WOOD AND FIBER SCIENCE

The Sustainable Natural Materials Journal

Volume 55, Number 1_2023 *(ISSN 0735-6161)*

Open Access

JOURNAL OF THE

SWST – International Society of Wood Science and Technology

SOCIETY OF WOOD SCIENCE AND TECHNOLOGY

2022–2023 Officers of the Society

President: Henry Quesada, Purdue University, Indiana, USA

Immediate Past President: Rupert Wimmer, BOKU Vienna, Austria

President-Elect: Jeffrey Morrell, University of the Sunshine Coast, Australia

Vice President: Ilona Peszlen, North Carolina State University, Raleigh, NC

Executive Director: Victoira Herian, Society of Wood Science and Technology, P.O. Box 6155, Monona, WI 53716-1655, vicki@swst.org

Executive Director: ANGELA HANEY, Society of Wood Science and Technology, P.O. Box 6155, Monona, WI 53716-6155, execdir@swst.org

Directors:

Omar Espinoza, University of Minnesota, St. Paul, MN 55108
Judith Gisip, Universiti Teknologi MARA, Malaysia
Hongmei Gu, USDA Forest Products Laboratory, Madison, WI 53726
Francesco Negro, DISAFA – University of Torino, Italy

Editor, Wood and Fiber Science: Jeffrey Morrell, University of the Sunshine Coast, Australia, jmorrell@usc.edu.au

Associate Editor, Wood and Fiber Science: Arijit Sinha, Oregon State University, Corvallis, OR 97331, arijit.sinha@oregonstate.edu

Digital Communication Coordinator: Pipiet Larasatie, University of Arkansas at Monticello, AR, larasatie@uamont.edu

Editor, BioProducts Business Editor: Eric Hansen, Oregon State University, Corvallis, OR 97331, eric.hansen@oregonstate.edu

WOOD AND FIBER SCIENCE

Wood and Fiber Science is published quarterly in January, April, July, and October by the Society of Wood Science and Technology, P.O. Box 6155, Monona, WI 53716-6155

There are three classes of membership (electronic only) in the Society: Members – dues $150; Retired Members – dues $75; Student Members – dues $50. We also have membership category for individuals from Emerging Countries where individual members pay $30, individual students pay $10; Emerging Group of 10 pay $290, and Student Groups of 10 pay $90. Institutions and individuals who are not members pay $300 per volume (electronic only). Applications for membership and information about the Society may be obtained from the Executive Director, Society of Wood Science and Technology, P.O. Box 6155, Monona, WI 53716-6155 or found at the website http://wwww.swst.org.

Site licenses are also available with a charge of:

- $300/yr for single online membership, access by password and email
- $500/yr for institutional subscribers with 2–10 IP addresses
- $750/yr for institutional subscribers with 11–50 IP addresses
- $1000/yr for institutional subscribers with 51–100 IP addresses
- $1500/yr for institution subscribers with 101–200 IP addresses
- $2000/yr for institutions subscribers with over 200 IP addresses.

New subscriptions begin with the first issue of a new volume. All subscriptions are to be ordered through the Executive Director, Society of Wood Science and Technology.

The Executive Director, at the Business Office shown below, should be notified 30 days in advance of a change of email address.

Business Office: Society of Wood Science and Technology, P.O. Box 6155, Monona, WI 53716-6155.
Editorial Office: Susan LeVan-Green, sue.levangreen@gmail.com

EDITORIAL AND PUBLICATION POLICY

Wood and Fiber Science as the official publication of the Society of Wood Science and Technology publishes papers with both professional and technical content. Original papers of professional concern, or based on research of international interest dealing with the science, processing, and manufacture of wood and composite products of wood or wood fiber origin will be considered.

All manuscripts are to be written in US English, the text should be proofread by a native speaker of English prior to submission. Any manuscript submitted must be unpublished work not being offered for publication elsewhere.

Papers will be reviewed by referees selected by the editor and will be published in approximately the order in which the final version is received. Research papers will be judged on the basis of their contribution of original data, rigor of analysis, and interpretations of results; in the case of reviews, on their relevancy and completeness.

As of January 1, 2022, Wood and Fiber Science will be an online only, Open Access journal. There will be no print copies. Color photos/graphics will be offered at no additional cost to authors. The Open Access fee will be $1800/article for SWST members and $2000/article for nonmembers. The previous five years of articles are still copyright protected (accept those that are identified as Open Access) and can be accessed through member subscriptions. Once a previous article has reached its 5th anniversary date since publication, it becomes Open Access.

Technical Notes

Authors are invited to submit Technical Notes to the Journal. A Technical Note is a concise description of a new research finding, development, procedure, or device. The length should be **no more than two printed pages** in WFS, which would be five pages or less of double-spaced text (TNR12) with normal margins on 8.5 x 11 paper, including space for figures and tables. In order to meet the limitation on space, figures and tables should be minimized, as should be the introduction, literature review and references. The Journal will attempt to expedite the review and publication process. As with research papers, Technical Notes must be original and go through a similar double-blind, peer review process.

On-line Access to *Wood and Fiber Science* Back Issues

SWST is providing readers with a means of searching all articles in *Wood and Fiber Science* from 1968 to present Articles from 1968 to 2017 are available to anyone, but in order to see 2017 to 2021 articles you must have an SWST membership or subscription. SWST members and subscribers have full search capability and can download PDF versions of the papers If you do not have a membership or subscription, you will not be able to view the full-text pdf.

Visit the SWST website at http://www.swst.org and go to Wood & Fiber Science Online. Click on either SWST Member Publication access (SWST members) or Subscriber Publication access (Institution Access). All must login with their email and password on the HYPERLINK "http://www.swst.org" www.swst.org site, or use their ip authentication if they have a site license.

As an added benefit to our current subscribers, you can now access the electronic version of every printed article along with exciting enhancements that include:

- IP authentication for institutions (only with site license)
- Enhanced search capabilities
- Email alerting of new issues
- Custom links to your favorite titles

WOOD AND FIBER SCIENCE

JOURNAL OF THE SOCIETY OF WOOD SCIENCE AND TECHNOLOGY

VOLUME 55 JUNE 2023 NUMBER 1

EDITOR'S NOTE

G'DAY

As you will see, I have recently taken the helm as editor of Wood and Fiber Science. First, I would like to thank our outgoing editor Susan Levan-Green for her long, dedicated service to the journal as well as to the Society. She has worked tirelessly to ensure that our journal is first-rate and has successfully labored to improve our ranking among our peers.

By way of background, I currently direct the Center for Timber Durability and Design Life at the University of the Sunshine Coast in Australia. Prior to that, I led the durability program at Oregon State University from 1983 to 2018. I am a long-time member of the Society and I look forward to serving in this new role.

In most ways, you will see few tangible changes to the journal, but we have several important tasks to undertake. The first is increasing the number of quality manuscripts submitted. This will be critical as we move to open access. The second will be broadening our reach to incorporate other subject areas to encompass the broader areas related to lignocellulosic materials. We will begin by adding special focus sections to selected journal issues. This issue is an example of that process, but we need other topics and are seeking your input for potential topics.

The journal faces competition from a multitude of newer academic outlets. It will be critical for those in the field to consider Wood and Fiber Science as one of their primary outlets for their research results.

On behalf of the Society, I wish Susan the best as she tackles other adventures and am grateful for her dedication to our profession.

Jeffrey Morrell
Editor
E-mail: jmorrell@usc.edu.au

https://doi.org/10.22382/wfs-2023-01

FOREWORD

Illegal logging is a global problem that poses significant environmental, economic, and social consequences. It is defined as the harvesting, transportation, processing, buying, or selling of timber in violation of national and international laws. The trade in illegal timber is worth billions of dollars annually and is often associated with corruption, organized crime, and human rights abuses. In addition to contributing to deforestation and the loss of biodiversity, illegal logging also impacts the livelihoods of local communities who depend on forests for their survival. It can also contribute to climate change by releasing carbon into the atmosphere and reducing the capacity of forests to absorb carbon.

Efforts to combat illegal logging have increased in recent years, with international agreements, laws, and certification schemes aiming to improve forest governance and promote sustainable forest management. However, illegal logging persists, fueled by high demand for cheap timber products, weak law enforcement, and inadequate governance. Addressing illegal logging requires a multifaceted approach, including strengthening legal frameworks, enhancing enforcement mechanisms, promoting sustainable forest management practices, and reducing demand for illegal timber products through consumer awareness and responsible sourcing. In this issue of *Wood and Fiber Science*, we examine the latest developments in the identification and authentication of wood species to combat illegal logging.

The first article in this issue, by Raobelina et al (2023), presents a study on the use of a portable near-infrared (NIR) spectrometer for the identification of four *Dalbergia* species from Madagascar. The authors demonstrate that the portable NIR spectrometer is a reliable tool for the identification of wood species, and its portability makes it an ideal tool for fieldwork. The findings of this study have important implications for the monitoring and enforcement of regulations on the trade of *Dalbergia* species.

The second article, by Kim et al (2023), presents a novel method for distinguishing between native- and plantation-grown mahogany (*Swietenia macrophylla*) using chromatography and high-resolution quadrupole time-of-flight mass spectrometry. This study has important implications for the enforcement of regulations on the trade of mahogany, which is a valuable timber species that is often illegally harvested from natural forests.

The third article in this issue, by Richardson et al (2023), presents the "Global Wood Species Priority List," a living database of tree species that are most at risk for illegal logging, unsustainable deforestation, and high rates of global trade. The authors argue that such a database is crucial for prioritizing conservation efforts and ensuring the sustainability of global forest ecosystems. This article highlights the need for international collaboration and cooperation in the fight against illegal logging.

The fourth article, by Wang et al (2023), presents a study on the identification and classification of bamboo based on cross-sectional images using deep learning. The authors demonstrate that deep learning can be used to accurately identify and classify bamboo species based on their anatomical features. This study has important implications for the monitoring and enforcement of regulations on the trade of bamboo, which is a valuable raw material for various industries.

The fifth and final article in this issue, by Duchesne et al (2023), presents a study on the forensic identification of five Meliaceae (mahogany) species using a combination of gas chromatography × gas chromatography-time-of-flight mass spectrometry-based metabolomic profiling and wood anatomy. The authors demonstrate that this combined approach can be used to accurately identify and authenticate mahogany species, which are often illegally harvested and traded. This study has important implications for the enforcement of

Wood and Fiber Science, 55(1), 2023, pp. 2-3
https://doi.org/10.22382/wfs-2023-02

regulations on the trade of mahogany and for the conservation of natural forests.

Overall, the articles in this issue highlight the importance of developing and implementing effective strategies for identifying and authenticating wood species to combat illegal logging. The use of portable NIR spectrometers, high-resolution quadrupole time-of-flight mass spectrometry, deep learning of images, and metabolomic profiling are all promising tools for the identification of wood species. Additionally, the establishment of common databases, such as the "Global Wood Species Priority List," can help to prioritize conservation efforts and ensure the sustainability of global forest ecosystems. We hope that these articles will stimulate further research and discussion on this important topic and contribute to the development of effective solutions to combat illegal logging. As a community of scientists, the impact of our research in the use of wood products should not be siloed from the moral and ethical implication from the harvest and trade of these products.

CADY LANCASTER
Content Editor for Special Topics

USE OF A PORTABLE NEAR INFRARED SPECTROMETER FOR WOOD IDENTIFICATION OF FOUR *DALBERGIA* SPECIES FROM MADAGASCAR

Andry Clarel Raobelina†

PhD student
Doctoral School of Natural Resources Management and Development (GRND)
School of Agronomy—University of Antananarivo
BP 175, Antananarivo 101 Madagascar
E-mail: andryclarel@gmail.com

Gilles Chaix

Researcher
CIRAD, UMR AGAP Institut
F-34398 Montpellier, France
and
UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro
Montpellier, France
E-mail: gilles.chaix@cirad.fr

Andriambelo Radonirina Razafimahatratra

PhD student
Doctoral School of Natural Resources Management and Development (GRND)
School of Agronomy—University of Antananarivo
BP 175, Antananarivo 101 Madagascar
E-mail: andriambelo.radonirina@gmail.com

Sarobidy Pascal Rakotoniaina

Environmental and Social Vice Manager
Tozzi Green Madagascar sau
La Tour—26 Etage Rue Ravoninahitriniarivo—Ankorondrano
Antananarivo 101 Madagascar
E-mail: sarobidypascal@gmail.com

*Tahiana Ramananantoandro**†

Associate Professor
Department of Forestry and Environment
School of Agronomy—University of Antananarivo
BP 175 Antananarivo 101 Madagascar
E-mail: ramananantoandro@gmail.com

(Received July 2021)

Abstract. This study focused on the use of Near InfraRed (NIR) Spectroscopy to address the lack of tools and skills for wood identification of *Dalbergia* species from Madagascar. Two sample sets of 41 wood blocks and 41 wood cores belonging to four *Dalbergia* species (*D. abrahamii*, *D. chlorocarpa*, *D. greveana*, and *D. pervillei*) were collected in the northern and western regions of Madagascar. Sapwood and heartwood NIR spectra were measured on wood at 12% moisture content by using a portable VIAVI MicroNIR 1700 spectrometer. Four discrimination models corresponding to sapwood and heartwood of the two sample forms were developed using Partial Least Square Discriminant Analysis (PLSDA).

* Corresponding author
† SWST member

Good accuracy of 83.3% and 81.8% were obtained from the heartwood-based PLSDA models respectively for wood blocks and wood cores samples. All *D. chlorocarpa* samples were well-classified by the two models. Results highlighted the potential of portable NIR Spectroscopy as a helpful tool to support sustainable management and trade of Madagascar's *Dalbergia* species. Further studies are, however, needed for its operational use in identification routine.

Keywords: Discrimination, Near InfraRed Spectroscopy, portable spectrometer, PLSDA, *Dalbergia*, Madagascar.

INTRODUCTION

Dalbergia is a botanical genus comprising about 250 tree species, shrubs, and lianas (Yin et al 2018), widespread in tropical and subtropical regions (Saha et al 2013). Several *Dalbergia* tree species known under the trade names of rosewood and palisander provide valuable wood, which is harvested for making musical instruments (Wegst 2006; Perez and Marconi 2018) and furniture (Kaner et al 2013). Despite the Convention on International Trade of Endangered Species (CITES) restrictions, Madagascar's precious wood resources are illegally logged and traded to supply the illegal wood market (Patel 2007; Schuurman and Lowry 2009; Randriamalala and Liu 2010; Ratsimbazafy et al 2016; Weaber et al 2019).

There are 84 identified *Dalbergia* species from Madagascar, 83 of which are endemics. A total of 58 species can provide large trees with a diameter larger than 20 cm and serve as sources of precious wood (Phillipson et al 2022). There was an increase in their illegal logging after the 2009 due to political instability and the increase in demand for precious wood on the international market (Ratsimbazafy et al 2016). Nearly 40,000 tons of rosewood from Madagascar were exported illegally from 2008 to 2010, 64% of which were exported in 2009 only (Ratsimbazafy et al 2016). To enable a sustainable management and trade of these resources, all Madagascar's *Dalbergia* species were listed in the Appendix II of the CITES in 2013 under the Malagasy Government recommendation (Ratsimbazafy et al 2016).

The CITES implementation allowing to regulate international trade of wood species requires a good taxonomic knowledge based on wood anatomy to identify felled timbers (CITES 2019). However, in Madagascar as throughout the world, very few experts possess these skills. Several alternative methods are currently being tested and/or used to identify tree species or their geographical origin from wood specimens (Schmitz et al 2020). These include the use of mass spectrometry (Espinoza et al 2015; Mcclure et al 2015; Evans et al 2017; Zhang et al 2019; Brunswick et al 2021), analyses of stable isotope ratios of chemical elements (C, O, H, S, N, and Sr), characterizing the environmental conditions of the tree's growing site (Micha et al 2009; Rees 2015; Hajj et al 2017), genetic analysis based on molecular markers on wood DNA (Hassold et al 2016; Fatima et al 2019), Laser Induced Breakdown Spectroscopy (Celani et al 2019), and Near InfraRed (NIR) Spectrometry (Pastore et al 2011; Snel et al 2018). Among those existing techniques, wood anatomy, DNA barcoding, and NIR Spectroscopy were already tested and currently under development by Malagasy scientists for the identification of Madagascar's *Dalbergia* species.

NIR Spectroscopy with chemometrics, as an approach based on the signal processing, is described in the literature (Tsuchikawa and Kobori 2015). Pastore et al (2011) used NIR Spectroscopy to separate four anatomically similar wood species (*Swietenia macrophylla* King, *Carapa guanensis* Aubl, *Cedrela odorata* L. and *Micropholis melinoniana* Pierre) from wood powder by using Partial Least Square Discriminant Analysis (PLSDA). One published study only has been done so far in the literature regarding the discrimination of wood species belonging to the same *Dalbergia* genus using NIR Spectroscopy (Snel et al 2018). Snel et al (2018) used NIR Spectroscopy with PLSDA to discriminate six CITES-listed *Dalbergia* species from America and Asia (*D. decipularis* Rizzini and A. Mattos, *D. sisso* DC, *D. stevensonii* Standl, *D. latifolia* Roxb, *D. retusa* Hemsl, and *D. nigra*

[Vell.] Benth) from heartwood spectra by using a portable NIR Spectrometer. NIR Spectroscopy have not yet been carried out for *Dalbergia* from Madagascar. In light of the illegal trafficking of Madagascar's precious wood and their inclusion in CITES Appendix II, the development of a rapid and inexpensive tool to assist in the identification of these species is of crucial interest. Laboratory spectrometers are more efficient and stable than portable instruments, but their acquisition cost is relatively expensive (up to 70 times). Miniaturization of NIR Spectrometer instrumentation has improved its deployability resulting in its increased use both for laboratory and in-field measurements even with limited spectral resolution and range (Yan and Siesler 2018; Zhu et al 2021; Giussani et al 2022).

Species discrimination from wood material using NIR Spectroscopy is usually carried out on wood blocks (Pastore et al 2011; Snel et al 2018) or wood powder (Bergo et al 2016). Using wood cores is not very common despite it is a less invasive method of wood sampling (Van Mantgem and Stephenson 2004; Helcoski et al 2019). Wood cores also allow consideration of the variability of wood chemical properties radially through the wood. This work addresses method development for the classification of four *Dalbergia* species of Madagascar using a handheld VIAVI MicroNIR 1700 spectrometer and investigates the impact of sample form (block *vs.* core) and wood type (heartwood *vs.* sapwood).

MATERIALS AND METHODS

Wood Blocks and Wood Cores Sampling

Sampling sites were located inside and outside of protected forest areas in Boeny, Diana, Sofia, Betsiboka, and Menabe regions of Madagascar (Fig 1). It was not possible to sample each of the four species in the five regions due to geographic distributions. A total of 41 wood cores belonging to *D. chlorocarpa* R. Vig., *D. greveana* Baill., and *D. abrahamii* Bosser & R. Rabev. and 41 wood blocks belonging to *D. chlorocarpa*, *D. greveana*, and *D. pervillei* Vatke (Fig 2[a], Table 1) were collected between 2016 and 2019 from a total of 67 trees. Wood blocks and cores were collected from trees with a diameter larger than 20 cm. Wood cores with a diameter of 5 mm were collected at 1.30 m from the ground, from pith to bark and in a perpendicular direction to the tree axis using a Pressler borer. Wood blocks (4 cm×4 cm×4 cm) were sampled from the tree trunk just below 1.30 m from the ground. An herbarium voucher was also prepared for each tree for botanical verification. Species identification was carried out by taxonomists from Missouri Botanical Garden Madagascar and National Museum of Natural History France.

Wood Moisture Conditioning

To minimize moisture effects due to moisture variation, all samples were stabilized at 12% theoretical moisture content in a climatic chamber at 20°C and 65% relative humidity. The wood moisture content of 12% was reached when the difference in mass between two measurements spaced in 24-h intervals for four reference samples did not exceed ±5%.

NIR Measurements

NIR absorbance spectra were measured on moisture stabilized wood samples with a portable MicroNIR VIAVI 1700 spectrometer (Viavi Solution–Milpitas, CA). Heartwood and sapwood NIR spectra measurements (Fig 2[b]) were made in diffuse reflection mode from 900 to 1700 nm. Each measurement was an average of 100 scans and the integration time was set to 10 ms. A Spectralon (99% reflectance) was used as reference background. NIR measurements were made on the transversal face for wood blocks, whereas they were made on unidentified face for wood cores since it was difficult to detect the wood ligneous plane.

Three spectra were collected radially from pith to bark for the sapwood and heartwood of each sample (n = 420) and then averaged (n = 140). These averaged spectra were divided in four spectral datasets (Table 2) according to sample form (block or core) and wood type (sapwood or heartwood). Each dataset consists of a spectral data

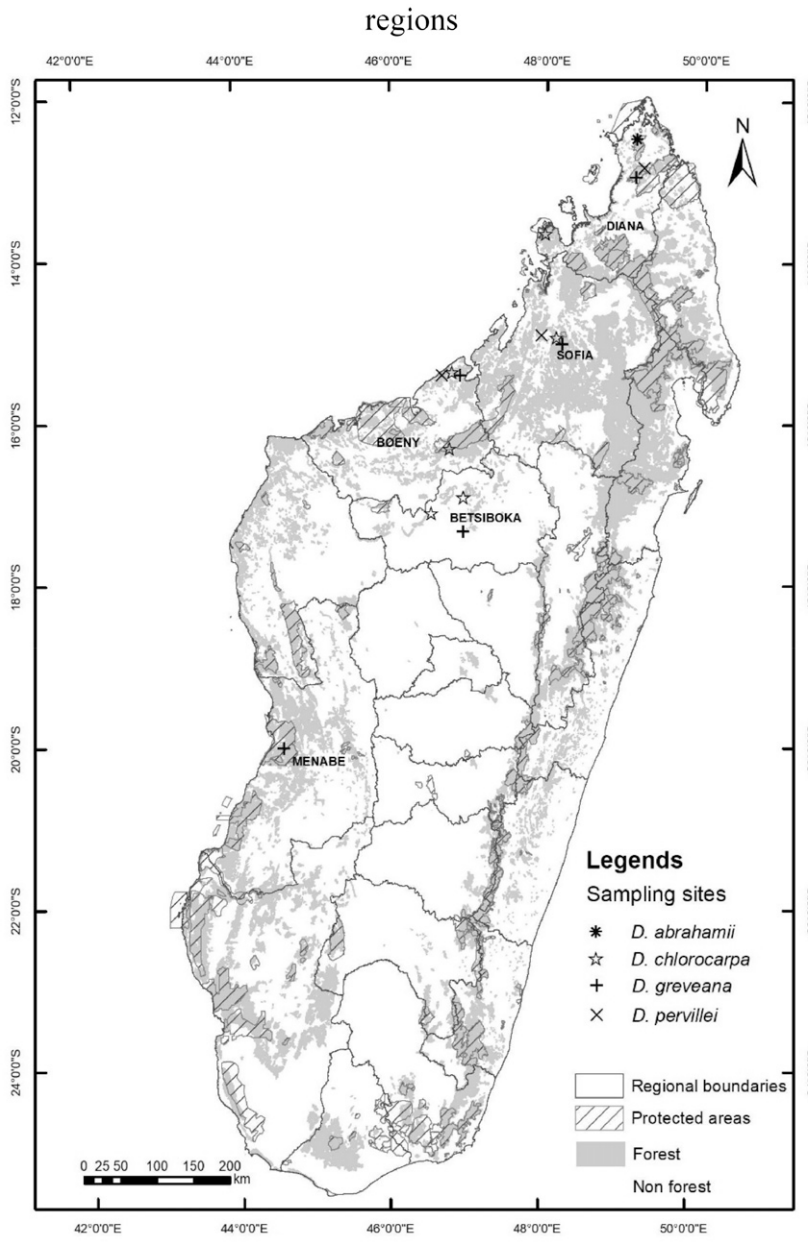


Figure 1. Location of wood sampling sites in Diana, Sofia, Boeny, Betsiboka, and Menabe regions. Source: Institut Géographique et Hydrologique de Madagascar (FTM) BD500; Google Earth, 2020.
Geographic coordinate system: GCS International 1924.

$X_{sample\_form \times wood\_type}$ $(n \times p)$ and a reference data $Y_{sample\_form \times wood\_type}$ $(n \times q)$, where n, p, and q are respectively the number of spectra, independent variables (wavelengths), and species within a dataset.

Spectral Data Preprocessing

For each of the four spectral data, absorbances data in the wavelength ranges of 900-950 nm and 1650-1700 nm were removed from the raw

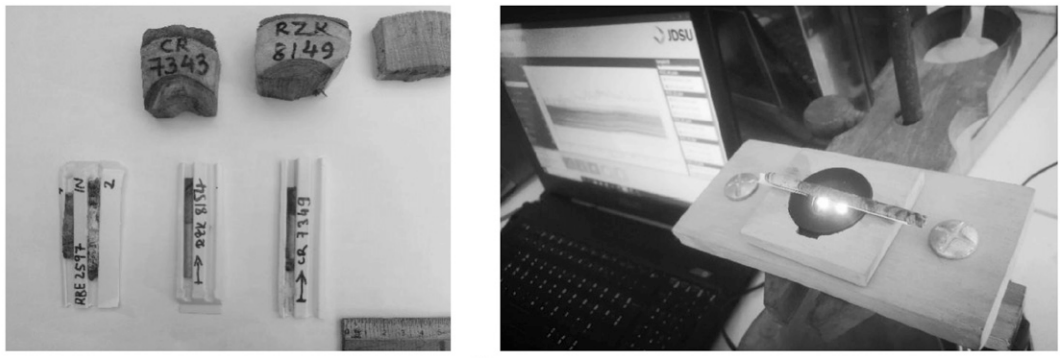


Figure 2. Wood blocks and cores samples (a) belonging to the four *Dalbergia* species: *D. abrahamii* (RBE2597), *D. chlorocarpa* (CR7343, RZK8154), *D. greveana* (RZK8149, CR7349) and *D. pervillei* (SH748); (b) NIR absorbance spectra measurements on core wood sample using VIAVI MocroNIR 1700 spectrometer.

spectral data to eliminate noise. Several preprocessing methods were used to improve the signal including the first (SG1) and second (SG2) derivative followed by the Savitzky–Golay smoothing with filter width from 5 to 25 points by 2 points of steps, second-order polynomial Detrending (Dt) and Standard Normal Variate (SNV). A combination of two, then three of the different preprocessing methods was applied to the spectra.

Principal Component Analysis

Principal Component Analysis (PCA) was carried out separately on each of the four preprocessed datasets (Table 2) to evaluate the clustering of spectral data $X_{(n\times 113)}$ and their variation in the space defined by the Principal Components (PCs) according to the qualitative variables $Y_{(n\times 3)}$, the corresponding *Dalbergia* species. Outlier spectra were identified and removed on the basis of the Hotelling distance (T^2) and the PCA residuals (Q) (Eriksson et al 2013) before the calibration of PLSDA models.

Partial Least Square Discriminant Analysis

The spectra were divided randomly into two datasets, 75% for training and 25% for validation, with the same proportionality of number of samples per species (Table 2). The PLSDA method was used based on the Non-Linear Iterative Partial Least Squares (NIPALS) algorithm for calibration (Wold et al 2001). Four PLSDA models were calibrated (Table 2): one model per combination of wood type (heartwood or sapwood) and sample form (wood block and core).

Each PLSDA model was calibrated with 20 Discriminant Variables (DVs). Due to the small number of available samples, Leave-One-OutCross-Validation (LOOCV) on the training datasets was used to select the best preprocessing methods and the optimal number of DVs

Table 1. Distribution of *Dalbergia* wood sample according to sample form and provenance from Madagascar.

		Region (cores/blocks)				
Species	Number of samples (cores/blocks)	Diana	Menabe	Sofia	Betsiboka	Boeny
D. abrahamii	9/–	9/–	–	–	–	–
D. chlorocarpa	16/14	–	–	1/–	1/–	14/14
D. greveana	16/17	4/1	1/–	2/–	3/–	6/16
D. pervillei	–/10	–/2	–	–/1	–	–/7

Table 2. Number and distribution of averaged spectra in the calibration and validation dataset according to species, sample form, sapwood, and heartwood.

				Number of spectra		
Sample form	Wood type	Species	Spectral and reference data	Calibration	Validation	Spectral outliers
Cores	Sapwood	*D. abrahamii*	$X_{cr \times spw(39\times 113)}$	6	3	0
		D. chlorocarpa	$Y_{cr \times spw(39\times 3)}$	11	4	0
		D. greveana		11	4	0
	Heartwood	*D. abrahamii*	$X_{cr \times hrt(41\times 113)}$	6	3	0
		D. chlorocarpa	$Y_{cr \times hrt(41\times 3)}$	12	4	0
		D. greveana		12	4	0
Blocks	Sapwood	*D. chlorocarpa*	$X_{blc \times spw(38\times 113)}$	9	3	0
		D. greveana	$Y_{blc \times spw(38\times 3)}$	13	4	0
		D. pervillei		7	2	0
	Heartwood	*D. chlorocarpa*	$X_{blc \times hrt(21\times 113)}$	6	2	1
		D. greveana	$X_{blc \times hrt(21\times 3)}$	5	2	0
		D. pervillei		4	2	0

cr, cores; blc, clocks; spw, sapwood; hrt: heartwood.

(Bertrand 2005). All data processing was performed using Chemflow (Rossard et al 2020).

The four most accurate PLSDA models according to wood type and sample form, which result from the optimal number of DVs and the best preprocessing methods were tested on the corresponding validation datasets. The classification results were shown in a confusion matrix Y, which is a k order square matrix generated by each discrimination model, where k is the number of species to be discriminated. The confusion matrix of each four models showed the predicted classes of the validation samples compared with their reference classes. A sample i is correctly classified if its predicted class $\hat{y}$ in the confusion matrix corresponds to its reference class y. The performance of each model was evaluated by precision, recall, and accuracy metrics defined as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (2)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \quad (3)$$

Where TP, FN, TN, and FP are respectively the number of true positives, false negatives, true negatives, and false positives. For one reference class, TP measures the number of correctly classified samples, whereas TN explains the number of correct classifications of other classes. FP measures the number of incorrectly classified samples of the reference class, whereas FN is the incorrect classification of others. The Recall metric evaluates the classification rate, which is the ratio of correctly classified positive samples to the total number of actual positive samples. The precision expresses the probability of certainty of the correct classification, which is the ratio of correctly classified positive samples to the total number of positive classifications. The accuracy evaluates how good the model is to classify all the classes. It is the ratio of total number of correct classifications to the total number of classification results.

RESULTS

NIR Absorbance Spectra

Figure 3(a) and (b) show the SG1 ($W = 7$ points) mean spectra of *D. abrahamii, D. chlorocarpa, D. greveana*, and *D. pervillei* according to the wood type and the sample form.

Blocks spectra show higher absorption than cores spectra in the second harmonic region, regardless of the species or wood part. Absorption peaks at the vicinity of 1410, 1180, and 1350 nm regions

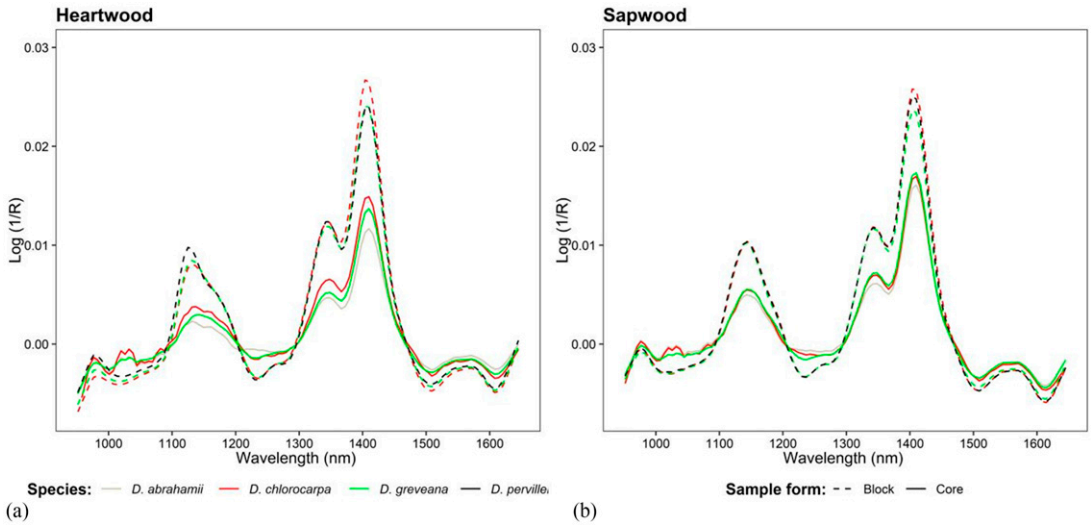


Figure 3. Mean SNV + SG1 (W = 7) heartwood (a) and sapwood (b) spectra of the four *Dalbergia* species according to wood type.

were detected (Fig 3[a] and [b]) that are more important for blocks than cores. Core spectra were also noisier than blocks spectra, especially before 1100 nm.

Principal Components Analysis

Figures 4(a), 5(a), 6(a), and 7(a) show the PCA scores plots from the heartwood and sapwood spectra of the blocks and cores samples in the PC1-PC2 space.

Separation between species is less significant based on sapwood than heartwood spectra (Fig 5[a], [b] and Fig. 7[a], [b]). For the heartwood of blocks, the preprocessing method based on the combination of the SG1 (W = 5 points), Dt, and SNV clustered spectra according to species in the PC1-PC2

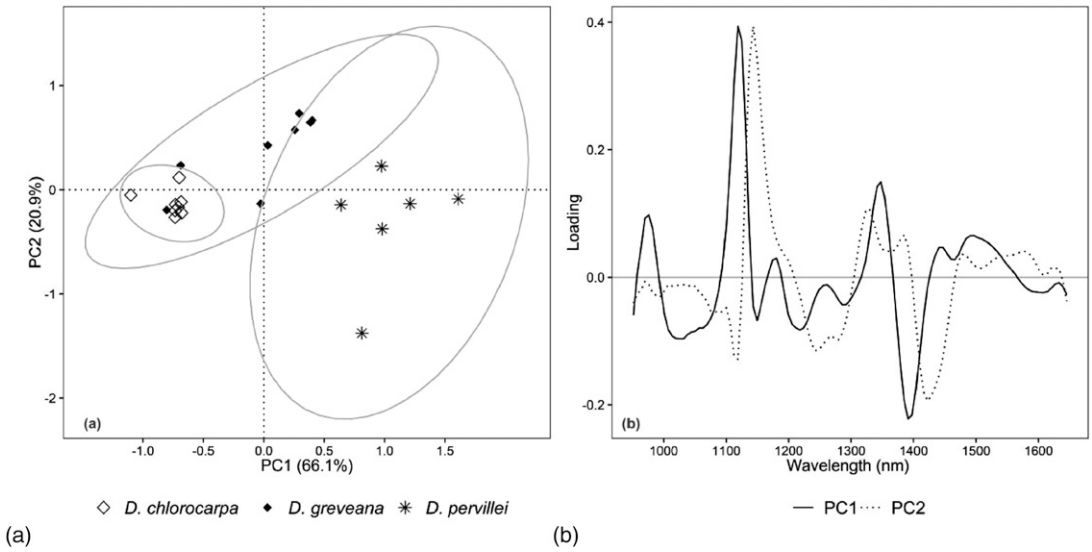


Figure 4. PCA results from the SNV + Dt + SG1 (W = 5 points) heartwood spectra of blocks: (a) scores plot of the first two PCs and (b) corresponding loading plot.

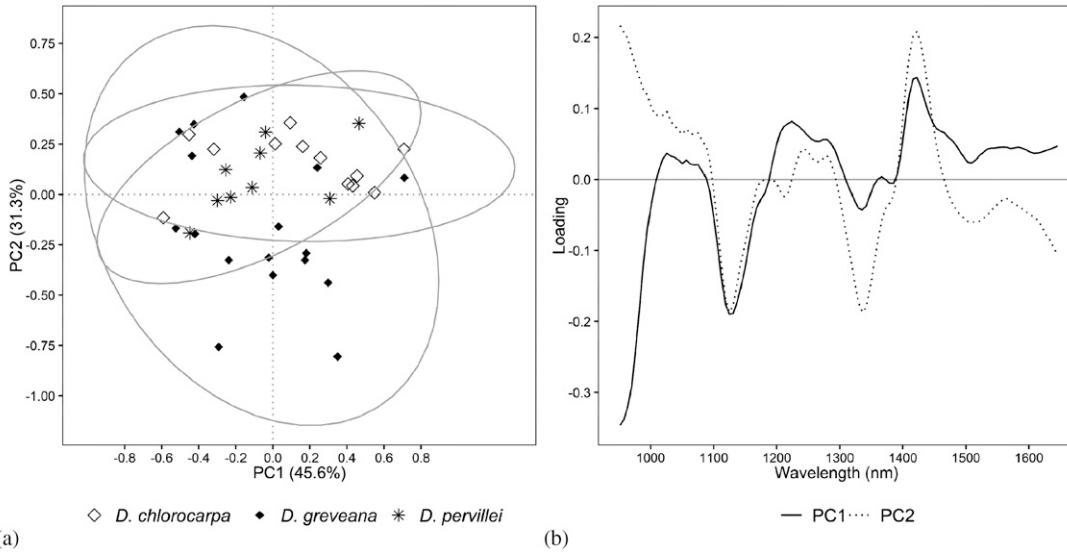


Figure 5. PCA results from the SG1 (W = 5 points) sapwood spectra of blocks: (a) scores plot of the first two PCs and (b) corresponding loading plot.

space (PC1 = 66.1%, PC2 = 20.9%) (Fig 4[a]). *D. pervillei* and *D. chlorocarpa* were clearly separated with a greater dispersion of *D. pervillei* spectra in the positive part of the PC1 axis (Fig 4[a]). The spectra of *D. greveana* were more scattered in the PC1-PC2 space. The PC1 and PC2 loadings (Fig 4[b]) showed high contribution of the wavelength regions at the vicinity of 975, 1180, and 1250 nm to discriminate the three *Dalbergia* species, especially between *D. chlorocarpa* and *D.*

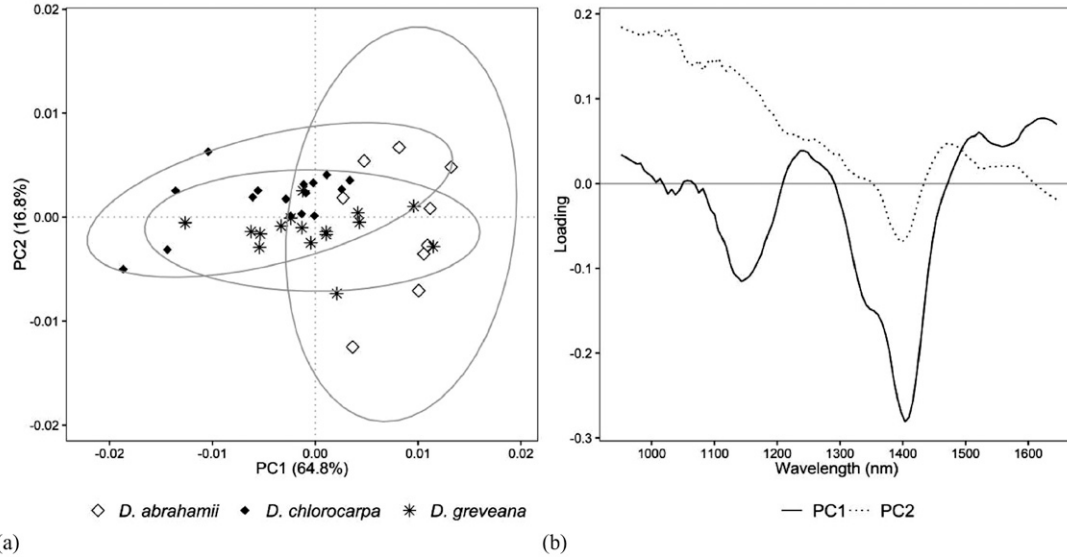


Figure 6. PCA results from the SG1 (W = 11 points) heartwood spectra of cores: (a) scores plot of the first two PCs and (b) corresponding loading plot.

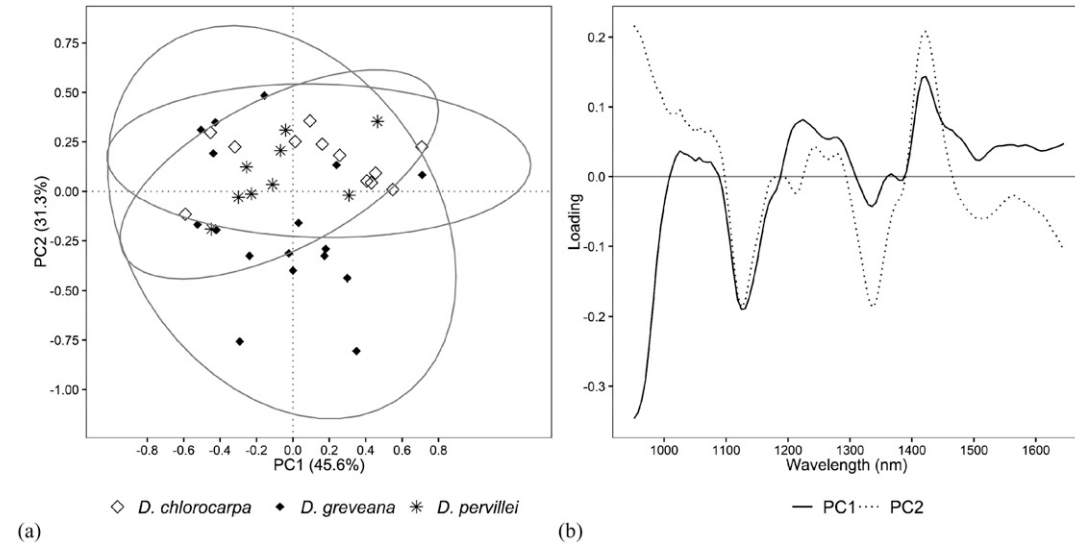


Figure 7. PCA results from the SG1 (W = 13 points) sapwood spectra of cores: (a) scores plot of the first two PCs and (b) corresponding loading plot.

pervillei, whose separation was significant. One sample of *D. chlorocarpa* was removed from the dataset because it had a high T^2 distance compared with the overall blocks' spectra (Table 2).

Spectra clustering according to species on the PCA score plot was less significant for the cores (Fig 6[a]) than blocks (Fig 4[a]). The use of the SG1 (W = 11 points) on the heartwood spectra of cores, however improved, the spectra clustering for the three species in PC1-PC2 space, which explains 81.6% of the spectral data variation (PC1 = 64.8%, PC2 = 16.8%) (Fig 6[a]). Several spectra of *D. abrahamii* and *D. chlorocarpa* were separated through PC1. *D. greveana* spectra were more scattered in the score plot. Corresponding loadings for PC1 and PC2 highlighted the contribution of the wavelength regions in the vicinity 1250 nm, and 1450 nm (Fig 6[b]) to discriminate the three species.

Partial Least Squares Discriminant Analysis

The performance of the preprocessing methods varies depending on the PLSDA models (Table 3). The SNV combined with the SG2 resulted in a higher accuracy in cross-validation for core PLSDA models. The addition of the Dt resulted in a higher discrimination accuracy in cross-validation for block PLSDA (Table 3). The optimal number of DVs corresponding to the highest correct classification rate in cross-validation ranged from 3 to 11 DVs.

The four *Dalbergia* species can be better discriminated from their heartwood spectra. Of the two heartwood models, the $PLSDA_{(Block \times heartwood)}$ show the higher accuracy of 83.3%. All samples of *D. chlorocarpa* and *D. pervillei* were well-classified. $PLSDA_{(Block \times heartwood)}$ classified *D. pervillei* with precision and recall of 100%. This finding shows that *D. pervillei* can be accurately separated from *D. chlorocarpa* and *D. greveana* from the heartwood spectra of blocks. The precision of the classification when the model identifies *D. chlorocarpa* is 67% because of a misidentification of a *D. greveana* sample, which was classified as *D. chlorocarpa* (Table 4[b]).

The $PLSDA_{(Core \times heartwood)}$ showed an accuracy of 81.8% (Table 3). All *D. chlorocarpa* samples were well-classified with recall of 100%. One sample of *D. abrahamii* and one sample of *D. greveana* were misclassified as *D. chlorocarpa* (Table 5), decreasing the precision of classification at 67% when the model identifies *D. chlorocarpa*.

Table 3. Discrimination and classification results of the four *Dalbergia* species according to the four PLSDA models.

PLSDA model names	Best preprocessing method	Number of DVs	LOOCV	Classification results		
				Accuracy	Precision	Recall
PLSDA (Block × Heartwood)	–	–	–	–	–	–
D. chlorocarpa	SG1 + Dt + SNV	5	86.70	83.30	0.67	1.00
D. greveana	–	–	–	–	1.00	0.50
D. pervillei	–	–	–	–	1.00	1.00
PLSDA (Block × Sapwood)	–	–	–	–	–	–
D.chlorocarpa	SG1 + Dt + SNV	3	65.50	77.80	0.75	1.00
D. greveana	–	–	–	–	0.75	0.75
D. pervillei	–	–	–	–	1.00	0.50
PLSDA (Core × Heartwood)	–	–	–	–	–	–
D. abrahamii	SG2 + SNV	3	76.70	81.10	1.00	0.50
D. chlorocarpa	–	–	–	–	0.67	0.10
D. greveana	–	–	–	–	1.00	0.60
PLSDA (Core × Sapwood)	–	–	–	–	–	–
D. abrahamii	SNV + SG2	14	62.10	72.20	0.75	0.60
D. chlorocarpa	–	–	–	–	0.67	0.80
D. greveana	–	–	–	–	1.00	0.40

PLSDA, Partial Least Square Discriminant Analysis; DVs, Discriminant Variables; SNV, Standard Normal Variate.

For *D. chlorocarpa*, which were common to the blocks and cores, classification results at species level were consistent for the four PLSDA models. Precisions of classification were between 67% and 75% while recall values were between 80% and 100% (Table 3).

DISCUSSION

PLSDA Model Accuracy According to Wood Type

The PCA score plot (Figs 4-7) showed that heartwood spectra from blocks and cores clustered more tightly according to species than sapwood spectra. This trend was confirmed by the PLSDA results, which also had better discrimination of species when using heartwood spectra. The need for a low number of DVs for heartwood PLSDA models (Table 3) also explains why it is easier to separate these four species using heartwood spectra.

In terms of factor inherent to the chemistry of wood, this good separation could be explained by the number of extractives, which are significantly higher in heartwood than sapwood (Morais and Pereira 2012; Razafimahatratra et al 2019), and which also vary among species (Freire et al 2005). The separation of the *Dalbergia* species studied here are consistent with the work of others who obtained discrimination results above 80% from heartwood spectra of non-Malagasy *Dalbergia species* using a portable and handheld NIR Spectrometer (Snel et al 2018).

Table 4. Confusion matrix from the validation of the two wood blocks PLSDA models for *Dalbergia* species discrimination ([a] for sapwood samples, [b] for heartwood samples).

	Reference class (sapwood)			Number of well-classified samples	Reference class (heartwood)			Number of well-classified samples
	D. chlorocarpa	*D. greveana*	*D. pervillei*		*D. chlorocarpa*	*D. greveana*	*D. pervillei*	
Predicted classes								
D. chlorocarpa	3	1	–	3	2	1	–	2
D. greveana	–	3	1	3	–	1	–	1
D. pervillei	–	–	1	1	–	–	2	2
Accuracy (%)	–	–	–	77.8	–	–	–	83.3

Table 5. Confusion matrix from the validation of the two wood cores PLSDA models for species discrimination of (a) sapwood and (b) heartwood.

	Reference classes (sapwood)			Number of well-classified samples	Reference classes (heartwood)			Number of well-classified samples
	D. abrahamii	*D. chlorocarpa*	*D. greveana*		*D. abrahamii*	*D. chlorocarpa*	*D. greveana*	
Predicted classes								
D. abrahamii	3	1	–	2	2	–	–	2
D. chlorocarpa	–	4	2	4	1	4	1	4
D. greveana	–	–	2	2	–	–	3	3
Accuracy (%)	–	–	–	72.7	–	–	–	81.8

Effect of the Sample Form (Blocks vs Cores) on NIR Absorbance Spectra and PLSDA Models

Block spectra generally showed a higher absorbance than core spectra. This could be due to the difference in sample size for the two sample forms. Unlike blocks, the surface of cores was not flat but curved, which could change the focal area for NIR measurements. The diameter of cores was also not large enough to cover the entire measurement window of the MicroNIR spectrometer resulted in some random noise on the core spectra. The SG1 ($W = 5$ points) spectra of the two sample forms had indeed made it possible to highlight more intense absorbance peaks on flat than curved wood surface. Some random noise was also generated by the effect of external stray light on the core spectra. The absorbance spectra of the blocks could then be more faithful than those of the cores.

Other effects related to the wood sample probably came from the difference in surface state of the two sample forms because block samples have smoother surfaces than cores. Previous studies have shown that the presentation and preparation of the samples strongly influence the performance of the models (Hein et al 2010). Zhang et al (2015) also showed that the absorption of NIR radiations by wood samples decreases with increasing surface roughness, and it could influence the performance of NIR calibration models.

PLSDA models based on blocks were calibrated from the spectra measured on the wood transversal face, unlike the core PLSDA models that were calibrated from the spectra collected on unidentified faces between radial and transversal. The SG1 spectra showed more intense peaks for the blocks than the wood cores around 1200 nm, which are dominated by the second overtone of O-H and N-H bonds in the lignin and cellulose component of the wood, and 1470 nm region which corresponds to the first overtone of O-H bonds in the hemicelullose (Schwanninger et al 2011). NIR spectra of the blocks, which were collected only on the transversal face of wood, thus provided more chemical information. Costa et al (2018) also highlighted that NIR spectra of the wood transversal face are different from those of the radial and tangential surfaces due to the difference in anatomical arrangement of cells in the three wood sections. According to Braga et al (2011), models calibrated on a given wood face should not be used to analyze spectra obtained on another wood face.

Ability of the Portable and Handheld MicroNIR Spectrometer on *Dalbergia* Species Discrimination

The NIR wavelength range covered by the MicroNIR (900-1700 nm) is smaller than that of benchtop spectrometers used in the literature. Chemical compounds with covalent bonds whose vibrational frequencies are outside the wavelength range of 900-1700 nm could, however, be useful for better species discrimination (Pastore et al 2011; Snel et al 2018). For example, Snel et al (2018) discriminated seven CITES appendix *Dalbergia* species using a portable microPHAZIR RX spectrometer (Thermo Scientific, Boston, MA) covering 1595-2396 nm with a discrimination accuracy of 90%. A portable instrument covering the entire NIR range may therefore provide additional discrimination power.

Use of NIR Spectroscopy for a Sustainable Management of Forest Resources

This study demonstrates the potential of the portable MicroNIR as a tool for supporting sustainable management and trade of *Dalbergia* tree species of Madagascar. Handheld spectrometers are accessible resources for developing countries like Madagascar as they are affordable and can be used to enforce the national law against trafficking and illegal trade of precious woods. Advantageously, once the models have been calibrated, the identification process by NIR Spectroscopy is fast and requires minimal science background for classification of unknowns. Future studies should investigate model recalibration with more samples in the training set, the effects of the external parameters, such as wood water content and surface aging of the wood on the discrimination models, and including more trade-significant species.

