbour approach by genetic distance to the assignment of individual trees to geographic origin. *Forensic science international. Genetics* 27: 132–141, https://doi.org/10.1016/j.fsigen.2016.12.011.

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

Deklerck, V., Lancaster, C.A., van Acker, J., Espinoza, E.O., van den Bulcke, J. and Beeckman, H. (2020). Chemical fingerprinting of wood sampled along a pith-to-bark gradient for individual comparison and provenance identification. *Forests* 11: 107, https://doi.org/10.3390/f11010107.

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

Dinerstein, E., Vynne, C., Sala, E., Joshi, A.R., Fernando, S., Lovejoy, T.E., Mayorga, J., Olson, D., Asner, G.P. and Baillie, J.E.M., et al. (2019). A global deal for nature: Guiding principles, milestones, and targets. *Science advances* 5: eaaw2869, https://doi.org/10.1126/sciadv.aaw2869.

Diniz, C.G., Souza, A.A.d.A., Santos, D.C., Dias, M.C., Da Luz, N.C., Moraes, D.R.V. de, Maia, J.S.A., Gomes, A.R., Da Narvaes, I.S. and Valeriano, D.M., et al. (2015). DETER-B: The new Amazon near real-time deforestation detection system. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 8: 3619–3628, https://doi.org/10.1109/JSTARS.2015.2437075.

Dormontt, E.E., Boner, M., Braun, B., Breulmann, G., Degen, B., Espinoza, E., Gardner, S., Guillery, P., Hermanson, J.C. and Koch, G., et al. (2015). Forensic timber identification: It's time to integrate disciplines to combat illegal logging. *Biological Conservation* 191: 790–798,

Duchesne, I., Vincent, M., Wang, X.D., Ung, C.-H. and Swift, D.E. (2015). Wood mechanical properties and discoloured heartwood proportion in sugar maple and yellow birch grown in New Brunswick. In Bouffard, J.-F. (Ed.), *5th international scientific conference on hardwood processing, 15.-17. September 2015,* Québec City, Canada, FPInnovations, Pointe-Claire, Canada, pp. 64–71.

EIA (2017). *The rosewood racket report: China's billion dollar illegal timber trade and the devastation of Nigeria's forests*, Environmental Investigation Agency (EIA), Washington, USA, Available at: https://eia.org/report/the-rosewood-racket/ (Accessed 17 September 2024).

Elhardt, M., Min, X. and Lu, G. (2018). *How China is reshaping the global pulp and paper industry: Webinar*, Fisher International Inc., Norwalk, USA, Available at: https://www.fisheri.com/hubfs/Fisher%20White%20Papers/Fisher_Analysis_How_China_Is_Reshaping_the_Global_Pulp_and_Paper_Industry.pdf?utm_campaign=Fisher%20China%20Insights&utm_medium=email&_hsenc=p2ANqtz-97FnHOA938TrVx6pSjZ4zaWPtKL8mt3529wQRzMqZRTTmV53GF-0VtaHWD-hSzkZwhKFMUCfkFmnLI32P-3Ztnh_AfOPJoxwSP7LtdYxrnDwbUmME&_hsmi=84975695&utm_content=84975695&utm_source=hs_automation (Accessed 16 September 2024).

Espinoza, E.O., Lancaster, C.A., Kreitals, N.M., Hata, M., Cody, R.B. and Blanchette, R.A. (2014). Distinguishing wild from cultivated agarwood (Aquilaria spp.) using direct analysis in real time and time of-flight mass spectrometry. *Rapid Commun Mass Spectrom.* 28: 281–289, https://doi.org/10.1002/rcm.6779.

Espinoza, E.O., Wiemann, M.C., Chavarria, G.D., Barajas-Morales, J. and McClure, P.J. (2015). Forensic analysis of CITES-protected Dalbergia timber from the Americas. *IAWA J.* 36: 311–325, https://doi.org/10.1163/22941932-20150102.

EU (2010). Regulation (EU) No 995/2010 of the European Parliament and of the Council of 20 October 2010 laying down the obligations of operators who place timber and timber products on the market: legislation 295/23 with EEA relevance, European Union (EU). *Official Journal of the European Union* 53: 23–34, https://doi.org/10.3000/17252555.L_2010.295.eng.

EU (2023). Regulation (EU) No 2023/1115 of the European Parliament and of the Council of 31 May 2023: on the making available on the Union market and the export from the Union of certain commodities and products associated with deforestation and forest degradation and repealing regulation (EU) no 995/2010: legislation 150/206, Deforestation Regulation (EUDR). *Official Journal of the European Union* 150: 206–247, https://doi.org/10.2779/9252.

EU Parliament and EU Council (2024). Regulation (EU) 2024/1991 of the European Parliament and of the Council of 24 June 2024 on nature restoration and amending regulation (EU) 2022/869, Publications Office of the European Union L-2985 Luxembourg, Luxembourg. *Official Journal of the European Union* 1991: 1–93.

European Commission (2018). Commission notice — Guidance document on steps to be taken by EU Member States in the case of doubts as to the legality of timber from CITES-listed species imported into the EU. *Official Journal of the European Union* 376: 1–7.

European Commission (2019). *The European green deal: communication from the Commision to the European Parliament, the European Council, the European economic and sozial committee and the committee of the regions*, Brussels, Belgium, Available at: https://eur-lex.europa.eu/legal-content/DE/TXT/?qid=1588580774040&uri=CELEX%3A52019DC0640 (Accessed 30 September 2024).

European Commission (2020). *EU biodiversity strategy for 2030 - Bringing nature back into our lives: Communication from the Commision to the European Parliament, the European Council, the European economic and sozial committee and the committee of the regions*, Brussels, Belgium, Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0380 (Accessed 30 September 2024).

European Commission (2021). *Commission staff working document executive summary of the firness check: on regulation (EU) no 995/2010 of the European Parliament and of the Council of 20 October 2010 laying down the obligations of operators who place timber and timber products on the market (the EU Timber Regulation)*, Brussels, Belgium, Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021SC0329 (Accessed 30 September 2024).

Exposito-Alonso, M., Booker, T.R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C.C., Lang, P.L.M., Leventhal, L., Nogues-Bravo, D. and Pagowski, V., et al. (2022). Genetic diversity loss in the Anthropocene. *Science* 377: 1431–1435, https://doi.org/10.1126/science.abn5642.

FAO (2020). *Global forest resources assessment 2020: main report*, (CC BY-NC-SA 3.0 IGO), Food and Agriculture Organization of the United Nations, Rome, Italy.

FAO (2021). Forestry production and trade, Food and Agriculture Organization of the United Nations, Available at: https://www.fao.org/faostat/en/#data/FO (CC BY 4.0) (Accessed 12 September 2022).

