Exploring the potential of field spectroscopy for tree species identification in different Amazonian forest ecosystems

Hilana Louise Hadlich [a,*,1], Jochen Schöngart [a,2], Florian Wittmann [a,b,3], Caroline C. Vasconcelos [a,4], Caroline L. Mallmann [c,5], Maíra L.G. Conde [a,6], Priscila Amaral de Sá [d,7], Layon O. Demarchi [a,8], Gisele Biem Mori [a,d,9], Maria T.F. Piedade [a,10], Flavia Machado Durgante [a,b,**,11]

[a] *National Institute for Amazonian Research (INPA), Ecology, Monitoring and Sustainable Use of Wetlands (MAUA Group), Manaus, Amazonas, Brazil*
[b] *Karlsruhe Institute of Technology (KIT), Department of Wetland Ecology, Rastatt, Germany*
[c] *Federal University of Santa Maria (UFSM), Department of Geoscience, Santa Maria, Rio Grande do Sul, Brazil*
[d] *University of the State of Mato Grosso (UNEMAT), Postgraduate Program in Ecology and Conservation, Nova Xavantina, Mato Grosso, Brazil*

ARTICLE INFO

ABSTRACT

Species identification in Amazonian forest inventories is challenging due to a shortage of taxonomists, high biodiversity, and morphological similarities leading to taxonomic errors. Full-range spectroscopy (350–2500 nm) is a promising tool for improving species identification efficiency and reliability. This study assessed the effectiveness of full-range spectroscopy in discriminating against 26 abundant tree species across three Amazonian ecosystems: upland forest, white-sand ecosystems, and floodplain forest, using spectral data from different tree tissues—outer bark, inner bark, and fresh leaves. Each tissue was tested using Linear Discriminant Analysis (LDA) spectral models with two cross-validation methods: leave-one-out and holdout (70/30), as well as external validation with samples from different sites. Results showed high discrimination accuracy for all tissues and ecosystems. The general models achieved 86 % accuracy for outer bark, 97 % for inner bark and 98 % for fresh leaves. The most informative spectral bands varied by tissue type: SWIR I (1300–1900 nm) for outer bark, and VIS (400–700 nm) + SWIR I (1300–1900 nm) for inner bark and fresh leaves. A general model

* Corresponding author.
** Corresponding author at: Karlsruhe Institute of Technology (KIT), Department of Wetland Ecology, Rastatt, Germany.
E-mail addresses: hilana@posgrad.inpa.gov.br (H.L. Hadlich), flavia.durgante@kit.edu (F.M. Durgante).
[1] https://orcid.org/0000-0002-3054-8574
[2] https://orcid.org/0000-0002-7696-9657
[3] https://orcid.org/0000-0001-9180-356X
[4] https://orcid.org/0000-0002-7850-6672
[5] https://orcid.org/0000-0002-2967-1313
[6] https://orcid.org/0000-0003-1074-3687
[7] https://orcid.org/0009-0000-4555-9658
[8] https://orcid.org/0000-0001-8441-2106
[9] https://orcid.org/0000-0002-6825-9508
[10] https://orcid.org/0000-0002-7320-0498
[11] https://orcid.org/0000-0002-5517-8821

integrating species across ecosystems confirmed full-range spectroscopy as an effective tool for in-field species identification. These findings highlight the potential of spectroscopy to Amazonian biodiversity inventories, contributing to more accurate species identification, refining forest management and conservation efforts.

1. Introduction

Conserving biodiversity of tropical forests is a challenge, especially given the lack of knowledge about the totality of species that occur in the vast mosaic of habitats and ecosystems that compose these forests. Although there have been many efforts to find out how many tree species exist in the Amazon (Cardoso et al., 2017; Oliveira-Filho et al., 2021; ter Steege et al., 2019, 2016, 2013) and it is estimated that there are more than 15,000 tree species (ter Steege et al., 2020), but only around 6000 have been cataloged (Cardoso et al., 2017). The challenges of delimiting and quantifying species persist due to taxonomic uncertainties (Draper et al., 2020; Stropp et al., 2022) and lack of botanical collections in remote regions (Hopkins, 2019, 2007; Nelson et al., 1990).This uncertainty directly affects biodiversity inventories that monitor tropical forests and compromises subsequent activities that depend on this information, such as forest management and conservation, as well as scientific research and ecological modeling. Although the problem of species identification in the Amazon is well-known in forest management (Cysneiros et al., 2018; Procópio and Secco, 2008), it is still neglected due to the lack of efficient and quick inspection techniques. The consequences of misidentification can be both economic and ecological, including the overestimation of commercially valuable species populations, the compromise forest product transaction integrity (Baraloto et al., 2007; Gaui et al., 2019; Procópio and Secco, 2008) and facilitation of rare species exploitation and illegal logging, by identifying a rare species in a group of species, usually called hyperdominant. Therefore, mitigating identification errors in forest inventories is an urgent priority. It is essential to focus on training new parataxonomists and developing innovative technologies and methodologies to optimize the process.