CONCLUSIONS

This study demonstrates the identification of Malagasy *Dalbergia* with a portable NIR Spectrometer. Significantly, this study showed discrimination within a genus, where there is less anatomical variation than between genera. Discrimination results from heartwood PLSDA using flat surface wood with full spot size were the most accurate. However, using wood core samples with round surface is less invasive. The correct classification rates for *D. greveana* and *D. abrahamii* were generally lower than those of *D. pervillei* and *D. chlorocarpa*. Some perspectives, such as the recalibration of PLSDA models, with more wood samples covering the natural variability related to wood chemical properties for the 58 Madagascar *Dalbergia* species that are potentially harvestable, as well as the testing of the models on samples completely independent of the training set are important to develop MicroNIR as a tool to assist in the identification of Madagascar's *Dalbergia* species. Due to its rapidity and portability, the availability of handheld NIR Spectrometer for forest law enforcement officers and custom officials is of crucial interest to support CITES in the management and sustainable trade of these resources. In-depth research on the external variables, which affect the performance of the models for on-field identification, must also be carried out to understand the accuracy limits of NIR Spectroscopy to use the tool to fight against illicit trafficking of the valuable wood species of Madagascar. Taxonomic clarification of Madagascar's *Dalbergia* with the objective of establishing a stable reference database must also be pursued.

ACKNOWLEDGMENT

Authors would like to thank the Missouri Botanical Garden of Madagascar (MBG), the Savaivo and Biodev Madagascar consulting firms for their support in the field work, the MBG-Madagascar and the National Museum of Natural History Paris (MNHN) for taxonomic identification, the laboratory of wood and plant anatomy (LABAP) at the Plant Biology and Ecology (MBEV) department of the Faculty of Sciences of the University of Antananarivo for providing wood blocks and some wood core samples, the Ministry of Environment and Sustainable Development (MEDD) for having enabled the implementation of this study and for their support in the administrative procedures. This study was financed by two projects: 1) project on precious wood of Madagascar, funded by Deutsche Gesellschaft für Internationale Zusammenarbeit—Programme d'Appui à la Gestion de l'Environnement (GIZ-PAGE, 2017) and 2) Sustainable Management of Precious Wood *Dalbergia* and *Diospyros* from Madagascar funded by the European Union (G3D, 2018-2022, research permits [280/18/MEEF/SG/DGF/DSAP/SCB.RE], [189/19/MEDD/SG/DGEF/DGRNE], [122/20/MEDD/SG/DGEF/DGRNE], [033/21/MEDD/SG/DGGE/DAPRE/SCBE.Re] and [092/22/MEDD/SG/DGGE/DAPRE/SCBE.Re]).

REFERENCES

Bergo MCJ, Pastore TCM, Coradin VTR, Wiedenhoeft AC, Braga JWB (2016) NIRS identification of Swietenia macrophylla is robust across specimens from 27 countries. IAWA J 37(3):420-430.

Bertrand D (2005) Étalonnage multidimensionnel: Application aux données spectrales. Extr de: Techniques de l'ingénieur, analyse et caractérisation. Pages 1-21. Bibliogr., Paris: Techniques de l'ingénieur, cop. 2005, Paris: Techniques de l'ingénieur, cop. 2005.

Braga JWB, Pastore TCM, Coradin VTR, Camargos JAA, Silva AR (2011) The use of near infrared spectroscopy to identify solid wood specimens of *Swietenia macrophylla* (CITES appendix II). IAWA J 32(2):285-296.

Brunswick P, Cuthberston D, Yan J, Chua CC, Duchesne I, Isabel N, Evans PD, Gasson P, Kite G, Bruno J, Aggelen G, Shang D (2021) A practical study of CITES wood species identification by untargeted DART/QTOF, GC/QTOF and LC/QTOF together with machine learning process and statistical analysis. Environ Adv 5(2021):1-10.

Celani C, Lancaster CA, Jordan JA, Espinoza EO, Booksh KS (2019) Assessing utility of laser induced breakdown spectroscopy as a mean of *Dalbergia* speciation. Analyst 144(17):1-23.

CITES (2019) Identification materials. https://cites.org/eng/imp/identification_materials/index.php (24 April 2021).

Costa EVS, Rocha MFV, Hein PRG, Amaral EA, Santos LM, Brandao LEVS, Trugilho PF (2018) Influence of spectral acquisition technique and wood anisotropy on the statistics of predictive near infrared-based models for wood density. J Near Infrared Spec 26(2):106-116.

Eriksson L, Byrne T, Johansson E, Trygg J, Vikstrom C (2013) Multi and megavariate data analysis basic principles and applications. Umetrics AB, Malmö, Sweden. 521 pp.

Espinoza EO, Wiemanm MC, Barajas-Morales J, Chavarria GD, McClure PJ (2015) Forensic analysis of CITES-protected *Dalbergia* timber from the Americas. IAWA J 36(3):311-325.

Evans PD, Mundo IA, Wiemann MC, Chavarria GD, McClure PJ, Voin D, Espinoza EO (2017) Identification of selected CITES-protected Araucariaceae using DART TOFMS. IAWA J 38(2):266-281.

Fatima T, Srivastava A, Somashekar PV, Hanur VS, Rao MS (2019) Development of DNA-based species identification and barcoding of three important timbers. Bull Natl Res Cent 43(1):1-17.

Freire CSR, Silvestre AJD, Neto CP (2005) Lipophilic extractives in Eucalyptus globulus kraft pulps. Behavior during ECF bleaching. J Wood Chem Technol 25(1-2):67-80.

Giussani B, Gorla G, Riu J (2022) Analytical chemistry strategies in the use of miniaturised NIR instruments: An overview. Crit Rev Anal Chem 0(0):1-33.

Hajj F, Poszwa A, Bouchez J, Guérold F (2017) Radiogenic and "stable" strontium isotopes in provenance studies: A review and first results on archaeological wood from shipwrecks. J Archaeol Sci 86(2017):24-49.

Hassold S, Lowry PP, Bauert MR, Razafintsalama A, Ramamonjisoa L, Widmer A (2016) DNA barcoding of Malagasy rosewoods: Towards a molecular identification of CITES-listed *Dalbergia* species. PLOS ONE 11(6):1-17.

Hein PRG, Lima JT, Chaix G (2010) Effects of sample preparation on NIR spectroscopic estimation of chemical properties of Eucalyptus urophylla S. T. Blake wood. Holzforschung 64(1):45-54.

Helcoski R, Tepley AJ, McGarvey JC, Gonzalez-Akre E, Meakem V, Thompson JR, Anderson-Teixeira KJ (2019) No significant increase in tree mortality following coring in a temperate hardwood forest. Tree-Ring Res 75(1):67-72.

Kaner J, Jiufang L, Yongji X, Ioras F (2013) A reevaluation of woods used in Chinese historic furniture. Bull Transilv Univ Brasov 6(55):9-18.

Mcclure PJ, Chavarria GD, Espinoza E (2015) Metabolic chemotypes of CITES protected *Dalbergia* timbers from Africa, Madagascar, and Asia. Rapid Commun Mass Sp 29(9):783-788.

Micha H, Michael J, Krehan H (2009) Control of origin of larch wood: Discrimination between European (Austrian) and Siberian origin by stable isotope analysis. Rapid Commun Mass Spectrom 23(2009):3688-3692.

Morais MC, Pereira H (2012) Variation of extractives content in heartwood and sapwood of Eucalyptus globulus trees. Wood Sci Technol 46(4):709-719.

Pastore TCM, Braga JWB, Coradin VTR, Magalhaes WLE, Okino EYA, Camargos JAA, Muniz GIB, Bressan OA, Davrieux F (2011) Near infrared spectroscopy (NIRS) as a potential tool for monitoring trade of similar woods: Discrimination of true mahogany, cedar, andiroba, and curupixá. Holzforschung 65(1):73-80.

Patel ER (2007) Looging af rare rosewood and palissandre (*Dalbergia* spp) within Marojejy National Park, Madagascar. Madag Conserv Dev 2(1):11-16.

Perez M, Marconi E (2018) Wooden musical instruments different forms of knowledge. La librairie, Paris. 415 pp.

Phillipson PB, Wilding N, Crameri S, Andriambololonera S, Rakotonirina N, Rabarimanarivo M, Manjato N (2022) Large tree species of *Dalbergia* from Madagascar. Catalogue of the Plants of Madagascar. http://www.tropicos.org/NamePage.aspx?nameid=40021450&projectid=17 (14 July 2022).

Randriamalala H, Liu Z (2010) Rosewood of Madagascar: Between democracy and conservation. Madagascar Conserv Dev 5(1):11-22.

Ratsimbazafy C, Newton DJ, Ringuet S (2016) The rosewood and ebony trade of Madagascar. Timber Island, Traffic Report. 125 pp.

Razafimahatratra AR, Belloncle C, Chaix G, Raobelina AC, Ramananantoandro T (2019) Variabilité des propriétés chimiques (extractibles et phénols totaux) de quelques espèces de bois précieux de Madagascar (*Dalbergia* spp. et *Diospyros* spp.). Pages 92-95 *in* Romain Rémond, Arnaud Besserer, Emmanuel Fredon, Philippe Gérardin and Joseph Gril, eds. 8èmes journées du GDR Sciences du bois, 18-20 November, Epinal-France.

Rees OG (2015) Verifying the declared origin of timber using stable isotope ratio and multi-element analyses. MSc by Research, University of York, Heslington, UK. 169 pp.

Rossard V, Boulet JC, Gogé F, Latrille E, Roger JM (2020) ChemFlow, chemometrics using galaxy. Galaxy Community Conference-GCC2016, Indiana University, June 2016, Bloomington, IN. 10.7490/f1000research. 1112573.1. hal-01837798.

Saha S, Shilpi JA, Mondal H, Hossain F (2013) Ethnomedicinal, phytochemical, and pharmacological profile of the genus *Dalbergia* L. (Fabaceae). Phytophramacology 4(2):291-346.

Schmitz N, Beeckman H, Blanc-Jolivet C, Boeschoten L, Braga JWB, Cabezas JA, Chaix G, Crameri S, Degen B, Deklerck B, Dormontt E, Espinoza E, Gasson P, Haag V, Helmling S, Horacek M, Koch G, Lancaster C, Lens F, Lowe A, Martínez-Jarquín S, Nowakowska JA, Olbrich A, Paredes-Villanueva K, Pastore TCM, Ramananantoandro T, Razafimahatratra AR, Ravindran P, Rees G, Soares LF, Tysklind N, Vlam M, Watkinson C, Wheeler E, Winkler R, Wiedenhoeft AC, Zemke VT, Zuidema P (2020) Overview of current practices in data analysis for wood identification. A guide for the different timber tracking methods. Global Timber Tracking Network. Großhansdorf, Germany: GTTN Secretariat, European Forest Institute and Thunen Institute. 141 pp.

Schuurman D, Lowry II P (2009) The Madagascar rosewood massacre. Madagascar Conserv Dev 4(2):98-102.

Schwanninger M, Rodrigues JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. J Near Infrared Spec 308 (19):287-308.

Snel FA, Braga JWB, da Silva D, Wiedenhoeft AC, Costa A, Soares R, Coradin VTR, Pastore TCM (2018) Potential field-deployable NIRS identification of seven *Dalbergia* species listed by CITES. Wood Sci Technol 52(5):1411-1427.

Tsuchikawa S, Kobori H (2015) A review of recent application of near infrared spectroscopy to wood science and technology. J Wood Sci 61(3):213-220.

Van Mantgem PJ, Stephenson NL (2004) Does coring contribute to tree mortality? Can J Forest Res 34(11): 2394-2398.

Weaber PO, Schuurman D, Ramamonjisoa B, Langrand M, Barber CV, Innes JL, Lowry II PP, Wilmé L (2019) Uplisting of Malagasy precious woods critical for their survival. Biol Conserv 235(2019):89-92.

Wegst UGK (2006) Wood for sound. Am J Bot 93(10): 1439-1448.

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: A basic tool of chemometrics. Chemometr Intell Lab 58(2):109-130.

Yan H, Siesler HW (2018) Hand-held near-infrared spectrometers: State-of-the-art instrumentation and practical applications. NIR News 29(7):8-12.

Yin X, Huang A, Zhang S, Liu R, Ma F (2018) Identification of three *Dalbergia* species based on differences in extractive components. Molecules 23(9):1-11.

Zhang M, Liu Y, Yang Z (2015) Correlation of near infrared spectroscopy measurements with the surface roughness of wood. BioResources 10(4):6953-6960.

Zhang M, Zhao G, Guo J, Wiedenhoeft AC, Liu CC, 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(11):975-985.

Zhu C, Fu X, Zhang J, Qin Kai, Wu C (2021) Review of portable near infrared spectrometers: Current status and new techniques. J Near Infrared Spec 0(0):1-16.

DISTINGUISHING NATIVE AND PLANTATION-GROWN MAHOGANY (*SWIETENIA MACROPHYLLA)* TIMBER USING CHROMATOGRAPHY AND HIGH-RESOLUTION QUADRUPOLE TIME-OF-FLIGHT MASS SPECTROMETRY

*Joseph Doh Wook Kim**

Graduate Student
E-mail: joseph.dw.kim@alumni.ubc.ca

Pamela Brunswick

Physical Scientist, Research
E-mail: pamela.brunswick@ec.gc.ca

Dayue Shang

Chemist
Environment and Climate Change Canada
2645 Dollarton Highway
North Vancouver, BC, V7H 1B1, Canada
E-mail: dayue.shang@ec.gc.ca

Philip D. Evans

Professor
Department of Wood Science
University of British Columbia
Vancouver, BC, V6T 1Z4, Canada
E-mail: phil.evans@ubc.ca

(Received March 2023)

Abstract. Plantation-grown mahogany (*Swietenia macrophylla*) from Fiji has been preferred as a sustainable wood source for the crafting of electric guitars because its trade is not restricted by Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), unlike *S. macrophylla* sourced from native forests. Ability to differentiate between the two wood types would deter sale of illegally harvested native-grown *S. macrophylla* to luthiers and other artisans. The chemical composition of wood is influenced by cambial age and geographical factors, and there are chemical differences between *S. macrophylla* grown in different regions. This study tested the ability of high-resolution mass spectrometry to chemotypically differentiate plantation-grown Fijian *S. macrophylla* from the same wood species obtained from native forests. Multiple heartwood specimens of both wood types were extracted and chromatographically profiled using gas and liquid chromatography tandem high-resolution quadrupole time-of-flight mass spectrometry (GC/QToF, LC/QToF). Visual comparison of mass spectral ions, together with modern analytical data-mining techniques, were employed to screen the results. Principal component analysis scatter plots with 95% confidence ellipses showed unambiguous separation of the two wood types by GC/LC/QToF. We conclude that screening of heartwood extractives using high-resolution mass spectrometry offers an effective way of identifying and separating plantation-grown Fijian *S. macrophylla* from wood grown in native forests.

Keywords: Heartwood, identification, plantation, Fiji, mahogany, *Swietenia macrophylla*, chromatography, mass spectrometry.

INTRODUCTION

True mahogany belonging to the genus *Swietenia* consists of three species, *Swietenia macrophylla*, *Swietenia mahagoni*, and *Swietenia*

* Corresponding author

humilis (Helgason et al 1996). Mahogany has been widely exploited to produce a great variety of wood products, including furniture, boats, wooden floors, paneling, and musical instruments (Anderson 2015). It is the wood of choice for electric guitars because of its lightweight, good machinability, resonance, and aesthetic appeal. Some guitars made from mahogany during the golden era (1958-1960) are favored by the world's greatest guitarists and fetch astronomical prices (>$1 million) at auction (Martinez-Reyes 2015). As a result of their overexploitation, *S. humilis* and *S. mahagoni* were listed as "species that would be in danger of extinction without strict regulation" in 1975 and 1995, respectively (Convention on International Trade in Endangered Species of Wild Fauna and Flora [CITES Appendix II]) (Cornelius et al 2004). CITES currently regards both *S. humilis* and *S. mahagoni* as commercially extinct due to their overexploitation (CITES 2021). *S. macrophylla* is still commercially important, but in 2003, it was also given the same status as *S. mahagoni* and *S. humilis* by CITES (Grogan and Barreto 2005). It is possible under CITES regulations to legally export *S. macrophylla* wood with a permit, but there is also ongoing illegal trade in the timber (Kometter et al 2004; Chimeli and Soares 2017). In addition, plantation-grown *S. macrophylla* is traded internationally. In particular, *S. macrophylla* from plantations in Fiji (where it is an introduced species) is widely available on international markets and represents the most important source of *Swietenia* wood imported into the USA (Flexport 2022). Hence, both native-grown *S. macrophylla* wood (legal and illegal) and plantation wood is potentially available to end users. One way of avoiding the possibility of using illegally logged *S. macrophylla* timber would be to simply utilize plantation-grown material. Such an approach would be assisted by the development of an effective way of identifying and separating the two wood types.

The wood identification of *Swietenia* and look-alike species received increased attention after individual species were listed by CITES, and as a result, it is now possible to separate *Swietenia* species from look-alike species. For example, *S. macrophylla* could be separated from *Carapa guianensis*, *Cedrela odorata*, and *Micropholis melinoniana* using near infrared (IR) spectroscopy (Pastore et al 2011). Both *S. macrophylla* and *S. mahagoni* could be separated from other neotropical Meliaceae with convolutional neural network analysis of end-grain images of wood (Ravindran et al 2018). All three *Swietenia* species could be distinguished from "mahogany" look-alike species belonging to the Meliaceae (*Khaya* spp. *Entandrophragma* spp, and *Lovoa trichilioides*) using direct analysis in real time (DART) ionization in combination with time-of-fight mass spectrometry (TOFMS) (Deklerck et al 2019). Furthermore, direct spray ionization injection mass spectrometry was able to distinguish *S. macrophylla* from six other hardwood species and African mahogany (Cabral et al 2012; Fasciotti et al 2015). The identification and separation of individual *Swietenia* species has received less attention and, as pointed out by He et al (2019), there is currently no direct CITES-related need for species-level separation now that all three species have been listed in CITES Appendix II. Nevertheless, identification of the three species is important for cultural and scientific reasons, and recent research has shown that they can be distinguished using machine learning models of quantitative wood anatomy data (He et al 2019) and DART-TOFMS of heartwood samples (Kane 2019). Most recently, *S. macrophylla* and *S. mahagoni* have been discriminated from each other using gas chromatography, TOFMS, and principal component analysis (PCA) (Lamichhane 2022). The same technique was also able to distinguish the two aforementioned *Swietenia* species from three look-alike species (*C. odorata, Khaya ivorensis*, and *Toona ciliata*). DART-TOFMS shows potential for identifying the provenance of individual wood species (Finch et al 2017), but to-date, research on identifying the geographic origin of *Swietenia* has relied on DNA fingerprinting (Degen et al 2013), or near IR spectroscopy (Bergo et al 2016). Both of these techniques were able to verify the geographic origin of *S. macrophylla* wood samples. Differences in the NIR spectra of *S. macrophylla* wood from

different countries reflect variation in the chemical composition of the wood specimens, suggesting that other analytical techniques, such as GC/QToF and LC/QToF, may be able to identify *S. macrophylla* wood samples from trees grown in different regions.

Recently, GC/QToF and LC/QToF in combination with machine learning proved to be very effective in distinguishing rosewoods (*Dalbergia* spp.) that are difficult to identify (Shang et al 2020; Brunswick et al 2021). We apply the same techniques here to identify and distinguish between plantation-grown Fijian *S. macrophylla* and the same wood species obtained from native forests.

MATERIALS AND METHODS

Wood Samples

Twenty different dressed-all-round Fijian mahogany samples were purchased from Paradise Timbers in Helensvale, Queensland, Australia. Each sample was 1800 mm (length) × 31 mm (width) × 11 mm (thickness). A 60-mm long specimen was cut from one end of each sample and shipped to Vancouver, Canada. Twenty-four native-grown *S. macrophylla* samples were retrieved from the University of British Columbia and FPInnovations wood collections (see Supplementary Material 1). The mahogany blocks varied in size and thickness. The average size was about 70 mm (length) × 50 mm (width) × 10 mm (thickness). The smallest were offcuts 10 mm (length) × 50 mm (width) × 10 mm (thickness) in size and the largest was 400 mm (length) × 200 mm (width) × 12 mm (thickness). All samples were conditioned in a constant climate control room at 20 ± 1°C and 65 ± 5% RH (r.h.), and they were only removed from the conditioning room immediately prior to analyses.

Preparation of Wood Samples for Chromatography

Each wood specimen was converted into chips and slivers using small carving tools, which were rinsed with 50% isopropanol among samples. Approximately 55-90 mg of wood chips were weighed on an analytical balance and transferred into 15-mL glass test tubes. A 2-mL aliquot of methanol containing 1% (v/v) formic acid was added to each test tube. All of the test tubes were vortex-mixed for 10 s, stored overnight in a fume hood, vortex-mixed again for 10 s, and then bench-top centrifuged at 5000 rpm for 2 min. The supernatant from each test tube was transferred via glass pipettes to glass vials for storage at −20 ± 5°C.

Chemicals

Acetic acid (LC-MS grade, LiChropur), formic acid (≥98% purity), daidzein, and poly(ethylene glycol) av. Mn 380-420 (PEG400) were supplied by Sigma-Aldrich (Oakville, Ontario, Canada), perfluorotributylamine (PFTBA) was purchased from Agilent Technologies Canada Inc. (Mississauga, Ontario, Canada), LC/MS grade acetonitrile and methanol obtained from Fisher Scientific (Ottawa, Ontario), and HPLC grade 2-propanol from Caledon Laboratories (Georgetown, Ontario). Caffeine-d9 (99%) was sourced from CDN Isotopes (Point-Claire, Quebec, Canada). Aqueous reagents were prepared in ultra-high purity water (MilliQ Plus).

GC/MS and GC/QToF Instrument Parameter and Analysis

GC amenable compounds were analyzed initially using an Agilent 7890A GC interfaced with an Agilent 7000C triple quadruple mass spectrometer (MS). GC separation was performed on a DB-5MS capillary column (Agilent Technologies, 30 m × 0.25 mm ID × 0.25 μm film thickness) column. The GC oven was programmed at 80°C and held for 2 min, heated to 310°C at 6°C/min, and kept at this temperature for 13 min to elute higher boiling point compounds (Table 1). The final selected conditions for GC/QToF analysis employed an Agilent 7890B GC interfaced to an Agilent 7250 QToF MS. GC separation was performed on a DB-5MS capillary column (Agilent Technologies, 30 m × 0.25 mm ID × 0.25 μm film thickness) column. The GC oven was programmed at 50°C and held for 2 min, heated to

310°C at 6°C/min, and kept at this temperature for 8 min to elute higher boiling point compounds (Table 1).

Detection of the compounds in EI + used centroid acquisition. Prior to each analytical sequence, an external mass calibration was performed using PFTBA. Additionally, a system suitability check using caffeine-d9 (100 µg/mL in acetonitrile) was performed to confirm instrument performance. Chromatographic peaks observed in total ion count scan mode (TIC) were reviewed for patterns and characteristic mass ions. For retention time locking, an extract of *Dalbergia latifolia* wood was employed with reference peak 268.073 m/z locked at 29 min. Peaks observed in blank control samples were excluded from the wood identification process.

LC/QToF Instrument Conditions and Analysis

LC/QToF analysis was performed using an Agilent Infinity 1290 LC system interfaced with 6550 iFunnelQToF MS, which was equipped with Jet Stream Technology Ion Source (AJS) controlled by MassHunter software. MS detection employed electrospray negative (ES−) mode ionization using the conditions listed in Table 2. The instrument used a fly mass correction function in the acquisition software algorithm. The time-of-flight lock mass ions for ES− at 119.03632 m/z (purine) and 980.016375 m/z (Agilent HP0 921 acetate adduct, sourced from the Agilent APITOF Reference Mass Solution) were employed for calibration. Analyte separation was performed by the reverse phase chromatographic mode with parameters listed in Table 3. System suitability (*Dalbergia latifolia*) confirmed that the signal-to-noise detection for the 267.066 m/z peak at approximately 12 min was greater than 10.

Statistical Analyses of GC/QToF and LC/QToF Data

Initially, visually observable chromatographic peaks in low-resolution GC/MS were reviewed by extraction and integration of peak areas. A canonical variate analysis (CVA) was used to formally examine this data in a multivariate way to determine differences between plantation-grown Fijian *S. macrophylla* and *S. macrophylla* wood grown in native forests. Statistical computation for CVA was performed using Genstat (v. 19). Further analysis of wood extracts was performed by LC and GC/QToF, with the collected data imported into Agilent's Unknowns Analysis software, deconvoluted using the SureMass

Table 1. Instrument parameters for operation of GC/QToF in EI + mode.

Component/Parameter	Details/Value
Column	Agilent DB-5MS + 10 m DG, 30 m, 0.25 mm id, 0.25 µm film
Injection	0.5 µL pulsed splitless
Purge flow to split vent	50 mL/min for 1 min
Initial temp.	250°C
Oven temp. program GC/MS	80°C hold for 2.0 min, 6°C/min to 310°C, hold for 13 min
—	total run time approximately 53 min
Oven temp. program GC/QToF	70°C hold for 2.0 min, 6°C/min to 310°C, hold for 12 min
—	total run time approximately 42 min
Carrier gas	1 mL/min helium
Interface temp.	280°C
EI+ (Electron ionization)	70 EV at 250°C
Quad temp.	150°C
TOF tune	1 GHz extended dynamic range
Scan	60-1000 m/z (GCMS) and 50-900 m/z (GC/QTOF)
Spectra acquisition	1 Hz in centroid
Scan rate	2.0 spectra/s
Microchannel plate (MCP)	750 V (variable to tune)
Photomultiplier tube (PMT)	500 V (variable to tune)

Table 2. Instrument parameters for LC/QToF analysis.

Component/Parameter	Details/Value
Ionization mode	ES−
Scan	90-1100 m/z
Data collect	1.5-32.0 min
Nebulizer	30 psig
VCap	3500
Nozzle V	1350
Fragmentor	365
Gas temp./flow	230°C (12 L/min)
Sheath gas/flow	350°C (11 L/min)
LC column	Agilent Poroshell 120 SB-C18 (2.7 μm, 2.1 × 100 mm)
LC flow	0.4 mL/min
LC post time	0.5 min
TOF lock mass (purine)	ES+ 121.0509 m/z; ES− 119.03632 m/z
TOF lock mass (HP0921)	ES+ 922.0098 m/z; ES− 980.016375 m/z

Algorithm, and converted to common event format (CEF) files for import into Mass Profiler Professional (MPP). The unbiased nature of the machine learning process of MPP was used to perform preliminary alignment frequency filtering of discriminating entities and ANOVA ($p > 0.05$) of data. PCA scatter plots were created to help visualize the output of the algorithms.

Total Extractive Contents and Light Microscopy

The preparation of wood and extraction of solvent soluble extractives were carried out according to TAPPI standards T 264 om-88 and T 204 om-88, respectively (Tappi 1993). Eight mahogany samples (four Fijian *S. macrophylla* samples and four native-grown *S. macrophylla samples*) were randomly selected to determine their total extractive contents. Wood samples were selected and split to create thin rod-like pieces of wood measuring approximately 2 mm^2 × 15 mm. Pieces of wood from each sample, approximately 3 g, were separately ground in a Wiley-mill to pass a 40 mesh sieve (0.4 mm). The Wiley-mill was thoroughly cleaned among samples. Wood flour, approximately 2 g from each sample, was placed in separate cellulose thimbles and Soxhlet-extracted with 1:2 ratio of ethanol toluene for not less than 24 extraction cycles over a 5-h period. The flasks were removed from the apparatus and the solvent allowed to partially evaporate in the extraction flask to a volume of 20 mL. The extract was transferred to a tared weighing dish with minimal amount of fresh solvent used. The dish was oven-dried for 1 h at 115 ± 5°C, cooled in a desiccator, and then weighed to the nearest 0.1 mg. The extractive content was calculated as follows: Total extractives, % = ([oven-dry weight of extract − oven-dry weight of blank residue]/oven-dry weight of wood) × 100.

Table 3. Conditions for the reverse phase LC/QToF separation.

Time (min)	% A: 0.1% v/v acetic acid in ultrapure water	%B: 0.1% v/v acetic acid, 2% v/v 2-propanol in acetonitrile
0.5	95.0	10.0
14.0	50.0	50.0
17.0	1.0	99.0
20.0	1.0	99.0
20.1	95.0	10.0
23.0	95.0	10.0

Wood blocks measuring 20 mm (length) × 8 ± 2 mm (width) × 8 ± 2 mm (thickness) and cut from plantation and native-grown *S. macrophylla* wood were placed in separate glass beakers containing ultrapure distilled water for 3 d. Water-saturated wood blocks were placed one-at-a-time in the microtome clamp of a sliding sledge microtome (Spencer Lens Co) and transverse sections, approximately 20 μm thick were cut from blocks using a microtome blade-holder (Feather® No. 160) containing a fresh disposable microtome blade (Feather® Type S35). Microtome sections were immersed in a 10% ethanolic solution of 1% safranin for approximately 5 min and dehydrated in 100% ethanol for 15 min. Sections were permanently mounted on glass slides using synthetic resin media (DPX Mountant, Aldrich) and each section was covered with a glass cover-slip. Slides were examined with a light microscope (Carl Zeiss Universal Fluorescent Microscope) equipped with a digital camera (Olympus DP71). Images of transverse surfaces were saved as TIFF files (Christy et al 2005).

RESULTS

Total Extractive Contents and Light Microscopy

The reference wood specimens of native-grown (sample #1, #13, #15, and #18 in supplementary material) and Fijian plantation *S. macrophylla* looked similar, but not identical under the microscope. Figure 1 shows transverse sections of both wood types. One difference between the two woods is that the lumens of vessels in plantation-grown *S. macrophylla* were largely free of resin-like deposits, whereas approximately 10% of vessels in native-grown wood contained deposits (Fig 1). However, there was no significant ($p > 0.05$) difference in extractive contents of the two wood types: the extractive content of plantation-grown Fijian *S. macrophylla* was 4.7% (6.9-1.6; 0.026; max-min, St Dev), whereas that of wood from native-grown trees was 5.5% (8.6-2.9; 0.024).

GC/MS

Preliminary analysis of heartwood extractives from the two wood types analyzed by GC/MS (EI+) showed close similarity in their TIC chromatography. The scans were reviewed visually and by instrument screening software for potential differences between the two wood groups. Selected peaks were extracted, integrated, and examined. The extracted ions that best discriminated between native-grown and Fijian *S. macrophylla* eluted at retention times 14.1, 19.2, and 22.2 min, corresponding to mass ions at 152.2, 182.3, and 181.2 m/z, respectively. Peak response areas for these ions from each wood sample were subject to CVA after log (ln) transformation using the statistical software package Genstat (Campbell and Atchley 1981). CVA has been used for identification and separation of two anatomically similar eucalyptus species (Evans et al 2008) and canonical correlation analysis has been used to model variation of wood properties of tepozán (*Buddleja cordata*) across its natural range in Mexico (Aguilar-Rodríguez et al 2006). Figure 2 plots the CVA scores for native-grown and Fijian *S. macrophylla.* A negative value for the CVA score indicates that the specimen is Fijian plantation-grown *S. macrophylla*, whereas a positive value indicates that the specimen is native-grown *S. macrophylla*. The larger the absolute values, the more certain the classification.

GC/QToF

Further analysis of wood extracts was performed by high-resolution GC/QToF (EI+). Overall, chromatographic results were not visually distinctive and did not easily separate native-grown and Fijian plantation-grown *S. macrophylla*. Peak responses varied within each study group and no

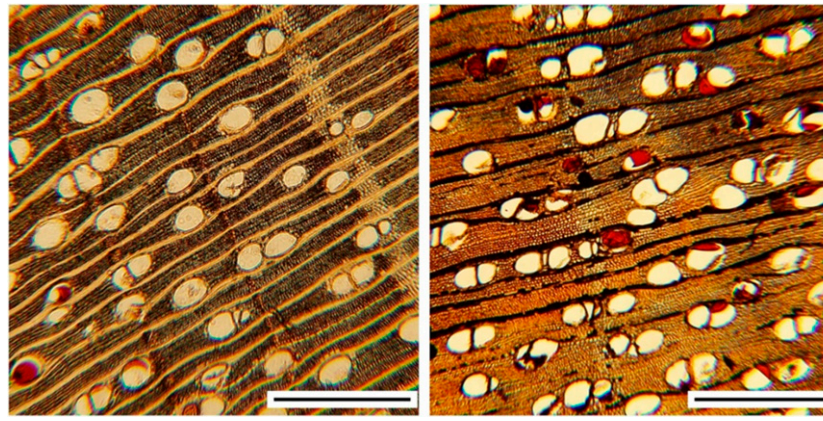


Figure 1. Transverse sections of plantation-grown Fijian mahogany (left) and native-grown *S. macrophylla* (right). Scale bars = 1 mm.

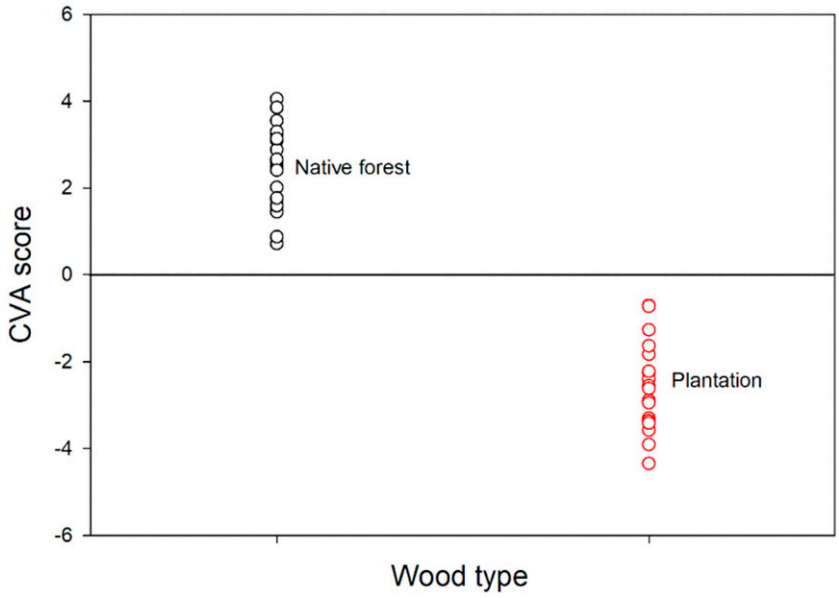


Figure 2. Canonical variate analysis (CVA) scores for native-grown and Fijian plantation-grown *S. macrophylla* analyzed by GC/MS.

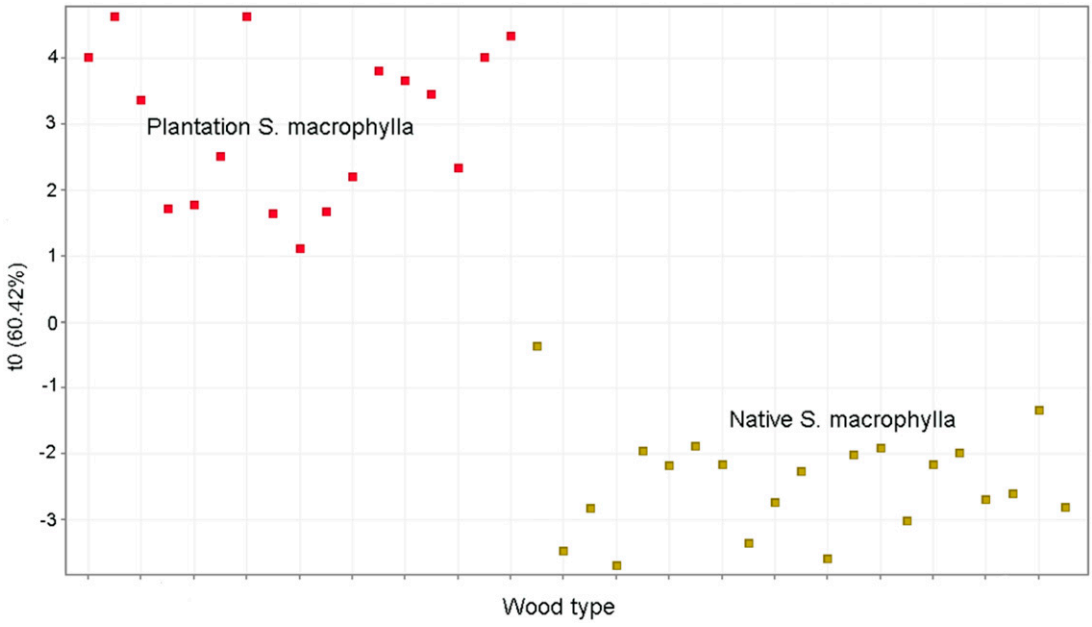


Figure 3. Three-dimensional principal component analysis (PCA) 2D scatter plots of EI + GC/QToF data for plantation and native-grown *S. macrophylla.*

signature peak or ion was found to be definitive for identification purposes. However, the original mass ions observed by low-resolution GC/MS were supported by further observation under different conditions during GC/QToF analysis, at the adjusted retention times of 9.5, 13.1, and 15.5 min, with corresponding mass ions at 151.039, 182.057, and 181.049 m/z. Statistical analysis of these peaks again showed a similar distinction between the Fijian plantation-grown *S. macrophylla* and all native-grown specimens.

Application of instrument machine learning software was able to exploit the collected data to select discriminate entities in an unbiased manner. The identification of the source of the specimens involved deconvolution of complex chromatographic data in Agilent Unknowns analysis, followed by frequency filtering and statistical analysis in MPP. Results showed discrimination between the two wood types.

A partial least squares (PLS) scores-plot of data-mined GC/QToF results showed a distinction between native-grown *S. macrophylla* samples from samples grown in plantation in Fiji (Fig 3). A two-dimensional PCA plot from GC/QToF data clearly demonstrated greater variability within the native-grown *S. macrophylla* wood compared with plantation wood (Fig 4). However, 95% confidence ellipses associated with each wood type, indicate clear separation of native-grown *S. macrophylla* samples and those grown in plantation in Fiji, which show little within-group variation.

LC/QToF

Heartwood extracts from native- and plantation-grown *S. macrophylla* were further analyzed by LC/QToF in electrospray negative mode (ES−). Both two-dimensional (Fig 5) and three-dimensional (not shown) PCA plots distinguished native-grown *S. macrophylla* wood from Fijian

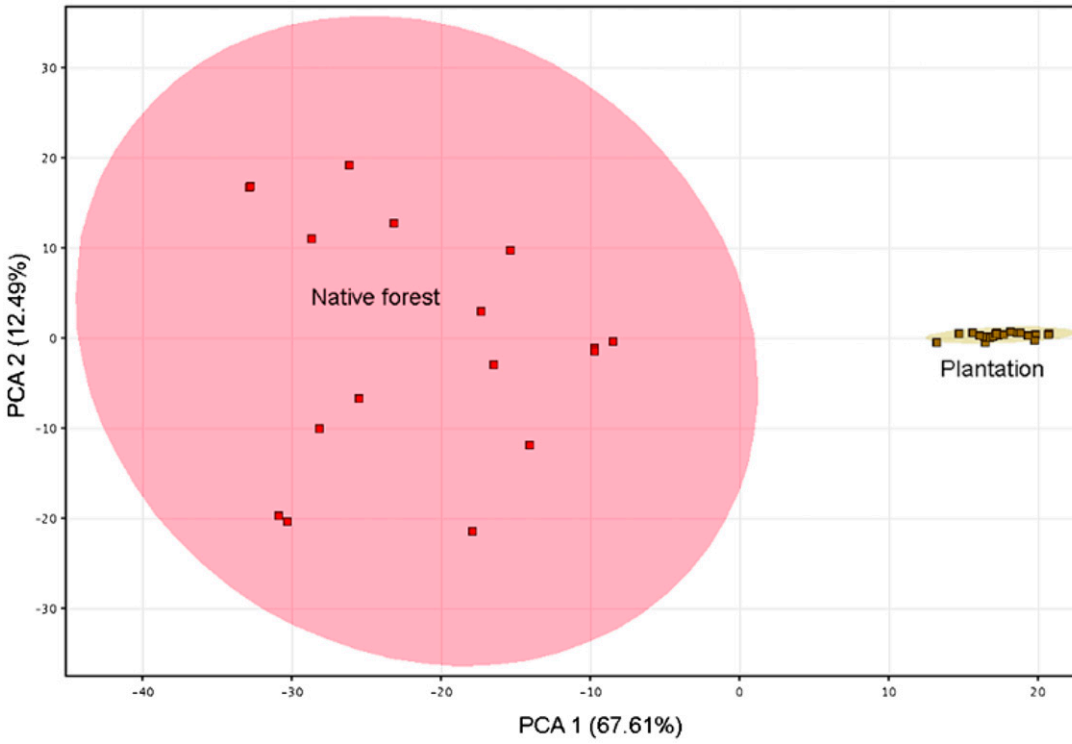


Figure 4. PCA 2D scatter plots of EI + GC/QToF data showing 95% confidence limits for data from plantation and native-grown S. *macrophylla*.

plantation-grown *S. macrophylla* wood. Again, greater variability within the native wood type was observed.

An unvalidated PLS discriminant analysis (PLS-DA) model also distinguished native-grown from Fijian plantation *S. macrophylla* using the set of reference samples (Fig 6).

DISCUSSION

The aim of this research was to determine whether gas/liquid chromatography combined with high-resolution mass spectrometry could offer a robust approach to distinguishing *S. macrophylla* grown in plantation in Fiji from wood sourced from native forests. Chromatography has long been used to help identify commercially important wood species. Initial research focused on identifying woods that were difficult to separate using anatomical features, eg *Shorea* spp. (Samapuddhi 1957; Fanega and Mule 1973; Kato and Hishiyama 2007), *Pinus* spp. (Seikel et al 1965; Swan 1966), *Acer* spp. (Park and Kim 1984), and three *Khaya* species, one of which caused dermatitis in furniture factories in the United Kingdom (Morgan and Orsler 1967). Application of gas chromatography and mass spectrometry initially focused on identifying wood with commercially important characteristics related to heartwood extractives, eg wood color (Swan 1966), aroma (Nonier et al 2006), and the flavor of woods used as ice-cream sticks (Sudarat and Harper 2004). Some of the early combined GC/MS studies demonstrated the potential of the technique to identify different wood genera (Challinor 1995; Nonier et al 2006) and such studies were the antecedents to the more advanced techniques used to identify endangered woods listed by CITES, including *Swietenia* (Lamichhane 2022), *Dalbergia* (Hou et al 2018; Shang et al 2020; Brunswick et al 2021), and *Pterocarpus* (Zhang et al 2019).

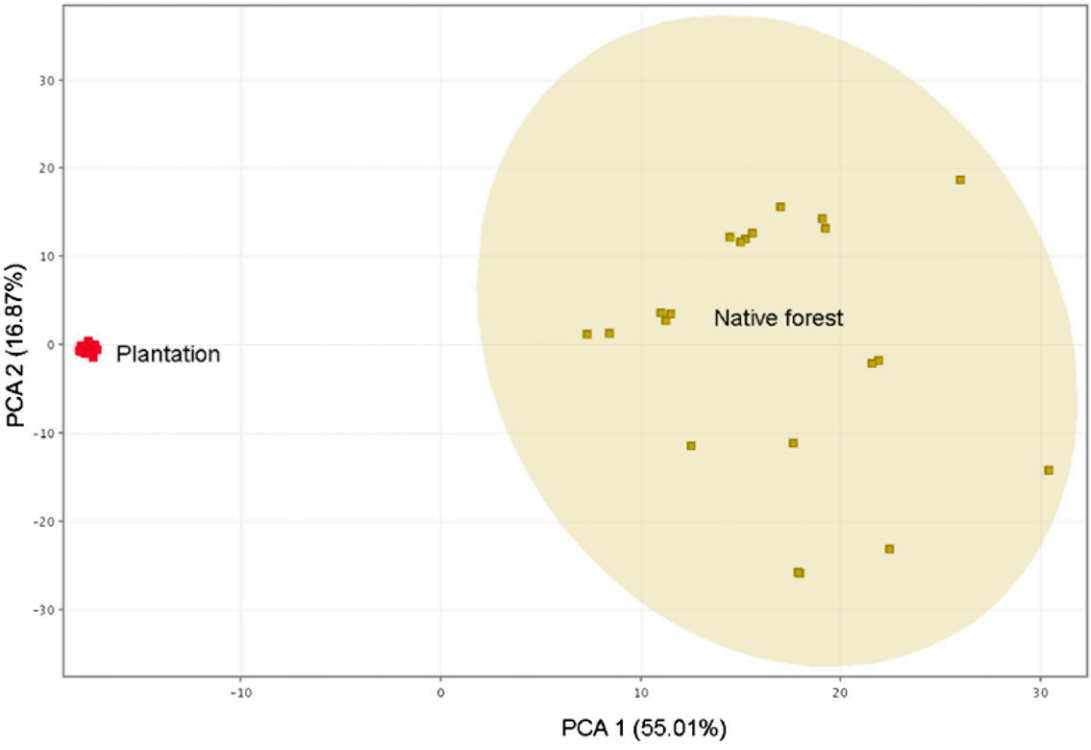


Figure 5. PCA 2D scatter plots of ES− LC/QToF data showing 95% confidence limits for wood samples from plantation and native-grown *S. macrophylla*.

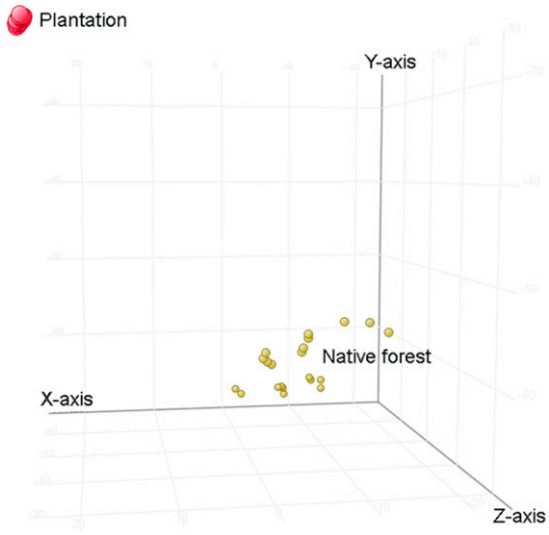


Figure 6. Partial least squares discrimination model of plantation and native-grown *S. macrophylla* using LC/QToF ES− data.

LC/MS is a more recent branch of the field and has been used to identify wood to the genus and species levels and to determine the geographical origin of Spanish cedar (*Cedrela odorata*) and Norway spruce (*Picea abies*) (Surowiec et al 2004; Kite et al 2010; Buche et al 2021; Creydt et al 2021, 2022). Our finding that LC/QToF and multivariate data analysis can identify and separate native *S. macrophylla* and plantation-grown Fijian *S. macrophylla* accords with previous research by Creydt et al (2021, 2022) on the geographical separation of Spanish cedar and Norway spruce. In addition, we show that similar separation of native and plantation-grown Fijian mahogany can also be achieved using GC/QToF. Overall, it was encouraging to find that the two very different approaches, one (LC) screening for the more hydrophilic, negative ion-forming compounds, and the other (GC) screening for the more volatile, hydrophobic, positive ion-forming compounds, were both able to distinguish the two *Swietenia* wood types from different sources. This finding should encourage further research on the use of this approach, ie chromatographic separation followed by mass spectrometry, to identify the geographical origin of wood species.

The listing of *S. macrophylla* by CITES and associated trade regulations led to a flourishing illegal trade in *S. macrophylla* wood exported under the category "other tropical timber species" (Chimeli and Soares 2017). This illegal trade coexists with a legal "certified" trade in *S. macrophylla* and also trade in plantation-grown *S. macrophylla*. In addition, small artisanal businesses making guitars may have stockpiles, albeit dwindling, of *S. macrophylla* wood accumulated prior to current trade restrictions (Martinez-Reyes 2015). These businesses, lacking the knowledge and control over supply chains of larger manufacturers, are afraid to use mahogany because they fear confiscations or fines (Dudley 2011). Our finding that GC/LC/QTOF can separate native and plantation-grown Fiji mahogany provides a potential solution to this issue because the machine learning software, which was very helpful in separating plantation-grown Fijian *S. macrophylla* from native-grown wood, is compatible with both low- and high-resolution instrumentation. In fact, single quadruple GC/MS instrumentation is a staple of analytical laboratories worldwide and the cost of GC/MS analysis is as little as $60 per sample, a faction of the cost of a custom-made mahogany electric guitar ($5000-$10,000) or the fines levied on companies using illegally harvested mahogany (Martinez-Reyes 2015). Further work to build a mahogany specific spectral library, validate the method with unknown testing samples, and transfer of technology to environmental labs is needed to develop such an analytical service, which would help the guitar industry and other end users of mahogany to avoid issues associated with the use of mahogany wood that is not sustainably harvested.

CONCLUSIONS

We conclude that screening of heartwood extractives by gas or liquid chromatography with high-resolution mass spectrometry, in combination with multivariate statistical analysis of discriminating ions, offers an effective way of separating plantation-grown Fijian *S. macrophylla* from the same wood species grown in native forests. Our findings offer a potential solution to environmentally conscious manufacturers of guitars who wish

to use plantation-grown Fijian mahogany rather than being exposed to the liability associated with using illegally logged native-grown mahogany. Further research is needed to extend our work to plantation-grown mahogany from other parts of the world, and to examine the feasibility and costs of developing an analytical service and associated databases to identify and separate plantation-grown mahogany from native-grown timber.

ACKNOWLEDGMENT

We thank Drs. Daniel Cuthbertson and Marcus Kim of Agilent Technologies Inc., Jeffrey Yan of ECCC, Dr. Chip (Robert) Cody of JEOL Inc., and Bill Leggate of DAF for material support or technical assistance. P.D.E. thanks Viance LLC, Tolko, FPInnovations, Faculty of Forestry (University of British Columbia), and the Government of British Columbia for supporting his BC Leadership Chair in Advanced Forest Products Manufacturing Technology.

REFERENCES

Aguilar-Rodríguez S, Terrazas T, Lopez-Mata L (2006) Anatomical wood variation of *Buddleja cordata* (Buddlejaceae) along its natural range in Mexico. Trees (Berl) 20(2):253-261.

Anderson JL (2015) Mahogany: The costs of luxury in early America. Harvard University Press, Cambridge, MA.

Bergo MC, Pastore TCM, Coradin VTR, Wiedenhoeft AC, Braga JWB (2016) NIRS identification of *Swietenia macrophylla* is robust across specimens from 27 countries. IAWA J 37(3):420-430.

Brunswick P, Cuthbertson D, Yan J, Chua CC, Duchesne I, Isabel N, Evans PD, Gasson P, Kite G, Bruno J, van Aggelen G, Shang D (2021) A practical study of CITES wood species identification by untargeted DART/QTOF, GC/QTOF and LC/QTOF together with machine learning processes and statistical analysis. Environ Adv 5:100089.

Buche G, Colas C, Fougère L, Giordanengo T, Destandau E (2021) Untargeted UHPLC-Q-TOF-HRMS based determination of discriminating compounds for oak species *Quercus robur* L. and *Quercus petraea* Liebl. identification. Phytochem Anal 32(5):660-671.