FAO (2022a). Forestry production and trade, Food and Agriculture Organization of the United Nations, Available at: https://www.fao.org/faostat/en/#data/FO (CC BY 4.0) (Accessed 28 August 2024).

FAO (2022b). *Pulp and paper capacities: survey 2021-2026*, Rome, Available at: https://www.fao.org/3/cc1985t/cc1985t.pdf (CC BY-NC-SA 3.0 IGO) (Accessed 16 August 2023).

FAO (2022c). *Yearbook of forest products 2020*, Rome, Available at: https://www.fao.org/forestry-fao/statistics/80570/en/ (Accessed 25 April 2024).

FAO and European Union (2022). *FAO-EU FLEGT Programme: results, impacts and lessons learned from designing and implementing timber legality verification systems*, Available at: https://www.fao.org/in-action/fao-eu-flegt-programme/en/ (Accessed 26 April 2024).

FAO and UNEP (2020). *The state of the World's forests 2020: forests, biodiversity and people*, FAO and UN enviornment programme, Rome, Italy.

FCPF (2013). *The carbon fund of the forest carbon partnership facility (FCPF)*, Carbon Finance at the World Bank, Available at: https://www.forestcarbonpartnership.org/sites/fcp/files/2013/june2013/Carbon%20Fund-web_1.pdf (Accessed 30 September 2024).

FCPF (2023). *Forest carbon partnership facility (FCPF): 2023 annual report*, The World Bank, Washington DC, USA, Available at: https://www.forestcarbonpartnership.org/sites/default/files/documents/_web_world_bank_2023_fcpf_annual_report_r01.pdf (Accessed 30 September 2024).

Finch, K., Espinoza, E., Jones, F.A. and 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, https://doi.org/10.3732/apps.1600158.

Flaig, M.L., Berger, J., Helmling, S., Olbrich, A., Schaffrath, H.J., Zahn, D. and Saake, B. (2024). Chemotaxonomic and anatomic wood species identification in bleached pulp: blind test and method validation. *Holzforschung* 78: 487–502, https://doi.org/10.1515/hf-2024-0025.

Flaig, M.L., Berger, J. and Saake, B. (2025). Tracing the geographic origin of wood in pulp and paper by GC-MS analysis of extractives. *Holzforschung* 79: 15–29, https://doi.org/10.1515/hf-2024-0080.

Flaig, M.L., Berger, J., Wenig, P., Olbrich, A. and Saake, B. (2023). Identification of tropical wood species in paper: a new chemotaxonomic method based on extractives. *Holzforschung* 77: 860–878, https://doi.org/10.1515/hf-2023-0048.

Forrest, G.I. (1975). Variation in polyphenol content within and between sitka spruce provenances at different sites. *Can. J. For. Res.* 5: 46–54, https://doi.org/10.1139/x75-006.

Forster, P.M., Smith, C.J., Walsh, T., Lamb, W.F., Lamboll, R., Hauser, M., Ribes, A., Rosen, D., Gillett, N. and Palmer, M.D., et al. (2023). Indicators of global climate change 2022: annual update of large-scale indicators of the state of the climate system and human influence. *Earth Syst. Sci. Data* 15: 2295–2327, https://doi.org/10.5194/essd-15-2295-2023.

Franklin, G.L. (1945). Preparation of thin sections of synthetic resins and wood-resin composites, and a new macerating method for wood. *Nature* 155: 51,

Fujii, T., Abe, H., Kagawa, A., Kato, A. and Yoshida, K. (2007). Trial identification of tree species and its origin of commercial veneer. In Fujii, T. (Ed.), *International symposium on development of improved methods to identify Shorea species wood and its origin, 25.-26. September 2007*, Forestry and Forest Products Research Institute, Tokyo, Japan, pp. 45–49.

G7ANPE (2023). *Terms of reference for the G7-alliance on nature positive economies (G7ANPE)*, G7 ministers' meeting on climate, energy and environment, Sapporo, Japan, Available at: https://g7anpe.com/ (Accessed 29 September 2024).

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

Geevandana, Y.A., Gunawardana, M., Sultanbawa, M., Uvais, S. and Balasubramaniam (1980). Distribution of some triterpenes and phenolic compounds in the extractives of endemic Dipterocarpaceae species of Sri Lanka. *Phytochemistry* 19: 1099–1102, https://doi.org/10.1016/0031-9422(80)83063-9.

Gerber, L., Eliasson, M., Trygg, J., Moritz, T. and Sundberg, B. (2012). Multivariate curve resolution provides a high-throughput data processing pipeline for pyrolysis-gas chromatography/mass spectrometry. *Journal of Analytical and Applied Pyrolysis* 95: 95–100, https://doi.org/10.1016/j.jaap.2012.01.011.

GI-TOC (2016). *Organized crime and illegally mined gold in Latin America*, The Global Initiative against Transnational Organized Crime (GI TOC), Geneva, Switzerland, Available at: https://globalinitiative.net/wp-content/uploads/2016/03/Organized-Crime-and-Illegally-Mined-Gold-in-Latin-America.pdf (Accessed 18 September 2024).

Gold, V. (2019). *The IUPAC compendium of chemical terminology gold book*, International Union of Pure and Applied Chemistry (IUPAC), Research Triangle Park, NC.

Google My Maps (2024). Rubroshorea spp. solid wood samples' provenances: created by Flaig, M.L., based on map data ©2024 Google, TMap Mobility, Available at: https://www.google.com/maps/d/edit?EN&mid=1QXAf22NYG7RiIgDJiEP4fecLmG2-U1o&ll=-0.6062574229907733%2C113.60337280000002&z=5 (Accessed 18 July 2024).

Government of Japan (2016). *Act on promotion of use and distribution of legally-harvested wood and wood products: the clean wood act*, Available at: https://www.rinya.maff.go.jp/j/riyou/goho/english/attach/pdf/english-index-3.pdf (Accessed 27 October 2024).

Government of Japan (2023). *National biodiversity strategy and action plan of Japan 2023-2030: roadmap for the realisation of nature positive (provisional translation for G7 participants)*, Tokyo, Japan, Available at: https://www.cbd.int/doc/nbsap/NBSAPJapan2023-2030-Flier-English.pdf (Accessed 28 September 2024).

Grützmann, N. (2013). Herstellung und Bleiche von Zellstoffen aus Tropenholz: als Beitrag zur Identifizierung in Zellstoff und Papier, Diplomarbeit, Universität Hamburg, Hamburg, Germany, 2013.

Guo, F.-Q., Liang, Y.-Z., Xu, C.-J., Huang, L.-F. and Li, X.-N. (2004). Comparison of the volatile constituents of Artemisia capillaris from different locations by gas chromatography–mass spectrometry and projection method. *Journal of Chromatography A* 1054: 73–79, https://doi.org/10.1016/j.chroma.2004.08.122.