Field monitoring of the Amazon rainforest is conducted through forest inventories, where the first step is to ensure species identification. This is a slow, laborious process and requires a high level of taxonomic knowledge. Identification begins in the field facing difficulties due to the high morphological similarity between tropical species, making this activity subjective and inaccurate (Gomes et al., 2013) and the scarcity of botanical taxonomists, since the samples collected in the field must be validated with reference collections in herbarium. However, the absence of reproductive material for accurate identification can lead to the accumulation of errors in herbarium, perpetuating misidentifications. This issue is evidenced by a mismatch of over 50 % in the names of tropical species (Gomes et al., 2013; Goodwin et al., 2015; Hopkins, 2007; Procópio and Secco, 2008). Therefore, those responsible for forest monitoring need species verification techniques beyond field identification to ensure quality control and reduce the errors associated with identification.

Some approaches have been developed to solve the taxonomic challenges, such as DNA barcoding, mass spectroscopy, visible and near-infrared spectroscopy (VIS-NIRS) in the field and the laboratory, and imaging spectroscopy. Each of these techniques has its advantages and limitations, and their integrated use is recommended to minimize identification errors (Damasco et al., 2019; Draper et al., 2020; Prata et al., 2018). Among the available tools, the *in situ* optical spectroscopy technique is a way to evaluate biodiversity and support forest management with the automation of data collection (Juola et al., 2023) and is suitable for optimizing forest inventories (Hadlich et al., 2018). Spectral reflectance operates by directing a light source onto organic samples and measuring the intensity of the light that is reflected or absorbed, producing characteristic spectral profiles in the visible and near-infrared regions (350–2500 nm, full-range spectroscopy - Cavender-Bares et al., 2025). These spectra, influenced by the chemical composition, cellular structure, and internal morphology of the samples (Ponzoni, 2002), can be analyzed using multivariate statistical techniques to extract qualitative and quantitative information (Foley et al., 1998; Pasquini, 2018, 2003; Tsuchikawa, 2007).

Over the years, spectral reflectance methods have been developed to discriminate between species, employing various data sources and both laboratory and field equipment. The technique has already been tested on herbarium leaf samples (Botelho, 2017; Durante et al., 2013; Lang et al., 2015; Meireles et al., 2020); laboratory-dried woods (Bergo et al., 2016; Braga et al., 2011; da Silva et al., 2024; Pastore et al., 2011; Pigozzo, 2011), fresh leaves in the field (Asner et al., 2014; Mallmann et al., 2023) and bark of standing trees (Hadlich et al., 2018; Juola et al., 2023, 2022), demonstrating remarkable potential for these applications. In an integrative context, spectral signature has also been effective in assisting the delimitation of species with cryptic morphological variation (Damasco et al., 2019; Prata et al., 2018), even hypothesizing new species (Vasconcelos et al., 2021; Vasconelos et al., 2020), which shows the importance of using multiple tools to solve eventual taxonomic errors, especially in the highly diverse Amazonian flora.