Cabral EC, Simas RC, Santos VG, Queiroga CL, da Cunha VS, de Sá GF, Daroda RJ, Eberlin MN (2012) Wood typification by Venturi easy ambient sonic spray ionization mass spectrometry: The case of the endangered mahogany tree. J Mass Spectrom 47(1):1-6.

Campbell NA, Atchley WR (1981) The geometry of canonical variate analysis. Syst Biol 30(3):268-280.

Challinor JM (1995) Characterisation of wood by pyrolysis derivatisation—Gas chromatography/mass spectrometry. J Anal Appl Pyrolysis 35(1):93-107.

Chimeli AB, Soares RR (2017) The use of violence in illegal markets: Evidence from mahogany trade in the Brazilian Amazon. Am Econ J Appl Econ 9(4):30-57.

Christy AG, Senden TJ, Evans PD (2005) Automated measurement of checks at wood surfaces. Measurement 37(2):109-118.

CITES (2021) CITES trade controls to take effect for mahogany. https://cites.org/eng/news/pr/2003/031111_mahogany.shtml (22 May 2022).

Cornelius JP, Wightman KE, Grogan JE, Ward SE (2004) Tropical Ecosystems. *Swietenia* (American Mahogany). Encyclopedia of Forest Sciences, Elsevier. 1720-1726 pp.

Creydt M, Lautner S, Fromm J, 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. J Chromatogr A 1663:462737.

Creydt M, Ludwig L, Köhl M, Fromm J, 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. J Chromatogr A 1641:461993.

Degen B, Ward SE, Lemes MR, Navarro C, Cavers S, Sebbenn AM (2013) Verifying the geographic origin of mahogany (*Swietenia macrophylla* King) with DNA-fingerprints. Forensic Sci Int Genet 7(1):55-62.

Deklerck V, Mortier T, Cody R, Goeders N (2019) A protocol for automated timber species identification using metabolome profiling. Wood Sci Technol 53(4):953-965.

Dudley KM (2011) Luthiers: The last endangered species. *The New York Times* op-ed, 25 October. www.nytimes.com/2011/10/26/opinion/are-guitar-makers-an-endangered-species.html (12 July 2022).

Evans P, Heady R, Cunningham R (2008) Identification of yellow stringybark (*Eucalyptus muelleriana*) and silvertop ash (*E. sieberi*) wood is improved by canonical variate analysis of ray anatomy. Aust For 71(2):94-99.

Fanega SM, Mule EI (1973) Chemotaxonomic contribution to the question of the identity of Tiaong, *Shorea agsaboensis* Stern (Dipterocarpaceae). Philippine Lumberman 19(7):14, 16, 18.

Fasciotti M, Alberici RM, Cabral EC, Cunha VS, Silva PR, Daroda RJ, Eberlin MN (2015) Wood chemotaxonomy via ESI-MS profiles of phytochemical markers: The challenging case of African versus Brazilian mahogany woods. Anal Methods-UK 7(20):8576-8583.

Finch K, Espinoza E, Jones FA, Cronn R (2017) Source identification of western Oregon Douglas-fir wood cores using mass spectrometry and random forest classification. Appl Plant Sci 5(5):1600158.

Flexport (2022) U.S. import and export data. Mahogany (*Swietenia* spp.) HS Code 440721. https://www.flexport.com/data/hs-code/440721-mahogany-sawn-chipped-lengthwise-sliced-peeled (22 May 2022).

Grogan J, Barreto P (2005) Big-leaf mahogany on CITES Appendix II: Big challenge, big opportunity. Conserv Biol 19(3):973-976.

He T, Marco J, Soares R, Yin Y, Wiedenhoeft AC (2019) Machine learning models with quantitative wood anatomy data can discriminate between *Swietenia macrophylla* and *Swietenia mahagoni*. Forests 11(1):36.

Helgason T, Russell SJ, Monro AK, Vogel JC (1996) What is mahogany? The importance of a taxonomic framework for conservation. Bot J Linn Soc 122(1):47-59.

Hou S, Huang X, Ma Y, Huang F, Lin X, Ma Y, Zhu Z, Luo H, Wu H (2018) Identification of *Dalbergia odorifera* and *Dalbergia* spp by solid phase microextraction-gas chromatography-mass spectrometry. Chinese J Chromat 36(5):493-498.

Kane PE (2019) Mahogany: New research on the wood of choice in early Rhode Island furniture. Yale University Art Gallery Bulletin. 69-77 pp.

Kato A, Hishiyama S (2007) Identification of red meranti based on the heartwood extractives. Pages 20-22 *in* International Symposium on Development of Improved Methods to identify Shorea species wood and its origin, IUFRO, Tokyo (14 September 2007).

Kite GC, Green PWC, Veitch NC, Groves MC, Gasson PE, Simmonds MSJ (2010) Dalnigrin, a neoflavonoid marker for the identification of Brazilian rosewood (*Dalbergia nigra*) in CITES enforcement. Phytochemistry 71(10):1122-1131.

Kometter RF, Martinez M, Blundell AG, Gullison RE, Steininger MK, Rice RE (2004) Impacts of unsustainable mahogany logging in Bolivia and Peru. Ecol Soc 9(1):1-20.

Lamichhane DD (2022) Identification of five wood species of the Meliaceae family on the basis of their wood anatomical and chemical characteristics. M.Sc. Thesis, Université Laval. Quebec City, Quebec, Canada. 154 pp.

Martinez-Reyes J (2015) Mahogany intertwined: Enviromateriality between Mexico, Fiji, and the Gibson Les Paul. J Mater Cult 20(3):313-329.

Morgan JW, Orsler RJ (1967) A simple test to distinguish *Khaya anthotheca* from *K ivorensis* and *K grandifoliola*. J Inst Wood Sci 1(18):61.

Nonier MF, Vivas N, Vivas de Gaulejac N, Absalon C, Soulié P, Fouquet E (2006) Pyrolysis—Gas chromatography/mass spectrometry of *Quercus* sp. wood: Application to structural elucidation of macromolecules and aromatic profiles of different species. J Anal Appl Pyrolysis 75(2):181-193.

Park KW, Kim SS (1984) A study on the wood identification of the genus Acer in Korea—Especially on the method by thin layer chromatography of lipid in heartwood. J Korean Soc For Sci 65(1):60-67.

Pastore TCM, Braga JWB, Coradin VTR, Magalhães WLE, Okino EYA, Camargos JAA, de Muñiz GIB, Bressan OA, Davrieux F (2011) Near infrared spectroscopy (NIRS) as a potential tool for monitoring trade of similar woods: Discrimination of true mahogany, cedar, andiroba, and curupixá. Holzforschung 65(1):73-80.

Ravindran P, Costa A, Soares R, Wiedenhoeft AC (2018) Classification of CITES-listed and other neotropical Meliaceae wood images using convolutional neural networks. Plant Methods 14(1):1-10.

Samapuddhi K (1957) A note on preliminary studies in some methods of identifying the timbers of *Pentacme siamensis*, Kurz., *Shorea obtusa* Wall., and *Shorea robusta*, Gaertn. f. Publication Royal Forest Department, Ministry of Agriculture, Bangkok.

Seikel MK, Hall SS, Feldman LC, Koeppen RC (1965) Chemotaxonomy as an aid in differentiating wood of eastern and western white pine. Am J Bot 52(10):1046-1049.

Shang D, Brunswick P, Yan J, Bruno J, Duchesne I, Isabel N, van Aggelen G, Kim M, Evans PD (2020) Chemotyping and Identification of protected *Dalbergia* timber using gas chromatography quadrupole time of flight mass spectrometry. J Chromatogr A 1615:460775.

Sudarat J, Harper WJ (2004) Identification of volatile flavor compounds in wooden ice cream sticks originated from different geographical locations. Milchwissenschaft 59(9-10):530-532.

Surowiec I, Nowik W, Trojanowicz M (2004) Identification of "insoluble" red dyewoods by high performance liquid chromatography–photodiode array detection (HPLC-PDA) fingerprinting. J Sep Sci 27(3):209-216.

Swan EP (1966) Chemical methods of differentiating the wood of several western conifers. Forest Prod J 16(1): 51-54.

Technical Association for the Pulp and Paper Industry (TAPPI) (1993) TAPPI standards and suggested methods, TAPPI T264OM-88; TAPPI T204 OM-88. Tech Assoc Pulp Pap Ind, New York, NY. 754 pp.

Zhang M, Zhao G, Guo J, Liu B, Jiang X, Yin Y (2019) A GC-MS protocol for separating endangered and non-endangered *Pterocarpus* wood species. Molecules 24(4): 799.

SUPPLEMENTARY MATERIAL

Supplementary Material 1. Dimensions and origins of native-grown *Swietenia macrophylla.*

Sample	UBC ID #	Dimensions (mm)	Origin
1	7522	59.5 × 14.5 × 2.1	Honduras
2	7698	54 × 19 × 1.3	Honduras
3	7711	54 × 19 × 1.3	Honduras
4	928	14.5 × 10 × 2.5	Honduras
5	2110	14 × 8 × 1	Honduras
6	1505	12 × 6.5 × 1.3	Honduras
7	494	8 × 3 × 1	Honduras
8	2111	10.5 × 7.5 × 1	Honduras
9	2001	10 × 7.5 × 2.5	Honduras
10	2988	14.5 × 7 × 1	Honduras
11	3027	10 × 6 × 1.3	Honduras
12	3346	9.5 × 5 × 1.5	Honduras
13	8583	10 × 6 × 2.4	Honduras
14	8584	10 × 6 × 1.5	Honduras
15	8841	13.7 × 6 × 4	Central America
16	9005	14 × 5 × 1	Honduras
17	9281	10.5 × 7.5 × 1	Honduras
18	9676	10.5 × 7 × 1	Honduras
19	10,046	5.8 × 5 × 1	Honduras
20	7826	20 × 12.5 × 0.5	Central America
21	7831	20.5 × 15 × 0.5	Central America
22	FPInnovations 1	11 × 7 × 1	Central America
23	FPInnovations 2	11 × 6.5 × 1	Mexico
24	FPInnovations 3	14 × 6.5 × 1	Mexico

THE GLOBAL WOOD SPECIES PRIORITY LIST: A LIVING DATABASE OF TREE SPECIES MOST AT RISK FOR ILLEGAL LOGGING, UNSUSTAINABLE DEFORESTATION, AND HIGH RATES OF TRADE GLOBALLY

S. B. Richardson

World Forest ID
Thomas Circle NW, Suite 700, Washington DC 20005, USA
E-mail: sarah.richardson@worldforestid.org

J. C. Simeone

Simeone Consulting, LLC
Littleton, NH
E-mail: simeoneconsulting@gmail.com

*V. Deklerck**

Royal Botanic Gardens Kew
Richmond, Surrey
E-mail: v.deklerck@kew.org

(Received July 2022)

Abstract. The illegal timber trade is one of the most impactful natural wildlife crimes, affecting the livelihood of local communities, natural resource availability, and the associated carbon storage and biodiversity. Many timber species are highly sought after and are at risk of exhaustion and subsequent extinction. Although several initiatives exist to indicate tree species risk and conservation status, there is no single resource, or prioritized list, that qualifies the most high-risk and highly traded species across the globe. Organizations end up creating their own priority species lists to meet this lack of aggregated information, requiring hours of independent research and resulting in the recreation of similar lists. To provide a one-stop-shop for similar initiatives, World Forest ID developed the Global Priority Wood Species List (GPWSL) to synthesize existing information. Currently, the GPWSL harbors 270 species most at risk for illegal logging, unsustainable deforestation, and high rates of international trade. The database contains relevant information on each species; such as natural distribution, conservation listings, and countries of import. Here, we present the list, the methods used in its development, and its potential applications for the wood industry as a whole.

Keywords: Illegal timber trade, deforestation, conservation, tree species priority, illegal logging.

INTRODUCTION

Forests cover 31% of the Earth's land surface (4.06 billion hectares) (FAO and UNEP 2020), with approximately half of the world's forests being at high-risk for deforestation or degradation by 2030 (WWF 2022). Since 1990, 420 million hectares of forests have been lost worldwide, which equates to around 10% of the world's remaining forest coverage (FAO 2020). From 2015 to 2020, global deforestation averaged 10 million hectares each year (FAO and UNEP 2020). The conversion of forests to other land use forms (whether human-induced or not) has caused 420 million hectares of forest to be lost by deforestation since 1990 (FAO and UNEP 2020). The degradation of forests leads to a decline in ecological function and ecosystem services provided to humans and the planet. Society both benefits from and is highly dependent on forest ecosystems, both in monetary and nonmonetary terms. It is estimated that over half of the world's Gross

* Corresponding author

Domestic Product (GDP) depends upon ecosystem services, with the forestry sector specifically contributing more than 1.52 trillion USD (directly, indirectly, and induced, in 2015) to the world's GDP and employing 33 million people globally (FAO 2022). Additionally, forests are critical for mitigating climate change, both directly through storing an estimated 662 billion tons of carbon, which is more than half the global carbon stock in soils and vegetation, and indirectly through their contribution to ecosystem processes (FAO 2022).

Following the 26th meeting of the UN Climate Change Conference of the Parties (COP26), also known as the "Glasgow Climate Pact," there is an increased focus on the critical role forests play in modulating the Earth's climate. Wood makes up the largest part of a forest's biomass, and the global estimate of carbon stored in these lignified tissues is upwards of 400 petagrams (Chave et al 2009; Beeckman 2016). At COP26, forests were recognized for this crucial role as global carbon sinks. Deforestation, however, can convert forests from carbon sinks to carbon sources as the carbon once stored in tree biomass and forest soil enters the atmosphere (Gatti et al 2021). Deforestation across the globe contributes 12-15% of worldwide greenhouse emissions, according to 2017 estimates (May 2017). Several commitments to forest conservation were made at COP26, such as the "Declaration on Forests and Land Use" calling for the halting and reversal of forest loss by 2030 (GOV.UK 2021). Additionally, the "Forest, Agriculture, and Commodity Trade Statement" was designed to deliver sustainable trade and reduce pressure on forests (GOV.UK 2021). This includes climate-conscious supply chain action by the largest companies trading in forest-risk commodities (GOV.UK 2021).

Despite these COP26 commitments, unsustainable deforestation and illegal logging remain global issues that threaten the Earth's climate and people. Forest legality remains one of the greatest and most complex barriers to conserving the world's forests. Understanding and mapping trade flows of timber and wood-based products is difficult due to the complexity of determining the legality of a harvested tree. This determination can depend on a multitude of factors, such as species identification, export quotas, concession boundaries, and legal ownership of land. On top of this, once harvested, legal and illegal trees are often mixed, transported to a processing plant in a different country, and exported as forest products to yet another country. Additionally, trade documents can be falsified and products can be mislabeled to avoid legal repercussions. Such issues do not occur at an insignificant scale, as the annual trade value of illegal forest products is estimated at 52-157 billion USD (May 2017). Thus, the capacity to verify harvest origin and species is essential given the length and complexity of the supply chain.

It remains a challenge to verify the sustainability of internationally harvested forest products. International regulatory policies, such as the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), the United States Lacey Act, and European Union Timber Regulation (EUTR) require targeted efforts to ensure the legality of timber supply chains. Consequently, entities involved in the global timber trade and wood industry need to understand which species and supply chains are threatened by illegality and unsustainable deforestation. A wood products' species identity and harvest origin can be determined using a variety of scientific methods. These techniques include wood anatomical analysis (manual and machine vision, see Gasson [2011] and Hermanson et al [2019]), Direct Analysis in Real Time—Time-of-Flight Mass Spectrometry (DART-TOFMS, see Deklerck & Price et al [2021]), DNA analysis (barcoding, fingerprinting, etc., see Jiao et al [2020]), and Stable Isotope Ratio Analysis (SIRA, see Watkinson et al [2022]). Each of these techniques, however, require reference data against which a product can be matched. World Forest ID (worldforestid.org) is a nonprofit organization building the world's largest georeferenced library of tree samples. World Forest ID aims to independently enable the identification of a forest product's species and harvest origin (Gasson et al 2021). As this reference library develops, there is a need to prioritize the species that are

most at risk of entering the supply chain illegally and face high rates of deforestation.

A key question, thus, for those in the forest legality field is, "which timber species should be prioritized for protection and close monitoring?" Existing publications and databases outline a general risk status (IUCN Red List of Threatened Species), high-risk genera in specific geographic regions (WWF 2015), and comprehensively list every timber species traded internationally (Mark et al 2014). However, there is no single resource quantifying the most high-risk and highly traded species across the globe. High-risk and highly traded species are those that require the most protection to avoid forest exhaustion and species extinction. Although the IUCN Red List is a crucial resource for quantifying the world's most threatened wildlife species, some IUCN tree assessments can be outdated. Thus, it is possible that some tree species are in a worse state than their IUCN-assessed threat level. Additionally, the IUCN Red List only takes natural extinction risk into account, factoring out plantation efforts.

To meet the need for an aggregated list of priority timber species across organizations, World Forest ID began a database to collect and organize those species most at risk for illegal logging, unsustainable deforestation, and high rates of trade globally. The list will not only benefit World Forest ID work streams, but also other organizations in the field with adjacent interests.

METHODOLOGY

The effort began with compiling information from peer-reviewed publications, web databases, and experts in the field. It became clear that the compilation of up-to-date risk species is a common need amongst many organizations and forest legality experts. The database then spread to other actors with a need for prioritizing timber species and became a collaborative effort. Now, the Global Priority Wood Species List (GPWSL) harbors nearly 300 species and contains relevant information on each, such as natural distribution, conservation listings, and countries of import. The list relies on scholarly trusted sources, such as government and non-governmental organization (NGO) reports to provide this information. Priority species must meet one or more of the following criteria to be added to the list: 1) illegal logging risk, 2) high trade internationally, or 3) threat of unsustainable harvesting. Whether a species fits one or more of the criteria is determined by reviewing the literature in which the species is found to be a priority; WWF (2015), Groves and Rutherford (2015), Environmental Investigation Agency (2017), Cramm and Van Brusselen (2019), Crowley et al (2020), Preferred by Nature (2020) and Bartholomew et al (2021). All three criteria were set in the interest of forest conservation and ensuring legality in the global forest product supply chain. Given that prioritization can be looked at through many lenses, it is important that the information in the GPWSL reflects this diversity in interests within the field.

The following information is filled in for each priority species, to the best of a contributor's ability and resource availability: scientific name, common name, country of interest, vulnerable forest(s), natural distribution, commercial plantation(s), CITES listing, IUCN Red List category, alternate scientific name(s), internationally traded (Yes/No/Banned), likely product form traded, and import countries. The details and explanation of each of these can be found in Table 1. For each of these categories, the source, where the information was found, is noted. Sometimes, there is just one source from which all the inputted information was derived. If this is the case, then only that source will be noted for the species in question. For example, *Abies guatemalensis* was found to be a priority species from the CITES and Timber report published in 2015 (Groves and Rutherford 2015). This report also provides additional information on natural distribution and commercial plantations. Thus, the CITES and Timber report was indicated as the sole source for *Abies guatemalensis*. It is critical that each contributor to the list indicates the source from which the information inputted for a species came. Given that users come from many perspectives, one entity may look into a source and

Table 1. The fields of the Global Priority Wood Species List; Note: This working list is constantly updated so some sources may not be included.

Name	Unique six-character species identifier composed of the first three letters of the genus, followed by the first three letters of the species. (e.g., Cedrela odorata is CEDODO)
Genus	First word of scientific name (e.g., *Cedrela*).
Species	Second word of scientific name (e.g., *odorata*). If information was only available to the genus level, spp. will be noted for the species field.
Common name	A name, or names, other than the scientific name, that is commonly used to describe the species. Common names often differ by country or region. Sources include, but are not limited to IUCN Red List (2021); WWF (2015); and expert knowledge.
Country of interest	The country, or countries, in which the particular species is either 1) highly exported from, 2) at risk for illegal logging, or 3) at risk for unsustainable deforestation (using ISO Alpha-3 Country Codes). From WWF (2015), Preferred by Nature (2020), Groves and Rutherford (2015), Bartholomew et al (2021), and Crowley et al (2020).
Vulnerable forests	The country, or countries, in which the particular species' conservation is threatened. The species is either 1) at risk for illegal logging or 2) at risk for unsustainable deforestation based on literature evidence (using ISO Alpha-3 Country Codes). From WWF (2015), Preferred by Nature (2020), Groves and Rutherford (2015), Bartholomew et al (2021), and Crowley et al (2020).
Natural distribution	The country, or countries, in which the species naturally grows (using ISO Alpha-3 Country Codes). From WWF (2015), Cramm and Van Brusselen (2019), and Groves and Rutherford (2015). Verified using Plants of the World Online | Kew Science.
Commercial plantations	The country, or countries, in which there are commercial plantations outside of the natural range (using ISO Alpha-3 Country Codes). From Groves and Rutherford (2015), Bartholomew et al (2021), and Crowley et al (2020).
CITES listing	Appendix in CITES from https://checklist.cites.org/#/en
IUCN red list	IUCN red list category from https://www.iucnredlist.org/
Alternate scientific name	Alternative scientific names that are no longer accepted, based on Plants of the World Online | Kew Science determinations.
Internationally traded	Assessment whether this species is/has been known to be internationally traded (Y = Yes; N = No; B = Banned); Data on whether the species is traded internationally was based on findings of Mark et al (2014), but also incorporates independent non-governmental organizations (NGOs) and industry expert understandings.
Likely product form(s) traded	If "Internationally_Traded_Y_N_B" = Y, then this field provides a subjective assessment or evaluation of the known product forms the species is often traded in (e.g., it may be exported in raw log form to a processing country, but then after being processed in the intermediate country, it may be exported as veneer or finished flooring, so this field would read: logs, veneer, flooring).
Import countries	If "Internationally_Traded_Y_N_B" = Y, then this is a manually curated field that provides the list (using ISO Alpha-3 Country Codes) of the known countries that import the species in any of its potential product forms. Sources for this type of information include NGO reports (like Norman and Zunino 2022), and ITTO's Tropical Timber Market Reports (https://www.itto.int/market_information_service/)
Source	Source of the information in preceding fields
Notes	Additional notes
Reason for inclusion	Answers the question; why was the species included? (e.g., is it commonly illegally logged, highly traded globally, or does it grow in a country facing high deforestation? . . .)
Date last edited	Date any field in the row was last edited by a contributor.
Contributor	Person(s) who added information to any of the columns in the species of interests' row and the organization they are from (can be abbreviated after first input; e.g., Sarah Richardson, World Forest ID, will then be abbreviated to SR).

agree that the species is a priority for their purposes while another may choose to omit this species for their purposes.

SUMMARY OF THE GLOBAL PRIORITY WOOD SPECIES LSIT

As of the writing of this article, the list is made up of 270 species, comprising 112 genera (Table 2). The species on the list have natural distributions covering 180 countries (Fig 1). The *Shorea* genus is leading the list, with a total of 23 species included. Second is the genus *Dalbergia*, with 19 species included. Southeast Asia is the region with the highest number of species included, largely due to the number of *Shorea* species that grow there. Other regions with a high number of species included are Amazonia and the Congo Basin. Unsurprisingly, these regions match up with the presence of tropical forests, known to be the hotspots of illegal and unsustainable logging. It is important to note that a large number of priority species is not the only factor in establishing hotspots. A region may have few priority species while exporting these species at large volumes. Such is the case in Eastern Europe, indicating that the number of priority species in a country does not equal trade flow volume. Figures 2 and 3 show a percentage distribution comparison between the IUCN Red List Status/CITES Appendices and the GPWSL. Most species included in the GPWSL are indicated as Least Concern or Endangered under the IUCN Red List. This is likely due to the limited trade in Critically Endangered species as there are little individuals left. Looking at the comparison with the CITES Appendices, we see that most species included in the GPWSL are not CITES listed. There are important distinctions in the evaluation criteria for inclusion in CITES and IUCN, particularly with respect to whether the threat to a species can be linked to international trade (Challender et al 2019). The GPWSL seeks to identify species that are traded regularly and are either not yet of relative concern such that they have yet to be CITES listed and are evaluated as Least Concern under the IUCN Red List, or are still traded and are of such high priority concern that they are listed as Endangered under the IUCN Red List.

APPLICATIONS

The purpose of the GPWSL is to provide a general overview of each timber species deemed a global priority. An organization working on a specific project may then use the information in the list to create a more specified list for their purposes. This allows for the simplification of the list and for further detailing if desired for individual project requirements and questions being asked by an organization.

Perhaps the most important component of the GPWSL is its current and future applications within the field of timber legality and forest conservation. The array of information provided for each species on the list is intended to meet a broad assortment of needs. One of those practical applications recalls the original purpose of the list, to prioritize the species and locations sampled in World Forest ID collections. However, the list can also be used to answer questions related to timber trade and business. Additionally, there is a potential for the list to be used as a tool by other tree species databases, and vice versa. Following are several already utilized, and potential, applications of the GPWSL.

World Forest ID

Given that the list was born out of a need to prioritize species for World Forest ID collections, it is important to note the number of the listed species that have been sampled by World Forest ID at this point. Out of the 270 species listed, World Forest ID has collected 78 (at time of writing - June 2022). The remaining species, and those that may be added to the list in the future, can be used to further prioritize World Forest ID sample collections. World Forest ID typically organizes collections by country, meaning that the list can be used to pull all species naturally growing in the country or region of interest.

Table 2. The species included in the Global Priority Wood Species List (current as of June 2022).

Genus	Species	Genus	Species
Abies	*guatemalensis, nordmanniana*	*Hopea*	*aequalis, altocollina, centipeda, depressinerva, enicosanthoides, ferrea, helferi, longirostrata, megacarpa, micrantha, obscurinerva, odorata, rudiformis, vacciniifolia*
Acacia	*auriculiformis, mangium, mearnsii, melanoxylon*	*Huberodendron*	*patinoi*
Acer	*binzayedii, fenzelianum, mazandaranicum, pictum, pseudosieboldianum*	*Humiriastrum*	*procerum*
Aesculus	*hippocastanum*	*Hura*	*crepitans*
Afzelia	*africana, bella, bipindensis, pachyloba, quanzensis, xylocarpa*	*Hymenaea*	*courbaril, oblongifolia, parviflora*
Allantoma	*decandra*	*Intsia*	*bijuga, palembanica*
Amburana	*cearensis*	*Julbernardia*	*pellegriniana*
Anadenanthera	*Colubrina*	*Khaya*	*anthotheca, grandifoliola, ivorensis, senegalensis*
Androstachys	*johnsonii*	*Leplaea*	*cedrata*
Aniba	*perutilis, rosodora*	*Liriodendron*	*tulipifera*
Anisoptera	*costata, reticulata*	*Lophira*	*alata*
Apuleia	*leiocarpa*	*Lovoa*	*trichilioides*
Aquilaria	*malaccensis*	*Machaerium*	*scleroxylon*
Araucaria	*angustifolia, araucana*	*Manilkara*	*bidentata, huberi, zapota*
Aspidosperma	*excelsum, macrocarpon*	*Mezilaurus*	*itauba*
Aucoumea	*klaineana*	*Microberlinia*	*bisulcata, brazzavillensis*
Autranella	*congolensis*	*Milicia*	*excelsa, regia*
Bagassa	*guianensis*	*Millettia*	*laurentii*
Baillonella	*toxisperma*	*Myroxylon*	*balsamum*
Berlinia	*confusa*	*Nauclea*	*diderrichii*
Calophyllum	*brasiliense, inophyllum, peekelii, soulattri*	*Neobalanocarpus*	*heimii*
Campnosperma	*brevipetiolata*	*Ormosia*	*coccinea, macrocalyx*
Cariniana	*pyriformis*	*Osyris*	*lanceolata*
Cedrela	*fissilis, odorata, salvadorensis, tonduzii*	*Paubrasilia*	*echinata*
Cedrelinga	*cateniformis*	*Peltogyne*	*purpurea*
Ceiba	*pentandra*	*Pericopsis*	*elata*
Centrolobium	*microchaete*	*Picea*	*abies*
Cotylelobium	*burckii*	*Pilgerodendron*	*uviferum*
Couma	*macrocarpa*	*Pinus*	*ayacahuite, koraiensis, leiophylla, montezumae, patula, pseudostrobus, sylvestris, teocote*
Couratari	*guianensis*	*Platymiscium*	*parviflorum*
Cunninghamia	*lanceolata*	*Plectrocarpa*	*sarmientoi*
Cylicodiscus	*gabunensis*	*Podocarpus*	*neriifolius, parlatorei*
Cyrtophyllum	*fragans*	*Pometia*	*pinnata*
Dacryodes	*buettneri*	*Porlieria*	*angustifolia*

(*continued*)

Table 2. The species included in the Global Priority Wood Species List (current as of June 2022). (cont.)

Genus	Species	Genus	Species
Dalbergia	*annamensis, assamica, bariensis, baronii, cochinchinensis, cultrata, fusca, greveana, latifolia, louvelii, madagascariensis, maritima, monticola, nigra, oliveri, pervillei, rimosa, sissoo, tonkinensis*	*Prioria*	*balsamifera, copaifera*
Didelotia	*africana*	*Prunus*	*africana*
Diospyros	*ferrea, mcphersonii*	*Pterocarpus*	*erinaceus, macrocarpus, santalinoides, santalinus, soyauxii, tinctorius*
Dipterocarpus	*alatus, costatus, cuspidatus, fusiformis, geniculatus, glabrigemmatus, lamellatus, littoralis, macrocarpus, ochraceus, tempehes, tuberculatus*	*Pterogyne*	*nitens*
Dipteryx	*ferrea, micrantha, odorata, oleifera*	*Pycnanthus*	*angolensis*
Dracontomelon	*dao*	*Quercus*	*alba, bicolor, frainetto, mongolica, montana, petraea, robur*
Dryobalanops	*aromatica, fusca, rappa*	*Shorea*	*alutacea, biawak, brunnescens, calcicola, cordata, dispar, domatiosa, elliptica, foraminifera, iliasii, inaequilateralis, induplicata, laevis, leprosula, pachyphylla, pallidifolia, platyclados, praestans, revoluta, rotundifolia, splendida, tenuiramulosa, woodii*
Entandrophragma	*angolense, candollei, cylindricum, utile*	*Staudtia*	*kamerunensis*
Erisma	*uncinatum*	*Swietenia*	*macrophylla*
Erythrophleum	*fordii, ivornese, suaveolens*	*Tabebuia*	*aurea, rosea*
Eucalyptus	*delegatensis, globulus, grandis, obliqua, regnans*	*Taxus*	*wallichiana*
Eusideroxylon	*zwageri*	*Tectona*	*grandis*
Fagus	*sylvatica*	*Terminalia*	*brassii*
Fitzroya	*cupressoides*	*Theobroma*	*cacao*
Fraxinus	*excelsior, mandshurica, pennsylvanica*	*Tieghemella*	*africana, heckelii*
Gilbertiodendron	*dewevrei*	*Tilia*	*amurensis, mandshurica*
Gonopterodendron	*sarmientoi*	*Triplochiton*	*scleroxylon*

(*continued*)

Table 2. The species included in the Global Priority Wood Species List (current as of June 2022). (cont.)

Genus	Species	Genus	Species
Guaiacum	*coulteri, officinale, sanctum, unijugum*	*Ulmus*	*parvifolia*
Guibourtia	*coleosperma, demeusei, ehie, pellegriniana, tessmanni*	*Vachellia*	*macracantha*
Handroanthus	*albus, heptaphyllus, impetiginosus, incanus, serratifolius*	*Vatica*	*adenanii, cauliflora, chartacea, congesta, endertii, globosa, patentinervia, pentandra, rotata, rynchocarpa*
Heritiera	*littoralis*	*Vitex*	*cooperi*
Hesperocyparis	*lusitanica*	*Xylia*	*xylocarpa*

International Trade

The GPWSL can also be used to inform and craft trade policies and regulations. The list can assist and educate the broader wood industry community in determining species risk and geographic origin. The species identification, along with additional information, such as likely product forms traded, can help inform sustainable development of timber trade and business.

Product forms differ between harvest and finished product, and this information has been used to assist prioritization of the US Harmonized Tariff Schedule (HTS) revisions. The revisions of the HTS product nomenclature codes maintain more genus and species-specific breakout categories. The GPWSL has been used by several NGOs to build consensus on which genera are highest priority to enumerate through the HTS, and in which specific product categories priority genera or

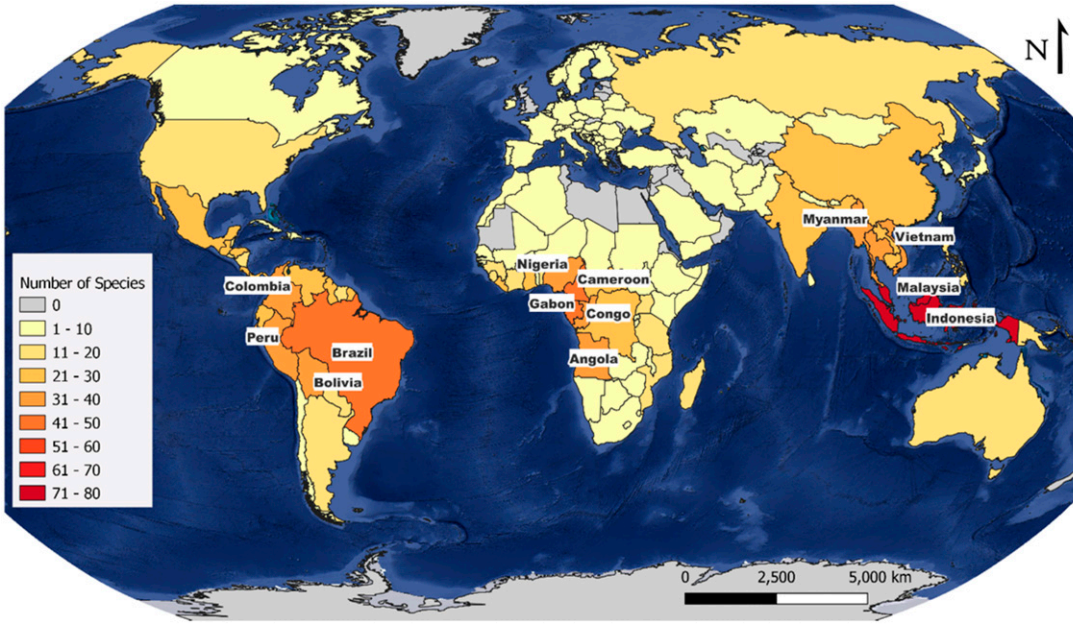


Figure 1. The number of species per country on the list, as represented by the natural distribution of each species. Countries with >31 species on the list are labeled.

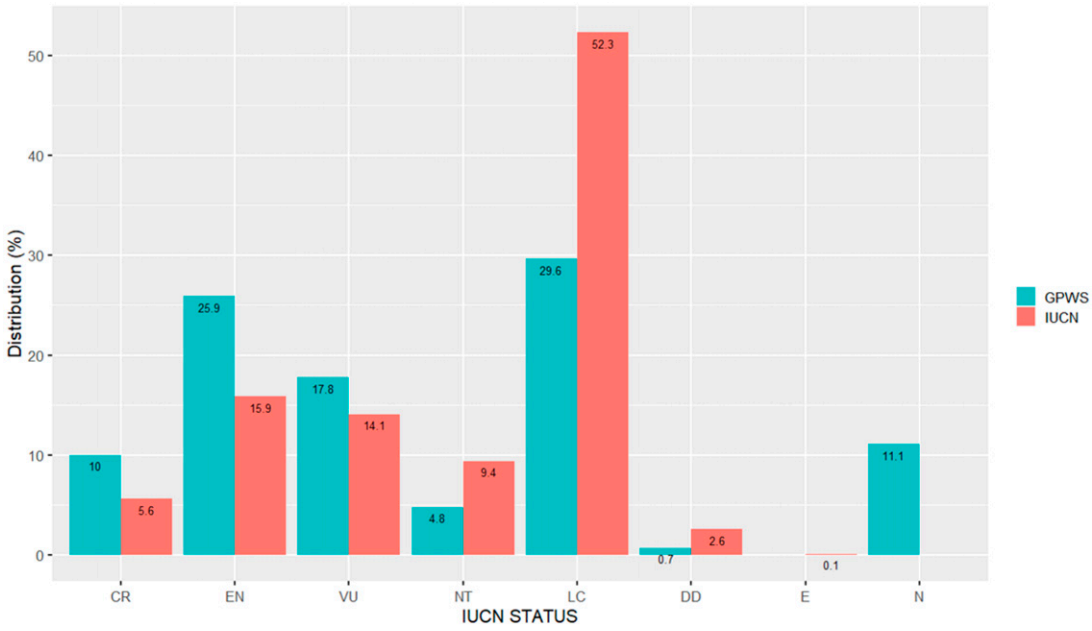


Figure 2. Distribution comparison of species included on IUCN and the Global Priority Wood Species List. CR = Critically Endangered, EN = Endangered, VU = Vulnerable, NT = Near Threatened, LC = Least Concern, DD = Data Deficient, E = Extinct, N = Not on IUCN Red List.

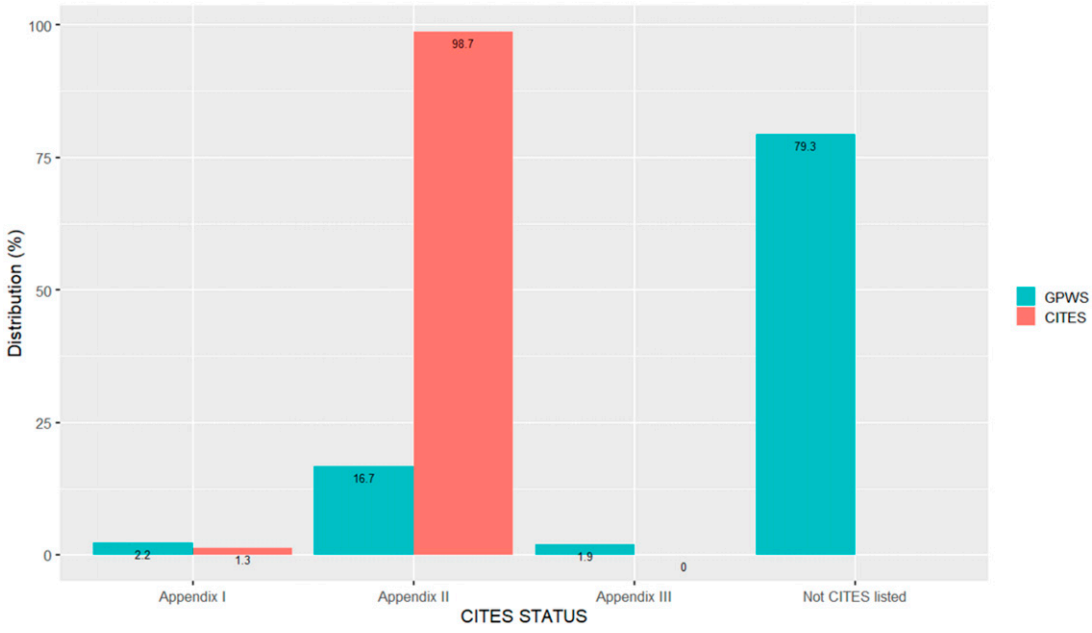


Figure 3. Distribution comparison of species included on CITES and the GPWSL.

species enumeration would be most impactful (Forest Trends 2022). Calls to increase species specificity in the global trade product nomenclature system, the harmonized system (HS), upon which the HTS is based, include high-risk wood species-product combinations (Norman and Zunino 2022), and extend beyond wood and plant species to include fish, seafood, and wildlife (Chan et al 2015; Cawthorn and Mariani 2017; Gephart et al 2019).

With respect to how this list can help the broader business community, the GPWSL is being used by organizations to build risk analysis tools and profiles. Companies can use the list to determine whether the species-countries they are sourcing from are of heightened risk or high conservation priority. Companies can also determine whether reference sample collections are present, or are prioritized for future collection efforts, thereby allowing them to know when sufficient reference data exists. This informs whether they can identify, verify, and track a given product's species and harvest origin through the supply chain using scientific identification methods and tools.

Database Harmonization

A variety of databases exist that provide detailed information on tree species used in timber. A harmonized database with all this information in one place would be a useful tool in the fields of forest legality and conservation. Thus, the GPWSL is created in the hopes that the information within can be utilized in other databases. Reversely, these relevant databases can ideally be utilized to supplement species information in the GPWSL. One example of such a database is Arbor Harbor, a developing platform that will compile reference information on tree species in the timber trade (https://woodid.info/). Comprehensive databases, such as Arbor Harbor, will be critical to improving the robustness of relevant information for each species in the GPWSL.

FUTURE OF THE LIST

The GPWSL is a living database, meaning that it is ever-changing and ever-adapting to the current state of the world's tree species. Ideally, the species on this list would fall to zero with the work of the organizations involved. The prioritization of these species is intended for organizations to take this knowledge and use it to inform future practices. Whether that be through the increased transparency of the timber supply chain or the development of robust reference samples, all efforts influence the fight against illegal logging. However, the reality is that while some tree species may reach a status that would merit their removal from the list, others may be added. Some species that are not deemed a global priority now may end up becoming a priority in the future because of overharvesting and legality issues. Hopefully, the removal of species will surpass the addition of species in the coming years as progress in the field of forest legality is made.

The list can also be made a better resource by utilizing Application Programming Interfaces (APIs) to facilitate real-time updates to the global priority list. This would relieve the obligation of an individual going in and manually updating each of the fields after a selected period of time. These APIs can be sourced from web databases (e.g., IUCN Red List 2021 and CITES 2022) and hopefully from collaborating organizations, such as World Forest ID and Arbor Harbor. The number of collaborators will ideally grow in the future as more become interested in contributing and using the list. A wide range of expertise in the field of wood products and illegal logging is critical to this list as it can help the vetting and addition of species. Thus, it is important that contributors are brought in who can meet this diversity and improve the robustness and accuracy of the priority species. This list depends on a balance of give and take, rather than only serving as a resource to extract information from.

Access

If you or your organization would like to contribute to the GPWSL, please contact World Forest ID at info@worldforestid.org with; organization name, organization mission, organization's relation and relevancy to the purpose of the list, and the reason

why you or your organization would like to be a collaborator.

REFERENCES

Arbor Harbor (2022) Arbor harbor: A trees to trade reference system—overview. https://woodid.info/ (8 March 2022).

Bartholomew D, Barstow M, Randi A, Bodos V, Cicuzza D, Hoo, PK, Juling S, Khoo E, Kusumadewi Y, Majapun R, Mariyani A, Maycock CR, Nilus R, Pereira JT, Sang J, Robiansyah I, Sugau JB, Tanggaraju S, Tsen S, Ying LC (2021) The red list of Bornean endemic dipterocarps. Botanic Gardens Conservation International, Richmond, UK. ISBN-10: 1-905164-79-3.

Beeckman H (2016) Wood anatomy and trait-based ecology. IAWA J 37:127-151.

Cawthorn D-M, Mariani S (2017) Global trade statistics lack granularity to inform traceability and management of diverse and high-value fishes. Sci Rep 7(1):12852. https://doi.org/10.1038/s41598-017-12301-x

Challender DWS, Hoffmann M, Hoffmann R, Scott J, Robinson JE, Cremona P, Hilton-Taylor C, Jenkins RKB, Malsch K, Conde D, De Meulenaer T (2019) Criteria for CITES species protection. Science 364(6437): 247-248. https://doi.org/10.1126/science.aax1266

Chan H-K, Zhang H, Yang F, Fischer G (2015) Improve customs systems to monitor global wildlife trade. Science 348(6232):291-292. https://doi.org/10.1126/science.aaa3141

Chave J, Coomes D, Jansen S, Lewis SL, Swenson NG, Zanne AE (2009) Towards a worldwide wood economics spectrum. Ecol Lett 12:351-366.

CITES (2022) Species+/CITES checklist Application Programming Interface (API). https://api.speciesplus.net/ (16 March 2022).

Cramm M, Van Brusselen J (2019) List of priority tree species for the development of reference data for the identification of tree species and their geographic origin. Update of the GTTN priority tree species list of 2013. Global Timber Tracking Network, European Forest Institute, Joensuu, Finland. https://doi.org/10.13140/RG.2.2.10132.04483.

Crowley D, Barstow M, Rivers M, Harvey-Brown Y (2020) The red list of Acer. Botanic Gardens Conservation International, Richmond, UK.

Deklerck V, Price E, Vanden Abeele S, Lievens K, Espinoza E, Beeckman H (2021) Timber identification of *Autranella*, *Baillonella* and *Tieghemella* in the taxonomically challenging Sapotaceae family. Plant Methods 17:64.

Environmental Investigation Agency (2017) The rosewood racket: China's billion dollar illegal timber trade and the devastation of Nigeria's forests. https://rosewoodracket.eia-global.org/

FAO (2020) Global forest resources assessment 2020: Main report. FAO, Rome. https://doi.org/10.4060/ca9825en

FAO (2022) The state of the world's forests 2022. Forest pathways for green recovery and building inclusive, resilient and sustainable economies. FAO, Rome. https://doi.org/10.4060/cb9360en

FAO and UNEP (2020) The state of the world's forests 2020. Forests, biodiversity and people. FAO, Rome. https://doi.org/10.4060/ca8642en

Forest Trends (19 January 2022) Amendments to U.S. Harmonized Tariff Schedule will help capture better data on wood products entering the U.S. Press Release. https://www.forest-trends.org/pressroom/amendments-to-us-harmonized-tariff-schedule/

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(2):137-154.

Gasson PE, Lancaster CA, Young R, Redstone S, Miles-Bunch IA, Rees G, Guillery RP, Parker-Forney M, Lebow ET (2021) WorldForestID: Addressing the need for standardized wood reference collections to support authentication analysis technologies; a way forward for checking the origin and identity of traded timber. Plants People Planet 3(2):130-141.

Gatti LV, Basso LS, Miller JB, Gloor M, Dominques LG, Cassol HLG, Tejada G, Aragão LEO, Nobre C, Peters W, Marani L, Arai E, Sanches AH, Corrêa SM, Anderson L, Von Randow C, Correia CSC, Crispim SP, Neves RAL (2021) Amazonia as a carbon source linked to deforestation and climate change. Nature 595:388-393.

Gephart J, Froehlich H, Branch T (7 May 2019) Opinion: To create sustainable seafood industries, the United States needs a better accounting of imports and exports. PNAS 116(19):9142-9146. https://doi.org/10.1073/pnas.1905650116

GOV.UK (2021) COP26: World leaders summit on "Action on forests and land use." World leaders summit on "Action on forests and land use"—GOV.UK. www.gov.uk

Groves M, Rutherford C (2015) CITES and timber: A guide to CITES-listed tree species. Kew Publishing, Royal Botanic Gardens, Kew.

Hermanson JC, Dostal D, Destree JC, Wiedenhoeft AC (2019) The XyloScope: A field-deployable macroscopic digital imaging device for wood. Research Note FPL-RN-0367. U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, Madison, WI. 18 p.

IUCN Red List (2021) IUCN Red List of threatened species. Version 2021-3. www.iucnredlist.org and https://apiv3.iucnredlist.org/ (16 March 2022).

Jiao L, Yang L, He T, Guo J, Yafang Y (2020) DNA barcoding for wood identification: Global review of the last decade and future perspective. IAWA J 41(4):620-664. https://doi.org/10.1163/22941932-bja10041

Mark J, Newton A, Oldfield S, Rivers M (2014) The international timber trade: A working list of commercial timber tree species. Botanic Gardens Conservation International, Richmond, UK.

May C (2017) Transnational crime and the developing world. Global Financial Integrity. https://www.gfintegrity.org/wp-content/uploads/2017/03/Transnational_Crime-final.pdf

Norman M, Zunino AR (2022) Demand for luxury decks in Europe and North America is pushing Ipê to the brink of extinction across the Amazon basin and threatening the forest frontier. Forest Trends. https://www.forest-trends.org/publications/demand-is-pushing-ipe-to-brink-of-extinction-across-the-amazon-basin/

Preferred by Nature (2020) Timber legality risk assessment: Angola, Cambodia, Equatorial Guinea, Ghana, Laos, Nigeria.

Watkinson CJ, Rees GO, Cynel Gwenael M, Gasson P, Hofem S, Michely L, Boner M (2022) Stable isotope ratio analysis for the comparison of timber from two forest concession in Gabon. Frontiers in Forests and Global Change. (27 January 2022).

WWF (2015) Country Profiles—2015: Bolivia, Cameroon, China, Colombia, Indonesia, Lao PDR, Malaysia, Myanmar, Panama, Peru, Russian Far East, Vietnam. Global Forest & Trade Network. http://assets.worldwildlife.org/publications/677/files/original/March_2015_Country_Risk_Profile_NA_Booklet.pdf

WWF (2022) Deforestation and forest degradation. World Wildlife Fund. https://www.worldwildlife.org/threats/deforestation-and-forest-degradation

IDENTIFICATION AND RECOGNIZATION OF BAMBOO BASED ON CROSS-SECTIONAL IMAGES USING COMPUTER VISION

Ziwei Wang
Student
E-mail: treasure0807fire@gmail.com

Fukuan Dai
Student
E-mail: daifuk@163.com

Xianghua Yue
Senior Engineer
E-mail: yuexianghua@icbr.ac.cn

Tuhua Zhong
Associate Researcher
E-mail: zhongth@icbr.ac.cn

Hankun Wang
Researcher
E-mail: wanghankun@icbr.ac.cn

*Genlin Tian**
Associate Researcher
Institute of New Bamboo and Rattan Based Biomaterials
International Center for Bamboo and Rattan
Beijing 100102, China
and
Key Laboratory of National Forestry and Grassland Administration/
Beijing for Bamboo & Rattan Science and Technology
Beijing 100102, China
E-mail: tiangenlin@icbr.ac.cn

(Received October 2022)

Abstract. Identification of bamboo is of great importance to its conservation and uses. However, identify bamboo manually is complicated, expensive, and time-consuming. Here, we analyze the most evident and characteristic anatomical elements of cross section images, that's a particularly vital breakthrough point. Meanwhile, we present a novel approach with respect to the automatic identification of bamboo on the basis of the cross-sectional images through computer vision. Two diverse transfer learning strategies were applied for the learning process, namely fine-tuning with fully connected layers and all layers, the results indicated that fine-tuning with all layers being trained with the dataset consisting of cross-sectional images of bamboo is an effective tool to identify and recognize intergeneric bamboo, 100% accuracy on the training dataset was achieved while 98.7% accuracy was output on the testing dataset, suggesting the proposed method is quite effective and feasible, it's beneficial to identify bamboo and protect bamboo in coutilization. More collection of bamboo species in the dataset in the near future might make EfficientNet more promising for identifying bamboo.

Keywords: Bamboo, cross section, images, accuracy, generalization.

* Corresponding author

INTRODUCTION

Bamboo, one of the fastest-growing plants in the world, is a complex plant that is difficult to be identified or classified. Considering the ecological and economic importance of bamboo, correct identification is critical to its silviculture and utilization (Clark et al 2015). There are 100 genera and more than 1642 species of bamboo all around the world (Kumar et al 2021). Bamboo is usually classified in accordance with morphology-based taxonomy (Yang and Dezhu 2013), anatomical feature-based taxonomy (Grosser and Liese 1971), molecular biotechnology (Zhao et al 2015), computer-assisted technology, or the combination mentioned above.