Gutteres, A. (2020). *State of the planet, UN Secretary General's address*, New York, USA, Available at: https://www.un.org/en/climatechange/un-secretary-general-speaks-state-planet (Accessed 5 November 2024).

Haag, V., Zemke, V.T., Lewandrowski, T., Zahnen, J., Hirschberger, P., Bick, U. and Koch, G. (2020). The European charcoal trade. *IAWA J.* 41: 463–477, https://doi.org/10.1163/22941932-bja10017.

Haberl, H., Erb, K.-H. and Krausmann, F. (2014). Human appropriation of net primary production: patterns, trends, and planetary boundaries. *Annu. Rev. Environ. Resour.* 39: 363–391, https://doi.org/10.1146/annurev-environ-121912-094620.

HAC for N&P (2024). High ambition coalition for nature & people (HAC for N&P): 30x30 conservation initative, Available at: https://www.hacfornatureandpeople.org/ (Accessed 13 September 2024).

Hegnauer, R. (1986). Phytochemistry and plant taxonomy — an essay on the chemotaxonomy of higher plants. *Phytochemistry* 25: 1519–1535, https://doi.org/10.1016/S0031-9422(00)81204-2.

Helmling, S., Olbrich, A., Heinz, I. and Koch, G. (2018). Atlas of vessel elements: identification of Asian timbers. *IAWA J.* 39: 249–352, https://doi.org/10.1163/22941932-20180202.

Helmling, S., Olbrich, A., Tepe, L. and Koch, G. (2016). Qualitative and quantitative characteristics of macerated vessels of 23 mixed tropical hardwood (MTH) species: a data collection for the identification of wood species in pulp and paper. *Holzforschung* 70: 839–844, https://doi.org/10.1515/hf-2015-0195.

Hillis, W.E. (Ed.) (1962). *Wood extractives and their significance to the pulp and paper industries*, Academic Press, New York, USA.

Hirschberger, P., Jokiel, D., Plaep, C. and Zahnen, J. (2010). *Tropenwaldzerstörung für Kinderbücher: eine Analyse des Buchmarktes in Deutschland*, Frankfurt a. M., Germany, Available at: https://www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/wwf-kinderbuchstudie-2009.pdf (Accessed 12 September 2022).

Hlásny, T., König, L., Krokene, P., Lindner, M., Montagné-Huck, C., Müller, J., Qin, H., Raffa, K.F., Schelhaas, M.-J. and Svoboda, M., et al. (2021). Bark beetle outbreaks in Europe: state of knowledge and ways forward for management. *Curr Forestry Rep* 7: 138–165, https://doi.org/10.1007/s40725-021-00142-x.

HMT, UK Government (2021). *The economics of biodiversity: the Dasgupta review: government response*, Presented to Parliament by the Exchequer Secretary to the Treasury by command of Her Majesty (HMT), London, England.

Holmberg, M. (1999). Pitch and precipitate problems. In Gullichsen, J. and Paulapuro, H. (Eds.), *Papermaking chemsitry: Book 4*, Fapet Oy, Helsinki, Finnland, pp. 223–239.

Holmbom, B. (1999). Extractives. In Sjöström, E. and Alén, R. (Eds.), *Analytical methods in wood chemistry, pulping, and papermaking*, *Springer series in wood science*, Springer, Berlin, Heidelberg, Germany, pp. 125–148.

Holmes, T.P., Aukema, J.E., Holle, B. von, Liebhold, A. and Sills, E. (2009). Economic impacts of invasive species in forests: past, present, and future. *Annals of the New York Academy of Sciences* 1162: 18–38, https://doi.org/10.1111/j.1749-6632.2009.04446.x.

Iason, G.R., Lennon, J.J., Pakeman, R.J., Thoss, V., Beaton, J.K., Sim, D.A. and Elston, D.A. (2005). Does chemical composition of individual scots pine trees determine the biodiversity of their associated ground vegetation? *Ecology Letters* 8: 364–369, https://doi.org/10.1111/j.1461-0248.2005.00732.x.

IAWA Committee (1989). IAWA list of microscopic features for hardwood identification: with an appendix on non-anatomical information. *IAWA Bulletin n.s.* 10: 219–332, https://doi.org/10.1163/22941932-90000496.

Ilvessalo-Pfäffli, M.-S. (1995). *Fiber atlas: identification of papermaking fibers.*, Springer Verlag, Berlin Heidelberg.

InsideWood (2004 onwards). *InsideWood database: published online*, Available at: https://insidewood.lib.ncsu.edu/ (Accessed 19 April 2023).

INTERPOL (2021). *Forestry crime*, Available at: https://www.interpol.int/Crimes/Environmental-crime/Forestry-crime (Accessed 25 April 2024).

IPBES (2019). *Global assessment report on biodiversity and ecosystem service*, Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), IPBES secretariat, Bonn, Germany.

IPCC (Ed.) (2021). *Climate change 2021 – the physical science basis*, Intergovernmental Panel on Climate Change, Cambridge University Press.

IPCC (Ed.) (2022). *Climate change 2022 – impacts, adaptation and vulnerability*, Intergovernmental Panel on Climate Change, Cambridge University Press.

ISO (1994). *Accuracy (trueness and precision) of measurment methods and results: Part 1: Genral principles and definitions* No. 5725-1-3:1994, Internation Organization for Standardization.

IUCN (2024). *The red list of threatened species: version 2024-08*, International Union for Conservation of Nature (IUCN), Available at: https://www.iucnredlist.org/ (Accessed 3 September 2024).

IUCN and BMUV (2011). Bonn challange: restore our future, International Union for Conservation of Nature (IUCN), German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMUV), Available at: https://www.bonnchallenge.org/ (Accessed 12 September 2024).

Jiao, L., Lu, Y., He, T., Guo, J. and Yin, Y. (2020). DNA barcoding for wood identification: global review of the last decade and future perspective. *IAWA J.* 41: 620–643, https://doi.org/10.1163/22941932-bja10041.

Johanson, M. (2024). How the Interoceanic Highway ushered in a new era of deforestation and social upheaval in the Amazon. *WWF Stories,* 6 June, Available at: https://www.worldwildlife.org/stories/how-the-interoceanic-highway-ushered-in-a-new-era-of-deforestation-and-social-upheaval-in-the-amazon (Accessed 6 September 2024).

Jones, N.F., Pejchar, L. and Kiesecker, J.M. (2015). The energy footprint: how oil, natural gas, and wind energy affect land for biodiversity and the flow of ecosystem services. *BioScience* 65: 290–301, https://doi.org/10.1093/biosci/biu224.