The application of spectral reflectance to improve species identification has shown promising results in the field mainly in upland forests based on bark (Hadlich et al., 2018) and leaf spectra (Asner et al., 2014) in the Amazon. However, its applicability in different Amazonian ecosystems remains largely unexplored. The Amazon is made up of distinct ecosystems in terms of floristics, soil types, topography and water availability, which form mosaics of landscapes, each with specificities that must be considered when monitoring biodiversity. The predominant vegetation is non-flooded upland forest (also known as *terra-firme* forest), which covers about 70 % of the basin. The remaining 30 % consists predominantly of wetland complexes, most of them forested, and which experience seasonal flooding (Junk et al., 2014, 2011). Additionally, other ecosystems are interspersed throughout the biome, including savannas, coastal vegetation, montane forests, and the white-sand ecosystem (locally known as *campinarana*), the latter covering approximately 5 % of the basin area (Adeney et al., 2016). Given this environmental heterogeneity, spectral-based identifications must be tailored to the specific conditions of each ecosystem to improve the accuracy of species identification as the spectral variation between tree species is
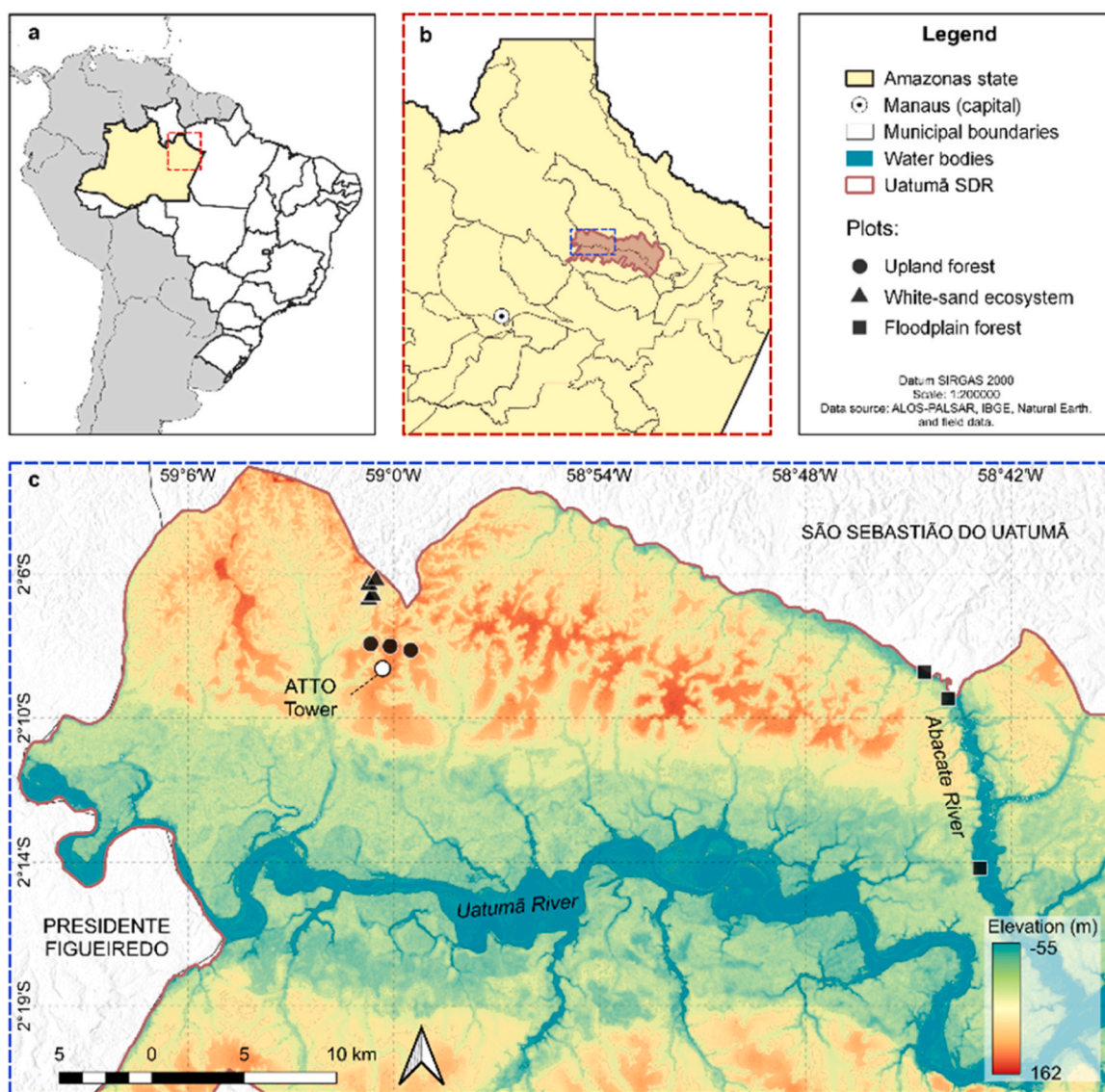
influenced by morphology, structural traits, chemical composition of tissues and environmental conditions (Clark and Roberts, 2012).

Inventorying Amazonian ecosystems is a complex task that requires strategic planning to minimize cost and maximize efficiency. Considering the specific limitations of each ecosystem - such as the inundation of flooded forests, phenology of most species, density of trees, accessibility to samples - during the planning phase can significantly optimize the process. To improve forest inventories using full-range spectroscopy for in-field species identification, four key questions have been addressed in this study: How effective are different plant tissues in predicting species identification in different Amazonian ecosystems? What is the efficiency of a local model combining all ecosystems? Is the local model effective in predicting species from other locations? Which wavelengths are the most informative for species recognition in different tissues? The findings will provide a practical framework for improving inventory methodologies, ensuring greater accuracy and reducing costs. This research will contribute to more reliable biodiversity assessments and better-informed conservation and management strategies by shifting from uncertain identifications to a high-precision approach.

2. Materials and methods

2.1. Study sites

Spectral data were systematically collected from established permanent plots as part of the Long-term Ecological Research Program

Fig. 1. Map of the study area in Central Amazon, showing the sampling permanent plots where spectral data were collected from tree species. Solid black markers represent plots in upland (circles), white-sand (triangles), and floodplain forests (squares).