The morphology-based taxonomy is established on the morphological and structural characteristics of bamboo reproductive organs (flowers, fruits, and seeds) and vegetative organs (roots, stems, and leaves). Geng et al summarized the morphologic characteristics of more than 500 bamboo species including both domestic species and foreign introduced species in China (Yang and Dezhu 2013). However, there is a remarkable limitation on morphology-based taxonomy due to infrequent and unpredictable flowering events of bamboo (Zhao et al 2015). Researches systematically studied and classified the vascular bundles in bamboo, providing the morphology of vascular bundles represents the evolutionary direction of bamboo, and also an important basis for the identification of different genera and species (Liese 1998; Wen et al 1984, 1985). Numerous vascular bundles are existed in the cross sectional of each individual bamboo ring, which are important identifying information. However, the anatomical characteristics associated with the cross, radial, and longitudinal sections of bamboo were complicated to be prepared and collected. Besides, it usually required a large number of bamboo slices to be prepared for the anatomical characteristic collection, which was time-consuming. The molecular biotechnology used for the classification of bamboo was based on the extraction of genetic information of bamboo with molecular marker and genes testing, Zhao et al found that the majority of the species classified in accordance with DNA information were consistent with their current taxonomic classification, which helped to avoid the homonym and synonym (Zhao et al 2015). However, the molecular biotechnology-based taxonomy demanded high quality of sampling and preservation of bamboo samples. Moreover, the extraction and analysis of genetic information were complicated and expensive.

Recently, as computer technology developed, especially the emergence of computer vision technology and digital image processing technology, machine learning has been used to identify and classify plants intelligently through extracting morphological or anatomical characteristics of plants (Chao et al 2018; Kaya et al 2019; Sun et al 2017; Yusof et al 2013), as well as animals like marine mammals (Shiu et al 2020). Now the method is widely used in wood science fields, Chao et al used the K-means clustering algorithm to make preclassification of wood according to anatomical features, and then the K-Nearest Neighbor classifier identified and classified the wood based on the LBP features that were selected by genetic algorithm. USDA Forest Service built and evaluated models to classify woods at the species and genus levels, with image-level model accuracy ranging from 87.4% to 97.5% (Ravindran et al 2018, 2019). Kobayashi et al had done a lot of researches in this respect and obtained considerable results, their experimental objects included wood used for sculpture and scripture carving, which is mainly used for cultural relic identification under nondestructive conditions, others are Lauraceae and Fagaceae wood identification for exploring the difference between the genus or species levels according the quantities of computer-based properties (Hwang et al 2018; Kobayashi et al 2017, 2019a, 2019b). These examples proved that artificial intelligence can be combined with Wood science field to help human recognition and classification.

There are also plentiful examples of successful research in the field of bamboo. YOLOv3 was trained for recognize vascular bundles to realize the coordinate of the vascular bundles to this location, the number of vascular bundles and calculation of its area (Li et al 2021), this experiment has solved the problem of manual measurement, it's time and

labor consuming. As a major feature and advantage of bamboo, the gradient structure is also used in conjunction with computer technology, the results provide a theoretical basis for the functional improvement of bamboo materials and the development of new bamboo-based smart materials (Xu et al 2021). As an important tissue and research hotspot of bamboo plants, whether the morphology and distribution of organization in cross section can be combined with modern computer vision technology to play its recognition value at this level. For computers, an image is a combination of pixels, different bamboo species have different images of anatomical elements, and the composition of the elements causes distinctions in pixel intensity, arrangement, and distribution. Such differences can be detected learned by convolutional neural network.

In this paper, we propose a practical solution to bamboo recognition, we took the cross-sectional images of bamboo as the research object, experiments were carried out at the level of identification, and computer vision technology was applied to the recognition of bamboo species, EfficientNet (Tan and Le 2019) was exploited as the identification engine. We also studied two diverse transfer learning strategies for the learning process, namely fine-tuning with fully connected layers and all layers. The approaches' performance has been validated on a bamboo dataset with 8000 images for training 2000 images for validation and 2000 images for testing. The experimental results demonstrate that the application of EfficientNet on the dataset considerably improves the overall performance, and thus achieves a better outcome.

MATERIALS AND METHODS

Sample Collection and Processing

Four-year-old bamboo was collected from Fujian province, China. The detailed characteristics was listed in Table 1. The bamboo ring was cut from the bamboo culm with a length of 2 mm at 1.3 m and then these rings were polished with sandpaper of 320 mesh to expose the vascular bundles and parenchyma for clear observation and investigation.

Original Dataset

In the experiment, black flocking cloth was used as the background to avoid the influence of other complex environments, the cross section of the bamboo ring was scanned by a high-resolution scanner (Epson Perfection V850 Pro, Epson, Japan) in 16 Gy modes with a resolution of 9600 ppi and uniformly saved the JPG format. The whole cross section images are shown in Fig 1.

Measurement of Characteristics at the Organizational Level

The number of vascular bundles, the distribution density of vascular bundles, the area of fiber sheath, and the volume fraction of fiber of these 10 genera bamboo were measured by using the bamboo structure analysis software studied in the laboratory (Li et al 2021; Xu et al 2021) (Table 2).

Computer Vision Technology

Transfer learning. Transfer learning refers to the transfer of the weight information in the pretrained neural network to other neural networks

Table 1. Information of 10 bamboo species.

Bamboo species	Wall thickness (mm)	Outer diameter (mm)	Outer circumference (mm)	Cross section area (mm^2)
Acidosasa edulis	6.18	26.02	81.75	403.79
Bambusa emeiensis	4.60	76.20	239.39	1026.75
Dendrocalamus latiflorus	8.29	84.63	265.87	2395.80
Gigantochloa apus	9.29	82.61	259.54	2169.84
Indosasa sinica	6.81	45.39	142.59	813.93
Oligostachyum sulcatum	4.61	41.62	130.76	473.45
Phyllostachys edulis	5.96	48.78	153.25	2254.54
Schizostachyum funghomii	6.32	34.86	109.52	519.74
Sinobambusa tootsik	4.68	34.47	108.28	388.51
Thyrsostachys siamensis	6.80	25.85	81.22	1104.25

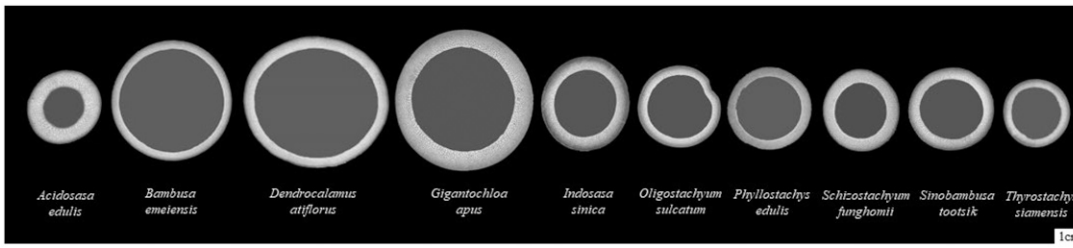


Figure 1. The cross-sectional scanner images of 10 genera bamboo rings.

(Dai et al 2009). At present, the feature distributions of training dataset and test dataset are assumed to be consistent when most machine learning algorithms conduct experiments, but there are differences in real scenes. For example, we are going to classify a new task, but the task of data is not enough, and need a lot of relevant training data, in the actual experiment, the training dataset and the corresponding validation and testing dataset of the experiment do not match the characteristics of distribution, the experimental study using appropriate migration methods can improve classification performance when insufficient samples. In this way, transfer learning allows for the reuse of existing parameters like convolution weights from a model trained on large datasets for training a new model with a relatively less number of labeled images. In this experiment, due to the different diameter of each bamboo ring, the number of images generated by step sampling is also various, and combined with the elimination of invalid images, the whole images of each bamboo are 1200, including training, validation, and testing dataset. In this work, we made use of weights pretrained from the ImageNet dataset as it contains all kinds of images and this is highly useful for classifying the bamboo dataset exploited in our evaluation. In addition, models pretrained on ImageNet has 1000 types of output, the experimental need to change the types of its output layers to 10, we study two different strategies for the learning process, namely fine-tuning with fully connected layers (fine-tuning-FC) and all layers (fine-tuning-AL), the corresponding weights are obtained after training.

Pretreated dataset. For the cross section of single bamboo ring, the experiment used Python sliding window to cut images by step sampling. On the one hand, the step covered all the vascular bundles of a single bamboo species from the outside and inside of the bamboo wall, and on the other hand, the number of images obtained by this method were enriched. The step size is the half size of the images (512×512 pixels), so as to ensure that valid information is not missed and the model can fully learn the organizational characteristics, the final images are shown in Fig 2(b). Since it is necessary to test the configuration of the model and whether the training degree

Table 2. Characteristics at the organizational level of 10 genera.

Bamboo species	Number of vascular bundles	Distribution density of vascular bundles	Area of fiber sheath	Volume fraction of fiber
Acidosasa edulis	1203	2.98	106.47	26.37
Bambusa emeiensis	3575	3.48	446.47	43.48
Dendrocalamus latiflorus	7704	3.22	732.91	30.59
Gigantochloa apus	4308	1.99	887.36	40.90
Indosasa sinica	1916	2.35	336.76	41.37
Oligostachyum sulcatum	1824	3.85	149.17	31.51
Phyllostachys edulis	6168	2.74	481.59	21.36
Schizostachyum funghomii	1230	2.37	181.58	34.94
Sinobambusa tootsik	1511	3.89	121.10	31.17
Thyrsostachys siamensis	3534	3.20	553.43	50.12

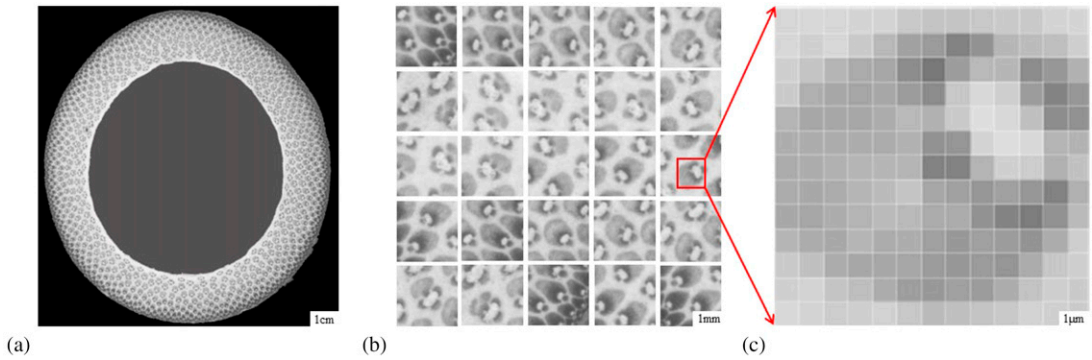


Figure 2. Schematic diagram of the images processing flow (a) the scanner image of *Schizostachyum funghomii* (b) the result images of step sampling (c) partial magnification of the cross-sectional image.

is overfitting or underfitting in the process of model construction, the dataset was divided into three parts: the training and validation dataset for training model, and the other is the testing dataset. The training and validation dataset is used to train the pretrained model, and then the test dataset is used to verify the effectiveness of the model and select the model with the best effect. In this experiment, the 12,000 images are divided into training, validation, and testing dataset by a ratio of 4:1:1.

Evaluation metrics. From the manually labeled dataset, we know exactly which category each image belongs to, and we are interested in how well the prediction label match with the true label. Thus, accuracy, precision, recall, and F1 score are utilized to measure the model performance.

Accuracy: The metric is defined as the ratio of correctly identified images to the total number of bamboo images in the dataset:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (1)$$

Precision and Recall: The metric is to measure how accurate the results for each label are, with respect to the corresponding true label.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (3)$$

F1 score: The metric is computed as the harmonic average of precision and recall by means of the following formula:

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \quad (4)$$

TP: The true label of positive that the model considers to be positive.
FN: The true label of positive that the model considers to be negative.
FP: The true label of negative that the model considers to be positive.
TN: The true label of negative that the model considers to be negative.

In addition, for a more intuitive view of the detailed recognition results, we derive the confusion matrix, where each genus of labels can be observed, which can help us to pursue deeper recognition mechanism.

RESULTS AND DISCUSSION

Characteristics at the Organizational Level

The progress of bamboo system classification needs multidisciplinary penetration, it is not only to the study of the external morphology of

bamboo nutrients and reproductive bodies, but also to cooperate with anatomy, pollen morphology, starch morphology, chromosome, biochemistry, quantitative classification, and other means. The morphology and distribution of bamboo vascular bundle are also an important means. The vascular bundles are a complex of the conductive tissue and strong tissue of bamboo, vessels, and sieve tubes in the vascular bundles connect the tip and the root, even branches and leaves, communicate of the whole plant body, to transport nutrient solution. Bamboo individual is relatively tall, to protect the smooth conduction tissue, there must be a relatively tough strong tissue in the outer edge of the conduction tissue to protect bamboo's growing, so the fiber sheath of bamboo is usually more developed, and the whole vascular bundle tissue usually accounts for 40-60% of the volume and 70-80% of the weight of bamboo (Liese 1998). To identify bamboo, computer vision is used to extract features of vascular bundles and fiber sheath, as well as their size and arrangement from bamboo sections, such as points, blobs, corners, and edges from images, and these are determined by the pixels (Fig 2[c]).

As shown in Table 2, there is a quick way to get the data in the table, 10 genera bamboo differ in their organizational characteristics, some of them belong to scattered bamboo, and some belong to cluster bamboo. In scattered bamboo, the vascular bundles from out of bamboo to inside of bamboo mainly transitions from semiopen to open type, while in cluster bamboo, the vascular bundles from growing species mainly transitions from tight waist type to broken waist type, on the other hand, it is caused by the different dimensions of its cross section. However, these data are enough to prove that the cross-sectional organizational characteristics of each genus of bamboo are different, so it is also a possible way for us to identify them from the organizational level. Using image techniques, anatomical statistical features such as the shape, size, number, and distribution of bamboo cells can be extracted from cross-sectional images. The vascular bundle and fiber are the most evident and characteristic anatomical element, so the difference of organizational characteristics data can prove its particularly is a vital breakthrough point.

Intergeneric Identification

After the two transfer learning models are trained for 100 epochs, the corresponding curve of loss and accuracy can be obtained (Figs 3 and 4). From the approximate trend in Fig 3(b), the accuracy of fine-tuning-FC fluctuates greatly, and fluctuates up and down with the increase of the number of epoch times, on the contrary, the overall trend of fine-tuning-AL is stable, and its loss decreases and

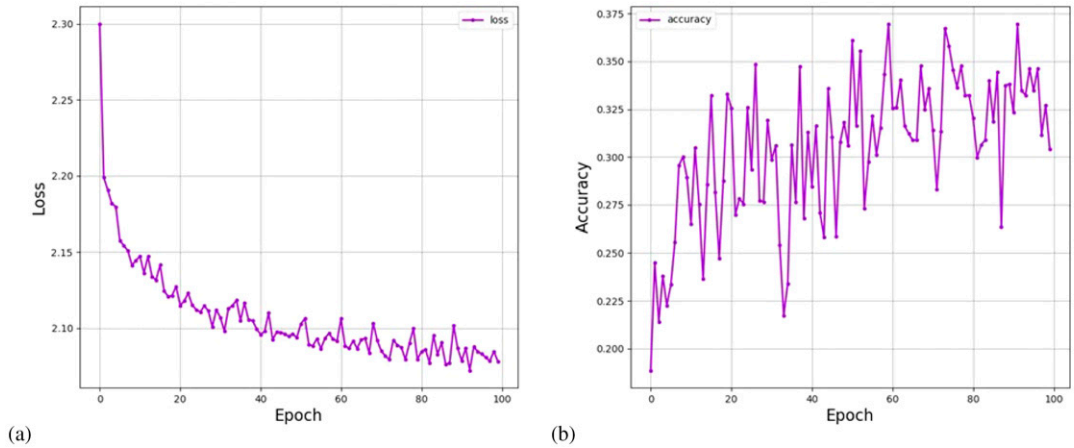


Figure 3. The curve of loss and accuracy of fine-tuning-FC model (a) the curve of loss of 10 genera bamboo (b) the curve of accuracy of 10 genera bamboo.

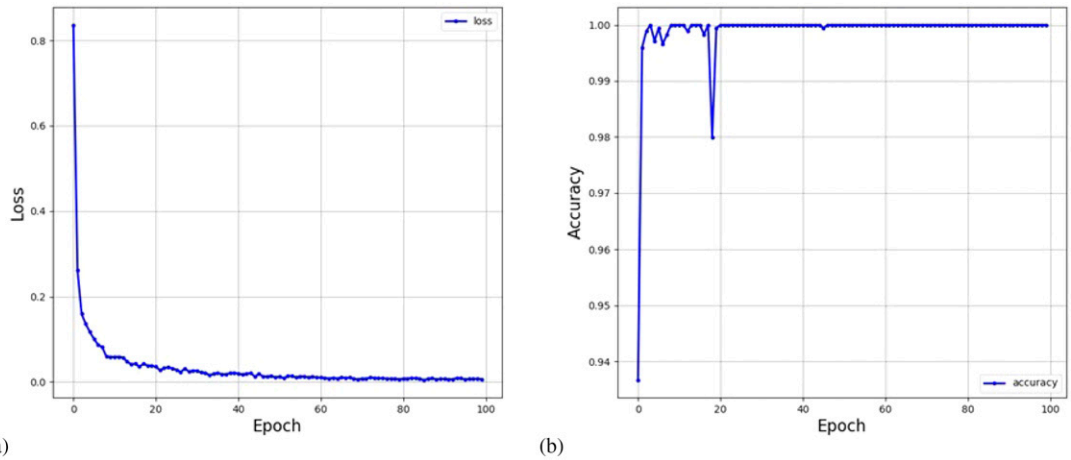


Figure 4. The curve of loss and accuracy of fine-tuning-AL model (a) the curve of loss of 10 genera bamboo (b) the curve of accuracy of 10 genera bamboo.

accuracy increases with the increase of epoch times, it's clear that fine-tuning-AL performs better than fine-tuning-FC. In particular, the loss value of fine-tuning-AL decreases rapidly from 0.84 to 0.07, then slowly approaches 0, and finally drops to the minimum value of 0.006, and the accuracy increases sharply at the beginning of the epoch, and the accuracy reaches 100% only by 20 epochs (Fig 4[a] and [b]).

Subsequently, the generalization ability of the model with the weights trained by fine-tuning-AL is tested. According to Table 3, the model achieves the two best performance with the accuracy of 100%, while the remaining eight genera have above 99.4% accuracy, this essentially means that EfficientNet with fine-tuning-AL yields the maximum prediction performance for those genera: all items in the test set are correctly classified to their real genera. In addition, we quantify the performance of the model with three additional dimensions of data as follows. From the predicted genera, the precision, recall, and F1 scores are calculated using Eqs 2-4, respectively, then the performance is shown in Table 3. A larger number in precision, recall, and F1 represents a better classification, with corresponding to the maximum performance, from these data, we can see that the trained model performs very well regardless of the dimension and has strong generalization ability, combining these individual genus accuracies yields an overall accuracy of 98.7%. In terms of confusion matrix, the blocks on the diagonal represent the number of correctly identified images, the horizontal coordinate is the true labels, and the vertical coordinate is the predicted labels, each genus of bamboo can be mapped one by one to observe which genera are misclassified. It can be seen from the Fig 5 that almost all of the 200 testing images of each genus of bamboo are correctly recognized. Only 188 images of *Gigantochloa apus* were identified correctly, and the remaining 12 images were identified as *Bambusa emeiensis* and *Oligostachyum sulcatum*, 8 images of *Indosasa sinica* were

Table 3. Accuracy of the fine-tuning-AL model to identify testing dataset (%).

Bamboo species	Accuracy	Precision	Recall	F1
Acidosasa edulis	99.9	99.0	100	99.5
Bambusa emeiensis	99.6	96.2	100	98.1
Dendrocalamus latiflorus	100	100	100	100
Gigantochloa apus	99.4	100	94.0	96.9
Indosasa sinica	99.5	100	95.0	97.4
Oligostachyum sulcatum	99.4	95.2	98.5	96.8
Phyllostachys edulis	100	100	100	100
Schizostachyum funghomii	99.9	99.0	100	99.5
Sinobambusa tootsik	99.8	99.0	99.5	99.2
Thyrsostachys siamensis	99.9	99.0	100	99.5

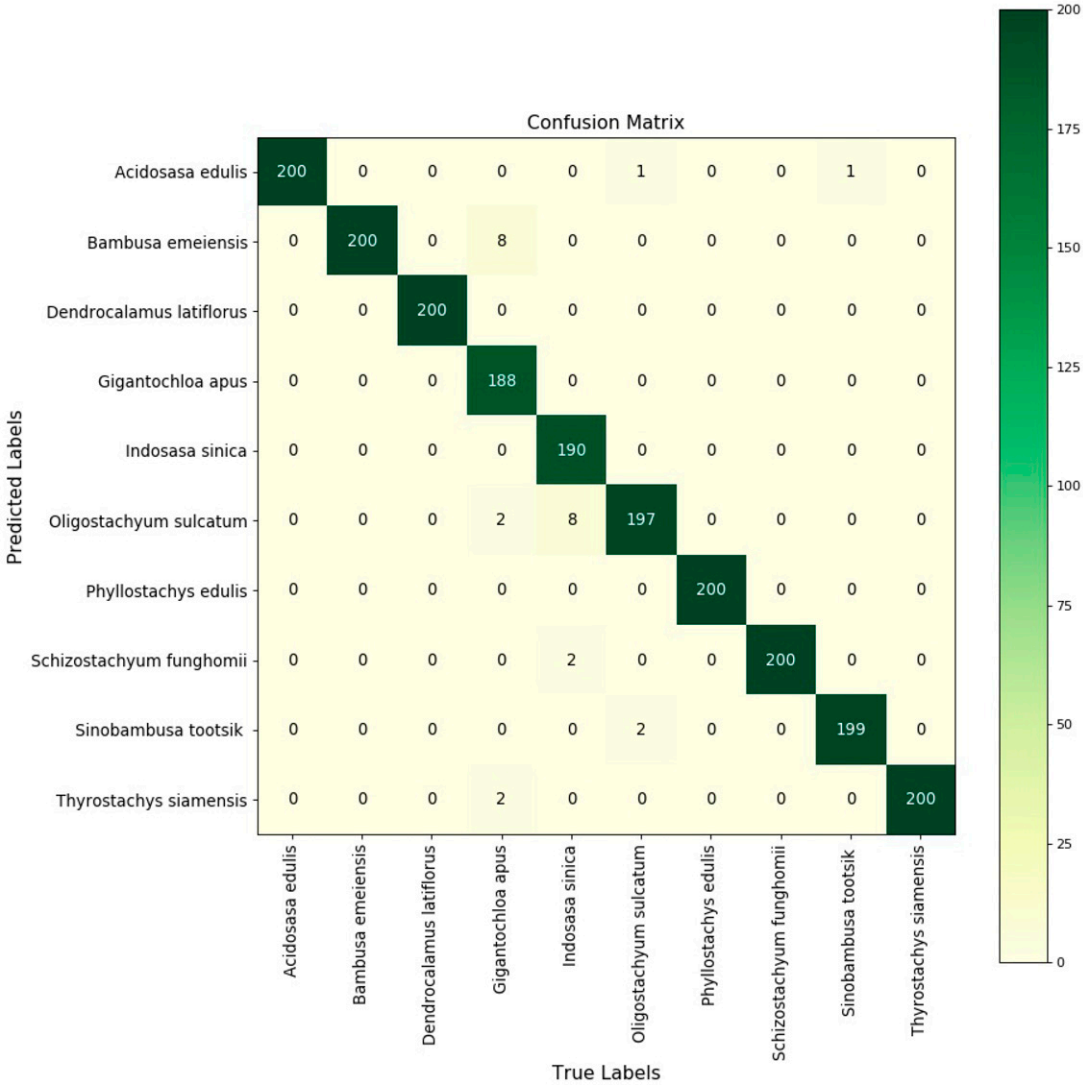


Figure 5. The confusion matrix of fine-tuning-AL model on testing dataset.

identified as *Oligostachyum sulcatum*, 2 as *Schizostachyum funghomii*, and 1 of the *Thyrsostachys siamensis* images was identified as *Acidosasa edulis*. For the testing dataset, it simulates the real situation, just like randomly input a cross-sectional scanner image into the trained weight model, it can output the bamboo genus category predicted by the computer. Through the confusion matrix, we can know which genera are easy to be misclassified, and realize which genera are misclassified, so that we can expand the database, for the genera that are easily misclassified, the number of its training dataset is increased and the model is continued to train, thus achieving best.

CONCLUSIONS

Classification of bamboo plays an important role in the bamboo industry. Automatically classification of bamboo is possible as computer technology developed in recent years. This study demonstrated the reliability and effectiveness of EfficientNet to

be performed for bamboo classification by identifying the vascular bundles in the cross-sectional of the bamboo culm. The model was trained by 10 bamboo species belonging to 10 genera, we confirmed that fine-tuning with all layers is useful comparing with fine-tuning with fully connected layer for the identification of bamboo as a better performance was obtained by both network families, the results indicated that the fine-tuning with all layers model can identify bamboo with the accuracy of 98.7%, computer vision and deep learning are a promising technology to identify bamboo according to the cross section of bamboo. For future work, the dataset will be expanded by including more bamboo species, we expect to promote the method and application in this paper and form a practical automatic bamboo recognition product. Last but not least, we are working to select the most compact network model and export to Android, making it an independent tool suitable to work on a smartphone.

ACKNOWLEDGMENTS

This work was supported by the Basic Scientific Research Funds of International Center for Bamboo and Rattan (1632022007), the National Natural Science Foundation (32071855) and the Basic Scientific Research Funds of the International Center for Bamboo and Rattan (1632022016).

REFERENCES

Chao X, Fan L, Cai C, He D (2018) Wood texture classification and identification based on multi-feature extraction and selection. Mod Agr Sci Technol 18:118-120.

Clark LG, Londoño X, Ruiz-Sanchez E (2015) Bamboo taxonomy and habitat. Springer, Cham, Switzerland. 30 pp.

Dai W, Jin O, Xue G, Yang Q, Yu Y (2009) EigenTransfer: A unified framework for transfer learning. Pages 193-200 *in* Proc. 26th Annual International Conference on Machine Learning, June 14-18, 2009, Montreal, Quebec, Canada.

Grosser D, Liese W (1971) On the anatomy of Asian bamboos, with special reference to their vascular bundles. Wood Sci Technol 5:290-312.

Hwang SW, Kobayashi K, Zhai S, Sugiyama J (2018) Automated identification of Lauraceae by scale-invariant feature transform. J Wood Sci 64(2):69-77.

Kaya A, Keceli AS, Catal C, Yalic HY, Temucin H, Tekinerdogan B (2019) Analysis of transfer learning for deep neural network based plant classification models. Comput Electron Agric 158:20-29.

Kobayashi K, Hwang S-W, Lee W-H, Sugiyama J (2017) Texture analysis of stereograms of diffuse-porous hardwood: Identification of wood species used in Tripitaka Koreana. J Wood Sci 63(4):322-330.

Kobayashi K, Hwang S-W, Okochi T, Lee W-H, Sugiyama J (2019a) Non-destructive method for wood identification using conventional X-ray computed tomography data. J Cult Herit 38:88-93.

Kobayashi K, Kegasa T, Hwang S-W, Sugiyama J (2019b) Anatomical features of Fagaceae wood statistically extracted by computer vision approaches: Some relationships with evolution. PLoS One 14(8):0220762.

Kumar M, Upadhyay SK, Kaur H, Verma R, Negi R, Sharma I, Singh R (2021) Taxonomical diversity, socio-economic and ethnomedicinal significance of Bambusa Schreber 1789 (Poaceae: Bambusoideae) in Forest Research Institute (FRI), Dehradun (Uttarakhand), India. Asian J Biol Life Sci 10(2):346-351.

Li J, Xu H, Yu Y, Chen H, Yi W, Wang H (2021) Intelligent analysis technology of bamboo structure. Part I: The variability of vascular bundles and fiber sheath area. Ind Crops Prod 174:114163.

Liese W (1998) The anatomy of bamboo culms. Brill, Leiden, The Netherlands. 208 pp.

Ravindran P, Costa A, Soares R, Wiedenhoeft AC (2018) Classification of CITES-listed and other neotropical Meliaceae wood images using convolutional neural networks. Plant Methods 14(1):25.

Ravindran P, Ebanyenle E, Ebeheakey AA, Abban KB, Lambog O, Soares R, Costa A, Wiedenhoeft AC (2019) Image based identification of Ghanaian timbers using the XyloTron: Opportunities, risks and challenges. arXiv preprint arXiv:1912.00296.

Shiu Y, Palmer KJ, Roch MA, Fleishman E, Liu X, Nosal EM, Helble T, Cholewiak D, Gillespie D, Klinck H (2020) Deep neural networks for automated detection of marine mammal species. Sci Rep 10:11-12.

Sun Y, Liu Y, Wang G, Zhang H (2017) Deep learning for plant identification in natural environment. Comput Intell Neurosci 2017:6.

Tan M, Le Q (2019) EfficientNet: Rethinking model scaling for convolutional neural networks. Pages 6105-6114 *in* International Conference on Machine Learning, June 10-15, 2019, Long Beach, California.

Wen T, Zhou W (1984) A report on the anatomy of the vascular bundle of bamboos from China. J Bamboo Res (1):1-21.

Wen T, Zhou W (1985) A report on the anatomy of the vascular bundle of bamboos from China. J Bamboo Res (1):28-43.

Xu H, Li J, Ma X, Yi W, Wang H (2021) Intelligent analysis technology of bamboo structure. Part II: The variability of radial distribution of fiber volume fraction. Ind Crops Prod 174:114164.

Yang L, Dezhu L (2013) Flora of China: Poaceae. Science Press, Beijing, China. 167-180 pp.

Yusof R, Khalid M, Khairuddin ASM (2013) Application of kernel-genetic algorithm as nonlinear feature selection in tropical wood species recognition system. Comput Electron Agric 93(2):68-77.

Zhao H, Yang L, Peng Z, Sun H, Yue X, Lou Y, Lou Y, Dong L, Wang L, Gao Z (2015) Developing genome-wide microsatellite markers of bamboo and their applications on molecular marker assisted taxonomy for accessions in the genus Phyllostachys. Sci Rep 5(1): 1-10.

COMPARING GC×GC-TOFMS-BASED METABOLOMIC PROFILING AND WOOD ANATOMY FOR FORENSIC IDENTIFICATION OF FIVE MELIACEAE (MAHOGANY) SPECIES

*Isabelle Duchesne**†*

Research Scientist
Natural Resources Canada
Canadian Forest Service
Canadian Wood Fibre Centre, Quebec, Canada
E-mail: isabelle.duchesne@nrcan-rncan.gc.ca

Dikshya Dixit Lamichhane

Former MSc Student
Département des sciences du bois et de la forêt
Centre de recherche sur les matériaux renouvelables
Université Laval, Quebec, Canada
Now with J. D. Irving Limited
E-mail: lamichhane.dikshya@jdirving.com

Ryan P. Dias

Graduate Research Assistant
E-mail: dias1@ualberta.ca

Paulina de la Mata

The Metabolomics Innovation Centre
Research Associate
Department of Chemistry
University of Alberta, Edmonton, Canada
E-mail: delamata@ualberta.ca

Martin Williams

Research Scientist
Natural Resources Canada
Canadian Forest Service
Atlantic Forestry Centre, Fredericton, Canada
E-mail: martin.williams@nrcan-rncan.gc.ca

Manuel Lamothe

Biologist
Natural Resources Canada
Canadian Forest Service
Laurentian Forestry Centre, Quebec, Canada
E-mail: manuel.lamothe@nrcan-rncan.gc.ca

James J. Harynuk

Professor
The Metabolomics Innovation Centre
Department of Chemistry

* Corresponding author
† SWST member

Wood and Fiber Science, 55(1), 2023, pp. 53-82
https://doi.org/10.22382/wfs-2023-07

University of Alberta, Edmonton, Canada
E-mail: james.harynuk@ualberta.ca

Nathalie Isabel
Research Scientist
Natural Resources Canada
Canadian Forest Service
Laurentian Forestry Centre, Quebec, Canada
E-mail: nathalie.isabel@nrcan-rncan.gc.ca

Alain Cloutier†
Professor
Département des sciences du bois et de la forêt
Centre de recherche sur les matériaux renouvelables
Université Laval, Quebec, Canada
E-mail: alain.cloutier@sbf.ulaval.ca

(Received March 2023)

Abstract. Illegal logging and associated trade have increased worldwide. Such environmental crimes represent a major threat to forest ecosystems and society, causing distortions in market prices, economic instability, ecological deterioration, and poverty. To prevent illegal imports of forest products, there is a need to develop wood identification methods for identifying tree species regulated by the Convention on International Trade in Species of Wild Fauna and Flora in Trade (CITES) and other look-alike species. In this exploratory study, we applied metabolomic profiling of five species (*Swietenia mahagoni, Swietenia macrophylla, Cedrela odorata, Khaya ivorensis,* and *Toona ciliata*) using two-dimensional gas chromatography combined with time-of-flight mass spectrometry (GC×GC-TOFMS). We also performed qualitative, quantitative (based on the measurement of vessel area, tangential vessel lumina diameter, vessel element length, ray height, and ray width), and machine-vision aided (XyloTron) wood anatomy on a subsample of wood specimens to explore the potential and limits of each approach. Fifty dried xylaria wood specimens were ground, extracted with methanol, and subsequently analyzed by GC×GC-TOFMS. In this study, the four genera could easily be identified using qualitative wood anatomy and chemical profiling. At the species level, *Swietenia macrophylla* and *Swietenia mahagoni* specimens were found to share many major metabolites and could only be differentiated after feature selection guided by cluster resolution (FS-CR) and visualization using Principal Component Analysis (PCA). Expectedly, specimens from the two *Swietenia* spp. could not be distinguished based on qualitative wood anatomy. However, significant differences in quantitative anatomical features were obtained for these two species. Excluding *T. ciliata* that was not included in the reference database of end grain images at the time of testing (2021), the XyloTron could successfully identify the majority of the specimens to the right genus and 50% of the specimens to the right species. The machine-vision tool was particularly successful at identifying *Cedrela odorata* samples, where all samples were correctly identified. Despite the limited number of specimens available for this study, our preliminary results indicate that GC×GC-TOFMS-based metabolomic profiles could be used as complementary method to differentiate CITES-regulated wood specimens at the genus and species levels.

Keywords: Mahogany, wood identification, GC×GC-TOFMS, wood anatomy, XyloTron.

INTRODUCTION

The Meliaceae family, often known as the mahogany family, comprises 50 genera and greater than 1400 species with approximately 500 species of economic importance, widely distributed in rainforests, mangrove swamps to semideserts (Mabberley et al 1995; Muellner et al 2006). Tree species of the Swietenioideae subfamily (see phylogeny in Muellner et al 2003) are particularly prized for fine furniture and musical instruments, ranking among the most economically significant species in the world (Danquah et al 2019) (Fig 1). Overexploitation is driving true mahogany (*Swietenia* species) toward extinction (White and Gasson 2008), and as

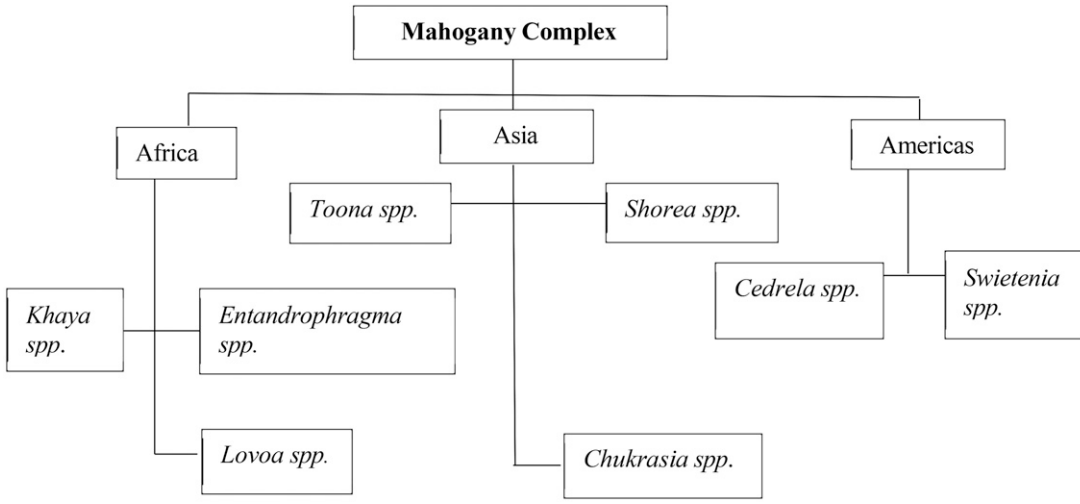


Figure 1. Economically significant mahogany species from Africa, Asia, and Americas (adapted from Danquah et al 2019).

stocks of true mahogany declines in the wild, other similar-looking species are becoming increasingly traded as substitute species, also threatening their survival (Rodan et al 1992; Verissimo et al 1998; Laurance 1999; Gullison et al 2000; O'Neill et al 2001; Kometter et al 2004; André et al 2008). To protect species survival and combat illegal logging, the genera *Swietenia*, *Cedrela* (of the Neotropics), and very recently, *Khaya* have been included in Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES 2021): trade in these tropical hardwoods has been restricted by the imposition of specific documentation and permits for importation and exportation of goods. As a result of these essential regulations, other closely related non-CITES species, including *Toona ciliata*, have been introduced as substitutes for these CITES-regulated species. Since Meliaceae species look very similar macroscopically and microscopically, their accurate identification represents a challenge for the application of CITES regulation worldwide (Gasson 2011). Hence, the development and application of efficient wood identification tools are greatly needed to control illegal logging and related trade worldwide.

Wood identification through microscopic analysis of the wood anatomy is the most common, recognized, and enduring method for wood identification available at the moment (Silva et al 2020). This methodology has a long history and is still the most frequently used on the front line for screening purposes (for instance, custom officers, in combination with macroscopic identification aides) and in the laboratory for diagnostic identification (Dormontt et al 2015; Koch et al 2015) in various contexts (archeology, architecture, and art). When the geographic origin and silvicultural history of a sample is known, wood anatomy may successfully determine tree species. If the origin is unknown, the method will allow the identification of wood at the genus or family level. In any case, genus or family confirmation by wood anatomy experts is very important for the development of novel spectrometric and genetic methods for forensic wood identification. As every method has limitations, a combination of approaches and techniques is generally recommended to corroborate identification results (Dormontt et al 2015). Such multidisciplinary approach facilitates the detection of labeling errors of wood specimens, which can occur in any curated xylarium. Indeed, taxonomic uncertainty remains in certain families or genera where species hybridize (eg Finch et al 2019; Bouka et al 2022). Automated wood anatomical analysis (machine-vision) using advanced image capture (Appendix A2) and processing algorithms is another novel area of research that has high potential in the field of wood identification to

identify plants down to the species level (Hermanson and Wiedenhoeft 2011; Hermanson et al 2019). The tool currently uses images captured from transverse wood sections only (end grain), and users need basic knowledge of wood structure to be able to use it as screening tool during field inspections. Confirmatory microscopic observations in the other two wood planes (ie in radial and tangential sections) are generally necessary to support legal cases (forensic evidence).

Apart from anatomy, another very effective approach used in forensic wood species identification is based on metabolite profiling of wood samples. The method is able to characterize species' metabolomes by pairing well-known chromatographic separation methods (ie gas chromatography [GC] and liquid chromatography [LC]) with highly sensitive detection methods using mass spectroscopy (Zhang et al 2019; Shang et al 2020; Brunswick et al 2021). Trees and other plants synthesize compounds called phytochemicals that are often species-specific and/or produced at different levels among species or higher taxonomic groups (Venkataraman 1972; Julkunen-Tiitto 1989). After separation using various chromatographic methods, mass spectrometry ionizes chemical compounds to produce charged molecules whose mass to charge proportions (m/z) that are then quantified and compared with various libraries (ie NIST, "etc.") for metabolite identification (De Hoffmann and Stroobant 2007; Zhang et al 2019). Depending on the natural variation in extractives present in the wood samples, various levels of identification may be possible, including genus and species. Although various mass spectrometric techniques have been used for wood identification (Kite et al 2010; Dormontt et al 2015), the most widely adopted by the research communities are Direct Analysis in Real Time-Time-of-Flight Mass Spectrometry (DART-TOFMS) (Cody et al 2005) and GC×GC-TOFMS (Pierce et al 2006).

DART-TOFMS is one of the most rapid and cost-effective methods for wood identification. However, the technique requires the prior development of an extensive spectral reference database of vouchered wood specimens to identify unknowns. Unlike GC-MS that is extensively used worldwide, only a few countries have currently access to the DART-TOFMS wood identification platform (eg United States and Canada). The DART-TOFMS chemotyping approach has shown potential application for wood identification, for instance; DART-TOFMS was successful in differentiating between two oak species, white oak (*Quercus alba*) from red oak (*Quercus rubra*) (Cody et al 2012), several *Dalbergia* and other commercial species (Lancaster and Espinoza 2012), African Madagascan *Dalbergia* and Asian *Dalbergia* (McClure et al 2015), Araucariaceae species (Evans et al 2017), Meliaceae species (Deklerck 2019), and two Fabaceae species (*Afzelia bipindensis* from *Afzelia pachyloba*) (Kitin et al 2021).

Combination of GC×GC with TOFMS is an important technique, which confirms the presence of large numbers of target metabolites and unknowns in one run (Beens and Udo 2005). GC×GC-TOFMS results in rapid accumulation of spectra, which leads to excellent reproducibility and better signal-to-noise characteristics and make full use of small quantities of samples. The GC×GC technology has been increasingly used for the analysis of petrochemicals and natural products, among others (Wu et al 2004). For example, application of two-dimensional GC-TOFMS method along with the Principal Component Analysis (PCA) in three different species of plants was successful to discover differences among plant samples based on separation of metabolites (Pierce et al 2006). Sun et al (2020) has recently applied two-dimensional GC quadrupole TOFMS to successfully determine and compare CITES-listed agarwood samples from eight different origins.

The objective of the study was to evaluate and compare three wood identification methods; 1) qualitative and quantitative analysis of wood anatomical structure by light microscopy, 2) anatomical analysis using machine-vision aided identification (XyloTron), and 3) wood metabolite profiling using GC×GC-TOFMS followed by feature selection guided by cluster resolution (FS-CR) (Sinkov and Harynuk 2011, 2013; Armstrong et al 2021) to identify and differentiate five easily confused wood species of the Meliaceae family

namely *Swietenia mahagoni*, *Swietenia macrophylla*, *Cedrela odorata*, *Khaya ivorensis,* and *Toona ciliata*. The five similar-looking species belong to the Swietenioideae subfamily and are highly prized for the manufacture of finest furniture products. Except for *Toona ciliata*, the international trade of these endangered species is currently regulated through CITES. A summary, including phylogeny and uses of the four selected genera, is provided as Appendix (A1). For additional information, Danquan et al (2019) wrote a review of the geographic distribution of the economically important Mahogany Complex, highlighting the threats to the sustainability of the species selected in this study.

MATERIALS AND METHODS

As a first step, we described the wood anatomy of 15 selected samples (three samples per species) among 50 available xylaria specimens to validate species' designation and determine the variability of recognized taxonomic characters (ie qualitative and quantitative wood anatomy). Thereafter, we used XyloTron on the same 15 samples as if they all were unknown samples to determine whether current algorithms could identify the wood to the species level. In parallel, we used the untargeted approach of the GC×GC-TOFMS platform to obtain metabolites profiles for all our 50 available xylaria specimens, with the objective to discover genus- or species-specific chemical compounds. When incongruency was detected between the taxonomic identification methods, a second round of verifications was conducted.

Wood Anatomy Using Light Microscopy

Fifteen wood specimens of five tropical hardwood species namely the two closely related species *Swietenia mahagoni* and *Swietenia macrophylla*, *Cedrela odorata* (native to America), *Khaya ivorensis* (native to Africa), and *Toona ciliata* (native to Asia), were sourced from two scientific collections located in Quebec City: the Canadian Forest Service—Canadian Wood Fibre Centre (CFS-CWFC) Xylarium located at the Laurentian Forestry Centre (LFC) and the Centre de recherche sur les matériaux renouvelables (CRMR) Xylarium, located at Université Laval (Table 1). To prepare permanent microscope slides, specimens were first cut into small blocks (1 cm^3) and boiled in water until they were saturated. The softened blocks were then sliced using a rotatory microtome (Leica HistoCore AUTOCUT Model) in cross, radial, and tangential sections at a thickness of 15-30 μm. The wood slices were then stained in laboratory grade, 1% aqueous safranin for 5 min and later with laboratory grade, Astra blue of 90% dye content for 3 min. After staining, slices were dehydrated in alcohol at three different concentrations (50% [v/v], 80% [v/v], and 100% [v/v]) each for 1 min and sections were mounted with Permount™ medium on glass slides. Wood features in the three planes were viewed out under an optical microscope (Nikon Eclipse, E600) connected to a computer where images were stored for subsequent image analysis (Pixel Link).

Anatomical identification of the 15 specimens of known geographic origin (except for sample 543940, see Table 1) to the genus level was performed using IAWA's lists of wood anatomical features (IAWA Committee 1989) through the Internet accessible Inside Wood tool (Wheeler 2011). Five anatomical measurements were further selected for quantitative analysis. Vessel area (μm^2), vessel element length (μm), and tangential vessel lumina diameter (μm) were measured on 30 randomly selected vessels per specimen (one slide per specimen × three specimens per species × five species) with a total magnification of 100× using image processing software Win CELL Pro 2018 e (Regent Instrument Inc., Canada). The tangential diameter of vessel lumina, excluding cell wall width, was measured at the widest part of the opening. The number of vessel elements with a total magnification of 40× were determined by counting the ones present within a field and expressed as a number of vessels per square millimeters (mm^2). All the vessels were counted as individuals, eg radial multiple of three vessels were counted as three individual vessels. For each specimen, the number of rays per linear unit (ie ray frequency per mm) was measured at

Table 1. Information about the 50 wood specimens used in this study. Three wood identification methods were compared on a subsample of 15 specimens (three per species).

Species	Sample Id	WA	XT	GC	Origin	Xylarium
Swietenia mahagoni[a]	539481	—	—	×	Unknown	CFS-CWFC
Swietenia mahagoni	539923	×	×	×	Barbados, North America	CFS-CWFC
Swietenia mahagoni	540152	—	—	×	North America	CFS-CWFC
Swietenia mahagoni	540153	×	×	×	North America	CFS-CWFC
Swietenia mahagoni	541145	—	—	×	Taiwan, Asia	CFS-CWFC
Swietenia mahagoni	543940	×	×	×	Unknown	CFS-CWFC
Swietenia mahagoni	773559	—	—	×	Unknown	CFS-CWFC
Swietenia mahagoni	773560	—	—	×	Unknown	CRMR
Swietenia macrophylla	539581	—	—	×	Central America	CFS-CWFC
Swietenia macrophylla	539857	×	×	×	Mexico, North America	CFS-CWFC
Swietenia macrophylla	539859	×	×	×	Mexico, North America	CFS-CWFC
Swietenia macrophylla	541143	—	—	×	Taiwan, Asia	CFS-CWFC
Swietenia macrophylla	541144	×	×	×	Taiwan, Asia	CFS-CWFC
Swietenia macrophylla	542318	—	—	×	Brazil, South America	CFS-CWFC
Swietenia macrophylla	543938	—	—	×	Central America	CFS-CWFC
Swietenia macrophylla	773561	—	—	×	Unknown	CRMR
Swietenia macrophylla	773562	—	—	×	Unknown	CRMR
Cedrela odorata	542126	×	×	×	Guyana, South America	CFS-CWFC
Cedrela odorata	542084	—	—	×	Unknown	CRMR
Cedrela odorata	542086	×	×	×	Guyana, South America	CFS-CWFC
Cedrela odorata	542085	—	—	×	Unknown	CFS-CWFC
Cedrela odorata	542087	×	×	×	Guyana, South America	CRMR
Cedrela odorata	773566	—	—	×	Unknown	CRMR
Cedrela odorata	773567	—	—	×	Unknown	CRMR
Cedrela odorata	773563	—	—	×	Unknown	CRMR
Cedrela odorata	773564	—	—	×	Unknown	CRMR
Cedrela odorata	773565	—	—	×	Unknown	CRMR
Cedrela odorata	544171	—	—	×	Unknown	CRMR
Khaya ivorensis	540727	—	—	×	Ghana (Gold Coast), Africa	CRMR
Khaya ivorensis	540728	×	×	×	Nigeria, Africa	CRMR
Khaya ivorensis	540729	×	×	×	Africa	CRMR
Khaya ivorensis	540730	×	×	×	West Africa, Africa	CRMR
Khaya ivorensis	540768	—	—	×	West Africa, Africa	CFS-CWFC
Khaya ivorensis	544225	—	—	×	Gabon, Africa	CRMR
Khaya ivorensis[b]	544226	—	—	×	Central African Republic, Africa	CRMR
Khaya ivorensis[b]	544227	—	—	×	Central African Republic, Africa	CRMR
Khaya ivorensis	539546	—	—	×	East Africa, Africa	CRMR
Khaya ivorensis	539547	—	—	×	East Africa, Africa	CFS-CWFC
Khaya ivorensis	773554	—	—	×	Unknown	CRMR
Khaya ivorensis	773555	—	—	×	Unknown	CRMR
Toona ciliata	539508	—	—	×	India, Asia	CFS-CWFC
Toona ciliata	541768	—	—	×	India, Asia	CFS-CWFC
Toona ciliata	543406	×	×	×	Papua New Guinea, Oceania	CFS-CWFC
Toona ciliata	541871	×	×	×	Pakistan, Asia	CFS-CWFC
Toona ciliata	543555	—	—	×	Thailand, Asia	CFS-CWFC
Toona ciliata	542956	—	—	×	Australia, Oceania	CFS-CWFC
Toona ciliata	541491	×	×	×	India, Asia	CFS-CWFC
Toona ciliata	541725	—	—	×	India, Asia	CFS-CWFC

(*continued*)

Table 1. Information about the 50 wood specimens used in this study. Three wood identification methods were compared on a subsample of 15 specimens (three per species). (cont.)

Species	Sample Id	WA	XT	GC	Origin	Xylarium
Toona ciliata	773557	—	—	×	India, Asia	CRMR
Toona ciliata	773558	—	—	×	Unknown	CRMR

[a]Note: Specimen 539481 initially labeled as *Swietenia mahagoni* and later verified through microscopic wood anatomy was found to belong to *Khaya* genus.

[b]Specimens 544226 and 544227 initially labeled as *Khaya ivorensis* and later verified through microscopic wood anatomy were found to belong to *Entandrophragma* genus. They were excluded from the chemical analysis.

WA denotes wood anatomy by light microscopy; **XT** denotes wood anatomy by means of XyloTron, a computer-aided wood identification tool used on the end grain surface of wood specimens; and **GC** refers to GC×GC-TOFMS, ie two-dimensional gas chromatography combined with time-of-flight mass spectrometry, to obtain metabolomic profiles.