Jong, H.N. (2022). Palm oil and pulpwood the usual suspects as Papua deforestation persists, Mongabay (U.S.-based non-profit conservation and environmental science news platform), Available at: https://news.mongabay.com/2022/03/palm-oil-and-pulpwood-the-usual-suspects-as-papua-deforestation-persists/ (Accessed 27 September 2024).

Kaakinen, S., Saranpää, P. and Vapaavuori, E. (2007). Effects of growth differences due to geographic location and N-fertilisation on wood chemistry of Norway spruce. *Trees* 21: 131–139, https://doi.org/10.1007/s00468-006-0103-1.

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

Kastens, J.H., Brown, J.C., Coutinho, A.C., Bishop, C.R. and Esquerdo, J.C.D.M. (2017). Soy moratorium impacts on soybean and deforestation dynamics in Mato Grosso, Brazil. *PLoS One* 12: e0176168, https://doi.org/10.1371/journal.pone.0176168.

Kaufmann, U., Wellendorf, H. and Hansen, M. (1974). Thin layer chromatography of fluorescent phenolic compounds in needles. Degree of genetic control in Picea abies L. *Forest Tree Improvement, Arboretet Hoersholm* 8: 3–32.

Khrisanfov, M. and Samokhin, A. (2022). A general procedure for rounding m/z values in low-resolution mass spectra. *Rapid communications in mass spectrometry RCM* 36: e9294, https://doi.org/10.1002/rcm.9294.

Kireta, D., Lowe, A.J., Guerin, G.R., Leijs, R. and Hogendoorn, K. (2024). Bee diversity and pollination services improve with revegetation effort. *Austral Ecology* 49: e13559, https://doi.org/10.1111/aec.13559.

Kite, G.C., Green, P.W.C., Veitch, N.C., Groves, M.C., Gasson, P.E. and Simmonds, M.S.J. (2010). Dalnigrin, a neoflavonoid marker for the identification of Brazilian rosewood (Dalbergia nigra) in CITES enforcement. *Phytochemistry* 71: 1122–1131, https://doi.org/10.1016/j.phytochem.2010.04.011.

Koch, G., Haag, V., Helmling, S., Heinz, I. and Olbrich, A. (2017). Fasern im Fokus: Holzartenbestimmung von Faserplatten - Erfahrungen aus den Prüfungen im Kontext der EUTR. *MDF-Magazin & Co,* 2017: 82–85, Available at: https://www.openagrar.de/receive/openagrar_mods_00033035 (Accessed 27 September 2024).

Köthke, M., Lippe, M. and Elsasser, P. (2023). Comparing the former EUTR and upcoming EUDR: some implications for private sector and authorities. *Forest Policy and Economics* 157: 103079, https://doi.org/10.1016/j.forpol.2023.103079.

Krogell, J., Holmbom, B., Pranovich, A., Hemming, J. and Willför, S. (2012). Extraction and chemical characterization of Norway spruce inner and outer bark. *Nordic Pulp & Paper Research Journal* 27: 6–17, https://doi.org/10.3183/npprj-2012-27-01-p006-017.

Lancaster, C. and Espinoza, E. (2012). Analysis of select Dalbergia and trade timber using direct analysis in real time and time-of-flight mass spectrometry for CITES enforcement. *Rapid Commun Mass Spectrom.* 26: 1147–1156, https://doi.org/10.1002/rcm.6215.

Lancaster, C.A., Newsome, G.A. and Campbell-Such, J. (2024). Interlaboratory and cross-platform accessibility of time-of-flight wood identification database. *Forensic science international* 363: 112201, https://doi.org/10.1016/j.forsciint.2024.112201.

Langbour, P., Gerard, J., Roda, J.M., Ahmad Fauzi, P. and Guibal, D. (2011). Comparison of wood properties of planted big-leaf mahogany (Swietenia macrophylla) in Martinique Island with naturally grown mahogany from Brazil, Mexico and Peru. *Journal of Tropical Forest Science* 23: 252–259.

Lazarescu, C., Hart, F., Pirouz, Z., Panagiotidis, K., Mansfield, S.D., Barrett, J.D. and Avramidis, S. (2017). Wood species identification by near-infrared spectroscopy. *International Wood Products Journal* 8: 32–35, https://doi.org/10.1080/20426445.2016.1242270.

Little, T. (2016). Establishing acceptance criteria for analytical methods. *BioPharm International* 29: 44–48.

Locke, H., Rockström, J., Bakker, P., Bapna, M., Gough, M., Hilty, J., Lambertini, M., Morris, J., Polman, P. and Rodriguez, C.M. et al. (2021). *A nature-positive world: the global goal for nature*, New York, USA, Available at: https://library.wcs.org/doi/ctl/view/mid/33065/pubid/DMX3974900000.aspx (Accessed 12 September 2024).

Low, M.C., Schmitz, N., Boeschoten, L.E., Cabezas, J.A., Cramm, M., Haag, V., Koch, G., Meyer-Sand, B.R., Paredes-Villanueva, K. and Price, E., et al. (2022). Tracing the world's timber: the status of scientific verification technologies for species and origin identification. *IAWA J.* 44: 63–84, https://doi.org/10.1163/22941932-bja10097.

Lowe, A.J. and Cross, H.B. (2011). The application of DNA methods to timber tracking and origin verification. *IAWA J.* 32: 251–262, https://doi.org/10.1163/22941932-90000055.

Lowe, A.J., Wong, K.-N., Tiong, Y.-S., Iyerh, S. and 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, https://doi.org/10.1515/sg-2010-0037.

LP4N (2020). *Leaders' pledge for nature: united to reverse biodiversity loss by 2030 for sustainable development*, Leaders' Pledge for Nature (LP4N), Available at: https://www.leaderspledgefornature.org/wp-content/uploads/2021/06/Leaders_Pledge_for_Nature_27.09.20-ENGLISH.pdf (Accessed 27 September 2024).

LP4N (2023). *New biodiversity initiatives and finance announced as leaders move to deliver on global nature agreements: press release*, Leaders' Pledge for Nature (LP4N), New York, USA, Available at: https://www.leaderspledgefornature.org/updates/press-release-new-biodiversity-initiatives-and-finance-announced-as-leaders-move-to-deliver-on-global-nature-agreements/ (Accessed 30 September 2024).

Lukmandaru, G. (2012). Chemotaxonomic study based on the variation of quinone compounds in the heartwood of Javanese teak using GC-MS. In Salih, B. and Çelikbıçak, Ö. (Eds.), *Gas chromatography in plant science, wine technology, toxicology and some specific applications*, IntechOpen, Rijeka, Croatia, pp. 31–38.