(PELD) of the Working Group Ecology, Monitoring and Sustainable Use of Wetlands (MAUA) within the Amazon Tall Tower Observatory (ATTO site). The ATTO site is situated in the Uatumã Sustainable Development Reserve (02°10′30′′ S; 59°00′30′′ W), approximately 150 km northeast of Manaus, Amazonas, Brazil (Fig. 1). The climate within the study area is the equatorial pluvial (Af) type, following the Köppen-Geiger classification system (Alvares et al., 2013). The region experiences an average annual rainfall of approximately 2030 mm and an average annual temperature of 28 °C. The dry season extends from June to October and is characterized by reduced rainfall in this period, with monthly precipitation below 100 mm (Andreae et al., 2015).

The present research was developed in three distinct forest ecosystems: upland forest (*terra-firme*), white-sand ecosystem (*campinarana*), and floodplain forest (*igapó*).

The upland plots are located on plateaus near the ATTO tower (Fig. 1), reaching a maximum altitude of approximately 130 m. The upland forest grows on clayey soil, Ferralsols (85.3 % clay), which have good moisture retention capacity (Andreae et al., 2015; Chauvel et al., 1987). It is a dense forest, with a mean height of 21 m and canopy trees reaching up to 45 m. The studied plots comprise 54 botanical families, with Fabaceae, Lecythidaceae, Sapotaceae, Burseraceae, and Euphorbiaceae being the most abundant. Based on previous forest inventory data, a total of 179 tree species were recorded, with *Protium apiculatum* and *Pouteria* sp. as the most abundant species (Andreae et al., 2015).

The white-sand ecosystem of the ATTO site is positioned between the ancient river terraces and the plateau slopes (Targhetta et al., 2015). White-sand ecosystems are characterized by nutrient-poor, highly acidic, and easily leached soils. The soil is predominantly sandy, with a sand fraction of 93.4 ± 1.5 % (Targhetta et al., 2015), resulting in low water retention capacity. According to Demarchi et al. (2022), the white-sand ecosystem of the RDSU are composed of six distinct phytophysiognomies, varying from open to dense white-sand ecosystem, in which the dense phytophysiognomie are differentiated by the reach of the water table to the surface. 167 taxonomically identified species were recorded in the ATTO site, 30 % of which are endemic. In this study, four plots were selected of the open to intermediate white-sand ecosystem (Fig. 1) that suffer the most significant water deficit, especially in the dry season. The mean height of the trees in this environment is between 15 and 20 m, with some emergent trees reaching 25–30 m high in the canopy. In these plots, 40 families were recorded, with Fabaceae, Humiriaceae, Sapotaceae, Chrysobalanaceae, and Rubiaceae being the most abundant. There were 146 species, of which *Sacoglottis guianensis* was the most abundant, followed by *Aldina heterophylla* and *Pradosia schomburgkiana*, both specialists in Amazonian white-sand ecosystem (Demarchi et al., 2022).

The floodplain forest plots are located along the Abacate River, the main tributary of the Uatumã River (Fig. 1). This river is classified as a blackwater (locally known as igapó) that originate in the Guiana Shield and the Central Brazilian Shield (Junk et al., 2011). These waters are poor in suspended sediments and carry few nutrients, so the soils associated with blackwater rivers are of low fertility and high acidity (Junk et al., 2015). The permanent plots used were those of the low and medium igapó of the Abacate River. These floodplain forests remain preserved, as their hydrological cycle has not been impacted by the Balbina Dam (Schöngart et al., 2021). The flooding regime varies from 1 to 3 m for 50–100 days a year in medium igapó. The species sampled from the low igapó such as *E. tenuifolia*, *M. acaciifolium*, *H. brasiliensis*, *M. tamaquarina* (Table 1) experience flood levels reaching up to 7 m in height, remaining inundated for 96–127 days each year (Householder et al., 2021; Lobo et al., 2019). The percentages of sand and clay are 57.0 % and 27.4 %, respectively (Mori, 2019). Therefore, the soil is characterized as clay-sandy. The trees in these environments have a mean height of 13.5 m and can reach a height of approximately 30 m. The flora of the plots is represented by 26 botanical families, with Lecythidaceae, Ochnaceae, Sapotaceae, and Fabaceae having the most abundance. A total of 65 species were recorded, with *Elvasia calophyllea*, *Eschweilera albiflora*, *Pouteria pachyphylla*, and *Eschwilera tenuifolia* as the most abundant species (Schöngart et al., 2021).