10 different locations using a magnification of 40×, whereas for ray height (μm) and ray width (μm), 30 randomly selected rays were manually determined from tangential section at a magnification of 100×.

Statistical analysis was performed with SPSS (version 25) (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.). The mean values of five anatomical features for five mahogany species were compared using an analysis of variance (ANOVA) followed by Duncan's multiple range test. Differences between means were considered significant when the *p*-value of the ANOVA was less than 0.05. Data were represented as Mean ± SD. For Duncan's multiple range test, the means carrying different letters indicate significant difference.

Machine-Vision Wood Anatomy Using XyloTron

The same 15 wood specimens identified to the genus level with Inside Wood were further evaluated using the XyloTron machine-vision tool (Hermanson et al 2019; Ravindran et al 2020, see Appendix (A2)) The computer vision system comprises multiple tree species-specific algorithms, which calculate a probability that an unknown specimen belongs to a given species present in the XyloTron reference database (Hermanson et al 2019; Ravindran et al 2020). To this end, the transverse surface of the selected wood specimens was sanded by grit sandpapers (120 [fine], 240 [very fine] and 1000 [super fine]), using compressed air and adhesive tape to remove the dust from the cell lumina between each grit. The imaging of sanded wood surfaces was done using a macroscopic handheld camera (10×). Rays of the wood were aligned in vertical position, as shown in Figs A2-2 to A2-6. Each image (with dimensions 2048 × 2048 pixels) from XyloTron represents 6.35 × 6.35 mm of wood tissue.

Metabolomic Profiling Using GC×GC-TOFMS

A total of 50 wood specimens of five tropical hardwood species were used for chemical analysis (Table 1). The heartwood part of each specimen was ground with a wood file at CFS-LFC. The powder was placed in 2 mL polypropylene tubes (Supplier: Eppendorf™ 022363352; Fisher Scientific, Canada) and sent to The Metabolomics Innovation Centre (TMIC) located at the University of Alberta, Edmonton, Canada, for chemical analysis. At the TMIC research platform, a quantity of 30 mg of wood powder was weighed in a 2 mL Eppendorf™ plastic centrifuge tube. And 1 mL of methanol containing 1% (v/v) formic acid was added to the centrifuge tube. Extraction solvent was prepared by dissolving an internal standard (IS) (20 mg/L, *n*-pentadecane-d32, CDN Isotopes) and 1% (v/v) formic acid (98%, Millipore-Sigma, Canada) in HPLC grade methanol (>99.9%, Millipore-Sigma, Canada). Tubes were vortexed for 10 min. Samples were left at room temperature and allowed to extract overnight (18.0 ± 0.5 h). After extraction, sample tubes were

centrifuged for 10 min at 10,000× g. An aliquot of 200 μL of the supernatant was transferred to a GC insert vial (Chromatographic Specialties, CA) for GC×GC-TOFMS analysis. Two replicate aliquots (R) and two replicate samples (RS) were also analyzed.

All GC×GC-TOFMS untargeted analyses were carried out on a LECO Pegasus 4D system (Leco Instruments, St. Joseph, MI) equipped with a four-jet dual stage modulator. The first-dimension column was a 60 m × 0.25 mm × 0.25 μm Rtx-5 and the second-dimension column was a 1.6 m × 0.25 mm × 0.25 μm Rtx-200MS (Chromatographic Specialties). Two-dimensional chromatographic separations were conducted with a constant flow rate of 2.0 mL/min utilizing helium as the carrier gas and a modulation period of 2.5 s. A GERSTEL MPS Autosampler was used for automated injection of 1 μL of sample. The oven was at first held at 40 °C for 4 min and warmed at 3.5 °C/min to a final temperature of 315 °C. The ultimate temperature was held for 10 min. The secondary oven and modulator temperature offset were constant at +10 °C relative to the GC oven temperature and +15 °C relative to the secondary oven temperature, respectively. Mass spectra were collected at an acquisition rate of 200 Hz over a mass range between 40 and 800 m/z. A relative voltage offset of 200 V was selected as the optimized detector voltage with an electron impact energy of −70 eV. The ion source temperature was 200 °C with a transfer line temperature of 250 °C. The total analysis time was 92.57 min.

Data processing was performed using Chroma TOF® (v.4.72; LECO), a commercial software from LECO. For the processing method, the baseline offset was set to 0.9 and the expected peak widths throughout the entire chromatographic run was set to 10 s for the first dimension and 0.12 s for the second dimension. The peak finding threshold of S/N was set to 50:1 with the minimum S/N ratio for subpeaks to be retained set at 6. All chromatographic peaks were searched against the NIST and Wiley MS libraries (2017). Peaks were tentatively identified based on forward and reverse mass spectral matches greater than 700 and retention index matching (±15 Kovat's RI).

Retention time shift of metabolites is common in metabolomics studies. ChromaTOF® Statistical Compare was performed for aligning the peak tables using retention times and mass spectral match. Tolerances for retention time shift were ±10 modulation period (PM = 2.5 s) in the first-dimension separation, and tolerances for the second-dimension separation were set to 0.2 s to account for the possible retention time shift across all samples. The minimum similarity for mass spectral match was set at 600 to combine if the peaks have a match score of 600 or greater for all m/z values with abundances greater than 10%. The Statistical Compare result was exported as a comma-separated values file (csv) for further data analysis. The exported aligned data from ChromaTOF® was imported in MATLAB® R2017a, Windows 64-bit version (The Mathworks Inc., Natick, MA). With the obtained results, multivariate statistical analysis, ie PCA were performed using PLS_Toolbox (Eigenvector Research, Manson, WA).

In this study, FS-CR was applied to reduce the number of variables required to model differences between the five species. Afterward, 45 different compounds were selected as the most significant and used to visualize the grouping of each species. A heat map was constructed in MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/MetaboAnalyst/home.xhtml) using the 45 selected variables.

RESULTS AND DISCUSSION

Efficient tools and techniques are essential to distinguish listed CITES species and their relatives to fight illegal logging and deforestation. In this study, we evaluated and compared three wood identification methods based either on wood anatomy or metabolic profiles. The latter approach is promising but it should be complemented by traditional wood anatomy method based on light microscopy to build reliable reference database.

Table 2. Qualitative wood anatomical characteristics and mean specific gravity of five Meliaceae species considered in this study.

	America			Africa	Asia
Anatomical features	*Swietenia mahagoni*	*Swietenia macrophylla*	*Cedrela odorata*	*Khaya ivorensis*	*Toona ciliata*
Growth rings	Distinct	Distinct	Distinct	Indistinct/absent	Distinct
Vessel distribution	Diffuse-porous	Diffuse-porous	Semiring porous	Diffuse-porous	Semiring porous
Vessel frequency per mm^2	7 to 10	5 to 7	2 to 3	3 to 4	3 to 4
Perforation plates	Simple	Simple	Simple	Simple	Simple
Intervessel pitting	Alternate	Alternate	Alternate	Alternate	Alternate
Axial parenchyma	Diffuse, scanty paratracheal, vasicentric, marginal bands	Diffuse, scanty paratracheal, vasicentric, marginal bands	Diffuse, scanty paratracheal, vasicentric, marginal bands	Scanty paratracheal, vasicentric, marginal bands (variable/rare)	Diffuse, diffuse-in-aggregates, vasicentric, marginal bands
Prismatic crystals	Present	Present	Present	Present	Present
Rays in radial section	Heterocellular	Heterocellular	Partially heterocellular	Heterocellular	Heterocellular
Ray width (number of cells)	Multiseriate 1-6	Multiseriate 1-6	Multiseriate 1-4	Two distinct sizes: short uniseriate and multiseriate 1-8	Multiseriate 1-6
Rays structure	Storied	Storied	Unstoried	Unstoried	Unstoried
Average ray frequency per mm	6 (4-10)	5 (4-8)	5 (3-6)	5 (4-7)	5 (4-7)
Specific gravity	0.66 ± 0.85	0.50 ± 0.06	0.45 ± 0.01	0.49 ± 0.03	0.42 ± 0.13

Note: Density (specific gravity) is also an important factor for wood identification. According to consensus on phylogeny at the angiosperm family level, strong phylogenetic signals are recorded in wood density (Chave et al 2006) which means that the most closely related species share a more similar wood density value rather than low closely related species. In this study, all five closely related species have a medium wood density value and fall within the range 0.3-0.7. Hence, they cannot be separated based on this criterion.

Wood Anatomy Using Light Microscopy

The observations of taxonomic/anatomical wood features of all five species of interest (subsample of three specimens per species) are summarized in Table 2, whereas the quantitative analysis of anatomical features and statistical results are shown in Tables 3 and 4.

Swietenia mahagoni. Figure 2(a) revealed that the wood samples of *Swietenia mahagoni* showed a tendency to diffuse porosity, contained distinct growth rings with apotracheal axial parenchyma banded and marginal. Vessels were solitary and in radial multiples of up to five (usually two and three); they were round or slightly oval in transverse view and often contained red gum deposits. Apotracheal axial parenchyma was diffused, whereas paratracheal axial parenchyma was scanty and occasionally vasicentric. In the radial section, rays were heterocellular as shown in Fig 2(b). Prismatic crystals were observed in axial parenchyma and in upright or procumbent ray cells. In tangential view, uniseriate rays were rare, whereas multiseriate rays were three to six cells wide as in Fig 2(c). The range of values reported for different quantitative wood anatomical characters for *Swietenia mahagoni* (Table 4) were similar to the values reported by Panshin (1933), except for vessel element length (150-500 μm vs 49-229 μm in this study).

Table 3. Analysis of variance for selected parameters among five different species of Meliaceae family.

ANOVA					
Dependent variables	Sum of squares	df	Mean square	*F* value	*p*-Value
Vessel area	68,944,541,090	4	17,236,135,273	67	0.000[a]
Tangential diameter of vessel lumina	728,259	4	182,065	80	0.000[a]
Vessel length	743,599	4	185,899	69	0.000[a]
Ray's height	1,068,812	4	267,203	14	0.000[a]
Ray's width	21,949	4	5487	13	0.000[a]

[a]Means are significantly different at 0.05 level of significance.

Swietenia macrophylla. The qualitative wood anatomical characters of *Swietenia macrophylla* (Fig 3) were similar to that of *Swietenia mahagoni* (Fig 2). Hence, *Swietenia mahagoni* and *Swietenia macrophylla* were indistinguishable based on qualitative wood anatomy (Table 2). Panshin (1933) detailed that "it is nearly impossible to isolate the woods of the *Swietenia* species anatomically." Donaldson (1984) reported roughly comparable values for vessel element lengths (100-600 μm), tangential vessel lumina diameter (29-126 μm), eight vessels/mm^2 (range 3-16 vessels/mm^2), ray's height (100-600 μm), and five rays/mm (2-9 rays/mm), respectively.

Cedrela odorata. According to Fig 4(a), vessels were predominantly solitary but also in radial multiples up to four. The wood showed tendency to have semi-ring porosity, vessels were round-shaped in transverse view. Apotracheal axial parenchyma were banded and diffuse, whereas paratracheal axial parenchyma were scanty and vasicentric (Fig 4[a]). Rays were slightly heterocellular consisting mainly of procumbent cells as shown in Fig 4(b). Multiseriate rays (Fig 4[c]) were generally two to four cells in width. The vessel measurement results shown in Table 4 were found to be in close proximity to the research conducted by Richter (2000).

Khaya ivorensis. Figure 5(a) revealed that the wood of *Khaya ivorensis* was diffuse-porous, with indistinct growth rings and absence of marginal parenchyma band. These characteristics were similar

Table 4. Quantitative wood anatomical characters of the five Meliaceae species achieved by Duncan's multiple range tests (the means carrying the same letter along a column are not significantly different at 0.05 level of significance). The first line represents the mean value followed by standard deviation, the second line gives the standard error, and the third line gives the range of values (minimum-maximum).

Species	Vessel area (μm^2)	Tangential diameter of vessel lumina (μm)	Vessel element length (μm)	Ray's height (μm)	Ray's width (μm)
Swietenia	12,374 ± 7129[a]	115 ± 40[a]	144 ± 40[a]	343 ± 112[a]	58 ± 28[b]
mahagoni	752	4	4	12	3
	1006-25,321	32-191	49-229	162-797	19-121
Swietenia	20,159 ± 9096[b]	148 ± 35[b]	172 ± 38[b]	417 ± 150[b]	65 ± 26[c]
macrophylla	959	4	4	16	3
	5037-38,956	53-236	68-254	91-1086	29-190
Cedrela odorata	48,873 ± 24,799[d]	236 ± 64[d]	260 ± 69[e]	369 ± 128[a]	49 ± 13[a]
	2614	7	7	13	2
	6510-115,713	76-400	90-408	133-785	22-76
Khaya ivorensis	31,465 ± 12,503[c]	182 ± 33[c]	206 ± 37[c]	484 ± 163[c]	67 ± 19[c]
	1318	4	4	17	2
	2754-66,071	83-254	96-309	219-987	27-107
Toona ciliata	35,722 ± 20,616[c]	185 ± 58[c]	225 ± 64[d]	379 ± 127[ab]	67 ± 14[c]
	2173	6	7	13	2
	5172-106,512	40-322	89-403	167-788	36-98

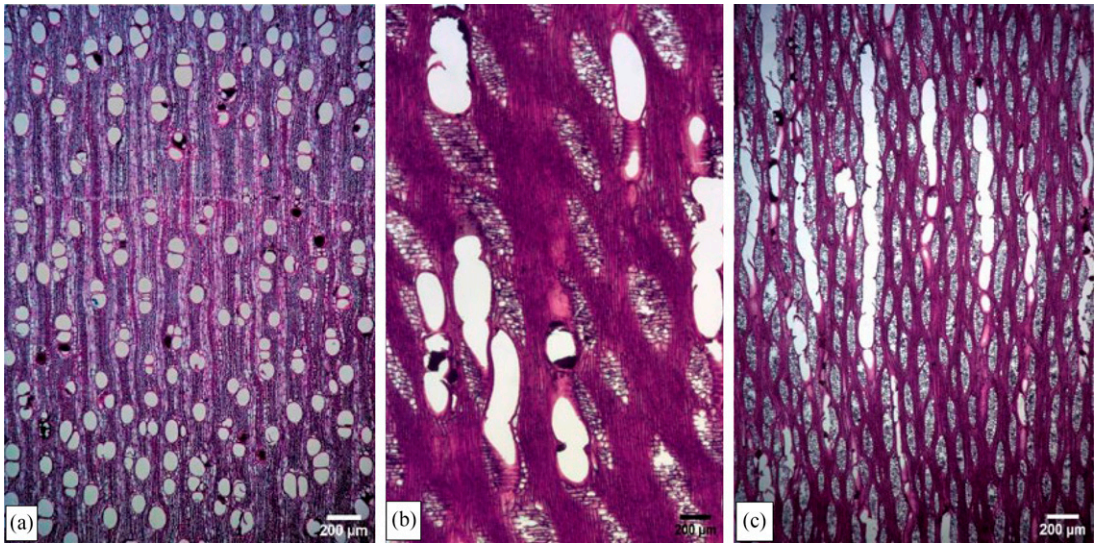


Figure 2. *Swietenia mahagoni* (Sample Id_539923). (a) Transverse section, 20× magnification, showing marginal parenchyma band, and scanty/vasicentric paratracheal axial parenchyma. (b) Radial section, 20× magnification, and heterocellular rays. (c) Tangential section, 20× magnification, multiseriate, and mostly storied rays.

to those reported by Inside Wood (2004 and onwards). Prismatic crystals were observed in upright and/or square ray cells, which is in agreement with Panshin (1933) and White and Gasson (2008). Paratracheal axial parenchyma were mainly vasicentric and also scanty. Rays were heterocellular in radial view (Fig 5[b]) and of two distinct sizes (ie uniseriate and multiseriate up to eight cells wide) in tangential view (Fig 5[c]). Two specimens initially labeled as "*Khaya ivorensis*"

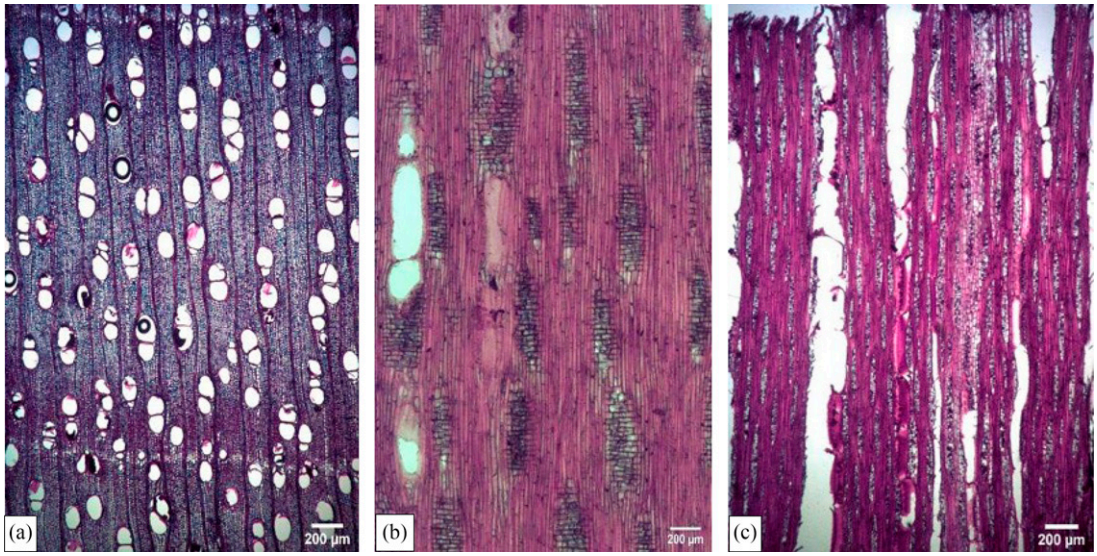


Figure 3. *Swietenia macrophylla* (Sample Id_539859). (a) Transverse section, 20× magnification, showing marginal parenchyma band, scanty/vasicentric, and diffuse axial parenchyma. (b) Radial section, 20× magnification, and heterocellular rays. (c) Tangential section, 20× magnification, multiseriate (≥tetraseriate), and mostly storied rays.

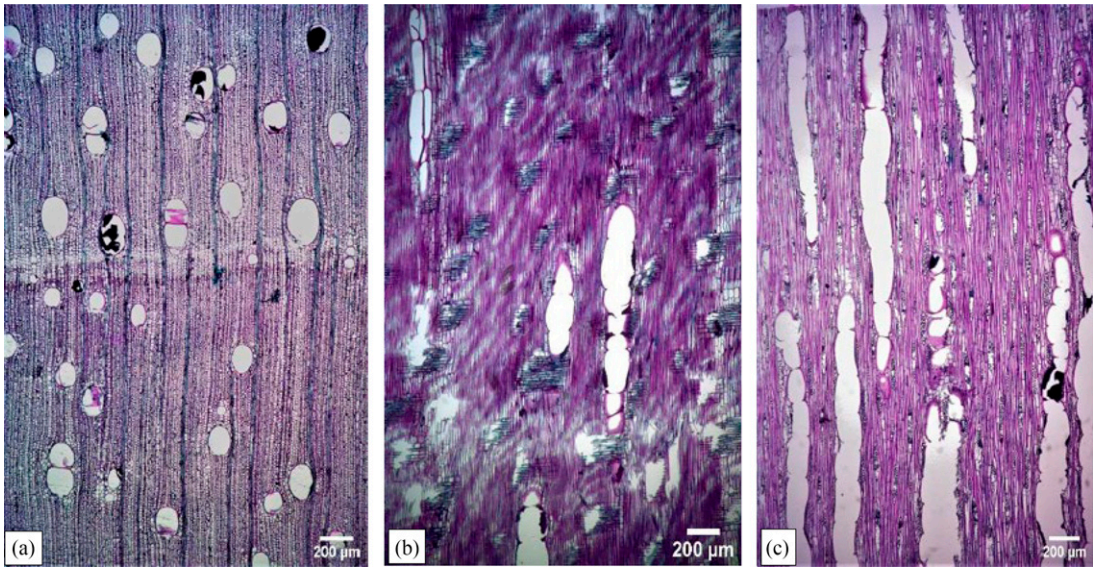


Figure 4. *Cedrela odorata* (Sample Id_542126). (a) Transverse section, 20× magnification, semiring-porous wood, marginal parenchyma band, diffuse apotracheal axial parenchyma, and vasicentric paratracheal axial parenchyma. (b) Radial section, 20× magnification, weakly heterocellular rays consisting mainly of procumbent cells, and simple perforation plate. (c) Tangential section, 20× magnification, commonly biseriate to triseriate rays.

were later found to belong to closely related *Entandrophragma* genus (based on microscopic wood anatomy observations). They were, therefore, removed from subsequent chemical analyses. This result highlights the importance of meticulous curation and the need for species verification using complementary techniques.

Toona ciliata. *Toona ciliata* was characterized by semiring-porous vessel distribution, with distinct growth ring boundaries delineated by marginal parenchyma bands and large earlywood vessels (Fig 6[a]). Perforation plate was simple, intervessel pit was alternate type, rays were heterocellular in radial section (Fig 6[b]), prismatic crystals were observed in axial parenchyma and ray cells (Fig 6[b]), rays were multiseriate in tangential view (Fig 6[c]), one to six cells in width, apotracheal axial parenchyma were diffused to diffuse-in-aggregates, and paratracheal axial parenchyma were vasicentric. Our quantitative results were similar to a previous study by Wood (2004), where mean tangential vessel lumina diameter of *Toona ciliata* was reported as 100-200 µm, mean vessel element length was less than or equal to 350 µm, vessel frequency per mm^2 was 5-20, and the average number of rays per mm was 4-12.

Our microscopic observations for the five Meliaceae species studied confirmed a close relationship among them as they share a large number of similar wood anatomical characteristics (Table 2). The characteristics agreed with previous studies (Kribs 1930; Pennington and Styles 1975; White and Gasson 2008; Oyedeji-Amusa et al 2020). Anatomical results were also consistent with the features documented in Inside Wood (2004 onwards) and Description Language for Taxonomy (DELTA) (https://www.delta-intkey.com), except for two features of *Cedrela odorata.* According to Inside Wood, the mean tangential vessel lumina diameter of five species falls within the range of 100-200 µm and this corresponds to our study except for *Cedrela odorata* (mean 236 µm). Further, in our study, the result for vessel element length (mean 260 µm) of *Cedrela odorata* complies with Delta intkey, however, slightly differs with Inside Wood (350-800 µm). These differences could be explained by age and environmental conditions of the sample tree, which are unknown for our xylaria samples.

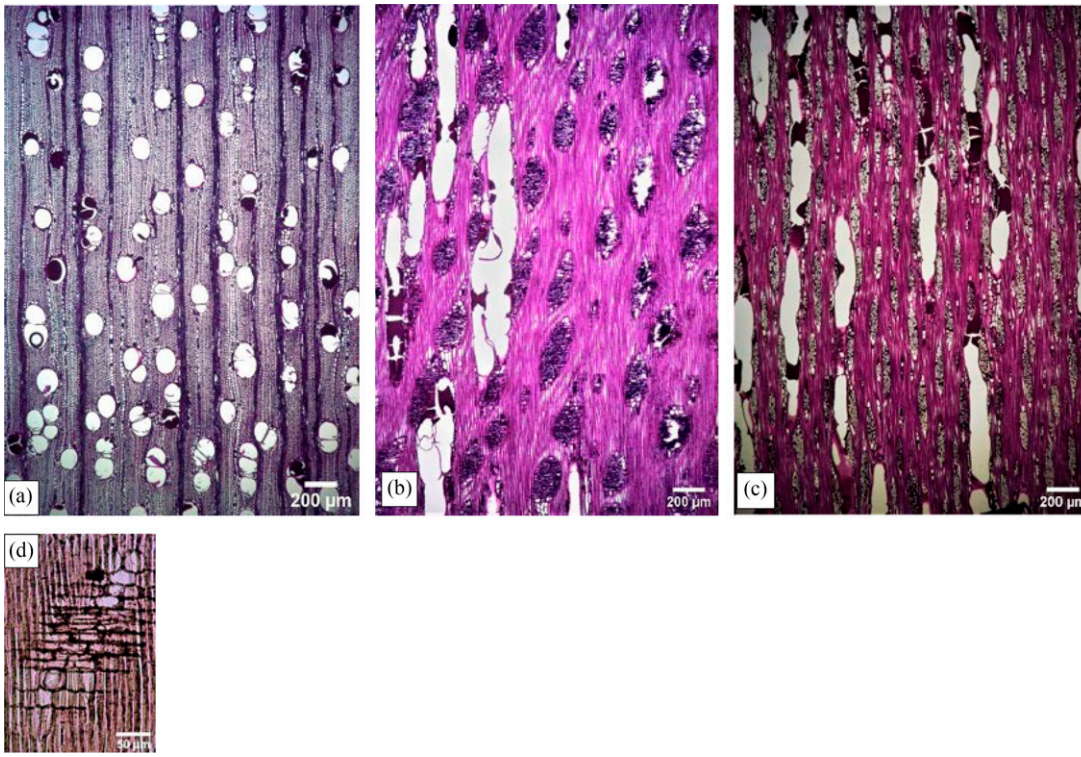


Figure 5. *Khaya ivorensis* (Sample Id_ 540728). (a) Transverse section, 20× magnification, and scanty/vasicentric paratracheal parenchyma. (b) Radial section, 20× magnification, heterocellular rays, and simple perforation plate. (c) Tangential section, 20× magnification, rays of two distinct sizes, uniseriate and multiseriate rays, and not storied. (d) Radial section, 40× magnification, prismatic crystals present in upright ray cells.

Species comparisons. According to our study, *Khaya ivorensis* can successfully be separated from the other four species by the presence of prismatic crystals present only in upright cells, and rays of two distinct sizes. The lack of banded apotracheal axial parenchyma in *Khaya* spp. also helps to differentiate it from *Swietenia* spp. as reported by White and Gasson (2008). *Cedrela odorata* and *Toona ciliata* can be distinguished from *Swietenia* species and *Khaya ivorensis* by their vessel distribution: *Cedrela odorata* and *Toona ciliata* have semiporous vessel arrangement while the other studied species have diffused porous vessel distribution. In addition, *Cedrela* and *Toona* woods are highly aromatic with a strong cedar-like scent, whereas the other species are much less aromatic in their chemical structure and thus, relatively odorless.

ANOVA among the five Meliaceae species showed that species are statistically different ($p < 0.05$) based on five measured wood anatomical characteristics (Table 3). The Duncan's multiple range test revealed that *Swietenia mahagoni* was significantly different from *Swietenia macrophylla* based for each of the five wood anatomical characteristics (Table 4). Moreover, *Cedrela odorata* was significantly different from *Toona ciliata* on the basis of four wood anatomical characteristics namely vessel area, tangential vessel lumina diameter, vessel element length, and ray's width (ie except for ray's height). *Khaya ivorensis* was not significantly different from *Toona ciliata* based on vessel area, tangential vessel lumina diameter, and ray's width (except for vessel element length and ray's height). However, *Khaya ivorensis* was significantly different from *Swietenia*

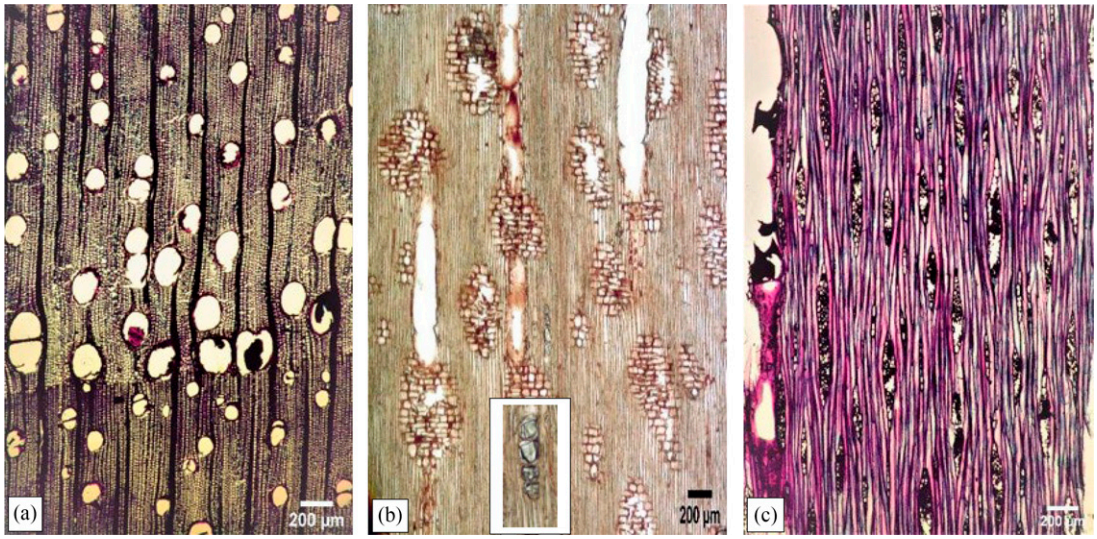


Figure 6. *Toona ciliata* (Sample Id_543406). (a) Transverse section, 20× magnification, semiring-porous wood, vasicentric axial paratracheal parenchyma, and diffuse and diffuse-in-aggregates apotracheal axial parenchyma. (b) Radial section, 20× magnification, heterocellular rays, and presence of prismatic crystals (top right). (c) Tangential section, 20× magnification, multiseriate rays, and simple perforation plate.

mahagoni, *Cedrela odorata*, and *Swietenia macrophylla* (except for ray's width).

White and Gasson (2008) reported that unless the geographic origin is known, differentiating *Cedrela odorata* from *Toona ciliata* is probably impractical. In our study, *Cedrela odorata* and *Toona ciliata* could not be separated based on qualitative wood anatomy, but they could be separated based on quantitative analysis (except for ray's height, Table 4). Similarly, our qualitative wood anatomical observations could not separate the two *Swietenia* species; however, our quantitative results could. Given the small number of wood specimens used in this study and the presence of substantial within species variability depending on environmental conditions and genotype (Baas and Miller 1985; Kite et al 2010; Gasson 2011; Tsumura et al 2011), this result should be taken with caution.

Machine-Vision Wood Anatomy Using XyloTron

A total of 15 wood samples were tested through XyloTron (Table 5). The three wood specimens of *Swietenia mahagoni* were identified as *Swietenia macrophylla* while two wood specimens of *Swietenia macrophylla* were identified as *Cedrela odorata* and only one wood sample was identified as *Swietenia* spp. All three wood specimens of *Cedrela odorata* were accurately identified. The two specimens of *Khaya ivorensis* were assigned to *Khaya ivorensis* with probability score of more than 50%, while the third one was assigned to *Swietenia macrophylla*. However, the tool associated *Toona ciliata* wood anatomy to *Cedrela odorata*, which is correct as these two species cannot be distinguished anatomically unless the geographic origin is known (White and Gasson 2008). At the time of analysis, the XyloTron image database comprised between 21 and 79 reference wood specimens per species (Table 5 personal communication with J. Hermanson 2021). There was no algorithm specifically developed for *Toona ciliata* (no images were available in the reference database). Nevertheless, we presented our 15 wood samples as "unknowns" to the XyloTron (Table 5) to evaluate if identification responses would be similar to that obtained through traditional wood anatomy. At present, the open-source system is in development and additional images of

Table 5. Wood specimens tested through the XyloTron machine-vision system of CFS (and traditional wood anatomy). For each sample, the XyloTron algorithms provided a conformity score (%) with the anatomically most similar tree species found in the image reference database.

Species	Sample Id	XyloTron taxonomic prediction (probability in %)	*N*
Swietenia mahagoni	539923	*Swietenia macrophylla* (98%)	21
Swietenia mahagoni	540153	*Swietenia macrophylla* (87%)	
Swietenia mahagoni	543940	*Swietenia macrophylla* (93%)	
Swietenia macrophylla	539857	*Swietenia macrophylla* (63%)	59
Swietenia macrophylla	539859	*Cedrela odorata* (78%), *Swietenia macrophylla* (3.5%), and species of other family (18.5%)	
Swietenia macrophylla	541144	*Cedrela odorata* (76%), *Swietenia macrophylla* (2.6%)	
Cedrela odorata	542126	*Cedrela odorata* (68%)	41
Cedrela odorata	542086	*Cedrela odorata* (98%)	
Cedrela odorata	542087	*Cedrela odorata* (95%)	
Khaya ivorensis	540730	*Swietenia macrophylla* (82%), *Cedrela odorata* (16%) and *Khaya ivorensis* (2%)	79
Khaya ivorensis	540729	*Khaya ivorensis* (53%)	
Khaya ivorensis	540728	*Khaya ivorensis* (74%)	
Toona ciliata	543406	*Cedrela odorata* (91%)	0[a]
Toona ciliata	541871	*Cedrela odorata* (97%)	
Toona ciliata	541491	*Cedrela odorata* (64%)	

[a]Note: *Toona ciliata* was not included in the XyloTron reference database when the test was carried out.
N indicates the current number of reference wood specimens used by the XyloTron algorithm for species identification (reference: Dr. J. Hermanson, XyloTron developer, pers. comm. 2021).

reference (vouchered) wood samples are being continuously acquired by different laboratories around the world to improve existing models and expand their ability to identify new species. The tool is also being tested for field inspections. With a larger number of reference wood specimens, XyloTron would be more well-suited to capture macroscopic images of wood sample with interesting macroscopic variation. Indeed, it is desirable to have a minimum of 20-30 samples for wood identification purposes (personal communication with J. Hermanson 2021), but the larger the number of individuals sampled, the better the wood identification will be.

Metabolomic Profiling Using GC×GC-TOFMS

The GC×GC-TOFMS analysis revealed the presence of nearly 1831 chemical compounds in the wood samples of five Meliaceae species (data not shown). Considering the diversity of biological compounds, there is a need for multivariate statistical analysis to separate chemically similar wood species.

Figure 7 shows the scores plot and biplot obtained from PCA. The PCA model was built using 45 selected features and allowed the visualization of the differences among clusters of *Swietenia mahagoni*, *Swietenia macrophylla*, *Cedrela odorata*, *Khaya ivorensis*, and *Toona ciliata* (Fig 7). The list of compounds of the biplot are shown in Table A1, whereas the heat map of the relative abundance of these compounds is shown in Fig 8. Results revealed the presence of two main groups (see cladogram at the top of Fig 8): on the left, Group I that includes the *Khaya* (green cell) and *Swietenia* species (dark blue and light blue cells); and on the right, Group II with the *Toona* (pink cell) and *Cedrela* (orange cell) species. This cluster is congruent with the phylogeny of Meliaceae, which is based on chloroplast and nuclear DNA regions (Muellner et al 2003). Effectively, the two groups of species (Group I and Group II)

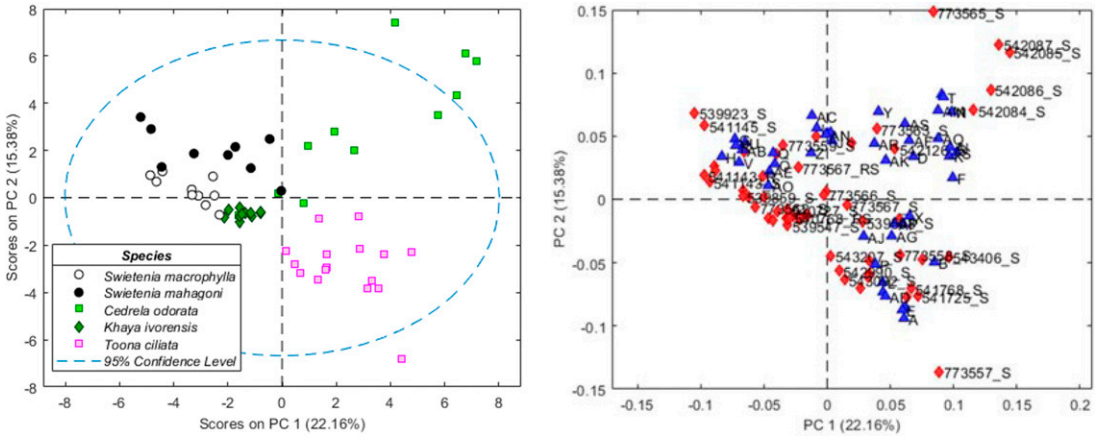


Figure 7. Scores plot (left) and biplot (right) obtained from principal component analysis of 51 wood samples (S), two replicate aliquots (R), and replicate samples (RS). The PCA model was built using 45 selected features to visualize the differences between clusters of *Swietenia mahagoni, Swietenia macrophylla, Cedrela odorata, Khaya ivorensis*, and *Toona ciliata*.

observed in our study and based on metabolic profiles correspond well to the two Swietenioideae subgroups. The compounds that distinguish the two groups are Sesquiterpenol C and alpha-Muurolene.

Chemical Profiles Among Genera

Group I (Khaya vs Swietenia). We observe that *Swietenia* samples comprised much larger amounts of triterpenoid (B, C, and E) and C28 sterol A compared with *Khaya* samples. Other metabolites (ie Metabolites 173 and 1273) were also more abundant in *Swietenia*, and may be selected as potential markers for future identification. One sample initially labeled as *Swietenia* (ie 539481 in Table 1) grouped as *Khaya* based on metabolite profile. We verified the wood anatomy of the specimen using light microscopy and Inside Wood (2004 and onwards) and confirmed its belonging to the *Khaya* genus (ie sample was mislabeled). Gasson (2011) detailed that customs officers usually have trouble distinguishing a shipment of reddish-brown-colored wood and mark it as *Swietenia* sp., *Khaya*, or *Entandrophragma* from Africa or a *Dipterocarp* from Southeast Asia

Group II *(Toona* vs *Cedrela).* We observe that *Toona* samples comprised much larger amounts of "C28 sterol C" and "C27 sterol B" compared with *Cedrela*. In this study, triterpenoid A was very abundant in *Toona* (practically unique) and could potentially be used to distinguish from *Cedrela* (and from the other study species). According to Muellner et al (2003), morphological, phytochemical, and molecular studies indicate a close relationship between the genera *Cedrela* and *Toona*, leading to establish a sister taxon forming a monophyletic clade within Swietenioideae, justifying their positioning within the same tribe.

In comparison with *Toona ciliata*, *Cedrela odorata* was found to be rich in sesquiterpenes, sesquiterpenols, sesquiterpene oxides, and sterols (see example in Figs A3-3 and A3-5). Suarez et al (2018) also stated that although the chemical compositions of *Cedrela odorata* are remarkably different, they all are dominated by sesquiterpene hydrocarbons regardless of the plant tissue or geographical location. Sesquiterpenes are secondary metabolites mainly found in essential plant oil. Further, several studies have revealed that sesquiterpenes were found in abundance in oil composition of *Cedrela* species (Campos et al 1991; Nogueira et al 2020).

Cedrela odorata wood specimens were scattered suggesting the presence of an outlier (Fig 7 left).

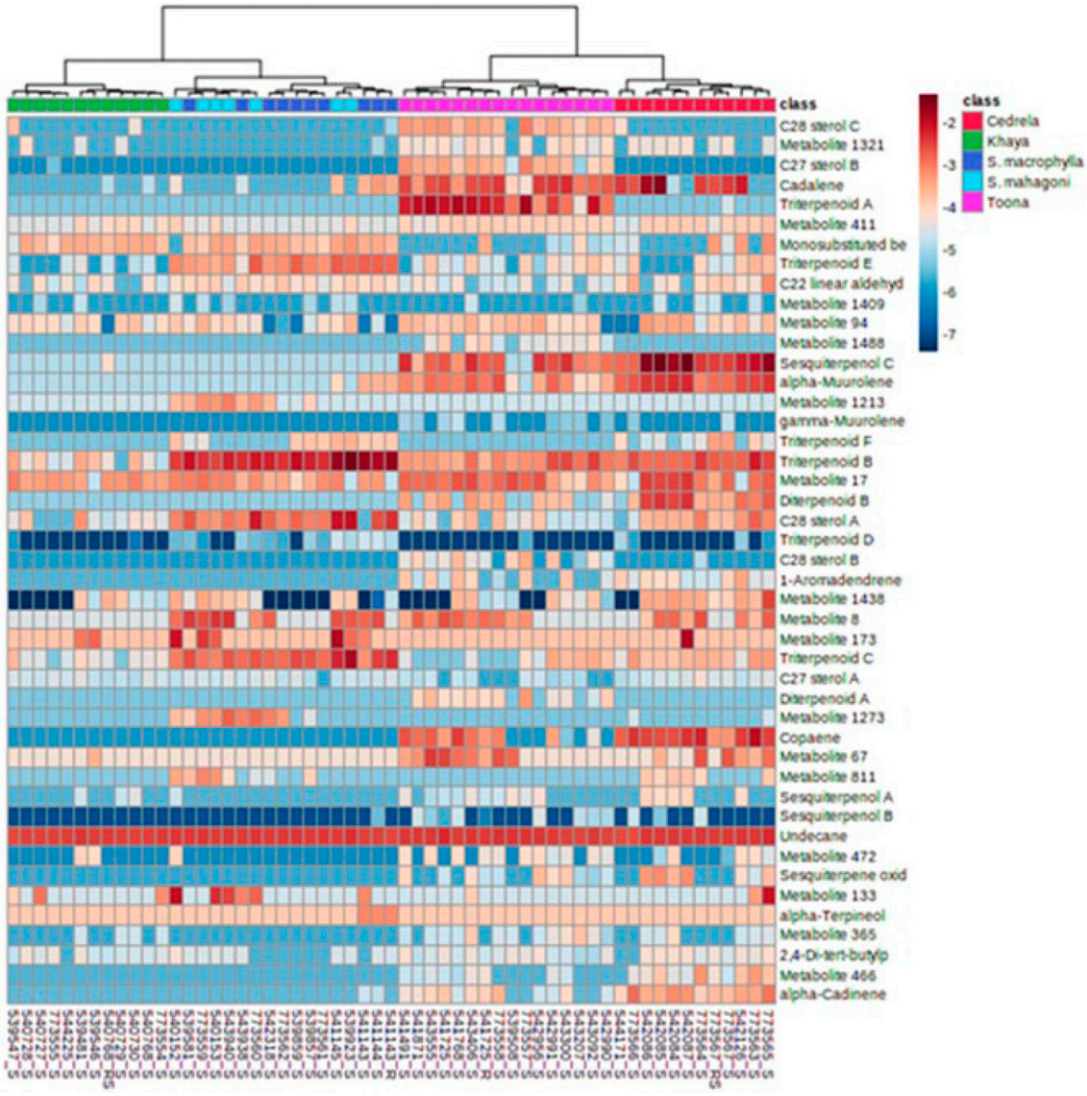


Figure 8. Heat map using 45 selected features from chemical profiles of *Swietenia mahagoni*, *Swietenia macrophylla*, *Cedrela odorata*, *Khaya ivorensis*, and *Toona ciliata*. Data were log-scaled to emphasize differences in the relative amounts of compounds across different wood species.

We verified the wood anatomical characteristics and found them similar to *Cedrela odorata* and *Cedrela fissilis* according to Inside Wood (2004 and onwards). However, we found that the *Cedrela odorata* specimen had been mistakenly sampled in sapwood (light beige) instead of heartwood (brown color), which may explain the difference in chemical signatures. It has been reported that heartwood contains more extractives than sapwood (Hillis 1971; Miranda et al 2006) although the content may decrease from the outer to the inner heartwood (Wilkes 1984). Another possible explanation is that some of these outliers could be other closely related species or hybrids between the various *Cedrela* species inhabiting these regions (see S1).

Table 6. Comparison of resolution (genus vs species) and degree of convergence between the analytical methods tested in this wood identification study.

Species	Can wood anatomy results differentiate species? (Yes/No)		Correspond to Inside Wood description? (Yes/No)	Can XyloTron differentiate species? (Yes/No)	Can major chemotypes alone differentiate species? (Yes/No)	Can chemotypes (PCA) analysis differentiate species? (Yes/No)
	Qualitative evaluation	Quantitative evaluation[a]				
Swietenia mahagoni	No, similar to *Swietenia macrophylla*	Yes, significantly different from *Swietenia macrophylla* on basis of vessel area, tangential vessel lumina diameter, vessel element length, ray's height, and ray's width	Yes (except vessel element length)	No, only at the genus level	No, only to genus level (almost similar chemotype to *Swietenia macrophylla*)	Yes
Swietenia macrophylla	No, similar to *Swietenia mahagoni*, both are diffuse porous	Yes, significantly different from *Swietenia mahagoni* on basis of vessel area, tangential vessel lumina diameter, vessel element length, ray's height, and ray's width	Yes (except vessel element length)	Yes, but occasionally	No, only to genus level	Yes
Cedrela odorata	No, similar to *Toona ciliata*	Yes, significantly different from *Toona ciliata* on basis of vessel area, tangential vessel lumina diameter, vessel element length, and ray's width	Yes (except tangential vessel lumina diameter and vessel element length)	Yes	Yes, richer in sesquiterpenes, sesquiterpene oxides than other four species	Yes

(*continued*)

Table 6. Comparison of resolution (genus vs species) and degree of convergence between the analytical methods tested in this wood identification study. (cont.)

Species	Can wood anatomy results differentiate species? (Yes/No)		Correspond to Inside Wood description? (Yes/No)	Can XyloTron differentiate species? (Yes/No)	Can major chemotypes alone differentiate species? (Yes/No)	Can chemotypes (PCA) analysis differentiate species? (Yes/No)
	Qualitative evaluation	Quantitative evaluation[a]				
Khaya ivorensis	Yes, *Khaya ivorensis* differs from other species by presence of indistinct growth rings and prismatic crystals only in upright cells	Yes, significantly different from *Swietenia mahagoni*, *Cedrela odorata* on basis of vessel area, tangential vessel lumina diameter, vessel element length, ray's height, and ray's width; *Swietenia Macrophylla* on basis of vessel area, tangential vessel lumina diameter, vessel element length, and ray's height	Yes (except vessel element length)	Yes, but occasionally	No, usually to genus level	Yes
Toona ciliata	No, similar to *Cedrela odorata*, both are semiporous	No, not significantly different from *Khaya ivorensis* on basis of vessel area, tangential vessel lumina diameter, and ray's width	Yes	No[a] (reference samples not available in XyloTron)	No, usually to genus level	Yes

[a]Vessel area, tangential vessel lumina diameter, vessel element length, ray's height, and ray's width.

The chemical composition of any plant species is known to be influenced by the genotype, environmental, and agronomic conditions (Rota et al 2008; Kizil 2010). This highlights the importance of collecting physical, photographic and accurate geolocation data for each reference sample that are used to build the databases. Even with accurate sample data, multiple techniques might be necessary to correctly identify a wood sample, since despite all care, collection specimens may have been mislabeled during the acquisition or curation process.

Chemical Profiles Within Genus

Swietenia mahagoni vs *Swietenia macrophylla*. Figure 8 indicates that the two *Swietenia* species are chemically closely related. The heat map reveals that Metabolite 1213 and Metabolite 1273 could potentially be used to separate *Swietenia mahagoni* from *Swietenia macrophylla*. But the separation is not as clear as for the above "between genera" comparisons. Indeed, we see that the groupings of samples are variable, as illustrated by dark blue samples (*Swietenia mahagoni*) intermingled with light blue samples (*Swietenia macrophylla*). This can be attributed to the fact that *Swietenia* species can potentially hybridize. As such, species delineation becomes more difficult and the notion of species might be revised in a context of CITES enforcement (meaning that it might not always be possible to achieve species-level identification because the definition of species is diffuse). The within-genus similarity in the main chemical profiles is also illustrated in Figs A3-1 and A3-2. In this study, these two species were successfully separated through GC×GC-TOFMS with PCA visualization acquiring a total variance of 44.48% as represented (not shown). The PC1 and PC2 captured 30.69% and 13.79% variance, respectively, allowing for wood species separation.

Deklerck (2019) reported that *Swietenia mahagoni* and *Swietenia macrophylla* represented a similar chemotype. According to our study, cycloeucalenol acetate (9,19-cycloergost-24[28]-en-3-ol, 4,14- dimethyl-, acetate, [3β,4α,5α]) is one of the dominant compounds found in *Swietenia* species and represents between 0.3% and 18.75% of the total spectra in *Swietenia mahagoni* and between 0% and 11.2% in *Swietenia macrophylla* (data not shown). Amorós-Marín et al (1959) stated that cycloeucalenol was isolated for the first time from West Indian Mahogany wood through petroleum ether extraction process. Other chemical compounds detected from heartwood samples of *Swietenia mahagoni* were ketone (2-acetyl-3-methoxyphenol); aldehyde (eg syringaldehyde), and so on and this result was supported by Asmara (2018). Moreover, chemical constituents extracted from *Swietenia macrophylla* were ketones (eg 2-acetyl-3-methoxyphenol) and sterols (eg β-Sitosterol) (not shown).

The untargeted metabolomics approach that was applied on five mahogany species of interest aimed to identify as many chemical compounds as possible for taxonomic identification. However, in the context of operational wood identification, it is not necessary to identify all chemical compounds present in wood, but only those that are discriminately consistent for identifying a tree species. This approach using a GC×GC-TOFMS paired with cluster analysis has the potential to increase the number of metabolites measured due to its sensitivity. This strategy could lead to the discrimination of species using a targeted approach for some species of interest. To this end, there is a need to continue the research with additional reference samples to strengthen the link between various analytical identification methods, including untargeted and targeted chemotyping approaches and improve identification efficiency. An overview of the results obtained through the various analytical methods used in this wood identification research project can be found in Table 6.

CONCLUSIONS

In this study, we compared traditional wood anatomy with machine imaging (XyloTron) and chemical (metabolite chemotyping) identification methods. Fifty wood samples were sourced from two xylaria and 15 of them were analyzed with the three methods. Microscopic wood anatomy,

which is the most widely used method for wood identification, was conducted by observing and measuring qualitative and quantitative wood anatomical characteristics, respectively. Evaluation of qualitative anatomical features alone failed to show clear differences among *Cedrela odorata* and *Toona ciliata* and the two *Swietenia* species. *Khaya ivorensis* could be differentiated from the other four species by the presence of rays of two distinct sizes, and the presence of prismatic crystals only in upright cells.

Quantitative anatomy analysis could distinguish *Swietenia mahagoni* from *Swietenia macrophylla* and *Cedrela odorata* from *Toona ciliata*. However, *Khaya ivorensis* was not significantly different from *Toona ciliata* for three of the five measured anatomical features (vessel area, tangential vessel lumina diameter, and ray's width). These quantitative results are counterintuitive knowing that traditional wood anatomy can easily identify trees to the genus level. The considerable variations observed in the morphology of vessel area, tangential vessel lumina diameter, vessel element length, ray's height, and ray's width through quantitative evaluations may not always be genus- or species-specific, ie be more reflective of contrasted growth environments. Overall, the value of additional quantitative features remains uncertain when applied to only a few samples per species.

XyloTron analysis of macroscopic end grain images achieved species-level identification for *Cedrela odorata* and for *Khaya ivorensis* in some cases. Similar to the traditional wood anatomy method, the machine-vision method could not separate *Swietenia mahagoni* from *Swietenia macrophylla.* The fact that XyloTron identified the *Toona ciliata* specimens (not included in the image database) as *Cedrela odorata* (included in the database) makes sense since the two species are closely related and cannot be distinguished with certainty using traditional wood anatomy (White and Gasson 2008). Hence, knowing the geographic origin becomes very important in a context of forensic analysis.