Ma, T., Inagaki, T., Ban, M. and Tsuchikawa, S. (2019). Rapid identification of wood species by near-infrared spatially resolved spectroscopy (NIR-SRS) based on hyperspectral imaging (HSI). *Holzforschung* 73: 323–330, https://doi.org/10.1515/hf-2018-0128.

Macedo, M.N., DeFries, R.S., Morton, D.C., Stickler, C.M., Galford, G.L. and Shimabukuro, Y.E. (2012). Decoupling of deforestation and soy production in the southern Amazon during the late 2000s. *Proceedings of the national academy of sciences of the United States of America* 109: 1341–1346, https://doi.org/10.1073/pnas.1111374109.

Martin, J. and Haggith, M. (Eds.) (2018). *The state of the global paper industry: shifting seas: new challanges and opportunities for forests, people and the climate*, Environmental Paper Network (EPN), Asheville, USA.

May, C. (2017). *Transnational crime and the developing world: global financial integrity*, Available at: https://gfintegrity.org/report/transnational-crime-and-the-developing-world/ (Accessed 25 April 2024).

Ministry of the Environment, Government of Japan (2024). *Transition strategies toward nature positive economy: creating corporate value based on natural capital*, Japan, Available at: https://www.env.go.jp/content/000213036.pdf (Accessed 27 September 2024).

Nellemann, C., Henriksen, R., Kreilhuber, A., Stewart, D., Kotsovou, M., Raxter, P., Mrema, E. and Barrat, S. (Eds.) (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 (UNEP), Nairobi, Kenya.

Nellemann, C., Henriksen, R., Pravettoni, R. and Jesperson, S. (2020). Forestry crimes and our planet. In Nikolakis, W. and Innes, J.L. (Eds.), *The wicked problem of forest policy: a multidisciplinary approach to sustainability in forest landscapes*, Cambridge University Press, Cambridge.

Ng, C.H., Lee, S.L., Tnah, L.H., Ng, K.K.S., Lee, C.T., Diway, B. and Khoo, E. (2017). Geographic origin and individual assignment of Shorea platyclados (Dipterocarpaceae) for forensic identification. *PLoS One* 12: 1–18, https://doi.org/10.1371/journal.pone.0176158.

Niamké, F.B., Amusant, N., Augustin, A.A. and Chaix, G. (2021). Teakwood chemistry and natural durability. In Ramasamy, Y., Galeano, E. and Win, T.T. (Eds.), *The teak genome*, *Compendium of plant genomes*, 1st ed. 2021, Springer Nature, Cham, Switzerland, pp. 83–102.

Nieradzik, L., Sieburg-Rockel, J., Helmling, S., Keuper, J., Weibel, T., Olbrich, A. and Stephani, H. (2024). Automating wood species detection and classification in microscopic images of fibrous materials with deep learning. *Microscopy and Microanalysis* 38: 1–17.

Nikolakis, W. and Innes, J.L. (Eds.) (2020). *The wicked problem of forest policy: a multidisciplinary approach to sustainability in forest landscapes*, Cambridge University Press, Cambridge.

Nisgoski, S., Gonçalves, T.A.P., Sonsin-Oliveira, J., Ballarin, A.W. and Muñiz, G.I.B. (2021). Near-infrared spectroscopy for discrimination of charcoal from eucalyptus and native cerrado species—contribution to a database for forestry supervision. *Forest Science* 67: 419–432, https://doi.org/10.1093/forsci/fxab015.

NPI (2023). *The definition of nature positive*, Nature Positive Initiative (NPI), Available at: https://www.naturepositive.org/app/uploads/2024/02/The-Definition-of-Nature-Positive.pdf (Accessed 12 September 2024).

Ocko, I.B., Ramaswamy, V. and Ming, Y. (2014). Contrasting climate responses to the scattering and absorbing features of anthropogenic aerosol forcings. *Journal of Climate* 27: 5329–5345, https://doi.org/10.1175/JCLI-D-13-00401.1.

Olbrich, A. (2024). *Development of the demand for expert opinions on the identification of wood taxa in different products since 2000*, Hamburg, Germany, Available at: personal communication (Accessed 16 September 2024).

Onuorah E. O., Nwabanne J. T. and Nnabuife, E.L.C. (2015). Pulp and paper making potentials of elaeis guineensis (oil palm) grown in south east, Nigeria. *World Journal of Engineering* 12: 1–12,

Örsa, F. and Holmbom, B. (1994). A convenient method for the determination of wood extractives in papermaking process waters and effluents. *Journal of Pulp and Paper Science* 20: 361–366.

Pailler, S. (2018). Re-election incentives and deforestation cycles in the Brazilian Amazon. *Journal of Environmental Economics and Management* 88: 345–365, https://doi.org/10.1016/j.jeem.2018.01.008.

Pfaff, A., Robalino, J., Herrera, D. and Sandoval, C. (2015). Protected areas' impacts on Brazilian Amazon deforestation: examining conservation-development interactions to inform planning. *PLoS One* 10: e0129460, https://doi.org/10.1371/journal.pone.0129460.

Phang, A., Niissalo, M.A., Ruhsam, M., Pezzini, F.F., Neo, W.L., Burslem, D., Middleton, D.J., Wilkie, P. and Khew, G.S. (2024). Genetic erosion in a tropical tree species demonstrates the need to conserve wide-ranging germplasm amid extreme habitat fragmentation. *Biodivers Conserv* 33: 2527–2548, https://doi.org/10.1007/s10531-024-02870-5.

Piispanen, R., Willför, S., Saranpää, P. and Holmbom, B. (2008). Variation of lignans in Norway spruce (Picea abies [L.] Karst.) knotwood: within-stem variation and the effect of fertilisation at two experimental sites in Finland. *Trees* 22: 317–328, https://doi.org/10.1007/s00468-007-0186-3.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. and Sexton, J.O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science (New York, N.Y.)* 344: 1246752, https://doi.org/10.1126/science.1246752.

Pimm, S.L., Russel, G.J., Gittleman, J.L. and Brooks, T.M. (1995). The future of biodiversity. *Science* 269: 347–350, https://doi.org/10.1126/science.269.5222.347.

Ponnuchamy, V., Gordobil, O., Diaz, R.H., Sandak, A. and Sandak, J. (2021). Fractionation of lignin using organic solvents: a combined experimental and theoretical study. *International journal of biological macromolecules* 168: 792–805, https://doi.org/10.1016/j.ijbiomac.2020.11.139.

Porkka, M., Virkki, V., Wang-Erlandsson, L., Gerten, D., Gleeson, T., Mohan, C., Fetzer, I., Jaramillo, F., Staal, A. and te Wierik, S. et al. (2024). *Global water cycle shifts far beyond pre-industrial conditions – planetary boundary for freshwater change transgressed*, EarthArXiv, Available at: https://doi.org/10.31223/X5BP8F (Accessed 5 September 2024).