2.2. Sampling design

The bark and fresh leaf spectral data were collected directly in the field from mature trees with over 10 cm DBH, with leaves reaching the canopy. We selected a minimum of 10 individuals of 10 abundant species from each ecosystem (upland, white-sand, and floodplain forest) according to an *a priori* inventory to create the local models on the ATTO site (see Table 1). The spectral collections were carried out by ecosystems at different times: in the floodplain forest (*igapó*) it was in October 2022, when the plots were not flooded; in the white-sand ecosystem it was in March 2023 and in the upland forest in August 2023, following the phenology of the

Table 1

Selected tree species in each ecosystem. A minimum of ten individuals were used for each species. n is the number of individuals collected.

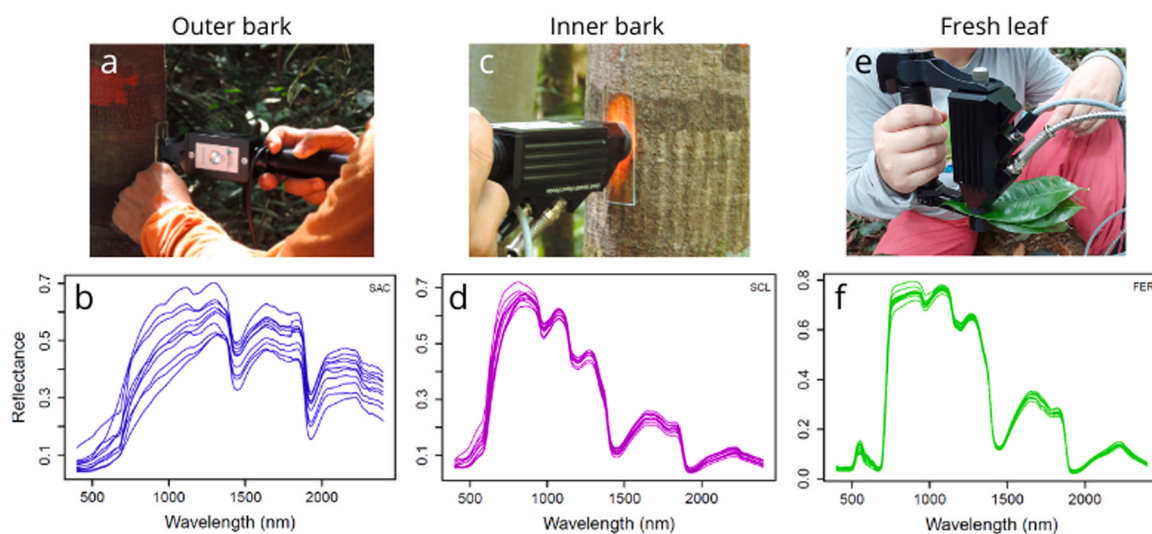| Upland forest			White-sand ecosystem			Floodplain forest		
Family	**Species**	**n**	Family	**Species**	**n**	Family	**Species**	**n**
Burseraceae	*Protium apiculatum*	11	Rubiaceae	*Ferdinandusa chlorantha*	10	Apocynaceae	*Malouetia tamaquarina*	10
Violaceae	*Rinorea guianensis*	10	Simaroubaceae	*Simaba guianensis*	10	Euphorbiaceae	*Hevea spruceana*	10
Euphorbiaceae	*Micrandropsis scleroxylon*	10	Primulaceae	*Cybianthus fulvopulverulentus*	10	Fabaceae	*Macrolobium acaciifolium*	10
Fabaceae	*Swartzia reticulata*	10	Fabaceae	*Aldina heterophylla*	11	Fabaceae	*Swartzia acuminata*	10
Humiriaceae	*Sacoglottis guianensis*	10	Humiriaceae	*Sacoglottis guianensis*	10	Humiriaceae	*Sacoglottis guianensis*	10
Lecythidaceae	*Corythophora alta*	10	Vochysiaceae	*Ruizterania retusa*	10	Lecythidaceae	*Couratari tenuicarpa*	10
Malvaceae	*Scleronema micranthum*	10	Malvaceae	*Scleronema micranthum*	11	Lecythidaceae	*Eschweilera tenuifolia*	10
Sapotaceae	*Ecclinusa guianensis*	10	Malvaceae	*Catostemma sclerophyllum*	11	Lauraceae	*Ocotea aciphylla*	10
Sapotaceae	*Manilkara elata*	10	Sapotaceae	*Manilkara bidentata*	10	Sapotaceae	*Manilkara bidentata*	11
Sapotaceae	*Pouteria anomala*	10	Sapotaceae	*Pradosia schomburgkiana*	12	Sapotaceae	*Pouteria pachyphylla*	11

trees to ensure that the leaves were in the mature leaf phase.