Chemical profiling using GC×GC-TOFMS was also used in this study for taxonomic identification. Separating closely related taxa based on chemical profile was challenging since some species shared similar chemotype, especially between the two congeneric species *Swietenia mahagoni* and *Swietenia macrophylla*. However, GC×GC-TOFMS, analyzed by cluster analysis and expressed using multivariate statistical analysis (PCA) was found to be an effective approach to separate these five closely related species. Hence, the GC×GC-TOFMS and PCA visualization method, in combination with wood anatomy, shows good potential to produce forensic identification of CITES-listed species, in support of law enforcement officials verifying the legality of timbers in trade. This research may contribute to the global effort to curb illegal logging and trade of forest products.

Limitations and Recommendations of Study

Taxonomically validated and geographically referenced wood standards are essential for wood forensics. In this study, we had a very limited number of reference collection samples for each species and some samples were only geo-referenced at the country level. Lack of accurate geo-referencing information may limit the interpretation of the presence of outliers amongst the samples tested, whose presence could be due to genetic factors (such as hybridization between closely related species) and/or environmental factors. Hence, there is a need for more research projects to support the continuation of wood collections (e.g. Williams et al 2020), updating botanical information of the xylarium specimens, coding the specimens according to the criteria of the International Association of Wood Anatomists, improving microscopic and macroscopic image material for identification keys and to assess the wood density, and its variability based on xylarium specimens. Finally, there is a need to develop adequate and complete reference databases including anatomical and metabolite profiling approaches that are supported by molecular phylogeny or population genetics approach. Then, this information can be widely used to combatting illegal logging by empowering law enforcement officers to make easy field triage of suspected wood shipments.

REFERENCES

André T, Lemes MR, Grogan J, Gribel R (2008) Post-logging loss of genetic diversity in a mahogany (*Swietenia macrophylla* King, Meliaceae) population in Brazilian Amazonia. For Ecol Mgmt 255(2):340-345.

Armstrong MS, de la Mata AP, Harynuk JJ (2021) An efficient and accurate numerical determination of the cluster resolution metric in two dimensions. J Chemometr 35(7-8):1-13.

Asmara AP (2018) A preliminary study of investigating of compound group contained in ethanolic extract of mahagony (*Swietenia mahagoni* L. Jacq.) seeds related to α-glucosidase inhibition. Jurnal Natural 18(2):49-56.

Baas P, Miller RB (1985) Functional and ecological wood anatomy some introductory comments. IAWA J 6(4): 281-282.

Beens J, Udo AT (2005) Comprehensive two-dimensional gas chromatography—A powerful and versatile technique. Analyst (Lond) 130(2):123-127.

Bouka GUD, Doumenge C, Ekué MRM, Daïnou K, Florence J, Degen B, Loumeto JJ, Doyle McKey D, Hardy OJ (2022) *Khaya* revisited: Genetic markers and morphological analysis reveal six species in the widespread taxon *K. anthotheca.* Taxon 71:1-19.

Brunswick P, Cuthbertson D, Yan J, Chua CC, Duchesne I, Isabel N, Evans PD, Gasson P, Kite G, Bruno J, van Aggelen G, Shang D (2021) A practical study of CITES wood species identification by untargeted DART/QTOF, GC/QTOF and LC/QTOF together with machine learning processes and statistical analysis. Environ Adv 5: 100089.

Campos AM, Oliveira FS, Machado MIL, Matos FJA, Braz-Filho R (1991) Triterpenes from *Cedrela odorata*. Phytochemistry 30(4):1225-1229.

Chave J, Muller-Landau HC, Baker TR, Easdale TA, ter Steege H, Webb CO (2006) Regional and phylogenetic variation of wood density across 2456 Neotropical tree species. Ecol Appl 16(6):2356-2367.

CITES (2021) CITES Trade Database 2021. https://cites.org/eng.

Cody RB, Dane AJ, Dawson-Andoh B, Adedipe EO, 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^{TM}) and time-of-flight mass spectrometry. J Anal Appl Pyrolysis 95:134-137.

Cody RB, Laramée JA, Durst HD (2005) Versatile new ion source for the analysis of materials in open air under ambient conditions. Anal Chem 77(8):2297-2302.

Danquah JA, Appiah M, Osman A, Pappinen A (2019) Geographic distribution of global economic important mahogany complex: A review. Annu Res Rev Biol 34(3): 1-22.

De Hoffmann E, Stroobant V (2007) Mass spectrometry: Principles and applications. John Wiley & Sons, West Sussex, England.

Deklerck V (2019) National treasure: Valorisation of the Federal Xylarium in Belgium for timber identification and wood technology. PhD thesis, Ghent University.

Donaldson LA (1984) Wood anatomy of five exotic hardwoods grown in Western Samoa. N Z J For Sci 14(3): 305-318.

Dordel J, Simard SW, Bauhus J, Seely B, Pozas LJ, Prescott C, Hampel H (2010) Trade-offs among establishment success, stem morphology and productivity of underplanted *Toona ciliata*: Effects of nurse-species and thinning density. For Ecol Mgmt 259(9):1846-1855.

Dormontt EE, Boner M, Braun B, Breulmann G, Degen B, Espinoza E, Gardner S, Guillery P, Hermanson JC, Koch G, Lee SL, Kanashiro M, Rimbawanto A, Thomas D, Wiedenhoeft AC, Yin Y, Zahnen J, Lowe AJ (2015) Forensic timber identification: It's time to integrate disciplines to combat illegal logging. Biol Conserv 191:790-798.

Estrada-Contreras I, Equihua M, Laborde J, Meyer ME, Sánchez-Velásquez LR (2016) Current and future distribution of the tropical tree *Cedrela odorata* L. in Mexico under climate change scenarios using MaxLike. PLOS ONE 11(10):e0164178.

Evans PD, Mundo IA, Wiemann MC, Chavarria GD, McClure PJ, Voin D, Espinoza EO (2017) Identification of selected CITES-protected Araucariaceae using DART TOFMS. IAWA J 38(2):266-281.

Figueroa Colon J (1994) An assessment of the distribution and status of big-leaf mahogany (*Swietenia macrophylla* King). Puerto Rico Conservation Foundation and USDA Forest Service, International Institute of Tropical Forestry, Rio Piodras, Puerto Rico.

Finch KN, Jones FA, Cronn RC (2019) Genomic resources for the Neotropical tree genus Cedrela (Meliaceae) and its relatives. BMC Genom 20:58:1-17.

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(2):137-154.

Grogan J, Barreto P (2005) Big-leaf mahogany on CITES Appendix II: Big challenge, big opportunity. Conserv Biol 19(3):973-976.

Gullison R, Rice R, Blundell A (2000) "Marketing" species conservation. Nature 404(6781):923-924.

Hansen CP, Treue T (2008) Assessing illegal logging in Ghana. Int Rev 10(4):573-590.

Hermanson JC, Dostal D, Destree JC, Wiedenhoeft AC (2019) The XyloScope—A field-deployable macroscopic digital imaging device for wood. Research Note FPL-RN-0367. 1-19.

Hermanson JC, Wiedenhoeft AC (2011) A brief review of machine vision in the context of automated wood identification systems. IAWA J 32(2):233-250.

Hillis W (1971) Distribution, properties and formation of some wood extractives. Wood Sci Technol 5(4):272-289.

IAWA Committee (1989). IAWA list of microscopic features for hardwood identification. Pages 219-332 *in* EA Wheeler, P Baas, and PE Gasson, eds., Volume 10.

IAWA Bull n.s, National Herbarium of the Netherlands, Leiden.

Inside Wood (2004) Inside Wood database. North Carolina State University. https://insidewood.lib.ncsu.edu/search;jsessionid=tR8Z76KkMC_AV24VsD0CjFKukY-D0uMixG2IClya?0

IUCN (2015) The IUCN red list of threatened species. https://www.iucnredlist.org/

IUCN (2021) The IUCN red list of threatened species. https://www.iucnredlist.org/

Julkunen-Tiitto R (1989) Phenolic constituents of *Salix*: A chemotaxonomic survey of further Finnish species. Phytochemistry 28(8):2115-2125.

Kite GC, Green PW, Veitch NC, Groves MC, Gasson PE, Simmonds MS (2010) Dalnigrin, a neoflavonoid marker for the identification of Brazilian rosewood (*Dalbergia nigra*) in CITES enforcement. Phytochemistry 71(10): 1122-1131.

Kitin P, Espinoza E, Beeckman H, Abe H, McClure PJ (2021) Direct analysis in real-time (DART) time-of-flight mass spectrometry (TOFMS) of wood reveals distinct chemical signatures of two species of *Afzelia*. Ann Sci 78(2):1-14.

Kizil S (2010) Determination of essential oil variations of *Thymbra spicata var. spicata L.* naturally growing in the wild flora of East Mediterranean and Southeastern Anatolia regions of Turkey. Ind Crop Prod 32(3):593-600.

Koch G, Haag V, Heinz I, Richter H-G, Schmitt U (2015) Control of internationally traded timber—The role of macroscopic and microscopic wood identification against illegal logging. J Forensics Res 6(6):1-4.

Kometter RF, Martinez M, Blundell AG, Gullison RE, Steininger MK, Rice RE (2004) Impacts of unsustainable mahogany logging in Bolivia and Peru. Ecol Soc 9(1):1-12. http://www.ecologyandsociety.org/vol9/iss1/art12/

Kribs DA (1930) Comparative anatomy of the woods of the Meliaceae. Am J Bot 17(8):724-738.

Lancaster C, 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(9):1147-1156.

Laurance WF (1999) Reflections on the tropical deforestation crisis. Biol Conserv 91(2-3):109-117.

Llerena S, Jadán M, Proaño K (2012) Filogeografía molecular de las especies Cedrela odorata, Cedrela fissilisy Cedrela montana (Meliaceae) del ecuador mediante el uso de genes cloroplastídicos cpADN y espaciadores internos transcritos (ITS). Tesis de pregrado, Laboratorio de Biotecnología Vegetal, Escuela Politécnica.

Louppe DA, Oteng-Amoako BM (2008) Plant resources of tropical Africa 7(1). Timbers 1. PROTA Foundation, Backhuys Publishers, Wageningen.

Mabberley DJ (2017) Mabberley's plant-book: A portable dictionary of plants, their classification and uses, 4th edition. Cambridge University Press, Wadham College, Oxford.

Mabberley DJ, Pannell CM, Sing A (1995) Flora Malesiana: Series I. Spermatophyta Volume 12, part 1. Rijksherbarium, Foundation Flora Malesiana, Meliaceae. Publications Department, Rijksherbarium / Hortus Botanicus, The Netherlands.

Mayhew JE, Newton AC (1998) The silviculture of mahogany: CABI Publishing, Wallington, UK.

McClure PJ, Chavarria GD, Espinoza E (2015) Metabolic chemotypes of CITES protected *Dalbergia* timbers from Africa, Madagascar, and Asia. Rapid Commun Mass Spectrom 29(9):783-788.

Miranda I, Gominho J, Lourenço A, Pereira H (2006) The influence of irrigation and fertilization on heartwood and sapwood contents in 18-year-old eucalyptus globulus trees. Can J Res 36(10):2675-2683.

Muellner AN, Samuel R, Johnson SA, Cheek M, Pennington TD, Chase MW (2003) Molecular phylogenetics of Meliaceae (*Sapindales*) based on nuclear and plastid DNA sequences. Am J Bot 90(3):471-480.

Muellner AN, Savolainen V, Samuel R, Chase MW (2006) The mahogany family "out-of-Africa": Divergence time estimation, global biogeographic patterns inferred from plastid rbcL DNA sequences, extant, and fossil distribution of diversity. Mol Phylogenet Evol 40(1):236-250.

Navarro C, Wilson J, Gillies A, Hernandez M (2003) A new Mesoamerican collection of big-leaf mahogany, in Big-Leaf Mahogany. Springer, New York, NY.

Nogueira TSR, Passos M, Nascimento LPS, Arantes MB, Monteiro NO, Boeno SI, de Carvalho Junior A, Azevedo O, Terra W, Vieira MGC, Braz-Filho R, Curcino Vieira IJ (2020) Chemical compounds and biologic activities: A review of *Cedrela* genus. Molecules 25(22):1-34.

O'Neill GA, Dawson I, Sotelo-Montes C, Guarino L, Guariguata M, Current D, Weber JC (2001) Strategies for genetic conservation of trees in the Peruvian Amazon. Biodivers Conserv 10(6):837-850.

Oyedeji-Amusa MO, Van Wyk B-E, Oskolski AA (2020) Wood anatomy of South African Meliaceae: Evolutionary and ecological implications. Bot J Linn Soc 193(2): 165-179.

Panshin AJ (1933) Comparative anatomy of the woods of the Meliaceae, sub-family Swietenioideae. Am J Bot 20(10):638-668.

Patiño Valera F, Kageyama PY, Linares BC, Navarro Pereira C (1996) Genetic resources of *Swietenia* and *Cedrela* in the Neotropics: Proposals for coordinated action. https://www.fao.org/3/AD111E/AD111E00.htm#TOC

Pennington TD (1981) A monograph of Neotropical Meliaceae (with accounts of the subfamily Swietenioideae by BT Styles and the chemotaxonomy by DAH Taylor).

Pennington TD, Styles BT (1975) A generic monograph of the Meliaceae. Blumea: Biodiversity, Evolution and Biogeography of Plants 22(3):419-540.

Pierce KM, Hope JL, Hoggard JC, Synovec RE (2006) A principal component analysis based method to discover chemical differences in comprehensive two-dimensional

gas chromatography with time-of-flight mass spectrometry (GC×GC-TOFMS) separations of metabolites in plant samples. Talanta 70(4):797-804.
Ravindran P, Thompson BJ, Soares RK, Wiedenhoeft AC (2020) The XyloTron: Flexible, open-source, image-based macroscopic field identification of wood products. Front Plant Sci 11:1-8.
Richter H (2000) Commercial timbers: Descriptions, illustrations, identification, and information retrieval. Version: 25 June 2009. http://delta-intkey.com/wood/index.htm
Rodan BD, Newton AC, Veríssimo A (1992) Mahogany conservation: Status and policy initiatives. Environ Conserv 19(4):331-338.
Rota MC, Herrera A, Martinez RM, Sotomayor JA, Jordan MJ (2008) Antimicrobial activity and chemical composition of *Thymus vulgaris*, *Thymus Zygis* and *Thymus Hyemalis* essential oils. Food Control 19(7):681-687.
Sexton GJ, Frere CH, Kalinganire A, Uwamariya A, Lowe AJ, Godwin ID, Prentis PJ, Dieters MJ (2015) Influence of putative forest refugia and biogeographic barriers on the level and distribution of genetic variation in an African savannah tree, *Khaya senegalensis* (Desr.) A. Juss. Tree Genet Genomes 11(5):1-15. https://link.springer.com/article/10.1007/s11295-015-0933-3
Shang D, Brunswick P, Yan J, Bruno J, Duchesne I, Isabel N, VanAggelen G, Kim M, Evans PD (2020) Chemotyping and identification of protected *Dalbergia* timber using gas chromatography quadrupole time of flight mass spectrometry. J Chromatogr A 1615:460775.
Silva JL, Bordalo R, Pissara J 2020 Wood identification: An overview of the current and past methods. ECR–Estudos de Conservação e Restauro 12:45-68.
Sinkov NA, Harynuk JJ (2011) Cluster resolution: A metric for automated, objective and optimized feature selection in chemometric modeling. Talanta 83(4):1079-1087.
Sinkov NA, Harynuk JJ (2013) Three-dimensional cluster resolution for guiding automatic chemometric model optimization. Talanta 103:252-259.
Suarez AV, Satyal P, Setzer WN (2018) Volatile components of the wood of Spanish cedar, *Cedrela odorata*, from Costa Rica. Am J Essent Oil Nat Prod 6(3):27-30.
Sun YA, Zhang H, Li Z, Yu W, Zhao Z, Wang K, Zhang M, Wang J (2020) Determination and comparison of agarwood from different origins by comprehensive two-dimensional gas chromatography—Quadrupole time-of-flight mass spectrometry. J Sep Sci 43(7):1284-1296.
Tsumura Y, Kado T, Yoshida K, Abe H, Ohtani M, Taguchi Y, Fukue Y, Tani N, Ueno S, Yoshimura K, Kamiya K, Harada K, Takeuchi Y, Diway B, Finkeldey R, Na'iem M, Indrioko S, Kit Siong Ng K, Muhammad N, Lee SL (2011) Molecular database for classifying *Shorea* species (Dipterocarpaceae) and techniques for checking the legitimacy of timber and wood products. J Plant Res 124:35-48.
Venkataraman K (1972) Wood Phenolics in the chemotaxonomy of the moraceae. Phytochemistry 11(5):1571-1586.
Veríssimo A, Júnior CS, Stone S, Uhl C (1998) Zoning of timber extraction in the Brazilian Amazon. Conserv Biol 12(1):128-136.
Wheeler EA (2011) Inside Wood—A web resource for hardwood anatomy. IAWA J 32(2):199-211.
White L, Gasson P (2008) Mahogany. Royal Botanic Gardens, Kew. 100 p.
Wilkes J (1984) The influence of rate of growth on the density and heartwood extractives content of eucalypt species. Wood Sci Technol 18(2):113-120.
Williams M, Isabel N, Ho V-M, Farr K, Duchesne I (2020) Cataloguing Canadian trees to help tackle illegal logging, Natural Resources Canada—Canadian Forest Service.
Wu J, Lu X, Tang W, Kong H, Zhou S, Xu G (2004) Application of comprehensive two-dimensional gas chromatography—Time-of-flight mass spectrometry in the analysis of volatile oil of traditional Chinese medicines. J Chromatogr A 1034(1-2):199-205.
Zhang M, Zhao G, Guo J, Liu B, Jiang X, Yin Y (2019) A GC-MS protocol for separating endangered and non-endangered *Pterocarpus* wood species. Molecules 24(422): 799.

APPENDIX

A1 DESCRIPTION OF THE FOUR GENERA AND FIVE SPECIES CONSIDERED IN THIS STUDY

Swietenia

The *Swietenia* genus is widely distributed in the Neotropics (Figueroa Colon 1994) and comprises three species: *Swietenia mahagoni*, *Swietenia macrophylla*, and *Swietenia humilis* and two natural hybrids (Pennington 1981). ***Swietenia macrophylla***, also known as big-leaf mahogany, has a wide geographic range from the southern tip of Florida extends through the southern parts of Mexico and Central America passing across the southern Amazon basin of Bolivia and Brazil (Pennington 1981; Danquah et al 2019). Since, the late 1700 s, *Swietenia macrophylla* has been heavily exploited (Rodan et al 1992). In Central America, populations have declined by 80% in the last 50 yr, and some are already extinct (Navarro et al 2003). In 2003, due to the high risks to population viability related to overexploitation and habitat destruction, *Swietenia macrophylla* was listed on CITES Appendix II (Grogan and Barreto 2005). ***Swietenia mahagoni***, also known as American mahogany, Cuban mahogany, or small-leaf mahogany, is widely distributed in Southern Florida in the United States and islands in the Caribbean, including the Bahamas, Cuba, Jamaica, and the Dominican Republic and Haiti (Mayhew and Newton 1998; Danquah et al 2019).

Cedrela

The genus *Cedrela* P. Browne is native to Neotropical region of America (Pennington 1981; Patiño Valera et al 1996; Estrada-Contreras et al 2016; Danquah et al 2019). The genus recognizes about 17 species and two of the 17 species (*Cedrela odorata*, *Cedrela fissilis*) are widespread lowland species occurring in both rainforests and drier areas, whereas the majority is either in wet mountainous forest or seasonally dry forest species scattered through Central and South America (Muellner et al 2009). *Cedrela odorata* wood is eminent for its magnificence, durability, and pest resistance. The heartwood of *Cedrela odorata* contains an aromatic and insect-repelling resin, which makes it resistant to insects, termite, and rot. *Cedrela* timber has been utilized for numerous finishes in colonial buildings in different countries of South America (Pennington et al 2010). Due to deforestation and the selective logging of *Cedrela odorata* individuals, the species has declined by 28.8% over the last 100 yr and it is estimated that it will decrease by 40.4% over the next 100 yr (CITES 2018). Selective overharvesting of *Cedrela odorata*, targeting the tallest and healthiest trees, has been shown to reduce populations across their range and genetic diversity (Llerena et al 2012). The *Cedrela* genus (populations of the Neotropics) has been placed in CITES Appendix II in August 2020. Finch et al (2019) recently described the genetic structure of *Cedrela*.

Khaya

The genus *Khaya*, traded as African mahogany, comprises about five species: *Khaya ivorensis*, *Khaya anthotheca*, *Khaya grandifoliola*, *Khaya senegalensis*, and *Khaya Madagascariensis*, which are native from Africa and the surrounding islands (Louppe et al 2008; Sexton et al 2015). *Khaya ivorensis*, also known as Lagos mahogany, is a very large evergreen tree attaining a height between 40 and 50 m. The wood is highly valued for a variety of uses, including boatbuilding and the making of furniture and musical instruments (Louppe et al 2008). The high-value timber of *Khaya* makes it, like many other tree species in tropical Africa, a victim of illegal logging (Hansen and Treue 2008). Throughout their ranges, natural regeneration after logging is seriously hindered by the low density of adult trees and low regeneration rates (Louppe et al 2008; IUCN 2015). Today, *Khaya* genus is under heavy exploitation pressure. Consequently, at the 19th meeting of the Conference of the Parties on CITES (CITES CoP19) held in November 2022 in Panama, the whole *Khaya* genus was added to CITES Appendix II, effective in February 2023.

Toona

Toona is a genus is native to Afghanistan moving southwards to India and east to North Korea, Papua New Guinea, and Eastern Australia (Mabberley 2017). It consists of six major commercially important timber species namely *Toona calantas* (Philippine mahogany), *Toona ciliata* (*Toona australis*—Australian red cedar or Indian mahogany), *Toona sinensis* (Chinese mahogany), *Toona sureni* (Vietnamese mahogany; Indonesian mahogany), and *Toona fargesii* (Danquah et al 2019). *Toona ciliata* is a timber tree species that is native from Southeast Asia. In the past, the genus was often incorporated within a wider circumscription of the related genus *Cedrela*, but that genus is now restricted to species from the Americas. *Toona ciliata* is red in color, easy to work, and very highly valued. The wood characteristics of *Toona ciliata* are similar to that of *Cedrela odorata* and its cultivation may reduce the harvesting pressure on *Cedrela odorata* in native forests (Dordel et al 2010). *Toona ciliata* is currently recorded as Least Concern by IUCN red list (IUCN 2021).

A2 LIST OF IMAGES OF FIVE MELIACEAE SPECIES ACQUIRED THROUGH XYLOTRON (MAGNIFICATION 10×)

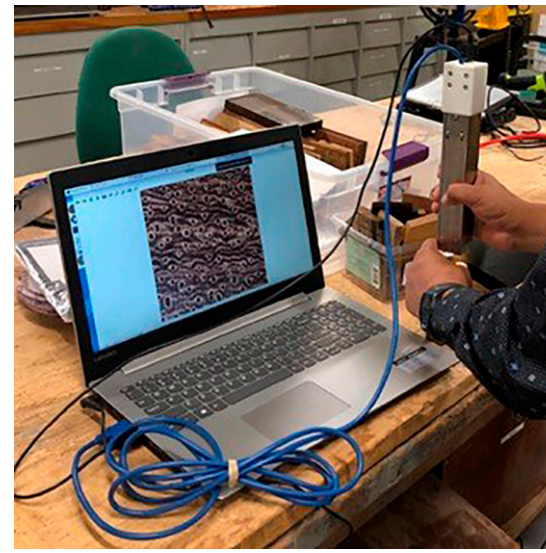


Figure A2-1. XyloTron is a machine-vision system for wood species identification (www.xylotron.org) (Hermanson et al 2019).

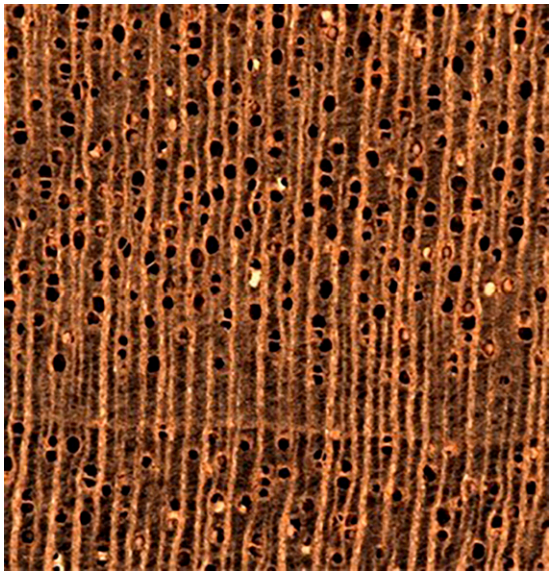


Figure A2-2. Macroscopic image of *Swietenia mahagoni* (Sample Id_539923).

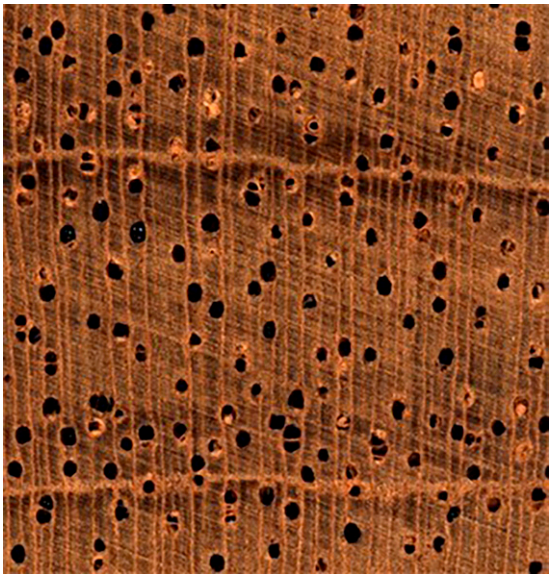


Figure A2-3. Macroscopic image of *Swietenia macrophylla* (Sample Id_539857).

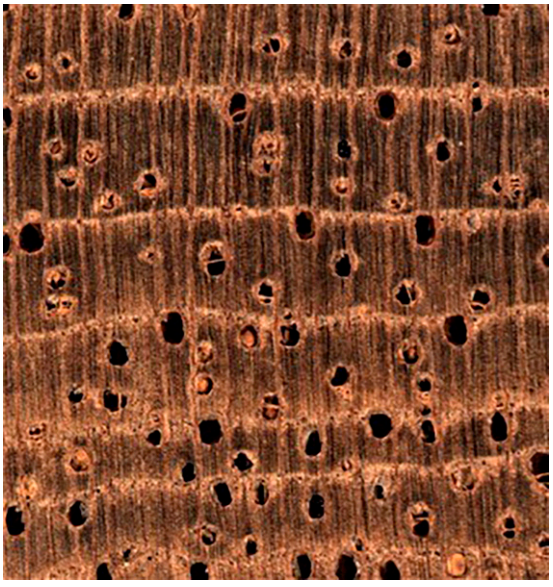


Figure A2-4. Macroscopic image of *Cedrela odorata* (Sample Id_542086).

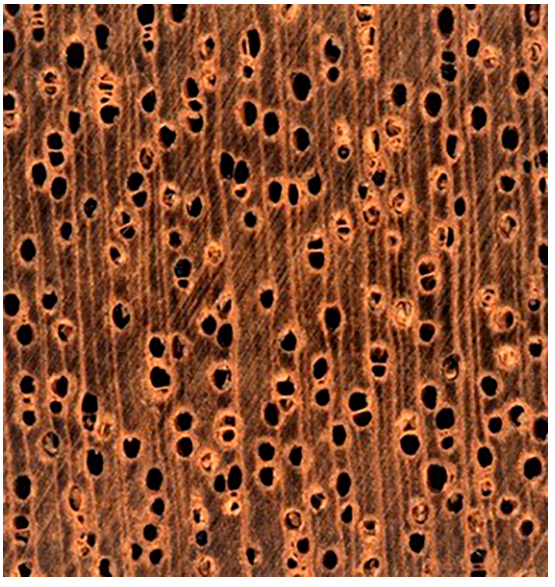


Figure A2-5. Macroscopic image of *Khaya ivorensis* (Sample Id_540729).

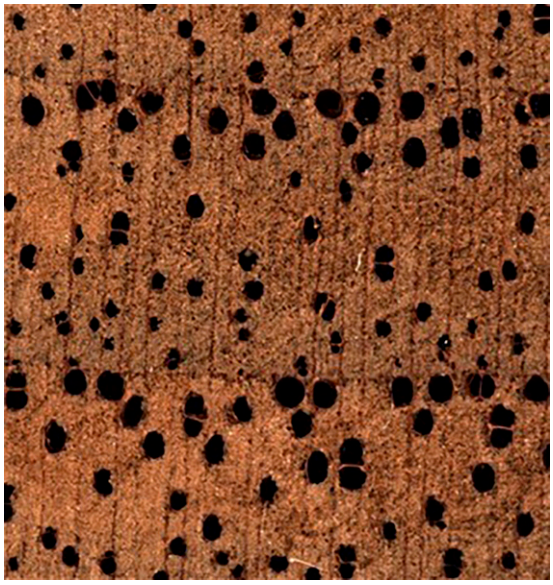


Figure A2-6. Macroscopic image of *Toona ciliata* (Sample Id_543406).

A3 LIST OF CONTOUR PLOTS/CHROMATOGRAPHIC PEAKS OF MELIACEAE SPECIES OBTAINED FROM GC×GC-TOFMS ANALYSIS

The following figures represent contour plot/chromatographic peak obtained after GC×GC-TOFMS analysis. The *x*-axis indicated first-dimension retention time, whereas *y*-axis indicated second-dimension retention time. Each colored "blob" or contour represents a peak and the color scale shows the intensity of the peaks. It can be observed that the two *Swietenia* species have fewer peaks and quite similar numbers of sterols, whereas *Cedrela odorata* and *Toona ciliata* have a greater variety of detected compounds and increased peak intensities compared with other species. Further, it is noteworthy that *Cedrela odorata* is found to be dominant in sesquiterpenes, sesquiterpenols, sesquiterpene oxides, and sterols.

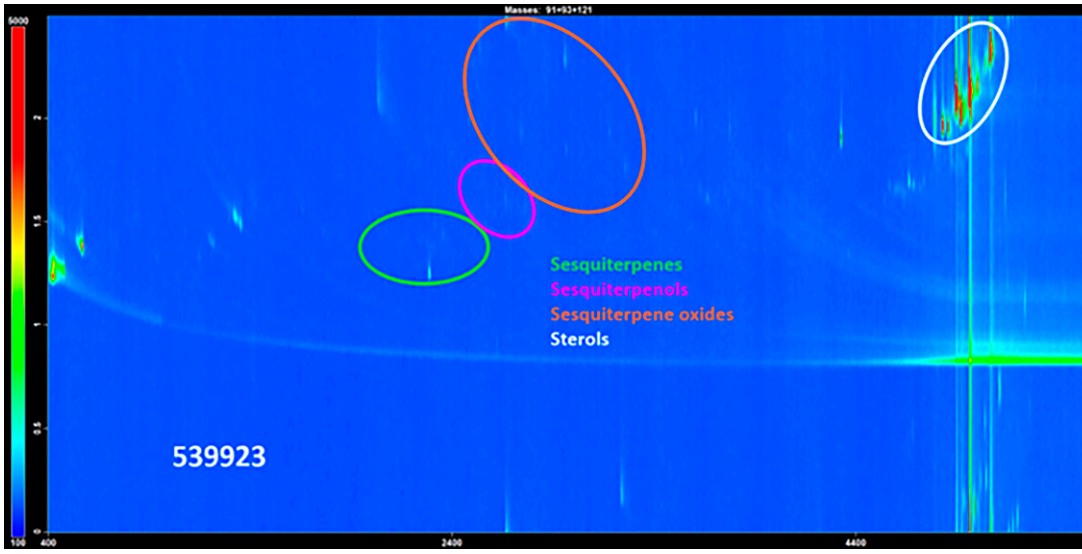


Figure A3-1. *Swietenia mahagoni* (Wood sample_Id: 539923).

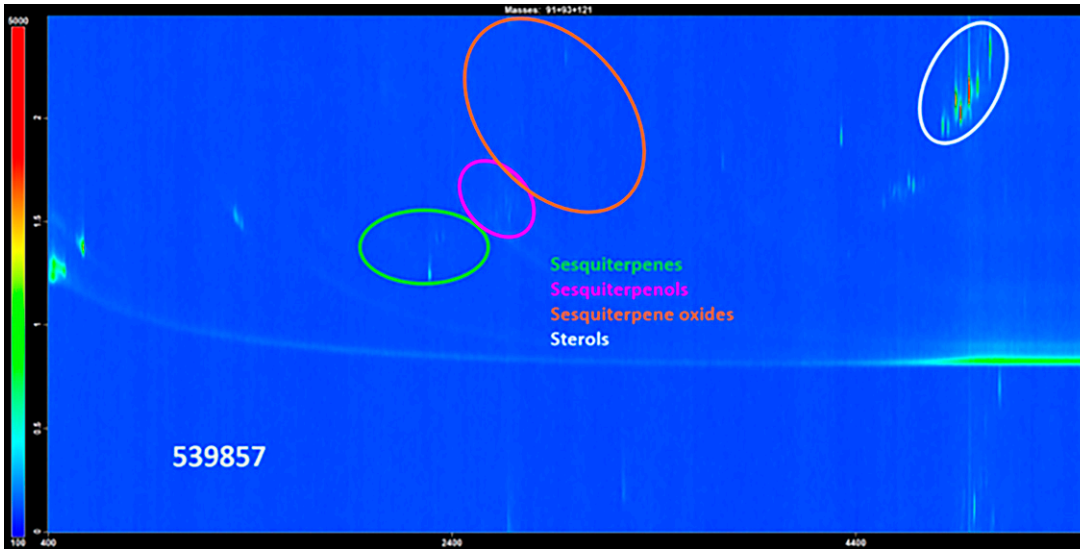


Figure A3-2. *Swietenia macrophylla* (Wood sample Id: 539857).

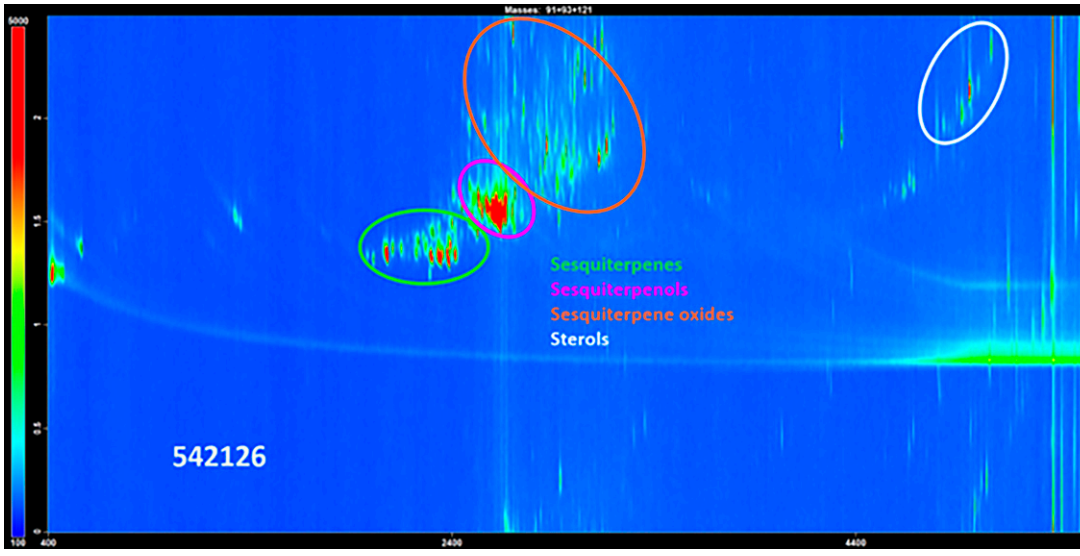


Figure A3-3. *Cedrela odorata* (Wood sample Id: 542126).

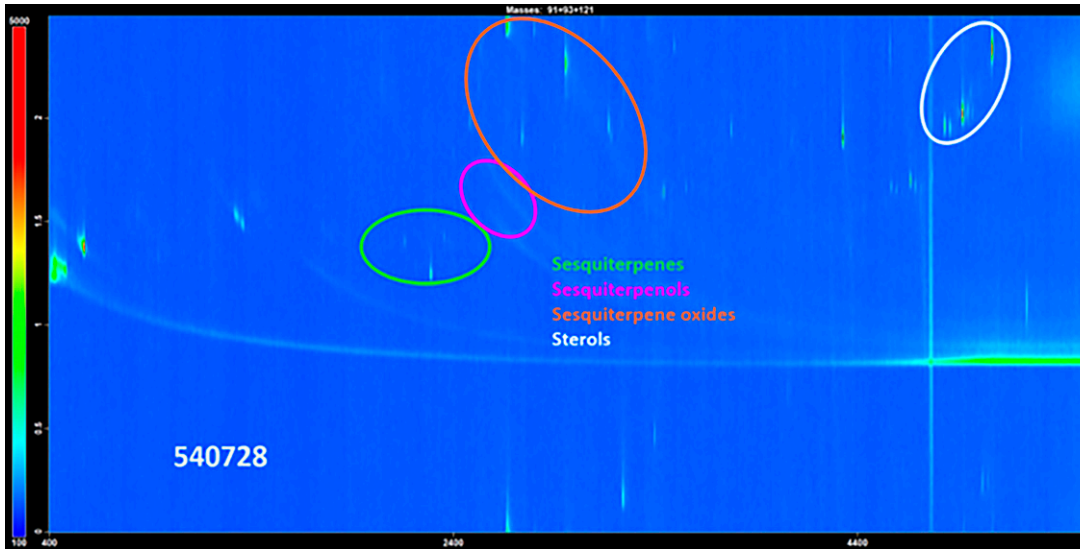


Figure A3-4. *Khaya ivorensis* (wood sample_ Id: 540728).

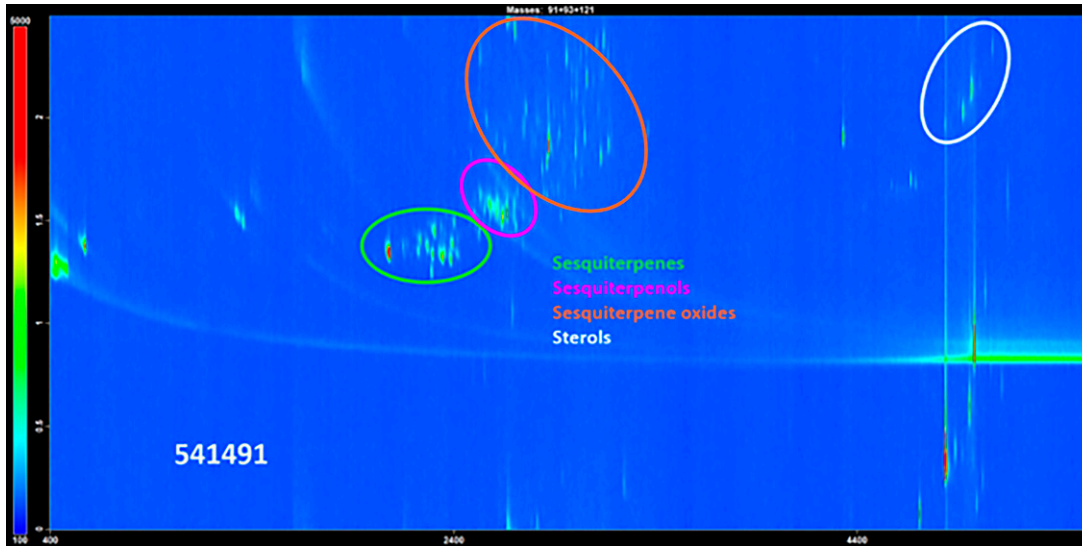


Figure A3-5. *Toona ciliata* (wood sample_Id: 541491).

Table A1. Selected features for principal component analysis of *Swietenia, Cedrela, Khaya,* and *Toona* samples.

Variable label	Compound name	Variable label	Compound name
A	C28 sterol B	AA	Metabolite 173
B	Metabolite 1321	AB	Triterpenoid C
C	C27 sterol B	AC	C27 sterol A
D	Cadalene	AD	Diterpenoid A
E	Triterpenoid A	AE	Metabolite 1273
F	Metabolite 411	AF	Copaene
G	Monosubstituted benzaldehyde	AG	Metabolite 67
H	Triterpenoid E	AH	Metabolite 811
I	C22 linear aldehyde A	AI	Sesquiterpenol A
J	Metabolite 1409	AJ	Sesquiterpenol B
K	Metabolite 94	AK	Undecane
L	Metabolite 1488	AL	Metabolite 472
M	Sesquiterpenol C	AM	Sesquiterpene oxide A
N	alpha-Muurolene	AN	Metabolite 133
O	Metabolite 1213	AO	alpha-Terpineol
P	gamma-Muurolene	AP	Metabolite 365
Q	Triterpenoid F	AQ	2,4-Di-tert-butylphenol
R	Triterpenoid B	AR	Metabolite 466
S	Metabolite 17	AS	alpha-Cadinene
T	Diterpenoid B	—	—
U	C28 sterol A	—	—
V	Triterpenoid D	—	—
W	C28 sterol B	—	—
X	1-Aromadendrene	—	—
Y	Metabolite 1438	—	—
Z	Metabolite 8	—	—

PRESERVATIVE TREATMENT OF TASMANIAN PLANTATION *EUCALYPTUS NITENS* USING SUPERCRITICAL FLUIDS

K. C. Wood†*
Research Fellow
Centre for Sustainable Architecture with Wood
School of Architecture and Design
College of Sciences and Engineering
University of Tasmania
Newnham Campus
T40 Building, Brooks Road
Newnham TAS 7248
E-mail: kyra.wood@utas.edu.au

A. W. Kjellow
Senior Project Manager
Building and Construction
Danish Technological Institute
Gregersensvej 3K, DK-2630 Taastrup
E-mail: awk@teknologisk.dk

M. J. Konkler
Senior Faculty Research Assistant
Department of Wood Science and Engineering
College of Forestry
Oregon State University
4700 SW Research Way
Corvallis, OR, 97333-5704
E-mail: matthew.konkler@oregonstate.edu

G. Presley
Assistant Professor
Department of Wood Science and Engineering
College of Forestry
Oregon State University
Corvallis, OR, 97333-5704
E-mail: gerald.presley@oregonstate.edu

J. J. Morrell†
Professor, Director of National Centre for Timber Durability and Design Life
University of the Sunshine Coast
41 Boggo Road, Ecosciences Precinct
Dutton Park QLD 4102
E-mail: jmorrell@usc.edu.au

(Received December 2022)

Abstract. Short rotation plantation forests in Tasmania, Australia, are dominated by *Eucalyptus nitens* (common name: shining gum). These forests were primarily planted to provide material for pulp and paper production, but the timber is increasingly sought after for higher value and more enduring applications.

* Corresponding author
† SWST member

Wood and Fiber Science, 55(1), 2023, pp. 83-93
https://doi.org/10.22382/wfs-2023-08

Plantation *E. nitens* has a high proportion of low-durability heartwood that resists penetration by conventional fluid preservatives. This limits its use to indoor applications. One approach to overcoming the refractory nature of *E. nitens* is to modify the treatment fluid. We investigated the use of supercritical carbon dioxide to deliver biocides deep into the wood. Timbers varying in thickness from 19 to 35 mm and 900-mm long were treated with a multicomponent biocide under supercritical conditions in a commercial facility in Denmark. The resulting timber was cut into zones inward from the surface. Wood from these zones was grounded and extracted for HPLC analysis for tebuconazole and propiconazole. Preservative was detected in the inner portion of every sample examined, indicating that the process resulted in treatment throughout the boards, with concentrations meeting and on average exceeding the targeted amounts.

Keywords: *Eucalyptus nitens*, preservative treatment, supercritical carbon dioxide, tebuconazole, propiconazole, wood preservation.

INTRODUCTION

Timber durability is a key consideration for architects and engineers concerned with reducing the environmental impact of buildings. Building construction contributes 39% of the world's carbon emissions, with building materials making up 11% of those emissions (Global ABC 2019). Structurally and aesthetically, wood can rival nonrenewable materials like steel and concrete that are also far more carbon intensive to produce (Churkina et al 2020; Allen et al 2021). In terms of durability, however, wood is a biological material that can be degraded by a number of abiotic and biotic agents under the proper conditions (Archer and Lebow 2006; Zabel and Morrell 2020). Designers can limit biodegradation risk, eg with water shedding designs, or by only specifying timber in interior applications. In instances where long-term wood contact with moisture is unavoidable, species that produce naturally durable heartwood may be specified, but the supply cannot always meet the demand for durable timbers. Alternatively, timber can be artificially impregnated with preservatives to protect against fungal, insect, and even marine borer attack.

In most regions, preservative treatment specifications address the easily treated sapwood of coniferous species, such as Scots pine, radiata pine, or the southern pines. Heartwood of these species is generally far more difficult to impregnate, and most treatment standards specify a high percentage treatment of the sapwood coupled with a shallow penetration of the heartwood (eg a 5-10 mm envelope). While softwoods are important in global commerce, many countries have expanding supplies of low-durability plantation hardwood timbers that require preservative treatment for exterior exposures. For example, Tasmania has extensive plantations of *Eucalyptus nitens* that were originally planted and managed for pulp and paper production, but are increasingly viewed as a potential resource for building construction including in exterior applications, where biodeterioration is likely (Wood et al 2020).

Plantation *E. nitens* is classified as Durability Class 4 in Australian Standard AS 5604 (Standards Australia 2005) and is characterized by shallow bands of sapwood surrounding a large core of lower durability heartwood that is highly resistant to conventional fluid preservative treatment. These refractory characteristics make it extremely difficult to meet the current Australian Standards for exposure in exterior Hazard Classes (H3, outside, above ground; or H4, in ground contact) as described in Standard AS/NZS 1604.1 (Standards Australia 2021), limiting its potential uses to indoor applications. It is also notoriously difficult to kiln dry without excessive defects and air-drying usually requires prolonged periods of a year or more depending on the thickness of the board, followed by reconditioning, final kiln dry, and planing or molding. Any preservative treatment that significantly rewets the wood after the material has been planned would potentially require another lengthy drying period, followed by a costly second reconditioning, final kiln dry, and planing or molding.

Globally, a variety of alternative methods have been explored for impregnating difficult to treat or refractory heartwoods, including modifying the solvent characteristics (Siau 1971), increasing

pressure, or using alternative pressure methods, but none of these can completely overcome the refractory nature of heartwood on their own (Hunt and Garratt 1967; Nicholas and Siau 1973; Cookson 2000). Timbers can be incised to increase the amount of cross-sectional area exposed to fluid flow, but this process mars the surface appearance, can reduce strength, and is mostly only effective to the depth of the incisions (Winandy et al 2022). Diffusible compounds, especially boron, have also been explored because of their ability to spread with moisture into normally refractory timber, but their water solubility renders them susceptible to leaching and, therefore, not suitable for exterior exposures (Lloyd 1998; Freeman et al 2009). Gaseous boron can also be used, but it has largely been limited to applications on composites (Burton et al 1990).

In Australia, the requirements for quality assurance of preservative treatments include visual assessment of the entire cross section of a treated board to ensure complete sapwood penetration as well as minimum heartwood treatment depths (AS/NZS 1604; Standards Australia 2021). This is usually done using colorimetric spot tests; however, a penetration tracer, such as zinc or copper, may be added when the preservative cannot be visually assessed (eg as in the case of azoles). The current Australian Standard (AS/NZS 1604) requires a minimum penetration of 5 mm in the heartwood of any hardwood board equal to or less than 35-mm thick, and complete penetration of the sapwood.

As noted previously, 5 mm of heartwood penetration of *E. nitens* is exceedingly difficult to consistently achieve as illustrated by two recent reports prepared for the Australian National Institute for Forest Products Innovation (Wood et al 2023a, 2023b), which outline a comprehensive suite of treatment strategies that were trialed to improve the durability of heartwood-dominant refractory Australian *Eucalyptus* species, including *E. nitens.* (The material used in both trials was from the same general source as the material used for the research described in this paper.) One of the strategies trialed was an extensive vacuum pressure impregnation trial that used varied scheduling, pressure, solution strengths, and chemical composition, including both copper-based and azole-based Australian-approved biocides. In the study of azole-based biocides (Wood et al 2023a), chemical retention analysis was not undertaken, and theoretical retention was unable to be assessed as the biocide used was a commercially available, ready-to-use solution and limited information was provided on the schedule and solution strengths. However, extremely low uptakes averaging less than 10 L/m^3 were recorded, which is far less than the typical targeted amount for a hardwood to meet exterior application requirements in Australia (ie approximately 40-45 L/m^3). In addition, because the biocide included trace amounts of copper, visual penetration assessments were completed on oven-dried biscuits using a preservative indicator (PAN [1-(2-pyridylazo)-2-napthol]) and evaluated against the criteria in the Australian Standards (AS/NZS 1604; Standards Australia 2021) using a grid analysis (Wood et al 2023a [see Fig 14], p. 22). The results indicated that azole-based biocides barely penetrated the *E. nitens* boards. Similar trials using alkaline copper quaternary (ACQ) compound and micronized copper azole (MCA) showed that samples could not consistently meet the required retention targets (Wood et al 2023a). Although some better penetration was achieved using novel process enhancements, treatments still failed to produce consistent preservative penetration demonstrating the difficulty of treating *E. nitens* using conventional vacuum pressure impregnation.

One potential alternative for impregnating refractory heartwoods like *E. nitens* is supercritical fluid (SCF) treatment. SCFs are defined as materials that are at a temperature and pressure where distinct gas and liquid phases do not exist. SCFs can behave like a liquid in terms of density, which can aid in solubility and viscosity that can facilitate movement through the wood matrix (Krukonis 1988; Kayihan 1992; Sahle-Demessie et al 1995a, 1995b; Kjellow et al 2010). SCF density and viscosity are easily adjustable by varying pressure or temperature. While a variety of solvents can be used for SCF treatments, carbon dioxide (CO_2) is

more commonly used because of its low cost, minimal toxicity, and low critical temperature/pressure. It can also be captured and recycled in a closed-loop treatment system.

Supercritical CO_2 (SC-CO_2) can solubilize a variety of organic preservatives at levels capable of protecting wood from decay and insect attack and has been shown to penetrate into a variety of wood species as well as wood-based composites (Hassan et al 1995; Kjellow and Henriksen 2009). Supercritical wood impregnation has been found to have no significant effects on timber properties despite the elevated pressures as long as pressure differentials are minimized (Smith et al 1993; Anderson et al 2000; Schneider et al 2003, 2005; Oberdorfer et al 2006). Limited field trials suggest that plywood impregnated with tebuconazole using SC-CO_2 performed well in a subtropical H3 exposure, but there have been few other public long-term durability studies (Acda et al 1997, 2001; Morrell et al 2005; Kjellow et al 2013).