Price, E.R., McClure, P.J., Huffman, A.N., Voin, D. and Espinoza, E.O. (2022). Reliability of wood identification using DART-TOFMS and the ForeST© database: a validation study. *Forensic Science International: Animals and Environments* 2: 100045, https://doi.org/10.1016/j.fsiae.2022.100045.

Price, E.R., Miles-Bunch, I., Gasson, P.E. and 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, https://doi.org/10.1016/j.fsiae.2021.100032.

Rachmayanti, Y., Leinemann, L., Gailing, O. and Finkeldey, R. (2009). DNA from processed and unprocessed wood: factors influencing the isolation success. *Forensic science international. Genetics* 3: 185–192, https://doi.org/10.1016/j.fsigen.2009.01.002.

Ramalho, F.M.G., Andrade, J.M. and Hein, P.R.G. (2018). Rapid discrimination of wood species from native forest and plantations using near infrared spectroscopy. *Forest Syst* 27: e008, https://doi.org/10.5424/fs/2018272-12075.

Rathore, A. and Jasrai, Y.T. (2013). Biodiversity: importance and climate change impacts. *International Journal of Scientific and Research Publications* 3: 536–540.

Ravindran, P. and Wiedenhoeft, A.C. (2020). Comparison of two forensic wood identification technologies for ten Meliaceae woods: computer vision versus mass spectrometry. *Wood Sci Technol* 54: 1139–1150, https://doi.org/10.1007/s00226-020-01178-1.

Richardson, K., Steffen, W., Lucht, W., Bendtsen, J., Cornell, S.E., Donges, J.F., Drüke, M., Fetzer, I., Bala, G. and von Bloh, W., et al. (2023). Earth beyond six of nine planetary boundaries. *Science advances* 9: eadh2458, https://doi.org/10.1126/sciadv.adh2458.

Richter, H.G. and Dallwitz, M.J. (2000 onwards). *Commercial timbers: descriptions, illustrations, identification, and information retrieval. In English, French, German, Portuguese, and Spanish.: Version: 9th April 2019*, Available at: https://www.delta-intkey.com/wood/index.htm (Accessed 19 April 2023).

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C. and Schellnhuber, H.J., et al. (2009). A safe operating space for humanity. *Nature* 461: 472–475, https://doi.org/10.1038/461472a.

Rounsevell, M.D.A., Harfoot, M., Harrison, P.A., Newbold, T., Gregory, R.D. and Mace, G.M. (2020). A biodiversity target based on species extinctions. *Science* 368: 1193–1195, https://doi.org/10.1126/science.aba6592.

Ryan, C.G., Clayton, E., Griffin, W.L., Sie, S.H. and Cousens, D.R. (1988). SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 34: 396–402, https://doi.org/10.1016/0168-583X(88)90063-8.

Sari, E.O., Utami, S.P., Nakagawa-Izumi, A. and Ohi, H. (2024). Lapachol from Indonesian teak (Tectona grandis) wood waste as a natural additive for alkaline cooking. *Wood Sci Technol* 58: 627–647, https://doi.org/10.1007/s00226-024-01537-2.

Sarukhán, J. and Whyte, A. (Eds.) (2005). *Ecosystems and human well-being: synthesis*, Millennium Ecosystem Assessment (MEA), Island Press, Washington DC, USA.

Savitzky, A. and Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36: 1627–1639, https://doi.org/10.1021/ac60214a047.

Scheffers, B.R., Joppa, L.N., Pimm, S.L. and Laurance, W.F. (2012). What we know and don't know about Earth's missing biodiversity. *Trends in ecology & evolution* 27: 501–510, https://doi.org/10.1016/j.tree.2012.05.008.

Schmitz, N., Beeckman, H., Blanc-Jolivet, C., Boeschoten, L., Braga, J., Cabezas, J.A., Chaix, G., Crameri, S., Degen, B. and Deklerck, V. et al. (2020). *Overview of current practices in data analysis for wood identification: a guide for the different timber tracking methods*, Global Timber Tracking Network (GTTN), Available at: https://www.researchgate.net/publication/342281916 (Accessed 15 May 2023).

Schmitz, N., Beeckman, H., Cabezas, J.A., Cervera, M.T., Espinoza, E. and Fernandez-Golfin, J. (2019). *The timber tracking tool infogram: overview of wood identification methods' capacity*, Global Timber Tracking Network (GTTN), Available at: https://www.researchgate.net/publication/332223451 (Accessed 15 May 2023).

Schroeder, H., Cronn, R., Yanbaev, Y., Jennings, T., Mader, M., Degen, B. and Kersten, B. (2016). Development of molecular markers for determining continental origin of wood from white oaks (Quercus L. sect. Quercus). *PLoS One* 11: e0158221, https://doi.org/10.1371/journal.pone.0158221.

Schwanninger, M., Rodrigues, J.C. and Fackler, K. (2011). A review of band assignments in near infrared spectra of wood and wood components. *Journal of Near Infrared Spectroscopy* 19: 287–308, https://doi.org/10.1255/jnirs.955.

Shah, V. (2016). Indonesian logging firm fined a record US$1.2 billion for deforestation. *Eco-Business,* 25 November, Available at: https://www.eco-business.com/news/indonesian-logging-firm-fined-a-record-us12-billion-for-deforestation/ (Accessed 27 September 2024).

Sieburg-Rockel, J. and Koch, G. (2020). Identification of wood species used in particleboard production. *IAWA J.* 41: 751–760, https://doi.org/10.1163/22941932-bja10018.

Silva, D.C., Pastore, T.C., Soares, L.F., Barros, F.A. de, Bergo, M.C., Coradin, V.T., Gontijo, A.B., Sosa, M.H., Chacón, C.B. and 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, https://doi.org/10.1515/hf-2017-0160.

Silva, J.L., Bordalo, R. and Pissarra, J. (2021). Wood identification: an overview of current and past methods. *Estudos de Conservação e Restauro* 12: 45–68, https://doi.org/10.34632/ECR.2020.10680.

Snel, F.A., Braga, Jez W. B., Da Silva, D., Wiedenhoeft, A.C., Costa, A., Soares, R., Coradin, Vera T. R. and Pastore, Tereza C. M. (2018). Potential field-deployable NIRS identification of seven Dalbergia species listed by CITES. *Wood Sci Technol* 52: 1411–1427, https://doi.org/10.1007/s00226-018-1027-9.

Solhaug, K.A. (1990). Stilbene glucosides in bark and needles from Picea species. *Scandinavian Journal of Forest Research* 5: 59–67, https://doi.org/10.1080/02827589009382593.