2.3. Spectral data

Spectral reflectance collections of different tree tissues were conducted in the field using a portable high-resolution spectroradiometer ASD FieldSpec® 4. This equipment has its own light source, allowing measurements to be taken independently of ambient lighting and ensuring consistent data quality across all samples. The contact probe accessory was used to take spectral measurements of bark directly from trunks at breast height. The leaf clip accessory was used to collect leaf spectra, with a black background, from freshly cut branches in the field. This portable instrument facilitated rapid data acquisition, capturing information for 0.1 s per spectrum. The spectrometer collected full-range reflectance data covered wavelengths from 350 to 2500 nm, encompassing the visible (VIS = 350–700 nm), near-infrared (NIR = 700–1300 nm), shortwave-infrared regions (SWIR I = 1300–1900; SWIR II = 1900–2500 nm) range of the electromagnetic spectrum. The instrument offered a spectral resolution of 3 nm for wavelengths from 350 to 700 nm, and 10 nm for the wavelengths between 701 and 2500 nm, measuring the sample's reflectance and generating 2151 data points.

The spectral signature of the species has distinct reflectance spectra between the three tissues evaluated: outer bark, inner bark, and fresh leaves (Fig. 2). To ensure the cleanliness of the instrument and prevent contamination with resins, latex, gum, sap, and dust, as well as to protect the lens of the contact probe, a glass slide (microscope slide) adapted to the optical reader was utilized to collect the trunk samples (outer and inner bark). Before each measurement, the equipment was calibrated with a white reference panel (Spectralon®). The white reference panel features material with approximately 100 % reflectance to convert the energy incident on the sample into spectral reflectance. Three spectral readings of the outer bark and three spectral readings of the inner bark were collected for each tree at breast height. The collection strategy involved obtaining spectra at different points around the tree to maximize the capture of individual variation (Hadlich et al., 2018). Spectra were collected from the most preserved part of the outer bark to ensure high-quality data. For inner bark collection, a machete delicately removes the outer bark until reaching the inner bark without passing through the cambium and the wood, i.e., without injuring the tree. The bark of the tree is formed by non-living outer bark and inner bark with living tissues, including all the tissues from the vascular cambium to the periderm, formed by three layers: suber, phellogen, and phelloderm (Esau, 1974; Ferri, 1999), this means the outer and inner bark have different spectral signatures (Hadlich et al., 2018). Visual inspection of the spectra collected in situ guided the accuracy of the spectral signature of each bark tissue according to the method used by Hadlich et al. (2018), where both tissues are separated by PCA. If the reflectance spectrum differed from the spectral signature of the tissue, especially in the inner bark due to not removing all the outer bark for example, the collection was redone by scraping more of the tissue until the fidelity of the data was guaranteed (see an example of the spectral signature of the barks in Fig. 2). Leaf spectral collections involved harvesting branches from the uppermost portions of the trees, which a professional climber facilitated. Approximately 20 mature, healthy leaves were selected from each individual, and readings were taken using four overlapping leaves on the adaxial side, based on two principles: (i) reducing intraspecific variability; (ii) strengthening the spectral signal by minimizing interference from the dark background and irregularities on the leaf surfaces (Jacquemoud and Ustin, 2019). In this way, the spectral signal tends to stabilize and more accurately reflect the structural and biochemical characteristics that are important for discriminating between species. By reducing the variability between leaves of the same individual, we increase the spectral discrimination capacity between species, the focus of the study. This approach produced five leaf spectral readings per individual. Calibration for each leaf measurement was carried out using the white side of the leaf clip. These meticulous and standardized procedures ensured the acquisition of high-quality spectral data, representative of different tree tissues, forming the foundation for the
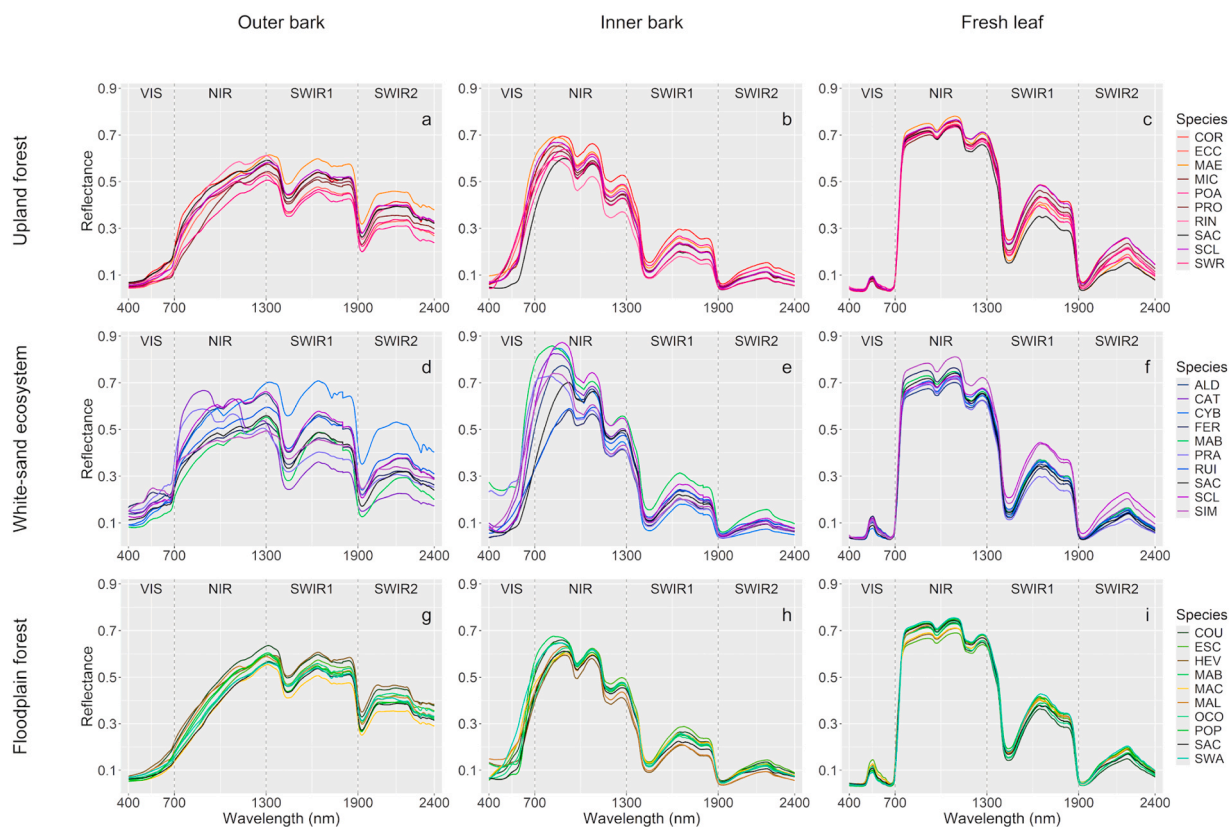