SCFs have been used to a lesser extent to treat hardwoods. Anderson et al (2000) explored the effects of treatment on flexural properties of white oak and sweetgum beams. Cookson (2009) reported on small-scale treatments of Australian hardwoods with permethrin for termite control, but the work was not commercialized. One major hurdle to the commercial use of SCFs is the large capital investment required for the treatment plant; however, the operating costs can be lower than regular treatment facilities and the ability to effectively impregnate species that are not currently suitable for exterior exposure creates a range of new market opportunities. There is already one commercial facility located in Denmark, which currently uses SCFs to impregnate timber with a biocide suitable for exterior exposure. Their treatment facility concentrates on Norway spruce (*Picea abies*) for above-ground applications, such as cladding and decking.

SCFs treatment might be an attractive alternative for properly impregnating the heartwood of Tasmanian plantation *E. nitens* with biocides for H3 or H4 exposures. The process also has some advantages from a processing perspective since it is nonswelling and can be used to treat finished and shaped timber without the need for posttreatment sanding or planing. However, the process can also induce collapse or internal checks as a result of excessive surface vs internal pressure differentials if pressurization and depressurization rates exceed the rate at which the wood can equilibrate pressure (Oberdorfer et al 2006). While SCF treatments have been explored for a number of timber-related applications, there is little in support of using this process to treat *E. nitens*. The objective of this study was to evaluate the effects of SCF treatment on preservative penetration and internal condition of *E. nitens* timber.

MATERIALS AND METHODS

E. nitens timber ($90 \times 35 \times 900$-mm long) with an average oven-dry density of 547 kg/m^3 was cut from trees in a 26-yr-old plantation that had been thinned and pruned over the rotation. The timber had been air-seasoned for approximately 6 to 9 mo, then reconditioned and kiln-dried to a target MC of 12% (mass % oven-dry basis) prior to final planing and cutting. On average, less than 5% of the boards contained any sapwood at all and of those boards, there was less than 2% sapwood on each sample, usually concentrated in an outer corner of the board due to the quartersawn cutting pattern. Fifteen samples each of 19, 25, or 35-mm thickness were all cut from the same parent material and the samples were sent to Danish commercial SCF facility, Superwood A/S (https://www.superwood.dk/), for inclusion in one of their commercial spruce treatment charges. Before treatment, the wood was conditioned to a constant weight at 85% RH and 20°C to a final MC of 19% MC. The higher MC was used because experience from commercial treatment of spruce shows that wood at this moisture level is less prone to cracking during treatment.

The treatment process consisted of pressurizing the treatment vessel with CO_2. During pressurization, the CO_2 was continuously circulated between the treatment vessel and a static mixer containing the biocide, gradually dissolving the biocides and bringing these to the wood. Treatment temperature

throughout the process was 45°C and maximum impregnation pressure was 15 MPa. Pressure was held to allow the biocides to diffuse inward, then gradually released back to atmospheric levels thereby depositing the biocides in the wood. Total treatment time was 7 h, which is longer than is normally used for spruce boards. The prolonged rates of pressure increase and decrease were used as previous tests on other dense species showed that they minimized the development of pressure gradients that induced internal stresses leading to either collapse or internal checking/splitting.

The samples were treated with SC200, a mixture containing tebuconazole, propiconazole, and iodopropynyl butylcarbamate (IPBC) in a relative ratio of 2:2:1. The target concentration for the spruce boards was 120 g/m^3 of total active ingredient, which is the biological reference value of SC200 against basidiomycetes determined by the biological test regime specified in EN 599-1:2009 + A1:2013 for Use Class 3 exposure (outside, above ground). In conventional treatments, it is possible to weigh the wood before and after treatment to determine net solution uptake; however, SCF treatments can solubilize wood extractives during the process while simultaneously depositing the biocides, potentially resulting in net weight losses. In addition, the target deposition of 120 g/m^3 corresponds to only about 0.00025 kg of actives per 1 kg of wood, so the uncertainty of a weight measurement would be too large. Chemical analysis was used instead to measure the amount of biocide through the cross sections of boards of each thickness.

After treatment, each 900-mm long parent board was cut into five sections with three 200 mm long planks retained for durability testing in a separate study, and two 150 mm cross sections for retention analysis. A 5-mm thick cross section was cut from each of the two 150-mm long cross sections and these were further divided into corresponding inner and outer zones for the 19- and 25-mm thick timbers and inner, middle, and outer zones for the 35-mm thick timber (Figs 1 and 2). Outer zones were 0-5 mm from the surface for the 19- and 25-mm thick timbers, while the inner zones corresponded to the zone 6-9 mm inward and 6-12 mm

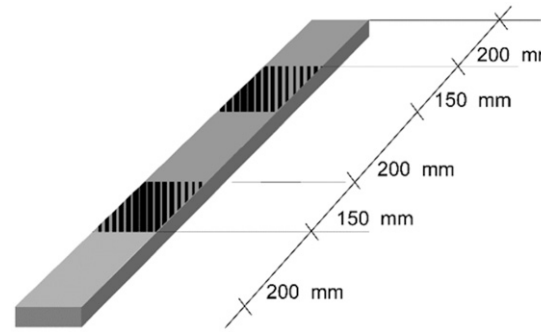


Figure 1. Schematic for 900-mm long SCF treated cut into 3 × 200 mm long planks and 2 × 150-mm long planks following treatment. Hatching indicates area from which a 5-mm biscuit was obtained for retention analysis.

inward for the 19- and 25-mm thick timbers, respectively (accounting for some loss due to the saw kerf). The 35-mm thick samples were cut into outer, middle, and inner zones corresponding to 0-5, 6-11, and 12-17 mm from the wood surface, respectively (Fig 2). Samples from a given zone were ground to pass a 20 mesh screen, oven-dried for 2 h at 80°C, and stored in a desiccator until cool. Approximately 0.5 g of ground wood from a given section and distance from the surface was added to a screw cap tube along with 25 mL of methanol and sonicated for 180 min at room temperature. The resulting liquid was filtered through a SAX SPE cartridge to remove particulate and interfering compounds.

The resulting filtrate was analyzed for tebuconazole and propiconazole on a Shimadzu High Performance Liquid Chromatograph equipped with an Intersil ODS-3 (150 × 4.6 mm, 3 µm) column. The mobile phases were 1) 55:45 acetonitrile:buffer, 2) acetonitrile, and 3) 2% methanol in HPLC water introduced at a flow rate of 1.00 mL/min. The two azoles were detected at 195 nm and quantified by comparison with similar analyses of prepared standards. The third preservative component (IPBC) was not analyzed in this experiment.

RESULTS AND DISCUSSION

There was no visible evidence of excessive crushing, splitting, or collapse in the samples following

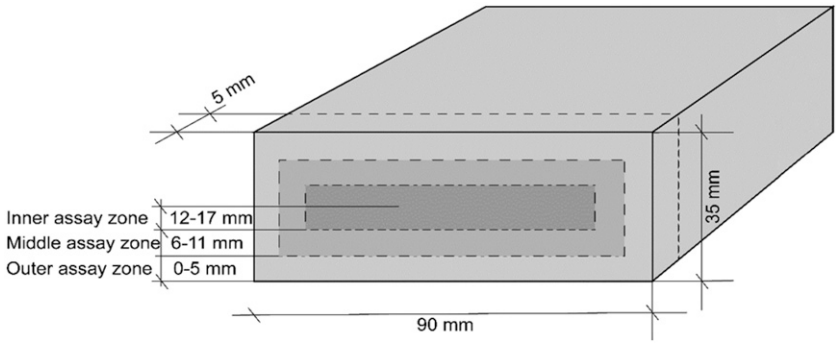


Figure 2. Schematic showing the cutting pattern for the three assay zones on a 35-mm thick sample.

treatment; however, some minor collapse was observed in the samples after several months stored under ambient conditions indoors (approximately 17-25°C) indicating that there may have been excess moisture in the timber that subsequently dried during the storage period indoors (Fig 3). This was particularly evident on the surface of the 35-mm thick samples and was more apparent when samples were cut into assay zones for retention analysis. The long lag between treating and collapse makes it unlikely that the collapse was due to pressure differentials during SCF treatment, but rather due to the reconditioning of the timber to 19% MC prior to the treatment. This conditioning was done to aid heat transfer through the wood and make it slightly more plastic to mitigate potential cracking. Oven-dry MC was not determined directly following treatment or in the months after the treatment; however, we do not believe that MC changed as a result of treatment because there was minimal potential for drying during treatment.

It is worth noting that interior and exterior timber in Australia is required to be at 9-14% MC at time of sale (AS 2796.1, Standards Australia 1999). This would require some redrying if the moisture conditioning to 19% was included in the treatment process. Further studies to determine the optimum moisture level that minimizes physical damage are recommended.

Chemical analyses revealed that both tebuconazole and propiconazole were detectable in all assay zones of the cross sections of every sample tested (Fig 4, Table 1), and all samples met and exceeded, on average, the targeted requirement of 120 g/m^3 for treatment of spruce using SC200. While retentions in the middle and inner zones of the three thicknesses of wood were much lower than those near the surface, both azoles were detected in every treated sample analyzed, indicating that SC-CO_2 was capable of carrying the biocides well into the heartwood. These analyses did not include the levels of IPBC in the wood.

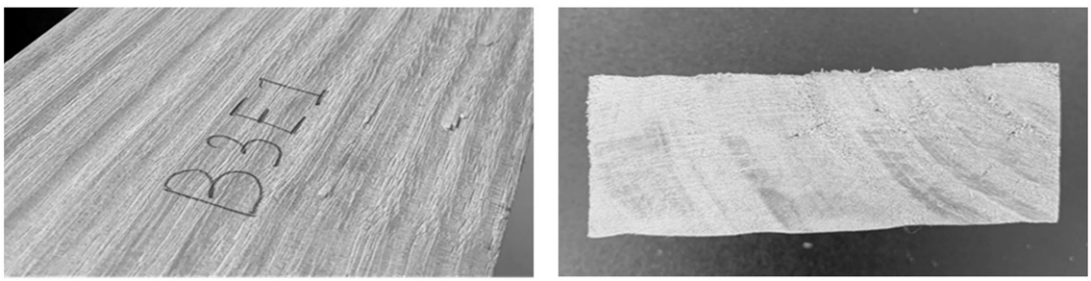
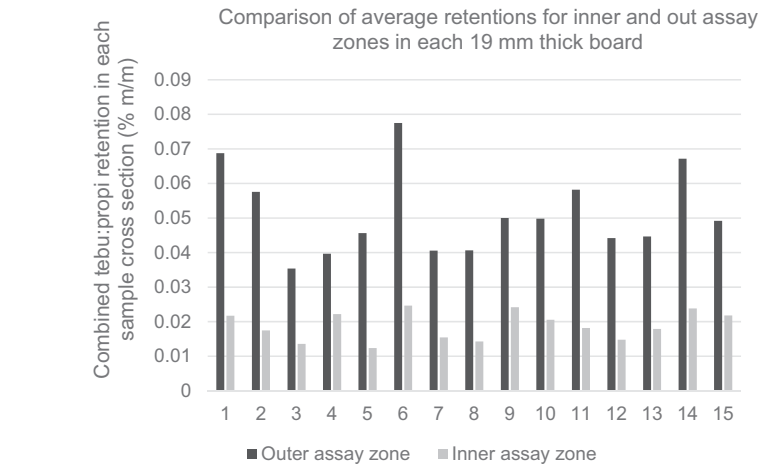


Figure 3. Examples of surface checking (left) and internal collapse (right) in 35-mm thick *Eucalyptus nitens* boards after SCF impregnation.



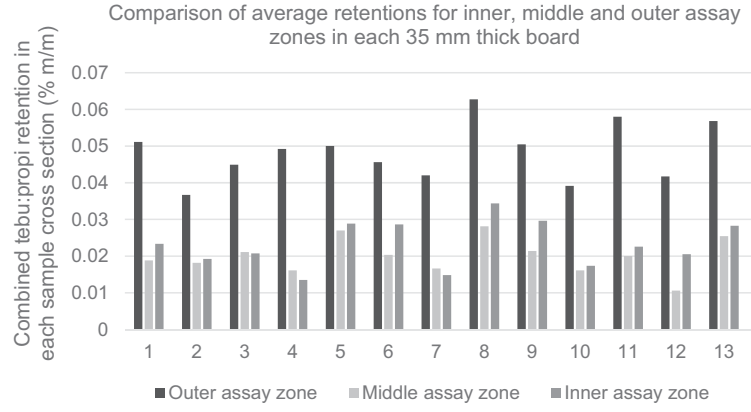


Figure 4. Retention of tebuconazole/propiconazole at selected depths of (top) 19-mm thick (middle), 25-mm thick and (bottom), 35-mm thick *E. nitens* timbers treated using supercritical carbon dioxide. Note: in the 19-mm thick sample set, two vessels were lost/broken during extraction which may affect the averages for the inner assay zones in boards 7 and 14. Also, although the middle assay zones appear lower than the inner assay zone in some of the 35-mm thick boards, this variation between the two averages was insignificant.

Table 1. Average g/m^3 of propiconazole/tebuconazole in *E. nitens* timbers treated using supercritical carbon dioxide[a].

Sample thickness (mm)	Assay zone	Average g/m^3 by assay zone[b]	Total g/m^3 in cross section[b]
19	Outer 0-5 mm	277 (68)	191
	Inner 6-14 mm	104 (26)	
25	Outer 0-5 mm	226 (77)	165
	Inner 6-19 mm	103 (28)	
35	Outer 0-5 mm	265 (52)	161
	Middle 6-11 mm	103 (32)	
	Inner 12-24 mm	116 (35)	

[a] Samples were treated to the spruce target retention of 120 g/m^3 of the azole/IPBC mixture.

[b] g/m^3 is a less precise treatment measure as SCF treatments can solubilize wood extractives during the process while simultaneously depositing the biocides, potentially resulting in net weight losses. Values represent analyses of 30 replicates per assay zone for three board thicknesses, and 60 or 90 analyses for the combined cross sections for the 19/25 mm and 35-mm thick samples, respectively. Values in parentheses represent one standard deviation.

IPBC is highly soluble in SC-CO_2 and would have further increased the retentions (Hassan et al 1995).

As noted in the introduction, Australian Standards for quality assurance of preservative treatments use visual assessments to ensure complete sapwood penetration as well as minimum heartwood treatment depths (AS/NZS 1604; Standards Australia 2021). Unfortunately, the chemicals used in the treatment of spruce at Superwood lacked either a specific color or a tracer making it difficult to determine whether the detected azole was uniformly distributed throughout a given zone. Establishing a method for visual evaluation of penetration will be critical for quality assessment of SCF azole treatment in an Australian context.

The Australian Standard also requires a minimum retention of 0.03% tebuconazole and 0.03% propiconazole mass/mass for H3 exposures based on the oven-dry mass each piece; a total of 0.06% m/m combined azoles. In this paper, analyzed retention % m/m was converted to g/m^3 for clarity using the following formula:

$$\text{Retention } (g/m^3) = \frac{\left\{ \begin{array}{l} retention\ (\%\ m/m) \times \\ oven\ dry\ timber\ density\ (g/m^3) \end{array} \right\}}{0.1}$$

Combined azole retentions in the SCF treated *E. nitens* boards were all below the Australian target level (Table 2); however, it is important to note that the system used for SCF treatment also contained IPBC, which was not measured in this study. In addition, the treatment process was designed for spruce and not specifically for this species. A further caveat is that the Australian Standard specifies analysis of the entire cross section. Average retentions for the combined inner, middle, and outer zones were 0.0348, 0.0330, and 0.0308% m/m, respectively, for the three thicknesses (Table 2).

Table 2. Retentions of propiconazole/tebuconazole in *E. nitens* timbers subjected to treatment using supercritical carbon dioxide (% m/m).

Sample thickness (mm)	Propiconazole/Tebuconazole retention (% m/m)[a]						
	Outer		Middle		Inner		Combined
	Avg	Range	Avg	Range	Avg	Range	Avg
19	0.0507 (0.0125)	0.0347-0.0853	—	—	0.0189 (0.0047)	0.0088-0.0292	0.0348
25	0.0453 (0.0140)	0.0224-0.0731	—	—	0.0207 (0.0052)	0.0118-0.0323	0.0330
35	0.0485 (0.0094)	0.0306-0.0687	0.0206 (0.0058)	0.0065-0.0323	0.0232 (0.0064)	0.0097-0.0361	0.0308

[a] Values represent 30 replicates per zone and sample thickness. Values in parentheses represent one standard deviation.

Figure 5. Example of extremely shallow and minimal penetration in the heartwood of conventional vacuum pressure treated *E. nitens* as indicated by darkened color on the corners of samples sprayed with PAN (1-[2-pyridylazo]-2-napthol) for the presence of copper.

The more important treatment parameter for this study was the retention away from the surface as an indicator of penetration, and, as noted above, it was clear from the chemical analysis that treatment was present in each assay zone, indicating that SC-CO_2 was able to carry biocides well into the heartwood of each sample. This result was in sharp contrast to those obtained using conventional vacuum treatment processes in a separate study (Fig 5).

While the levels of azoles in the treated materials were lower than the current requirements in the Australian Standards, the most obvious solution for achieving the Australian benchmarks would be to increase the amount of fungicide in the treatment vessel.

The use of longer treatment times would sharply reduce production capacity, but cosolvent addition to enhance azole solubility would increase the potential for diffusion. Previous studies have shown that solubilities of tebuconazole increased markedly with methanol addition and field trials of SC-CO_2 tebuconazole treated plywood in a subtropical environment showed that the resulting products performed well in above-ground exposures (Hassan et al 1995; Acda et al 1996, 1997a, 1997b, 2001).

CONCLUSIONS

SCF treatment resulted in far deeper preservative penetration in the heartwood of *E. nitens* than is typical for conventional preservative pressure treatments. The treatment target of 120 g/m^3 was reached and even exceeded in each board. While the retentions were below those required for azole treatment in an Australian context, this could be likely overcome by the addition of more fungicide in the treatment vessel, or by using modifiers to increase solubility. The results illustrate the potential for improved preservative protection of this species for above-ground applications and further research is planned.

ACKNOWLEDGMENT

The authors are grateful to Superwood A/S for their in-kind support of this research by including our material in one of their commercial spruce treatments. The authors are grateful to the National Institute for Forest Products Innovation, the Australian Government, the Tasmanian Government, Forest and Wood Products Australia, and related timber industry partners for their sponsorship of this research. The authors are grateful to the University of Tasmania, the University of the Sunshine Coast, the Danish Technological Institute, and Oregon State University for their support and the use of facilities and equipment to conduct this research.

REFERENCES

Acda MN, Morrell JJ, Levien KL (1996) Decay resistance of composites following supercritical fluid impregnation with tebuconazole. Mater Org 30:293-300.

Acda MN, Morrell JJ, Levien KL (1997a) Effects of supercritical fluid treatments on physical properties of wood-based composites. Wood Fiber Sci 29(2):121-130.

Acda MN, Morrell JJ, Levien KL (1997b) Effect of process variables on supercritical fluid impregnation of composites with tebuconazole. Wood Fiber Sci 29(3):282-290.

Acda MN, Morrell JJ, Levien KL (2001) Supercritical fluid impregnation of selected wood species with tebuconazole. Wood Sci Technol 35:127-136.

Allen C, Oldfield P, Teh SH, Wiedmann T, Langdon S, Yu M, Yang J (2021) Modelling ambitious climate mitigation pathways for Australia's built environment. Sustain Cities Soc 77:103554. doi:10.1016/j.scs.2021.

Anderson ME, Leichti RJ, Morrell JJ (2000) The effects of supercritical CO2 on bending properties of four refractory wood species. Forest Prod J 50(11/12):85-93.

Archer K, Lebow S (2006) Wood preservation. Pages 297-338 *in* JCF Walker, ed. Primary wood processing principles and practice, 2nd edition. Springer-Verlag, New York, NY.

Burton R, Bergervoet A, Nasheri K, Vinden P, Page DR (1990) Gaseous preservative treatment of wood. *In* Proc IRG Annual Meeting IRG/WP/3631, May 13-19, 1990, Rotorua, New Zealand.

Churkina, G, Organschi, A, Reyer, CPO, Ruff, A, Vinke, K, Liu, Z, Reck, BK, Graedel, TE, Schellnhuber, HJ (2020) Buildings as a global carbon sink. Nat Sustain 3:269-276.

Cookson LJ (2000) The preservation of eucalypts. Pages 248-255 *in* The Future of Eucalypts for Wood Products, Proc IUFRO Conference, March 19-24, 2000, Launceston, Tasmania.

Cookson LJ (2009) Treatment of timber with permethrin in supercritical carbon dioxide to control termites. J Supercrit Fluids 49(2):203-208.

European Standard CEN EN 599-1:2009 + A1:2013 (2013) Durability of wood and wood-based products—Efficacy of preventive wood preservatives as determined by biological tests—Part 1: Specification according to use class. European Committee for Standardisation, CEN, Brussels.

Freeman MH, McIntyre CR, Jackson D (2009) A critical and comprehensive review of boron in wood preservation. Pages 279-294 *in* Proc American Wood Protection Association, 105, April 1-21, 2009, San Antonio, Texas.

Global ABC, Alliance for Buildings and Construction, International Energy Agency, and the United Nations Environment Programme (2019) 2019 Global Status Report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector. United Nations Environment Programme.

Hassan A, Levien KL, Morrell JJ (1995) Phase behavior of binary and ternary mixtures of wood preservatives in supercritical carbon dioxide with co-solvents. Pages 402-414 *in* Innovations in Supercritical Fluids, ACS Symposium Series No. 608, Washington, D.C.

Hunt GM, Garratt GA (1967) Wood preservation. McGraw-Hill, New York, NY.

Kayihan F (1992) Method of perfusing a porous workpiece with a chemical composition using cosolvents. US patent 5,094,892.

Krukonis VJ (1988) Processing with supercritical fluids: Overview and applications. Pages 27-43 *in* ACS Symposium Series No. 366, American Chemical Society, Washington, D.C.

Kjellow AW, Henriksen O (2009) Supercritical wood impregnation. J Supercrit Fluids 50:297-304.

Kjellow AW, Henriksen O, Sørensen JC, Johannsen M, Felby C (2010) Partitioning of organic biocides between wood and supercritical carbon dioxide. J Supercrit Fluids 52:1-5.

Kjellow A, Imsgard F, Fernandes J, Wagner R, Delis J (2013) Field performance of wood impregnated with siloxanes using supercritical carbon dioxide. *In* Proc IRG Annual Meeting, IRG/WP/13-40632, June 16-20, 2013, Stockholm, Sweden.

Lloyd J (1998) Borates and their biological applications. *In* Proc IRG Annual Meeting, IRG/WP 98-30178, June 14-19, 1998, Maastricht, The Low Countries.

Morrell JJ, Acda MN, Zahora AR (2005) Performance of oriented strandboard, medium density fiberboard, plywood and particleboard treated with tebuconazole in supercritical carbon dioxide. *In* Proc IRG Annual Meeting, IRG/WP/05-30364, April 24-28, 2005, Bangalore, India.

Nicholas DD, Siau JF (1973) Factors influencing treatability. Pages 299-343 *in* DD Nicholas, ed. Wood deterioration and its prevention by preservative treatments, Vol 2, Syracuse University Press, Syracuse, New York, NY.

Oberdorfer G, Leichti RL, Morrell JJ (2006) Internal pressure development and deformation during supercritical fluid impregnation of selected wood-based materials. Wood Fiber Sci 38:190-205.

Sahle-Demessie E, Hassan A, Levien KL, Kumar S, Morrell JJ (1995a) Effect of supercritical carbon dioxide treatment on permeability of Douglas-fir heartwood. Wood Fiber Sci 27:296-300.

Sahle-Demessie E, Levien KL, Morrell JJ (1995b) Deposition of chemicals in semi-porous solids using supercritical fluid carriers. Pages 415-428 *in* Innovations in supercritical fluids. ACS Symposium Series No. 608, Washington, D.C.

Schneider PF, Levien KL, Morrell JJ (2003) Internal pressure measurement techniques and pressure response in wood during treating processes. Wood Fiber Sci 35: 282-292.

Schneider PF, Morrell JJ, Levien KL (2005) Internal pressure development during supercritical fluid impregnation of wood. Wood Fiber Sci 37(3):413-423.

Siau JF (1971) Flow in wood. Syracuse University Press, Syracuse, New York, NY.

Smith SM, Sahle-Demessie E, Morrell JJ, Levien KL, Ng H (1993) Supercritical fluid (SCF) treatment: Its effect on bending strength and stiffness of ponderosa pine sapwood. Wood Fiber Sci 25(2):119-123.

Standards Australia (1999) Standard AS2796. Timber: Hardwood: Sawn and milled products. Standards Australia, Melbourne, Australia.

Standards Australia (2005) Standard AS5604 Timber—Natural durability ratings. Standards Australia, Melbourne, Australia.

Standards Australia (2021) Standard AS/NZS1604 Preservative-treated wood-based products. Standards Australia, Melbourne, Australia.

Winandy JE, Hassan B, Morrell JJ (2022) Review of the effects of incising on treatability and strength of wood.

Wood Mater Sci Eng 17:1-12. doi:10.1080/17480272.2022.2028008.

Wood KC, Leggate W, Robinson R, Meldrum S, Belleville B, Wiesner F, Ghani R (2023a) Increasing the durability, and other material characteristics of Tasmanian hardwoods: Final report. Tech Rep. Forest and Wood Products Australia, Melbourne, Australia. 88 pp. ISBN: 978-1-922718-22-8.

Wood KC, Morrell JJ, Hassan B, Meldrum S, Leggate W, Vargas JR (2023b) New methods of reliably demonstrating species durability in commercially relevant timeframes. Tech Rep. Forest and Wood Products Australia, Melbourne, Australia. 84 pp. ISBN: 978-1-922718-23-5.

Wood KC, Morrell JJ, Leggate W (2020) Enhancing the durability of low durability eucalyptus plantation species: A review of strategies. *In* Proc IRG Annual Meeting, IRG/WP 20-40910, June 10-11, Online Webinar.

Zabel RA, Morrell JJ (2020) Wood microbiology. Academic Press, San Diego, CA.

CASE STUDY OF 3-PLY COMMERCIAL SOUTHERN PINE CLT MECHANICAL PROPERTIES AND DESIGN VALUES

*L. M. Spinelli Correa**
Graduate Research Assistant
E-mail: ls2090@msstate.edu

R. Shmulsky
Head and Warren S. Thompson Professor of Wood Science & Technology
E-mail: rs26@msstate.edu

F. J. N. França
Assistant Research Professor
Department of Sustainable Bioproducts
Mississippi State University
Mississippi State, MS 39762-9820
E-mail: fn90@msstate.edu

(Received January 2022)

Abstract. This work elucidates on a case study of industrially manufactured cross-laminated timber (CLT). Two methods are used to calculate specimens section modulus: S_{gross} and $S_{effective}$. The first assumes that specimens behave as a continuous material, whereas the second considers the cross laminations (shear analogy method). Although the shear analogy method is indicated for construction purposes, applications, such as trench shoring, matting, and work platforms, could benefit from a simpler calculation method. Therefore, the objective of this work was to conduct a case study of Modulus of Rupture (MOR) and Modulus of Elasticity (MOE) of southern pine CLT to compare the previously mentioned calculation methods. Both parametric and nonparametric fifth percentiles and associated F_b values are reported and were substantially higher than those of the constituent lumber. For MOE, empirical testing and calculation based on gross moment of inertia provided lower values as compared with the constituent lumber.

Keywords: Cross-laminated timber, bending design value, section modulus, strength, stiffness.

INTRODUCTION

The research reported herein provides a case study of industrially manufactured cross-laminated timber (CLT). Over the past decade, CLT has made significant advancement in the building construction sector. As a relatively new mass timber panel, CLT has demonstrated both potential and promise in various building construction applications. To enhance North American production and market acceptance, APA-The Engineered Wood Association (2018) has published a related product standard. Therein, among other items are minimum grade, strength, and stiffness, requirements for lumber to be used in layup laminations. It also contains information regarding moment capacity (strength) information ($F_b \cdot S$) as well as sectional stiffness, that is, the product of Modulus of Elasticity (MOE) times the moment of inertia ($E \cdot I$). These values are derived from the basic lumber lamination mechanical properties and effective section properties. Effective section properties are somewhat reduced from gross section properties to account for the cross lamination(s) in the inner ply or plies. The S_{gross} method assumes that the CLT panel behaves as a continuous composite material through its thickness, whereas the $S_{effective}$ method uses shear analogy applied to CLT. As such, the $S_{effective}$ is less than the S_{gross} because the rolling (across the grain) shear strength is taken as a fraction of parallel to grain shear. Although it's necessary to count for rolling shear strength for construction purposes, the shear analogy method can be seen as over conservative when applied to

* Corresponding author

other CLT uses. The shear analogy method not only requires more measurements, but also entails a more complex understanding of composite materials and strength calculations, which can act as a limitation to secondary CLT uses. Outside of the building construction industry, there are other opportunities for CLT use and adoption. Industrial applications, such as matting, trench shoring, other temporary shoring, and work platforms, are potential markets. In such cases, it is often helpful to have basic bending strength (Modulus of Rupture [MOR] and F_b) and stiffness, as MOE, properties of the manufactured panels. In such instances, the parameters calculation can still be seen as conservative, but will often be more easily assessed by quality control procedures already widely used by the wood products industry. Therefore, the objective of this work was to conduct a case study of MOR and MOE of southern pine CLT, along its major strength axis, to compare the previously mentioned calculation methods.

MATERIALS AND METHODS

In this case study, 24 2.44 m (8 ft) × 4.88 m (16 ft) 3-ply commercial CLT panels were acquired and defined as parent panels. Panels were made in accordance with PRG-320 (APA-The Engineered Wood Association 2018) from 5.08 × 20.3 cm (2 × 8 in.) nominal, 3.81 × 18.4 cm actual (1.50 × 7.25 in.) Number-2 southern yellow pine lumber and glued with polyurethane. According to PRG-320, this material is classified as V3 (Table 1). In addition, also in agreement with PRG-320, the basic bending design values for CLT are based directly on the material properties of the constituent lumber. In other words, PRG-320 uses basic, minimal lumber design values as the direct feedstock for CLT design value calculation. In that matter, one of the purposes of this study was to demonstrate that direct testing of CLT offers the possibility to derive or demonstrate superior properties. The process of lamination and development of a composite system routinely improves the allowable strength values.

Table 1. Design values (MPa) for laminations in longitudinal layers, per V3.

F_b[a]	Characteristic value[b]	F_b for #2 2x8 lumber[c]	MOE[c]
5.17	10.9	6.38	9650

[a] PRG-320, Table A1
[b] PRG-320, Table 1. (Note: F_b = Characteristic value/2.1)
[c] SPIB 2014

From each of 20 parent panels, one test specimen was ripped. From each of four parent panels, two test specimens were ripped. In sum, 28 unique test specimens were considered. This is the minimum number required, per American Society for Testing and Materials (ASTM) D2915 (2017a), for estimation of a nonparametric fifth percentile. In general, during the ripping process, the material from which any given specimen was ripped was not immediately adjacent to the material from which any other specimen was ripped. After ripping, some of the face lumber on the specimens was asymmetric, that is, it included two full 7.25-in. wide pieces, ripped strip to make up the full 18-in. wide test specimen. Test specimens were kept on an outside covered area previous to testing. Specimens presented an average density of 535.5 Kg/m^3 and average MC of 14%. As measured, each specimen was approximately 10.5-cm (4.13-in.) thick, 45.7-cm (18.0-in.) wide, and crosscut to 3.05-m (120-in.) long. The reason the specimens were 4.13-in. thick (rather than full 4.25-in.) is because each constituent piece was skim planed at the time of CLT manufacture.

The samples were destructive tested in the major direction (3.05-m), as arranged in a flatwise layup, via third-point bending over a 2.90-m (114-in.) span (Fig 1) and at a span to depth ratio of 27.6 according to a modified ASTM 5456 (2017b) and with consideration of PRG-320 "specimen width not less than 12 in. and the on-center span equal to approximately 30 times the specimen depth for the tests in the major strength direction. . . ." This relatively long specimen size minimizes the incidence of shear failure during the flexural test. The timber blocks between the machine fixture load heads and the specimen are approximately 6.5 × 10 × 60 cm (2.56 × 3.94 × 23.6 in.) in dimension. They are

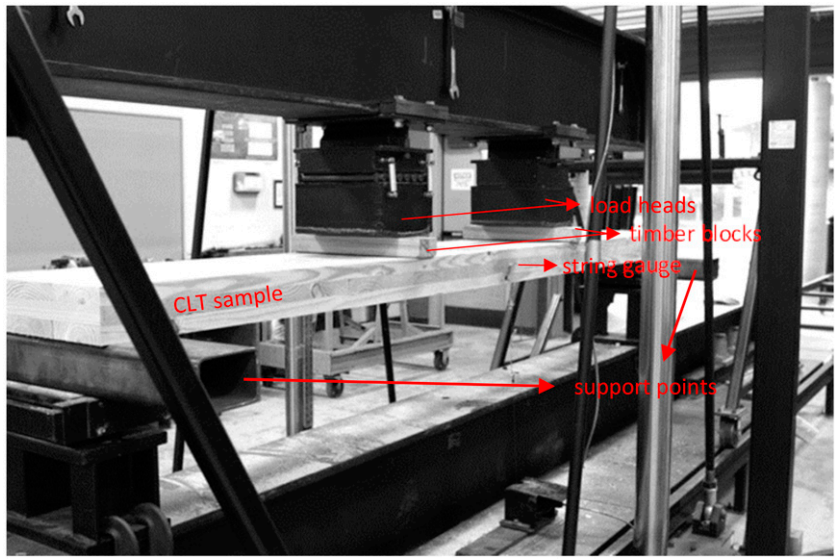


Figure 1. Test setup.

radiused per the standard at the points of specimen contact and provide line loads at the one-third points and are sufficiently wide to avoid indentation. A 600-kN capacity hydraulic universal test frame was used for testing. To record the deflection, a string gage deflectometer with the 0.001 ± 0.0005-in. accuracy was placed at midspan and at the panel's neutral axis. The test was displacement controlled with a rate of 0.0003 m/s (0.8 in./min). Load, deflection, testing rate, time to failure, and failure mode were recorded.

To calculate flexural stress (MOR), one must first calculate the section modulus of the panel. For CLT, calculation of section modulus for uniform rectangular sections is done in two ways and thus, yields two different MOR, and subsequently, F_b values. Either method might be acceptable, depending on the final use of the panel, as they ultimately equate to the same moment capacity. The S_{gross} method assumes that the CLT panel behaves as a continuous composite material, whereas the $S_{effective}$ method uses shear analogy applied to CLT, considering the orientation of the laminations. In the case of industrial applications, such as matting, it is often more practical to use gross section modulus (S_{gross}) for determination of F_b as it is readily calculable.

$$S_{gross} = \frac{b \cdot h^2}{6}$$

Where:

b = width; h = thickness.

$$S_{eff} = \frac{2EI_{eff}}{E_1 \cdot h}$$

Where:

EI_{eff} = Effective bending stiffness; E_1 = MOE of outermost layer (Characteristic value of 9.65 × 10^3 MPa [1.4 × 10^6 psi per SPIB 2014]); and h = Entire thickness of the panel (Karacabeyli and Douglas 2013).

$$EI_{eff} = \sum \left(E_i \cdot b_i \cdot \frac{{h_i}^3}{12} \right) + \sum \left(E_i \cdot A_i \cdot {z_i}^2 \right)$$

Where:

E_i = "i" layer's design value MOE (9.65 × 10^3 MPa [1.4 × 10^6 psi per SPIB 2014]); b_i = "i" layer's width; h_i = "i" layer's thickness; and

Table 2. Summary statistics for the flexural testing.

	Load (kN)	MOR gross (MPa)	MOR effective (MPa)	MOE gross (MPa)
N (number of specimens)	28	28	28	28
Average	65.4	34.9	35.9	8142
Maximum	81.6	44.9	45.9	10,469
Minimum	42.2	23.1	23.9	5755
Standard deviation	9.58	4.9	5.09	1068
Coefficient of variation	14.7%	14.1%	14.2%	13.1%
K factor*	1.88	1.88	1.88	1.88
Parametric fifth percentile	47.4	25.7	26.4	—
Order statistic	1	1	1	—
Non parametric 5th percentile	42.2	23.1	23.9	—
Factor for conservatism (NDS 2015)	—	0.85	0.85	—
Combined load duration and safety factor (ASTM 2017b)	—	2.1	2.1	—
F_b (parametric)	—	10.4	10.7	—
F_b (nonparametric)	—	9.34	9.67	—

*Estimated K factor for one sided tolerance limit, for $n = 28$, at 75% confidence per Table 3, (ASTM 2017a).

A_i = "A" layer's section area; z_i = distance from the neutral axis of the panel to the center of respective layer.

These two section moduli were then used to calculate panel stress values. With these two section moduli, two sets of stress values were calculated. F_b gross was calculated as maximum moment divided by S_{gross}. F_b effective was calculated as maximum moment divided by $S_{effective}$. Per associated guidance from PRG-320, effective moment capacities must be multiplied by a factor of 0.85 for conservatism. As such, one can either multiply the 0.85 factor times the F_b value, the section modulus value, or their product ($F_b \cdot S$). To calculate the stiffness of the panel, the traditional calculation method for lumber was applied (ASTM 2022).

$$E_{app\ (gross)} = \frac{23P \cdot l^3}{108b \cdot d^3 \cdot \Delta}$$

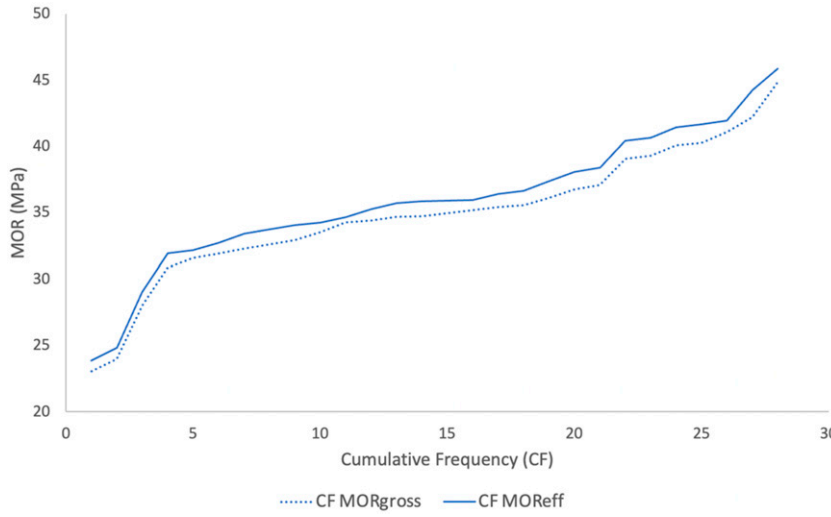


Figure 2. Cumulative frequency distribution of MOR values.

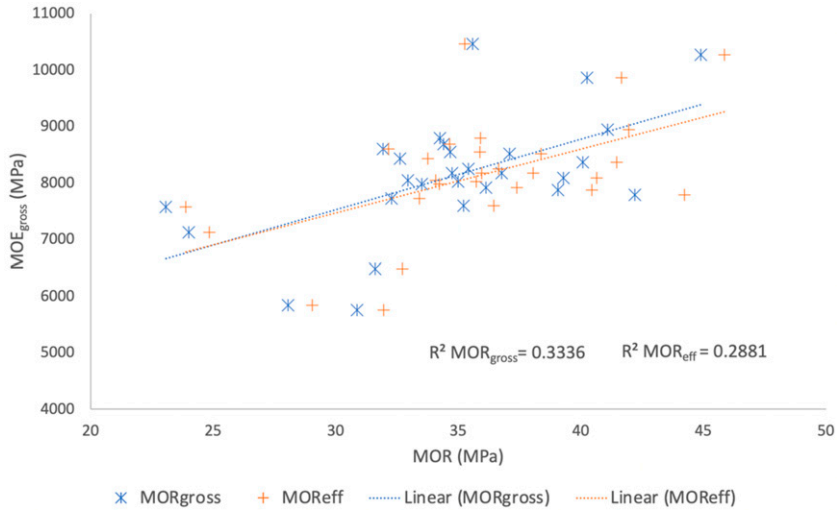


Figure 3. Relationship between MOE and MOR for the 28 specimens.

Where:

$E_{app\ (gross)}$ = Apparent MOE; P = load; l = span; b = width; d = panel thickness; and Δ = increment of deflection.

RESULTS AND DISCUSSION

From the test data, MOR was calculated by both gross and effective section moduli. For each of these methods, both parametric and nonparametric fifth percentiles (ASTM 2017a) and associated F_b values are reported. MOE was calculated based on the gross moment of inertia. The summary statistics are presented in Table 2. Both parametric and nonparametric F_b values (9.34 and 10.4 MPa, respectively) were substantially higher than those of the constituent lumber (6.38 MPa).

MOE gross is included because it can be readily calculated based on the direct physical measurements of the panel along with its observed deflection in response to a given load. MOE effective is not considered herein because it is generally calculated based on the published design value MOE of the constituent lumber rather than on the empirical observations. Figure 2 illustrates the cumulative frequency distribution of MOR values. Figure 3 illustrates the relationship between MOE and MOR. The MOR_{gross} R^2 value for this relationship is 0.33. This finding indicates that 33% of the variation in MOR_{gross} is explained by MOE.

CONCLUSIONS

- For MOR_{gross}, empirical testing provided favorable results as compared with currently assigned values derived as defined in PRG-320 based on the published values for the constituent lumber. This finding suggests that it is likely in a manufacturer or user's best interest to evaluate their specific material's flexural strength. In this manner, a manufacture can most accurately market their material based on its inherent properties and a user can derive the maximum possible potential utility and engineering value from said materials.
- For MOE, empirical testing and calculation based on gross moment of inertia provided lower values as compared with the constituent lumber. This result is likely due to the fact that the center ply was oriented perpendicular to the facial plies and as such, displayed predictably lower stiffness. This finding suggests that it is likely in a user's best interest to evaluate their specific material's flexural stiffness if deflection under load is an important use criteria.
- The relationship between MOR_{gross} and MOE was relatively weak. This finding indicates

that nondestructive evaluation based on MOE, for this material, may not be a particularly useful tool for evaluating ultimate or allowable strength characteristics.

- In the case of matting, heavier loads applied over softer soils require increasingly predicable strength and stiffness. Reliable strength values prevent mat breakage, potential equipment loss, and unsafe working conditions. Reliable stiffness values minimize rutting, enhance environmental protection, and increase safety particularly with respect to crane and other lifting operations. The information developed and reported herein can be useful for those who employ CLT mats in heavy construction, road building, powerline and pipeline operations, and so on.

ACKNOWLEDGMENTS

This publication is a contribution of the Forest and Wildlife Research Center, Mississippi State University. The authors wish to acknowledge the support of U.S. Department of Agriculture (USDA), Research, Education, and Economics (REE), Agriculture Research Service (ARS), Administrative and Financial Management (AFM), Financial Management and Accounting Division (FMAD), and Grants and Agreements Management Branch (GAMB), under Agreement No. 58-0204-6-001. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture.

REFERENCES

American Society for Testing and Materials (2022) ASTM D 198-22a: Standard test methods of static tests of lumber in structural sizes. https://compass.astm.org/document/?contentCode=ASTM%7CD0198-22A%7Cen-US.

American Society for Testing and Materials (2017a) ASTM D 2915-17: Standard practice for sampling and data-analysis for structural wood and wood based products. https://compass.astm.org/EDIT/html_annot.cgi?D2915+17.

American Society for Testing and Materials (2017b) ASTM D 5456-17: Standard specification for evaluation of structural composite lumber products. https://compass.astm.org/EDIT/html_annot.cgi?D5456+21e1.

American Wood Council (2015) National design specifications for wood construction. https://awc.org/codes-standards/publications/nds-2015.

APA-The Engineered Wood Association (2018) PRG 320-2018. Standard for performance-rated cross-laminated timber. https://www.apawood.org/ansi-apa-prg-320.

Karacabeyli E, Douglas B (2013) CLT handbook: Cross-laminated timber. https://web.fpinnovations.ca/clt/.

SPIB (2014) Standard grading rules for southern pine lumber. Table 1-c-s light framing, structural joists and planks, and studs—2″ to 4″ Thick (8″-wide only). Southern Pine Inspection Bureau, Pensacola, FL.

FIBER QUALITY PREDICTION USING NIR SPECTRAL DATA: TREE-BASED ENSEMBLE LEARNING VS DEEP NEURAL NETWORKS

*Vahid Nasir**

Postdoctoral Fellow
University of British Columbia
and
Courtesy Faculty
Department of Wood Science and Engineering
Oregon State University
E-mail: vahid.nasir@oregonstate.edu

Syed Danish Ali

Research Associate
USDA Forest Service
Forest Products Laboratory
Madison
and
Department of Biological Systems Engineering
University of Wisconsin-Madison
E-mail: syed.ali2@usda.gov

Ahmad Mohammadpanah

Assistant Professor
Department of Mechanical Engineering
Loyola Marymount University
E-mail: Ahmad.Mohammadpanah@lmu.edu

Sameen Raut

PhD Candidate
E-mail: sameen.raut@uga.edu

Mohamad Nabavi

Graduated Student
E-mail: m.nabavi84@gmail.com

Joseph Dahlen

Associate Professor
Warnell School of Forestry and Natural Resources
University of Georgia
E-mail: jdahlen@uga.edu

Laurence Schimleck

Professor
Department of Wood Science and Engineering
Oregon State University
E-mail: laurence.schimleck@oregonstate.edu

(Received March 2023)

* Corresponding author

Wood and Fiber Science, 55(1), 2023, pp. 100-115
https://doi.org/10.22382/wfs-2023-10

Abstract. The growing applications of near infrared (NIR) spectroscopy in wood quality control and monitoring necessitates focusing on data-driven methods to develop predictive models. Despite the advancements in analyzing NIR spectral data, literature on wood science and engineering has mainly utilized the classic model development methods, such as principal component analysis (PCA) regression or partial least squares (PLS) regression, with relatively limited studies conducted on evaluating machine learning (ML) models, and specifically, artificial neural networks (ANNs). This could potentially limit the performance of predictive models, specifically for some wood properties, such as tracheid width that are both time-consuming to measure and challenging to predict using spectral data. This study aims to enhance the prediction accuracy for tracheid width using deep neural networks and tree-based ensemble learning algorithms on a dataset consisting of 2018 samples and 692 features (NIR spectra wavelengths). Accordingly, NIR spectra were fed into multilayer perceptron (MLP), 1 dimensional-convolutional neural networks (1D-CNNs), random forest, TreeNet gradient-boosting, extreme gradient-boosting (XGBoost), and light gradient-boosting machine (LGBM). It was of interest to study the performance of the models with and without applying PCA to assess how effective they would perform when analyzing NIR spectra without employing dimensionality reduction on data. It was shown that gradient-boosting machines outperformed the ANNs regardless of the number of features (data dimension). All the models performed better without PCA. It is concluded that tree-based gradient-boosting machines could be effectively used for wood characterization utilizing a medium-sized NIR spectral dataset.

Keywords: Convolutional neural network (CNN), deep learning, ensemble learning, gradient-boosting (TreeNet), near infrared (NIR) spectroscopy, random forest, wood.

INTRODUCTION

Smart quality control (QC) and characterization of wood materials requires developing intelligent monitoring systems through combining nondestructive evaluation (NDE) methods with data-driven techniques. NDE tools play a crucial role in fast and reliable data acquisition when it comes to wood QC and properties monitoring. Common NDE methods applied to wood materials include color measurement, X-ray computed tomography (X-ray CT), thermography, wave propagation, and near infrared (NIR) spectroscopy, among which, wave propagation methods and NIR spectroscopy have been widely studied in the literature (Nasir et al 2022). NIR spectroscopy can be used to predict wood properties directly related to sample chemistry (Schimleck et al 2019) through analyzing absorption or reflectance in the NIR range. It has also been widely applied to predict wood density, MC, mechanical properties, such as modulus of elasticity and modulus of rupture (Tsuchikawa 2007; Tsuchikawa and Schwanninger 2013). Apart from predicting wood properties, NIR spectroscopy has been used for wood identification, wood classification, and QC of thermally modified wood (Tsuchikawa and Kobori 2015; Willems et al 2015).

Despite the growing opportunities that NIR spectroscopy offers for wood characterization and QC, there are still limitations associated with this technique. Apart from the need for a robust calibration model, spectral data acquisition in industrial settings imposes practical challenges because of noise related to temperature and humidity variations (Hein et al 2017). Specifically, there is a significant gap between the performance of NIR spectroscopy in laboratory vs industrial conditions, where the high performance achieved under laboratory-controlled conditions using multivariate data analysis may not necessarily account for the high variation in data acquired in real time in a manufacturing environment (Hein et al 2017).

NIR spectral data are typically processed using a pretreatment (bandwidth selection, smoothing, normalization, etc.), feature extraction and/or selection, and final decision-making, which mainly involves using classification or regression models. Most studies on the applications of NIR spectroscopy in wood materials have focused on dimensionality reduction for feature extraction through applying principal component analysis (PCA) regression, or partial least squares (PLS) regression (Sandak et al 2016). There is a gap in the literature about advanced data-driven

methods applied to NIR spectral data analysis. Feature engineering could be a crucial step toward designing a decision-making model with improved accuracy over classical techniques. The main feature selection techniques (filter, wrapper, and embedded methods) are reviewed in the literature (Chandrashekar and Sahin 2014), including an overview on the variable selection in multivariate analysis of NIR spectra (Yun et al 2019). PCA does use employ dimensionality reduction; however, feature extraction via PCA yields a new set of variables with no clear physical meaning (Mao 2005). An alternative approach could be feature selection in the paradigm of supervised learning. Examples include utilizing heuristic search methods, such as genetic algorithm for variable selection in predicting the wood properties, such as pulp yield using NIR spectral data (Ho et al 2022; Zhen et al 2022).

The performance of a data-driven predictive model depends on both the feature engineering (extraction or selection) technique and the choice of regression or classification model for final prediction. Moreover, it also significantly impacted by the size and complexity of the dataset (Nasir and Sassani 2021). Machine learning (ML), and specifically, artificial neural networks (ANNs), have been utilized to deal with datasets of higher complexity to better unveil hidden patterns within the data. Superior performance of ANNs over PLS regression is reported in the literature with a sample size in the range of 172-480 (Watanabe et al 2014; Costa et al 2019; Ayanleye et al 2021). Variable (feature) extraction was typically performed using PCA and the reduced datasets were then fed into ANNs (Costa et al 2019; Nasir et al 2019; Ayanleye et al 2021). While the reported accuracy could be impacted by the size and complexity of the data, it highlights the importance of employing ML for enhancing the prediction accuracy of NIR-based models. Yet, literature in wood science and engineering still lacks comparative studies on the predictive performance of models developed using different ML and feature engineering techniques.

While traditional ML models rely heavily on the knowledge of user(s) for feature extraction and selection, deep learning (DL) models exhibit built-in feature engineering characteristics (Miotto et al 2018; Wang et al 2018), enabling them to model complex nonlinear relationships in big data (Nasir and Sassani 2021) through extracting complex high-level abstractions as data representations (Najafabadi et al 2015). Convolutional neural networks (CNNs) are among the most commonly used types of DL models and have been combined with NIR hyperspectral imaging for wood species identification (Kanayama et al 2019). One-dimensional (1D) CNN has also been applied to softwood species classification (Yang et al 2020) and QC of Chinese zither panels (Huang et al 2019).