Spielman, D., Brook, B.W. and Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *Proceedings of the national academy of sciences of the United States of America* 101: 15261–15264, https://doi.org/10.1073/pnas.0403809101.

Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., Vries, W. de and Wit, C.A. de, et al. (2015). Planetary boundaries: guiding human development on a changing planet. *Science* 347: 1259855, https://doi.org/10.1126/science.1259855.

Stein, S.E. and Scott, D.R. (1994). Optimization and testing of mass spectral library search algorithms for compound identification. *American Society for Mass Spectrometry* 5: 859–866, https://doi.org/10.1016/1044-0305(94)87009-8.

Svahn, J. and Brunner, D. (2018). Did the soy moratorium reduce deforestation in the Brazilian Amazon?: a counterfactual analysis of the impact of the soy moratorium on deforestation in the Amazon biome, Master Thesis, Norwegian School of Economics, Bergen, Norway, 2018.

ter Steege, H., Pitman, N.C.A., Sabatier, D., Baraloto, C., Salomão, R.P., Guevara, J.E., Phillips, O.L., Castilho, C.V., Magnusson, W.E. and Molino, J.-F., et al. (2013). Hyperdominance in the Amazonian tree flora. *Science* 342: 1243092, https://doi.org/10.1126/science.1243092.

Traoré, M., Kaal, J. and Martínez Cortizas, A. (2018). Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood science and technology* 52: 487–504, https://doi.org/10.1007/s00226-017-0967-9.

Tsuchikawa, S. and Kobori, H. (2015). A review of recent application of near infrared spectroscopy to wood science and technology. *J Wood Sci* 61: 213–220, https://doi.org/10.1007/s10086-015-1467-x.

UK Government (2013). *The timber and timber products (placing on the market) regulations 2013 (UKTR): no. 233 environmental protection*, London, England, Available at: https://www.legislation.gov.uk/uksi/2013/233/contents (Accessed 1 October 2024).

UNEP (2022). *Kunming-Montreal global biodiversity framework: decision adopted by the conference of the parties (COP15) to the convention on biological diversity (CBD)*, United Nations Environment Programme (UNEP), Montreal, Canada / Kunming, China, Available at: https://www.cbd.int/doc/decisions/cop-15/cop-15-dec-04-en.pdf (Accessed 13 September 2024).

UNEP-FI, CBD, Business for Nature, Finance for Biodiversity and Principles for Responsible Investment (2021). *Financial sector guide for the convention on biological diversity: key actions for nature*, UN Environment Programme - Finance Initiative (UNEP-FI), Available at: https://www.cbd.int/doc/c/8e24/f151/326b69024f014a8fb9684a8d/cbd-financial-sector-guide-f-en.pdf (Accessed 30 September 2024).

UNFCCC (2015). *Paris agreement: CP/2015/L.9/Rev.1*, United Nations (UN) Framework Convention on Climate Change (FCCC), Paris, France, Available at: http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf (Accessed 27 September 2024).

UNODC (2016). *World wildlife crime report: trafficking in protected species, 2016*, United Nations Office on Drugs and Crime (UNODC), *UNODC research*, United Nations, New York, USA.

Uryu, Y., Mott, C., Foead, N., Yulinato, K. and Budiman, A. (2008). *Deforestation, forest degradation, biodiversity loss and CO2 emissions in Riau, Sumatra, Indonesia: technical report*, Jakarta, Indonesia, Available at: https://files.worldwildlife.org/wwfcmsprod/files/Publication/file/2i5fo6l9lz_WWF_Indo_27Feb08_Riau_Deforestation___English.pdf (Accessed 12 September 2022).

US Government (2008). *Food, conservation, and energy act of 2008: Public law 110–246 June 18, 2008*, Available at: https://www.govinfo.gov/content/pkg/PLAW-110publ246/pdf/PLAW-110publ246.pdf (Accessed 1 October 2024).

Valto, P., Knuutinen, J. and Alén, R. (2012). Overview of analytical procedures for fatty and resin acids in the papermaking process. *Bioresources* 7: 6041–6076, https://doi.org/10.15376/biores.7.4.6041-6076.

van Lierop, P., Lindquist, E., Sathyapala, S. and Franceschini, G. (2015). Global forest area disturbance from fire, insect pests, diseases and severe weather events. *Forest Ecology and Management* 352: 78–88, https://doi.org/10.1016/j.foreco.2015.06.010.

Wan Daud, W.R. and Law, K.-N. (2011). Oil palm fibers as papermaking material: potentials and challenges. *Bioresources* 6: 901–917, https://doi.org/10.15376/biores.6.1.901-917.

Wang-Erlandsson, L., Tobian, A., van der Ent, R.J., Fetzer, I., te Wierik, S., Porkka, M., Staal, A., Jaramillo, F., Dahlmann, H. and Singh, C., et al. (2022). A planetary boundary for green water. *Nat Rev Earth Environ* 3: 380–392, https://doi.org/10.1038/s43017-022-00287-8.

Waters, C.N., Zalasiewicz, J., Summerhayes, C., Barnosky, A.D., Poirier, C., Gałuszka, A., Cearreta, A., Edgeworth, M., Ellis, E.C. and Ellis, M., et al. (2016). The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science* 351: aad2622, https://doi.org/10.1126/science.aad2622.

Watkinson, C., Gasson, P., Rees, G. and Boner, M. (2020). The development and use of isoscapes to determine the geographical origin of Quercus spp. in the United States. *Forests* 11: 862, https://doi.org/10.3390/f11080862.

Watkinson, C.J., Rees, G.O., Gwenael, M.C., Gasson, P., Hofem, S., Michely, L. and Boner, M. (2022). Stable isotope ratio analysis for the comparison of timber from two forest concessions in Gabon. *Frontiers in Forests and Global Change* 4: 650257, https://doi.org/10.3389/ffgc.2021.650257.

Welling, J. and Liese, W. (2019). Wood, bamboo and palm wood - similarities and differences in research and technology Development. In El-Mously, H., Midani, M. and Wagih, M. (Eds.), *By-products of palm trees and their applications, 2018/12/15*, Materials Research Forum LLC, pp. 83–87.

Wenig, P. (2011). Post-optimization of Py-GC/MS data: A case study using a new digital chemical noise reduction filter (NOISERA) to enhance the data quality utilizing OpenChrom mass spectrometric software. *Journal of Analytical and Applied Pyrolysis* 92: 202–208, https://doi.org/10.1016/j.jaap.2011.05.013.

Wenig, P. and Mailänder, M. (2024). *OpenChrom® software handbook: version 0.1.4*, Lablicate GmbH, Hamburg, Germany.