Fig. 2. Reflectance spectral signature for each tissue: panels a, c, and e show examples of field measurements for the outer bark; inner bark; and fresh leaf respectively; panels b, d, and f show the corresponding reflectance spectrum.

subsequent analyses and model development.

2.4. Data analysis

To evaluate the potential of the *in situ* collected spectra of the bark (inner and outer) and leaves for species recognition, spectral models were constructed for different tree tissues within each ecosystem. The initial phase involved preprocessing the spectral data to assess the quality and incorporated careful visualization of spectral curves and principal component analysis (PCA) to identify spectral behavior and outliers. Subsequently, wavelengths within 350–399 nm and 2401–2500 nm were excluded to eliminate equipment-induced noise.

Models based on Linear Discriminant Analysis (LDA) were then developed for the different tree tissues in each ecosystem to assess their predictive capacity for species recognition. In the LDA models, the dependent variables were the species (data known *a priori*), and the independent variables comprised the reflectance obtained at each wavelength from individual readings (Hair et al., 2009). The predictive efficiency of the models was tested using tree cross-validation techniques: leave-one-out (LOO); holdout (70/30), in which the model (training set) was built with 70 % of the individuals of each species, and the test set was made up of the remaining 30 % (Kohavi, 1995) and was repeated 100 times with the model and test sets randomly selected each time. These results were calculated by averaging the 100 permutations, which is consistent with the methodologies employed by Durante et al. (2013) and Hadlich et al. (2018); and finally to test the efficiency of the general model in predicting independent data, external validation was applied. The individuals that make up the independent data set were collected in the same ecosystems (upland forest, white-sand ecosystem and floodplain forest- Igapó) at sites located on the Acaraí River, near the Xingu River, approximately 733 km from the ATTO site; the Trombetas River, approximately 300 km from ATTO site; and the Arapiuns River, near the Tapajós River, approximately 395 km from ATTO site, all located in Pará - Brazil (see fig. S1). A total of 7 species and 26 individuals were used to test the inner and outer bark models and 9 species with 43 individuals to test the fresh leaf model (see table S1). As the spectral data is highly collinear, the stepwise