This study aims to perform a comparative study between the performance of some of the ML and DL models for prediction of the properties of wood-based materials. Since most published studies on data-driven methods applied to NIR spectral data deal with small sample sizes, it also aims to investigate applying ML and DL to a medium-sized dataset. The target property for prediction is tracheid width, which in addition to other tracheid morphological characteristics (diameter, length, and wall thickness) largely determine pulp fiber quality and paper performance. Evaluation typically occurs on macerated samples that are examined with microscope or optical imaging system; however, more rapid techniques are required for routine incorporation into tree improvement programs. NIR spectroscopy offers an alternative approach and fiber/tracheid properties have been estimated in both hardwoods (Inagaki et al 2012; Pereira et al 2015) and softwoods (Schimleck and Evans 2004; Via et al 2004; Nabavi et al 2018; Dahlen et al 2021). Fiber/tracheid length models have generally demonstrated strong performance. Similarly, models for wall thickness measured on solid samples have also performed well; however, models for tracheid diameter, whether it be measured by SilviScan on solid samples (Evans 1994) or on macerated samples, have been noticeably weaker. Therefore, the study investigates options for enhancing the prediction accuracy of tracheid

width. Two neural network (NN)-based models, including the multilayer perceptron (MLP) and 1D-CNN, are utilized. Their performances will be compared with those obtained from tree-based ensemble methods; random forest, TreeNet gradient-boosting, extreme gradient-boosting (XGBoost), and light gradient-boosting machine (LGBM). The study evaluates the impact of applying PCA and compares that with the case of developing models using the full range of NIR data (no dimensionality reduction). The objective is to improve prediction performance by designing models capable of handling medium-sized NIR spectral datasets while analyzing the full range of NIR features with embedded feature selection characteristics.

EXPERIMENTS AND DATA ACQUISITION

ML study was performed on samples utilized by Nabavi et al (2018), which represents one of the-largest spectral datasets on wood materials and was comprised of 1842 loblolly pine (*Pinus taeda* L.) radial strips from 225 trees and 99 stands. The samples were taken from the Wood Quality Consortium (WQC) baseline study (Jordan et al 2008; Antony et al 2010), which totaled 134 stands representing six physiographic regions (Gulf Coastal Plain, Hilly Coastal, North and South Atlantic Coastal Plains, Piedmont, and Upper Coastal Plain) across the Southeastern United States. At each stand, three trees were felled and 25-mm thick disks were cut at 1.5 m intervals from the base to a top diameter of 25 mm. A breast height (1.37 m) disk was also collected and four book-matched pith-to-bark radial strips (12 mm longitudinally × 12 mm tangentially) were cut from these disks. Jordan et al (2008) and Nabavi et al (2018) provide additional detail regarding the collection and preparation of these samples. The stands were largely sampled in the early 2000s and many of the breast height disks had been used in various WQC research projects that followed their collection, hence some samples were unavailable for study (Nabavi et al 2018). Several of these disks also had blue stain and were excluded from the study (Nabavi et al 2018). Following the sample selection procedure, the selected sections were cut from the radial strips into individual 10-mm sections using a razor blade. The sections were further cut into 2-mm sections using a razor blade. The samples were then macerated with 20 mL of 50% hydrogen peroxide, 30 mL of water, and 50 mL glacial acetic acid at 60°C for 48 h (Franklin 1945). After 48 h, the tracheids and chemicals were cooled to room temperature, then the tracheids were separated from the pulping chemicals using a Buchner funnel, rinsed with approximately 1.9 L of water, and then the acidic spent chemicals and the rinse water were neutralized with sodium carbonate. The rinsed tracheids were then diluted with approximately 1 L of water prior to analysis. The macerated samples were analyzed using a TechPap MorFi Compact Fiber & Shive Analyzer with a 4 µm resolution camera (Techpap SAS, France) to assess tracheid properties. The equipment measured the number of tracheids, frequency of tracheids, tracheid length, and width using image analysis. For each sample, approximately 2500 tracheids were analyzed. The width-weighted tracheid width was used for the calculations instead of the mean value to minimize the effect of fines (Carvalho et al 1997). Width-weighted tracheid width is calculated by:

$$W_l = \frac{\sum_{n_i w_i^2}}{\sum_{n_i w_i}} \quad (1)$$

where W_l is width-weighted tracheid length, n_i is number of tracheids in the ith class, and w_i is the mean width of the ith class (Carvalho et al 1997). Prior to the collection of NIR spectroscopic data, each radial sample was marked from pith to bark into 10-mm sections on the transverse face. Subsequently, NIR diffuse reflectance spectra were collected from the radial–longitudinal surface of each pith-bark strip for the same 10-mm sections using a FOSS NIRSystems Model 5000 scanning spectrophotometer (FOSS NIRSystems, Inc., Laurel, MD) fitted with a diffuse reflectance static module. A custom-made autosampler consisting of a Parker servo motor and controller, and a linear stage provided precise sample motion and a Teflon mask with a window 5-mm high × 10-mm wide to limit the area scanned (Jones et al

2007). NIR spectra were collected at 2 nm resolution over the wavelength range 1100-2500 nm. The instrument reference was a ceramic standard provided by the manufacturer. A total of 11,428 NIR spectra were collected from the 1842 samples.

Data Analysis

The preliminary analysis indicated the effectiveness of applying second derivative to the NIR spectra with left and right gaps of 4 nm. Also, the data were smoothed using the Savitzky–Golay approach (Savitzky and Golay 1964). All models in this study were fit on the data after pretreatment. A total of 70% of the data ($N = 1404$) were used for model training while the remaining 30% ($N = 614$) were used to test the model. The test dataset was not used in the model development until the final model was selected and thus, maintained its independence.

Models here were fit using two approaches. The first approach extracted features by first applying PCA to reduce data dimensionality. The reduced dataset was then fed into the predictive ML models. The second approach fit the ML models using the full spectral data, and the two methods compared. Six different types of deep neural networks (DNNs) and tree-based ensemble algorithms were used in this stud. Python and Minitab statistical software were used in this study for the data analysis.

Deep neural networks. MLP known as a universal approximator (Hornik et al 1989) is one of the most widely used NNs for regression and classification. The model consists of input, hidden, and output layers (Fig 1). The number of hidden layers and neurons in each layer should be defined by the user or tuned using an optimization technique for hyperparameter tuning. Having several hidden layers could result in a deep model with higher complexity. MLP NN with a feedforward architecture is among the most basic types of deep models with a series of feedforward fully connected layers. The error backpropagation was used by applying the "Adam" optimization algorithm (with default parameter settings), which is a stochastic gradient descent optimization method

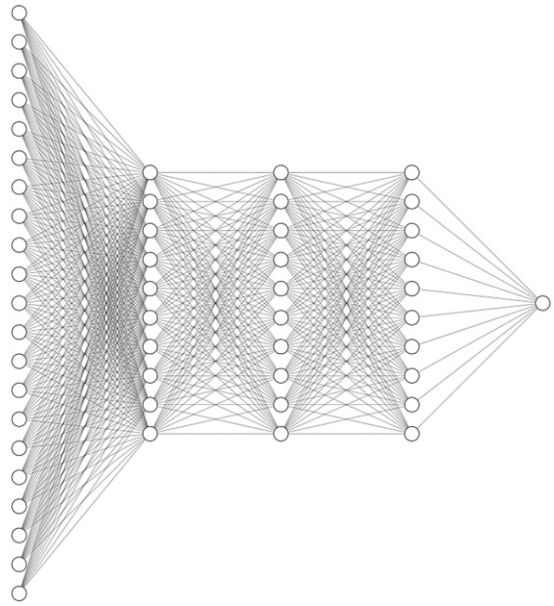


Figure 1. A schematic representation of the multilayer perceptron model used in this study with three hidden layers each having 10 neurons. The feed forward model with error backpropagation for hyperparameters tuning is among the most widely used ANNs. The shown sample network has 21 input parameters representing the case, in which PCA is applied to the dataset.

(Kingma and Ba 2014) to adjust the weight of neurons in the training process. The architecture of MLP (number of hidden layers and neurons) is shown to significantly impact the performance of the model. Three hidden layers were used in the MLP each having 10 neurons. The first hidden layer used "Swish" activation function while the remaining two used rectified linear unit ("ReLU"). Model training was done with a batch size of 32 and 25% of the training data were separately used as validation data to evaluate loss and model metrics at the end of each epoch. L2 regularization with a value of 1 e-3 was employed to prevent overfitting. Early stopping method was used to monitor the validation loss, which stopped model training if the validation loss did not improve over 40 epochs. Thus, the model training was stopped after 425 epochs on the dataset without applying PCA and after 521 epochs on the dataset reduced using PCA.

The performance of MLP NN was compared with that obtained from CNN. CNN is an effective ML algorithm that requires significantly fewer parameters than traditional DNNs. This is due to the convolution operation, sparse connectivity, and weight sharing in the network, resulting in much fewer preprocessing operations compared with other neural networks (Liu 2018). CNNs are a subset of DL models, which is a class of ML that uses multilayered ANNs to deliver state-of-the-art accuracy in tasks, such as object detection, speech recognition, language translation, image classifications, and so on (Alzubaidi et al 2021). 1D-CNNs are similar to 2-D CNNs, but are used mainly on 1D signals. In 1D-CNN, the convolutional kernel/filter moves in just one direction to calculate the output, and the output is a 1D signal. Modeling using 1D-CNN performs convolution operation on data and extracts significant features from the raw input data. Convolution involves sliding the kernel over the input signal, which is also known as a shift-compute procedure. The architecture of a typical CNN model is shown in Fig 2. The architecture of the 1D-CNN model consists of 72 filters, each with a size of 9, and rectified linear unit (ReLU) as an activation function for the convolution layer. The convolution layer's output is then flattened, followed by two dense layers having ReLU as an activation function with 64 and 32 neurons, respectively. To prevent overfitting, the first dense layer employs L2 regularization with a value of 1e-5. The model was constructed using the Adam optimizer with a learning rate of 0.0005 and a batch size of 27. During the training of the model, an early stopping technique is used to prevent overfitting by terminating the training if there is no improvement in validation loss for 30 epochs.

Tree-based ensemble learning. Ensemble learning is an approach to ML, which aims to enhance predictive performance through combining the prediction of multiple models that could be considered as weak learners. While it can work for a variety of algorithms, decision trees are commonly used in this method. The intuitive nature of decision trees make them easy for interpretation. Decision trees were used for predicting the mechanical properties and check formation in weathered timber (Nasir et al 2021a, c; van Blokland et al 2021a, b) using the classification and regression trees (CART) algorithm (Steinberg and Colla 2009). Decision trees have some advantages, such as using a white-box model that is easy to understand, require little preprocessing of data, and can handle both numerical and categorical data as well as missing data. However, they are prone to overfitting by creating very complex trees, which suffer from the generalization issue. This highlights the importance of the pruning phase

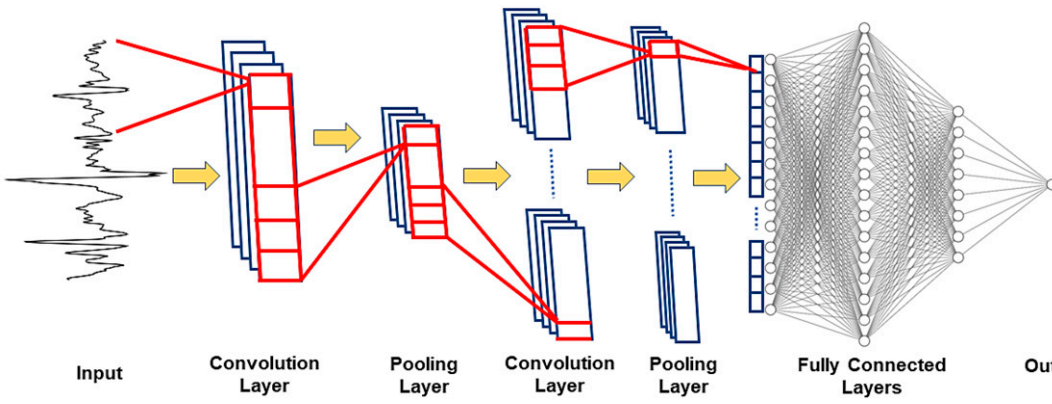


Figure 2. A sample representation of a typical 1D-CNN showing the convolution and pooling layers. The down-sampled data are fed into a fully connected layer(s) for predicting the output. The number of convolution and pooling layers along with its hyperparameters should be defined by the user or optimized.

and imposing conditions, such as defining the maximum number of terminal nodes or tree depth. Also, decision trees may become unstable due to variation in the data. That said, decision trees can benefit from ensemble learning algorithms. Two of the most common ensemble algorithms, bagging and boosting, were used in this study. Both methods had embedded feature selection characteristics that enables ranking of the relative importance of the features (wavelengths, in this study) through considering the contribution of the nodes having that feature when finding an optimal decision (Nasir et al 2021b).

Employing the bagging approach, random forest uses a subset of data to train multiple decision trees and aggregates the decisions of the trees known as weak learners through a voting scheme for the final prediction (Liaw and Wiener 2002). Random forest has been applied to the estimation of mechanical properties of wood fiber insulation boards (Schubert et al 2020). It has also used in wood machining for monitoring tool wear classification (Nasir et al 2021d) and identifying frozen and green lumber during circular sawing of logs (Nasir et al 2020). The random forest in this study contained 3000 trees and 80% of the training dataset was randomly chosen to identify bootstrap sample size. Additionally, a minimum number of three cases was specified to split an internal node. Finally, the number of predictors for node splitting was set to be the root mean root square of the total number of predictors. Figure 3 shows a schematic representation of a random forest model.

While the training process of weak learners in the bagging method is independent and in parallel, it is sequentially in the boosting method as the performance of a prior model is modified in the subsequent one (Fig 4). This approach has been used for wood species identification (Sun et al 2021), tool wear classification during wood machining (Nasir et al 2021d), predicting the properties of wood composites (Carty et al 2015), and developing an online color classification model for solid wood flooring (Zhuang et al 2021). In the TreeNet gradient-boosting machine, a random portion of the training data are utilized to calibrate a CART model with the maximum number of terminal

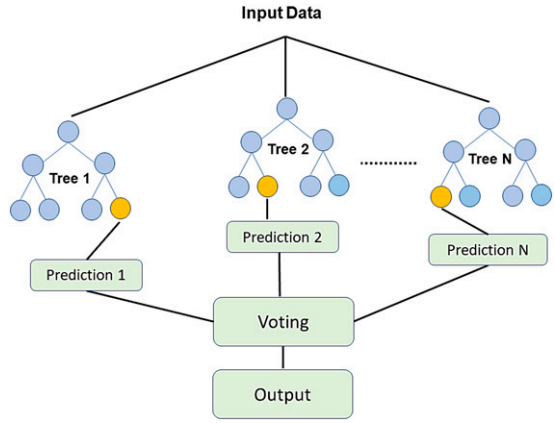


Figure 3. Schematic representation of a random forest model showing multiple decision trees trained separately using a random subset of data. Then a voting scheme is used for the final prediction.

nodes or the depth of the tree predefined. The maximum number of terminal nodes per tree was set to 12 and the minimum number of cases allowed for a terminal node was set to six. The subsample fraction in this study was set to 0.25. The CART model should then update based on the loss function. However, the update shrinks by a learning rate, which was set to 0.1 following a trial and error process during preliminary analysis. Subsequently, additional CARTs are added to improve the error for a specified number of iterations that defines the number of trees in the model (Modeler 2019) (set to 5000 in this study). Finally, the number of predictors for node splitting was set to be the root mean root square of the total number of predictors.

XGBoost is a scalable and distributed gradient-boosting algorithm for various ML problems (Chen and Guestrin 2016). It uses decision trees as base models and trains them in a stagewise manner, creating a strong composite model by combining multiple weak models. XGBoost is designed to improve the model generalization and handle sparse data efficiently. In this optimal model, the number of estimators was set to 1455, the maximum depth of the trees in the model was set to 5, the learning rate was defined as 0.027, the subsample rate was equal to 0.33, and

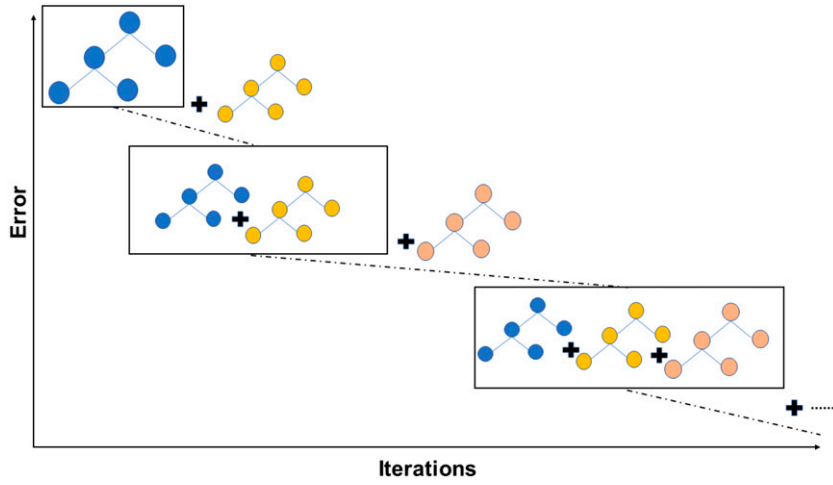


Figure 4. In TreeNet model, the gradient-boosting algorithm performs the training process sequentially, in which subsequent models correct the performance of prior models. The process starts with training a classification and regression tree (CART) model. It is then updated and CART models are sequentially added for a specified number of iterations lowering the error.

colsample_bytree and gamma were also set to 0.25 and 0.50, respectively. The colsample_bytree is set to 0.25 determines 25% of the columns are randomly chosen for each tree, and a gamma value of 0.5 means that a node will only be split if it results in a reduction of the loss function greater than or equal to 0.5, which potentially prevent overfitting and enhance the model's generalization ability.

LGBM is also an efficient, scalable, and optimized tree-based learning algorithm using a histogram-based decision tree learning algorithm that reduces communication overhead and minimizes memory usage, resulting in faster training times and improved performance on large datasets (Ke et al 2017). Its strong predictive power and efficiency have made it popular in various applied ML tasks. In the optimal model, the bagging_fraction and bagging_freq, which respectively determine the subsample ratio and frequency of bagging during training, were set to 0.42 and 0.53, respectively. It indicates that 42% of the data are sampled for each tree, and bagging is applied every 53 iterations in the training process. Also, the L1 and L2 regularization terms on the weights (by adding a penalty term to the loss function to reduce the magnitude of the weights and prevent overfitting) were 0.00028 and 0.043, respectively. The learning rate was set to 0.011, and no restriction was put on the maximum depth of the tree allowing the model to learn complex interactions between features until all leaves are pure or until the num_leaves parameter is reached; for instance, in this case, the number of leaves is set to 95. Finally, the number of estimators was equal to 1600. In this study, the optimal hyperparameters for XGBoost and LGBM were selected using the Python API Optuna (Akiba et al 2019). The Tree-Structured Parzen Estimator (TPE) sampler in Optuna was used, which is a Bayesian optimization-based sampler that uses a probabilistic model to guide the search process of hyperparameters. The TPE uses a tree-structured representation of the search space to model the probability distribution of the target function and generates new samples in regions likely to contain optimal solutions.

RESULTS AND DISCUSSION

Figure 5 shows the distribution for tracheid width for the 2018 samples measured. The histogram shows that the tracheid width of the samples had an average of 40.67 μm with a standard deviation of 2.75 μm. The normality test using the Ryan–Joiner method yielded a p-value < 0.01

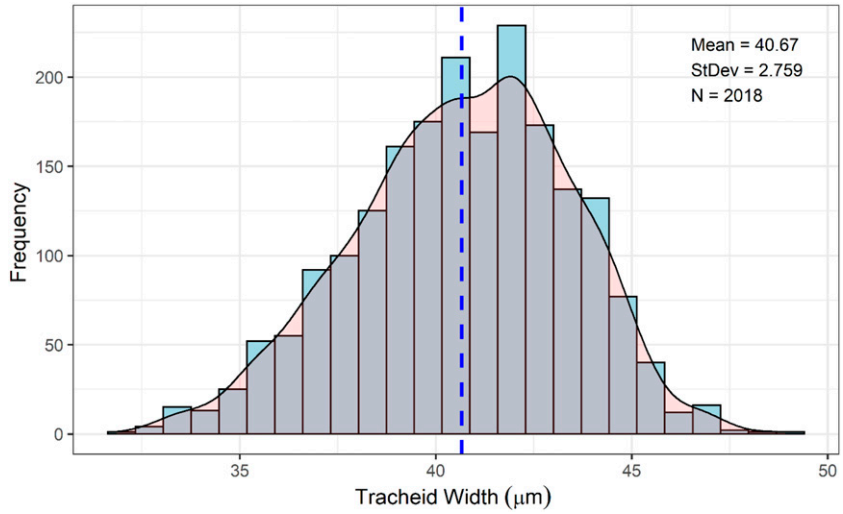


Figure 5. Histogram showing the distribution of tracheid width.

rejecting the null hypothesis of normal distribution. It can be seen that the measured tracheid width varies in the range of 31-49 μm. Table 1 summarizes tracheid width descriptive statistics for the training and test data. The average and standard deviation of the tracheid width in both the training and test data are very similar to that of the total dataset indicating appropriate data splitting.

Prior to any analysis, the input dataset was standardized by rescaling the data to show an average of 0 with standard deviation of 1. Subsequently, PCA was performed on the input NIR data for dimensionality reduction. Figure 6 indicates the contribution of PCs to the data and reveals that using 21 PCs can explain 95% of variation in the data. This reduced the size of the input dataset to a matrix of 2018×21. Table 2 shows the summary of the performance (mean square error [MSE] and coefficient of determination [R^2]) obtained from the deep and ensemble models.

The lowest R^2 (test data) was obtained using random forest model followed by the MLP NN. 1D-CNN showed a slightly higher R^2 than MLP NN. Compared with the DNN models used in this study, the gradient-boosting machines (TreeNet, XGBoost, and LGBM) showed superior performance resulting in higher R^2 and lower MSE. The lowest error and highest R^2 was obtained from the LGBM model followed by TreeNet. Figure 7 shows the R^2 variation with respect to the number of trees in the TreeNet model. Although the model was expanded to comprise 5000 trees, the optimal performance (considering overfitting) was chosen when having 2159 trees in the model. Compared with DNNs, the gradient-boosting machine algorithms show overfitting on the training data, yet, yielded the best performance on the test data.

The relatively higher performance of LGBM model compared with other boosting methods, such as XGBoost, could be linked to its tree growth strategy, which is leaf-wise (it is level-wise for XGBoost) resulting in growing the tree in a selective manner, leading to smaller and faster models as compared with XGBoost. It is

Table 1. Descriptive statistics of the tracheid width for the training and test data.

Dataset	*N*	% of *N*	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Training	1404	69.57	40.65	2.71	32.59	38.85	40.80	42.60	48.92
Test	614	30.43	40.71	2.86	31.86	38.81	40.82	42.77	47.58

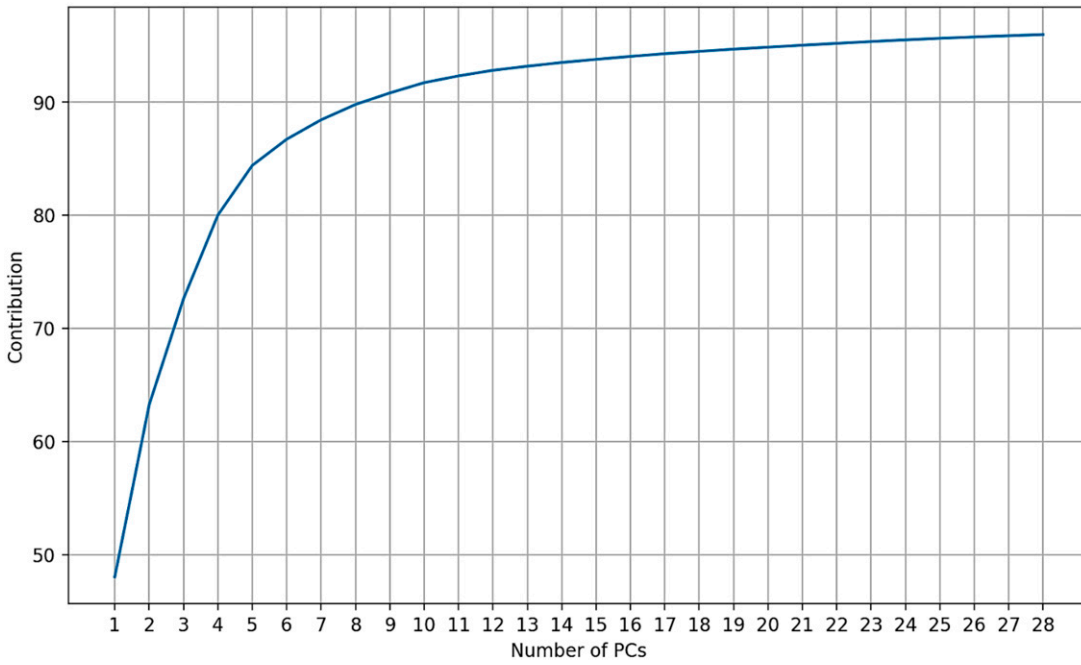


Figure 6. Contribution of PCs to the data.

observed that LGBM performs slightly better than XGBoost, which could be attributed to either a better hyperparameter tuning or a better model fitting. As the hyperparameters of both LGBM and XGBoost are different, it could be that the optimal hyperparameters for LGBM are more suitable for the given dataset. Additionally, the LGBM model is a better fit for the data and was able to capture more complex relationships within it.

Table 2. Performance summary of the deep and ensemble models for prediction (PCA is applied).

	Performance (R^2)			
	MSE		R^2	
Models	Train	Test	Train	Test
MLP NN	2.88	3.17	0.61	0.61
1D-CNN	2.53	3.45	0.63	0.62
Random Forest	3.05[a]	3.39	0.58[a]	0.58
TreeNet	0.54	2.94	0.93	0.64
XGBoost	0.11	2.99	0.99	0.63
LGBM	0.27	2.86	0.96	0.65

[a]Results indicating the out-of-bag performance.

The spectral data were also fed into the ML models without applying PCA. Table 3 summarizes MSE and R^2 obtained from the different models when full-range spectral data were directly used for prediction. The lowest R^2 and highest MSE (test data) corresponded to MLP NN. Random forest and 1D-CNN revealed similar performance while both slightly outperformed the MLP NN. Similar to Table 2, it can be seen that gradient-boosting algorithms outperformed the NNs and random forest. The highest R^2 of 0.69 was achieved using the LGBM followed by TreeNet and XGBoost, which showed similar performance ($R^2 = 0.68$). Figure 8 shows the variation in R^2 vs the number of trees in the TreeNet model. Optimal performance considering R^2 of 0.68 and overfitting was obtained using a model with 1959 trees. It can be seen that feeding the full range of NIR data into the models without applying PCA enhances the performance of the models.

All of the ML models fit here had improved fit statistics compared with the R^2 of 0.61 (test data) achieved by Nabavi et al (2018) for tracheid width

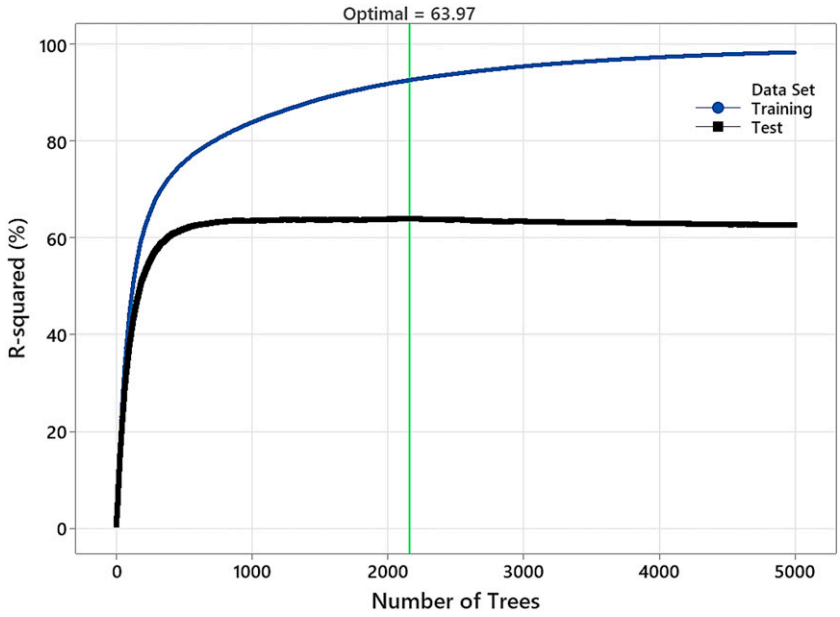


Figure 7. Variation in the R^2 vs the number of trees in the TreeNet model (PCA was applied to the input data).

prediction using PLS regression. The slightly higher performance of CNN compared with MLP could be associated with the high-level data abstraction capability of CNN. While MLP could be considered a DNN model having three hidden layers in its structure, it is one of the most basic types of NNs. Representation of raw data are learned in a DL model, such as CNN, through multiple levels of abstraction (LeCun et al 2015; Miotto et al 2018). Thus, one can conclude that high-level data abstraction and the built-in feature engineering capability of CNN during its learning process results in superior performance compared with a more traditional model, such as MLP NN.

Table 3. Performance summary of the deep and ensemble models for prediction (PCA is not applied and the entire spectral data were directly fed into the models).

	Performance			
	MSE		R^2	
Models	Train	Test	Train	Test
MLP NN	2.73	3.03	0.63	0.63
1D-CNN	1.73	2.95	0.77	0.64
Random Forest	2.81[a]	2.92	0.61[a]	0.64
TreeNet	0.36	2.64	0.95	0.68
XGBoost	0.16	2.57	0.99	0.68
LGBM	0.09	2.56	0.99	0.69

[a]Results indicating the out-of-bag performance.

The higher performance of gradient-boosting algorithms over NNs is aligned with the benchmark study of Grinsztajn et al (2022), in which tree-based models, such as XGBoost, outperformed the NNs on medium-sized data (approximately 10 K samples). The competitive performance of tree-based models even when dealing with an irregular pattern in the target function and uninformative features makes them powerful tools for regression and classification tasks on tabular datasets, while MLP-like architectures are noticeably affected by uninformative features (Grinsztajn et al 2022). This could explain the lower performance of MLP in this study. The superior performance of gradient-boosting model over random forest was also reported when classifying tool wear during wood machining (Nasir et al 2021d). However, further studies on datasets with different sample size and complexity is needed to benchmark this observation.

The embedded feature selection nature of tree-based ensemble learning models resulted in

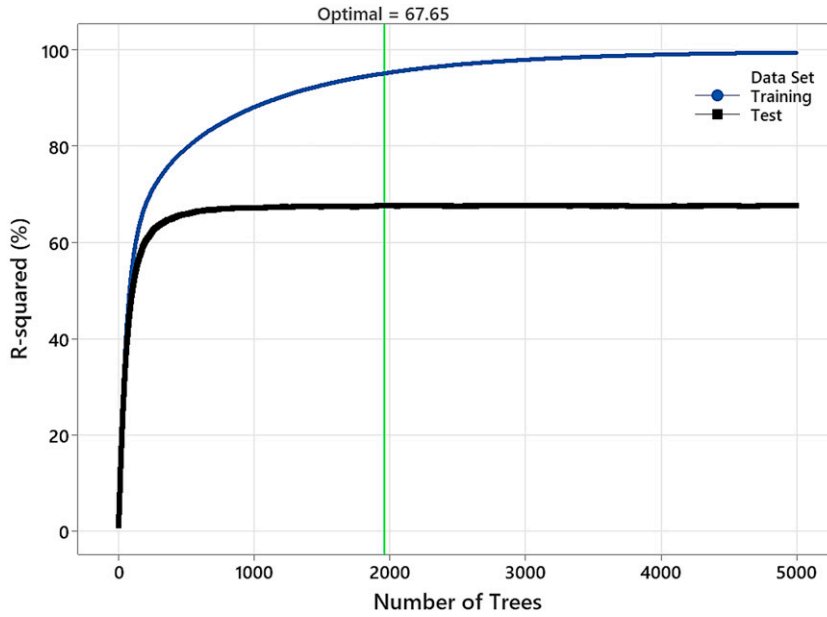


Figure 8. Variation in the R^2 vs the number of trees in the TreeNet model (PCA was not applied to the input data).

identifying the relative importance of features (wavelengths). For example, the results showed that all 692 wavelengths were important in the TreeNet model, while the random forest model had 516 important predictors. The top 10 wavelengths for the TreeNet and random forest models are summarized in Figs 9 and 10. In both, the relative importance of features are shown, where the most important was assigned a score of 100, with the remaining scaled accordingly.

Band assignments reported by Schwanninger et al (2011) for wood components were used to determine, where possible, the origins of the features shown in Figs 9 and 10. Four wavelengths identified for the TreeNet model (including the top three) were in the range 1540-1546 nm and this region is associated with first overtone O-H stretch vibrations in cellulose. Wavelengths at 1542 and 1544 nm were also identified as the top two important features in LGBM and XGBoost

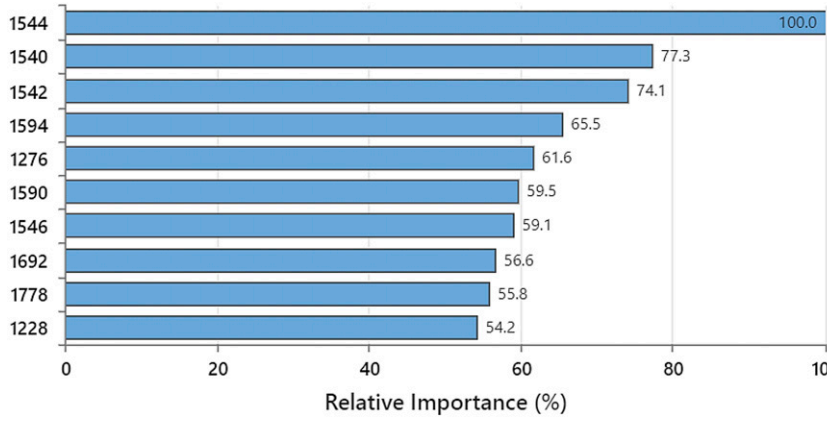


Figure 9. Top 10 important wavelengths contributing to the obtained predictive TreeNet model.

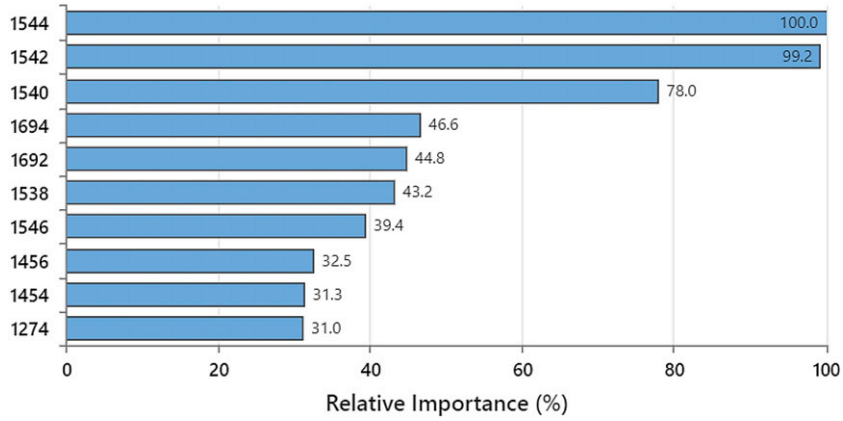


Figure 10. Top 10 important wavelengths contributing to the obtained predictive random forest model.

models. Wavelengths at 1590 and 1594 nm were also among the most important and bond vibrations in this region (1579-1597 nm) share the same origin. Wavelengths in both regions have been assigned to specific bond types in cellulose (Schwanninger et al 2011), which are summarized in Table 4.

Of the remaining four wavelengths identified by the TreeNet model, two (1692 and 1778 nm) occurred in regions associated with first overtone C-H stretch vibrations in lignin (1685 and 1698 nm) and cellulose (eg 1780, 1788, and 1790 nm). Similarly, 1228 nm is in a region (1212-1225 nm) assigned to second overtone C-H stretch vibrations in cellulose. The only wavelength without a recognized band assignment arising from a specific wood component was 1276 nm.

Table 4. The specific bond types in cellulose for different wavelengths.

Wavelength (nm)	Specific cellulose
1540	Adsorbed water (strong H-bond with microcrystalline cellulose)
1545	Intramolecular H-bond in cellulose
1548	Crystalline regions in C1. O(3)-H(3)$\cdots\cdots$O(5) intrachain H-bond
1588	Strong O(2)-H(2)$\cdots\cdots$O(6) of cellulose
1591	H-bonds of the cellulose 1_3 phase
1592	Crystalline regions in cellulose
1597	Strongly H-bonded O-H group in cellulose 1_α

The Random Forest model shared many frequently used features with models developed using TreeNet model and wavelengths (1544, 1542, and 1540 nm) associated with first overtone O-H stretch vibrations in cellulose were again the most important. The 1546 nm was also identified (ranked seventh for both models), but its relative importance was lower (39.4% vs 59.1% for TreeNet). Further, an additional wavelength (1538 nm) from this region was identified among the top 10 features. First overtone C-H stretch vibrations in lignin (1685 and 1698 nm) were again important, with two wavelengths (1690 and 1692 nm) identified (vs 1692 nm for TreeNet). Two additional wavelengths (1454 and 1456 nm) were important and occur close to 1447 and 1448 nm, wavelength which arise from first overtone O-H stretch vibrations in lignin. The only wavelength lacking a specific assignment was 1274 nm. The presence of recognized bond vibrations for the majority of important features arising from specific wood components indicate the importance of cellulose and lignin in the development of our tracheid width models. In general, the wavelengths identified for the model types were very similar; however, the random forest model appears to have greater emphasis on lignin owing to the presence of 1454 and 1456 nm and the corresponding absence of 1590 and 1594 nm (the cellulose-related wavelengths important for the TreeNet model).

Overall, methods such as tree-based ensemble models based on full-range NIR data could be

applied for property assessment of a variety of wood products. Not only may this enhance predictive performance of the developed model, but it provides an opportunity for NIR feature ranking and studying the relative importance of wavelengths in contributing to the performance of the predictive model. For example, this study highlighted the importance of wavelengths related to cellulose and lignin in developing a model for predicting tracheid width. Future studies should focus on the role of dataset size when employing different ML and DL models. Further expanding dataset size through employing synthetic data augmentation methods, such as generative adversarial network (GAN), could be an important topic of future studies.

CONCLUSIONS

Overall, the tree-based gradient-boosting machines outperformed the neural network models. LGBM yielded the highest R^2 and lowest error followed by XGBoost and TreeNet, whereas using MLP NN did not result in high performance. All models performed better without PCA when full-range NIR spectra were directly fed into them. The embedded feature selection characteristic of the tree-based ensemble models could be used to study the relative importance of features indicating the importance of cellulose and lignin in the development of tracheid width models. The developed models improved the reported results for tracheid width prediction using a PLS model, demonstrating the potential of DL, and specifically tree-based ensemble models, when applied to a medium-sized NIR dataset. The methodology could be applied to any wood product for which NIR spectral data can be collected and provides better predictive performance and improved insight into the relative importance of individual NIR wavelengths. Since the performance could be impacted by the size of data and its complexity, the proposed models should be tested on other NIR datasets, specifically those of small sample size, which is more typical when collecting data under laboratory conditions. Finally, the models could also be tested on big data; however, this requires large-scale data acquisition to better simulate industrial conditions and account for the high-variability of real-life situations. Future research could study employing synthetic data augmentation methods for increasing the size of datasets and how it impacts the performance of predictive models based on the NIR data spectral.

ACKNOWLEDGMENTS

This research was possible through support from the Wood Quality Consortium (WQC) at the University of Georgia, the National Science Foundation (NSF), Center for Advanced Forestry Systems (CAFS), and the NIFA McIntire-Stennis project 1023340. The authors wish to thank the WQC, CAFS, and NIFA for funding this project.

This work is partially supported by a cooperative agreement, 22-CO-11111137-017 between the US Department of Agriculture Forest Service Forest Products Laboratory and the University of Georgia through the US Department of Energy Award Number: DE-EE0008911 along with the University of Wisconsin.

REFERENCES

Akiba T, Sano S, Yanase T, Ohta T, Koyama M (2019) Optuna: A next-generation hyperparameter optimization framework. Pages 2623-2631 *in* Proceedings, 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, August 4-8, 2019, Anchorage, AK. Association for Computing Machinery, New York, NY.

Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Farhan L (2021) Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. J Big Data 8(1):1-74.

Ayanleye S, Nasir V, Avramidis S, Cool J (2021) Effect of wood surface roughness on prediction of structural timber properties by infrared spectroscopy using ANFIS, ANN and PLS regression. Eur J Wood Wood Prod 79:101-115.

Antony F, Schimleck LR, Daniels RF, Clark A III, Hall DB (2010) Modeling the longitudinal variation in wood specific gravity of planted loblolly pine (*Pinus taeda*) in the United States. Can J For Res 40(12):2439-2451.

Carty DM, Young TM, Zaretzki RL, Guess FM, Petutschnigg A (2015) Predicting and correlating the strength properties of wood composite process parameters by use of boosted regression tree models. For Prod J 65(7-8):365-371.

Carvalho MG, Ferreira PJ, Martins AA, Figueiredo MM (1997) A comparative study of two automated techniques for measuring fiber length. Tappi J 80(2):137-142.

Chandrashekar G, Sahin F (2014) A survey on feature selection methods. Comput Electr Eng 40(1):16-28.

Chen T, Guestrin C (2016). XGBoost: A scalable tree boosting system. Pages 785-794 *in* Proceedings, 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, NY.

Costa LR, Tonoli GHD, Milagres, FR, Hein PRG (2019) Artificial neural network and partial least square regressions for rapid estimation of cellulose pulp dryness based on near infrared spectroscopic data. Carbohydr Polym 224:115186.

Dahlen J, Nabavi M, Auty D, Schimleck L, Eberhardt TL (2021) Models for predicting the within-tree and regional variation of tracheid length and width for plantation loblolly pine. Forestry 94:127-140.

Evans R (1994) Rapid measurement of the transverse dimensions of tracheids in radial wood sections from *Pinus radiata.* Holzforschung 48(2):168-172.

Franklin G (1945) Preparation of thin sections of synthetic resins and wood-resin composites, and a new macerating method for wood. Nature 155(3924):51-55.

Grinsztajn L, Oyallon E, Varoquaux G (2022) Why do tree-based models still outperform deep learning on tabular data? arXiv preprint arXiv:2207.08815.

Hein PR, Pakkanen H, Santos AAD (2017) Challenges in the use of near infrared spectroscopy for improving wood quality: A review. For Syst 26(3):eR03.

Ho TX, Schimleck LR, Dahlen J, Sinha A (2022) Utilization of genetic algorithms to optimize loblolly pine wood property models based on NIR spectra and SilviScan data. Wood Sci Technol 56(5):1419-1437.

Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2(5):359-366.

Huang Y, Meng S, Zhao P, Li C (2019) Wood quality of Chinese zither panel based on convolutional neural network and near-infrared spectroscopy. Appl Opt 58(18): 5122-5127.

Inagaki T, Schwanninger M, Kato R, Kurata Y, Thanapase W, Puthson P, Tsuchikawa S (2012) *Eucalyptus camaldulensis* density and fiber length estimated by near-infrared spectroscopy. Wood Sci Technol 46(1-3):143-155.

Jones PD, Schimleck LR, So CL, Clark A, Daniels RF (2007) High resolution scanning of radial strips cut from increment cores by near infrared spectroscopy. IAWA J 28(4): 473-484.

Jordan L, Clark A, Schimleck LR, Hall DB, Daniels RF (2008) Regional variation in wood specific gravity of planted loblolly pine in the United States. Can J For Res 38(4):698-710.

Kanayama H, Ma T, Tsuchikawa S, Inagaki T (2019) Cognitive spectroscopy for wood species identification: Near infrared hyperspectral imaging combined with convolutional neural networks. Analyst (Lond) 144(21): 6438-6446.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Liu TY (2017) LightGBM: A highly efficient gradient boosting decision tree. Pages 3146-3154 *in* I Guyon, UV Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, and R Garnett, eds. Advances in neural information processing systems. Curran Associates, New York, NY.

Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436-444.

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3):18-22.

Liu YH (2018) Feature extraction and image recognition with convolutional neural networks. J Phys Conf Ser 1087(6):062032. IOP Publishing.

Mao KZ (2005) Identifying critical variables of principal components for unsupervised feature selection. IEEE Trans Syst Man Cybern B Cybern 35(2):339-344.

Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2018) Deep learning for healthcare: Review, opportunities and challenges. Brief Bioinform 19(6):1236-1246.

Nabavi M, Dahlen J, Schimleck L, Eberhardt TL, Montes C (2018) A regional calibration model for predicting loblolly pine tracheid properties using near-infrared spectroscopy. Wood Sci Technol 52(2):445-463.

Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E (2015) Deep learning applications and challenges in big data analytics. J Big Data 2(1):1-21.

Nasir V, Ayanleye S, Kazemirad S, Sassani F, Adamopoulos S (2022) Acoustic emission monitoring of wood materials and timber structures: A critical review. Constr Build Mater 350:128877.

Nasir V, Dibaji S, Alaswad K, Cool J (2021a) Tool wear monitoring by ensemble learning and sensor fusion using power, sound, vibration, and AE signals. Manuf Lett 30:32-38.

Nasir V, Fathi H, Fallah A, Kazemirad S, Sassani F, Antov P (2021b) Prediction of mechanical properties of artificially weathered wood by color change and machine learning. Materials (Basel) 14(21):6314.

Nasir V, Fathi H, Kazemirad S (2021c) Combined machine learning–wave propagation approach for monitoring timber mechanical properties under UV aging. Struct Health Monit 20(4):2035-2053.

Nasir V, Kooshkbaghi M, Cool J (2020) Sensor fusion and random forest modeling for identifying frozen and green wood during lumber manufacturing. Manuf Lett 26:53-58.

Nasir V, Kooshkbaghi M, Cool J, Sassani F (2021d) Cutting tool temperature monitoring in circular sawing: Measurement and multi-sensor feature fusion-based prediction. Int J Adv Manuf Technol 112(9):2413-2424.

Nasir V, Nourian S, Zhou Z, Rahimi S, Avramidis S, Cool J (2019) Classification and characterization of thermally modified timber using visible and near-infrared spectroscopy and artificial neural networks: A comparative study on the performance of different NDE methods and ANNs. Wood Sci Technol 53(5):1093-1109.

Nasir V, Sassani F (2021) A review on deep learning in machining and tool monitoring: Methods, opportunities, and challenges. Int J Adv Manuf Technol 115(9): 2683-2709.

Pereira H, Santos AJA, Anjos O (2015) Fibre morphological characteristics of kraft pulps of *Acacia melanoxylon* estimated by NIR-PLS-R models. Materials (Basel) 9(1):8.

Salford Predictive Modeler by Minitab (2019) Introducing TreeNet gradient boosting machine. Stage College, PA.

Sandak J, Sandak A, Meder R (2016) Assessing trees, wood and derived products with near infrared spectroscopy: Hints and tips. J Near Infrared Spectrosc 24(6): 485-505.

Savitzky A, Golay MJ (1964) Smoothing and differentiation of data by simplified least squares procedures. Anal Chem 36(8):1627-1639.

Schimleck L, Dahlen J, Apiolaza LA, Downes G, Emms G, Evans R, Wang X (2019) Non-destructive evaluation techniques and what they tell us about wood property variation. Forests 10(9):728.

Schimleck LR, Evans R (2004) Estimation of *P. radiata* D. Don tracheid morphological characteristics by near infrared spectroscopy. Holzforschung 58(1):66-73.

Schubert M, Luković M, Christen, H (2020) Prediction of mechanical properties of wood fiber insulation boards as a function of machine and process parameters by random forest. Wood Sci Technol 54(3):703-713.

Schwanninger M, Rodrigues JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. J Near Infrared 19(Spec): 287-308.

Steinberg D, Colla P (2009) CART: Classification and regression trees. Page 179 *in* X Wu and V Kumar eds. The top ten algorithms in data mining, Volume 9. CRC Press, Boca Raton, FL.

Sun Y, Lin Q, He X, Zhao Y, Dai F, Qiu J, Cao Y (2021) Wood species recognition with small data: A deep learning approach. Int J Comput Intell Syst 14(1):1451-1460.

Tsuchikawa S (2007) A review of recent near infrared research for wood and paper. Appl Spectrosc Rev 42(1): 43-71.

Tsuchikawa S, Kobori H (2015) A review of recent application of near infrared spectroscopy to wood science and technology. J Wood Sci 61(3):213-220.

Tsuchikawa S, Schwanninger M (2013) A review of recent near-infrared research for wood and paper (Part 2). Appl Spectrosc Rev 48(7):560-587.

van Blokland J, Nasir V, Cool J, Avramidis S, Adamopoulos S (2021a) Machine learning-based prediction of surface checks and bending properties in weathered thermally modified timber. Constr Build Mater 307:124996.

van Blokland J, Nasir V, Cool J, Avramidis S, Adamopoulos S (2021b) Machine learning-based prediction of internal checks in weathered thermally modified timber. Constr Build Mater 281:122193.

Via BK, Stine M, Shupe TF, So CL, Groom L (2004) Genetic improvement of fiber length and coarseness based on paper product performance and material variability—A review. IAWA J 25(4):401-414.

Wang J, Ma Y, Zhang L, Gao RX, Wu D (2018) Deep learning for smart manufacturing: Methods and applications. J Manuf Syst 48:144-156.

Watanabe K, Kobayashi I, Matsushita Y, Saito S, Kuroda N, Noshiro S (2014) Application of near-infrared spectroscopy for evaluation of drying stress on lumber surface: A comparison of artificial neural networks and partial least squares regression. Dry Technol 32(5): 590-596.

Willems W, Lykidis C, Altgen M, Clauder L (2015) Quality control methods for thermally modified wood. Holzforschung 69(7):875-884.

Yang SY, Kwon O, Park Y, Chung H, Kim H, Park SY, Yeo H (2020) Application of neural networks for classifying softwood species using near infrared spectroscopy. J Near Infrared Spectrosc 28(5):298-307.

Yun YH, Li HD, Deng BC, Cao DS (2019) An overview of variable selection methods in multivariate analysis of near-infrared spectra. Trends Analyt Chem 113:102-115.

Zhen Y, Ho TX, Roberts L, Schimleck LR, Sinha A (2022) On the selection of the weighting parameter value in optimizing *Eucalyptus globulus* pulp yield models based on NIR spectra. Wood Sci Technol 56(6):1835-1850.

Zhuang Z, Liu Y, Ding F, Wang Z (2021) Online color classification system of solid wood flooring based on characteristic features. Sensors (Basel) 21(2):336.

(ISSN 0735-6161)
CODEN WOOFAJ

WOOD AND FIBER SCIENCE

JOURNAL OF THE SOCIETY OF WOOD SCIENCE AND TECHNOLOGY

VOLUME 55 JUNE 2023 NUMBER 1

CONTENTS