Wenig, P. and Odermatt, J. (2010a). Efficient analysis of Py-GC/MS data by a large scale automatic database approach: an illustration of white pitch identification in pulp and paper industry. *Journal of Analytical and Applied Pyrolysis* 87: 85–92, https://doi.org/10.1016/j.jaap.2009.10.007.

Wenig, P. and Odermatt, J. (2010b). OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data. *BMC bioinformatics* 11: 405, https://doi.org/10.1186/1471-2105-11-405.

West, T.A. and Fearnside, P.M. (2021). Brazil's conservation reform and the reduction of deforestation in Amazonia. *Land Use Policy* 100: 105072, https://doi.org/10.1016/j.landusepol.2020.105072.

Westervelt, D.M., You, Y., Li, X., Ting, M., Lee, D.E. and Ming, Y. (2020). Relative importance of greenhouse gases, sulfate, organic carbon, and black carbon aerosol for South Asian monsoon rainfall changes. *Geophysical Research Letters* 47: e2020GL088363, https://doi.org/10.1029/2020GL088363.

Wheeler, E.A. (2011). InsideWood – a web resource for hardwood anatomy. *IAWA J.* 32: 199–211, https://doi.org/10.1163/22941932-90000051.

Willför, S., Hemming, J. and Leppänen, A.-S. (2006). *Analysis of extractives in different pulps – method development, evaluation, and recommendations: report B1-2006*, Turku, Finnland.

Wilson, M.C., Chen, X.-Y., Corlett, R.T., Didham, R.K., Ding, P., Holt, R.D., Holyoak, M., Hu, G., Hughes, A.C. and Jiang, L., et al. (2016). Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landscape Ecol* 31: 219–227, https://doi.org/10.1007/s10980-015-0312-3.

Windhagen, K., Moldenhauer, T., Burkard, A. and Geiger, G.A. (2019). *Papier 2019: ein Leistungsbericht (annual report)*, Bonn, Germany.

Wold, S., Esbensen, K. and Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 2: 37–52, https://doi.org/10.1016/0169-7439(87)80084-9.

Worth (2021). Knock on Wood: WWF Exposes Illegal Timber Practices in Germany. *Whistleblower Network News,* 13 December, Available at: https://whistleblowersblog.org/environmental-whistleblowers/knock-on-wood-wwf-exposes-illegal-timber-practices-in-germany/ (Accessed 24 September 2024).

WWF (2015). WWF-Germany files charges against paper retailers suspected of violating EU Timber Regulation, Available at: https://wwf.panda.org/wwf_news/?244971/WWF-files-charges-against-paper-retailers-suspected-of-violating-EU-Timber-Regulation (Accessed 27 September 2024).

WWF (2021). *EU forest crime initiative: EU summary of the gap analysis project countries: Belgium, Bulgaria, France, Romania, Slovakia, and Ukraine*, World Wide Fund for Nature (WWF), Available at: https://www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/EU-Forest-Crime-Initiative-EU-GAP-Analysis.pdf (Accessed 17 September 2024).

WWF (2022). *Living Planet Report 2022: building a nature-positive society*, World Wide Fund for Nature (WWF), Gland, Switzerland.

WWF (2024). Losing their homes because of the growing needs of humans, World Wide Fund for Nature (WWF), Available at: https://wwf.panda.org/discover/our_focus/wildlife_practice/problems/habitat_loss_degradation/ (Accessed 6 September 2024).

Yang, G., Liang, K., Zhou, Z., Wang, X. and Huang, G. (2020). UPLC-ESI-MS/MS-based widely targeted metabolomics analysis of wood metabolites in teak (Tectona grandis). *Molecules (Basel, Switzerland)* 25: 2189, https://doi.org/10.3390/molecules25092189.

Yang, L., Wen, K.-S., Ruan, X., Zhao, Y.-X., Wei, F. and Wang, Q. (2018). Response of plant secondary metabolites to environmental factors. *Molecules (Basel, Switzerland)* 23: 762, https://doi.org/10.3390/molecules23040762.

Yu, M., Liu, K., Zhou, L., Zhao, L. and Liu, S. (2016). Testing three proposed DNA barcodes for the wood identification of Dalbergia odorifera T. Chen and Dalbergia tonkinensis Prain. *Holzforschung* 70: 127–136, https://doi.org/10.1515/hf-2014-0234.

Zahnen, J., Hirschberger, P., Haag, V. and Lewandrowski, T. (2020). *2020 analysis of the EU charcoal market*, WWF Germany, Berlin, Germany, Available at: https://www.openagrar.de/receive/openagrar_mods_00066143 (Accessed 24 September 2024).

Zemke, V., Haag, V. and Koch, G. (2020). Wood identification of charcoal with 3D-reflected light microscopy. *IAWA J.* 41: 478–489, https://doi.org/10.1163/22941932-bja10033.

Zhang, M., Guo, J., Lu, Y., Jiao, L., He, T. and Yin, Y. (2022a). Similarity network fusion for aggregating headspace GC–MS and direct analysis in real time–mass spectrometry data from solid samples to enhance species identification efficiency of high–temperature heated wood. *J Wood Sci* 68: 38, https://doi.org/10.1186/s10086-022-02044-3.

Zhang, M., Zhao, G., Guo, J., Wiedenhoeft, A.C., Liu, C.C. and Yin, Y. (2019). Timber species identification from chemical fingerprints using direct analysis in real time (DART) coupled to Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS): comparison of wood samples subjected to different treatments. *Holzforschung* 73: 975–985, https://doi.org/10.1515/hf-2018-0304.

Zhang, Y., Wang, Z., Lu, Y. and Zuo, L. (2022b). Editorial: biodiversity, ecosystem functions and services: interrelationship with environmental and human health. *Front. Ecol. Evol.* 10: 1086408, https://doi.org/10.3389/fevo.2022.1086408.

Zheng, C., Liu, S., Wang, J., Lu, Y., Ma, L., Jiao, L., Guo, J., Yin, Y. and He, T. (2024). Opening the black box: explainable deep-learning classification of wood microscopic image of endangered tree species. *Plant methods* 20: 1–11, https://doi.org/10.1186/s13007-024-01191-6.

Zillén, L., Conley, D.J., Andrén, T., Andrén, E. and Björck, S. (2008). Past occurrences of hypoxia in the Baltic Sea and the role of climate variability, environmental change and human impact. *Earth-Science Reviews* 91: 77–92, https://doi.org/10.1016/j.earscirev.2008.10.001.

Zwanziger, H.W. and Sorkau, E. (2020). *Kalibration analytischer Methoden: Theorie & Techniken, eine Orientierungshilfe*, Hochschulverlag Merseburg, Merseburg.