Fig. 3. Mean spectral reflectance curves of tree species per tissue type across ecosystems. Species abbreviations are as follows: upland forest (COR = Corythophora alta, ECC = Ecclinusa guianensis, MAE = Manilkara elata, MIC = Micrandropsis scleroxylon, POA = Pouteria anomala, PRO = Protium apiculatum, RIN = Rinorea guianensis, SAC = Sacoglottis guianensis, SCL = Scleronema micranthum, and SWR = Swartzia reticulata); white-sand ecosystem (ALD = Aldina heterophylla, CAT = Catostemma sclerophyllum, CYB = Cybianthus fulvopulverulentus, FER = Ferdinandusa chlorantha, MAB = Manilkara bidentata, PRA = Pradosia schomburgkiana, RUI = Ruizterania retusa, SAC = Sacoglottis guianensis, SCL= Scleronema micranthum, and SIM = Simaba guianensis); floodplain forest (COU = Couratari tenuicarpa, ESC = Eschweilera tenuifolia, HEV = Hevea spruceana, MAB = Manilkara bidentata, MAC = Macrolobium acaciifolium, MAL = Malouetia tamaquarina, OCO = Ocotea aciphylla, POP = Pouteria pachyphylla, SAC = Sacoglottis guianensis, and SWA = Swartzia acuminata).

method was used to determine the most important wavelengths for discriminating between species and reduce the variables' collinearity, using a maximum of 1/3 of the number of samples (Williams and Titus, 1988). The models were built using the mean of the spectral readings per individual. Five spectral models were established: individual models for each ecosystem (terra firme forest, white sand and floodplain) based on the data set for each tissue (outer bark, inner bark and fresh leaf); a general model that included all the individuals from the three ecosystems for each tissue and a general model built only with the wavelengths selected by the stepwise method.

All data processing and analysis were performed using the open-source statistical software R (v 4.4.1; R Core Team, 2024), employing custom-made R scripts and add-on packages. The analysis of variance was carried out in Python 3 with Jupyter Notebook as the primary environment. The Python libraries pandas, NumPy and matplotlib.

3. Results

The variation of the spectra signals changes between plant tissues and ecosystems (Fig. 3 and Fig. S2). The outer and inner bark tissues showed a similar pattern of variation in the spectra of the species in all ecosystems. The lowest variation of both tissues was in the floodplain forest (Fig. S2.A) and the highest variation in the white-sand ecosystem, with the greatest difference in the VIS region, showing 111 % (CV) in the inner bark and 54.6 % (CV) in the outer bark. Inner bark showed very similar mean spectral curves of the species and CV in the upland and floodplain forest (Fig. 3 and S1.A). The fresh leaf spectral signal was the tissue with the lowest variation in all ecosystems (S2.B), with a maximum variation of 16.5 % (CV) in the upland forest, 23.9 % (CV) in the white-sand ecosystem and 29.9 % (CV) in the floodplain forest, in the VIS region. Overall, the region with the greatest variations was the VIS region, followed by SWIR II and then SWIR I and NIR.

Species discrimination was successfully achieved using spectral signatures across all tested tree tissues and ecosystems. Taxonomic prediction accuracy ranged from 78.5 % to 100 %, depending on the validation method (LOO and 70/30 holdout) and using mean spectral readings per individual (see Table 2). Among the tissues, outer bark had the lowest prediction accuracy, ranging from 71 % to 89 %, while inner bark models achieved accuracies between 91 % and 97 %. Fresh leaf models exhibited the highest accuracy, varying between 92.4 % and 100 %. In the outer bark models, the lowest predictions in the 70/30 holdout validation occurred in white-sand trees (79.6 %), while the highest was in the floodplain forest trees (86.5 %). Inner bark models had the lowest prediction in floodplain forests (92.2 %) and slightly higher accuracy in upland and white-sand ecosystems (95.5 %). For fresh leaf spectral models, the lowest accuracy was observed in floodplain forests (92.4 %), while white-sand trees showed the highest accuracy (99.9 %).

Fig. 4 presents confusion matrices highlighting misclassified species in each ecosystem based on the LOO validation. Correctly classified individuals are positioned along the diagonal, whereas misclassified trees appear outside the diagonal. In upland forests, the outer bark model misclassified 13 out of 101 individuals evaluated using the LOO validation. In contrast, the inner bark and fresh leaf models were misclassified by only three and two individuals, respectively. Notably, one misclassified individual was familiar across all three models belonging to *Pouteria anomala*. In the white-sand ecosystem, the outer bark misclassified 13 out of 105 individuals evaluated using the LOO validation, while the inner bark model misclassified only five trees. The fresh leaf model, however, attained 100 % accuracy. In the floodplain forest, the outer bark model misclassified 9 out of 102 individual trees, performing the best model among forest types. The inner bark model misclassified seven individuals, while the fresh leaf model misclassified only four. In this ecosystem, the LOO validation (Fig. 4) revealed that *Hevea spruceana* and *Couratari tenuicarpa* were commonly misclassified across the outer bark, inner bark, and fresh leaf models.
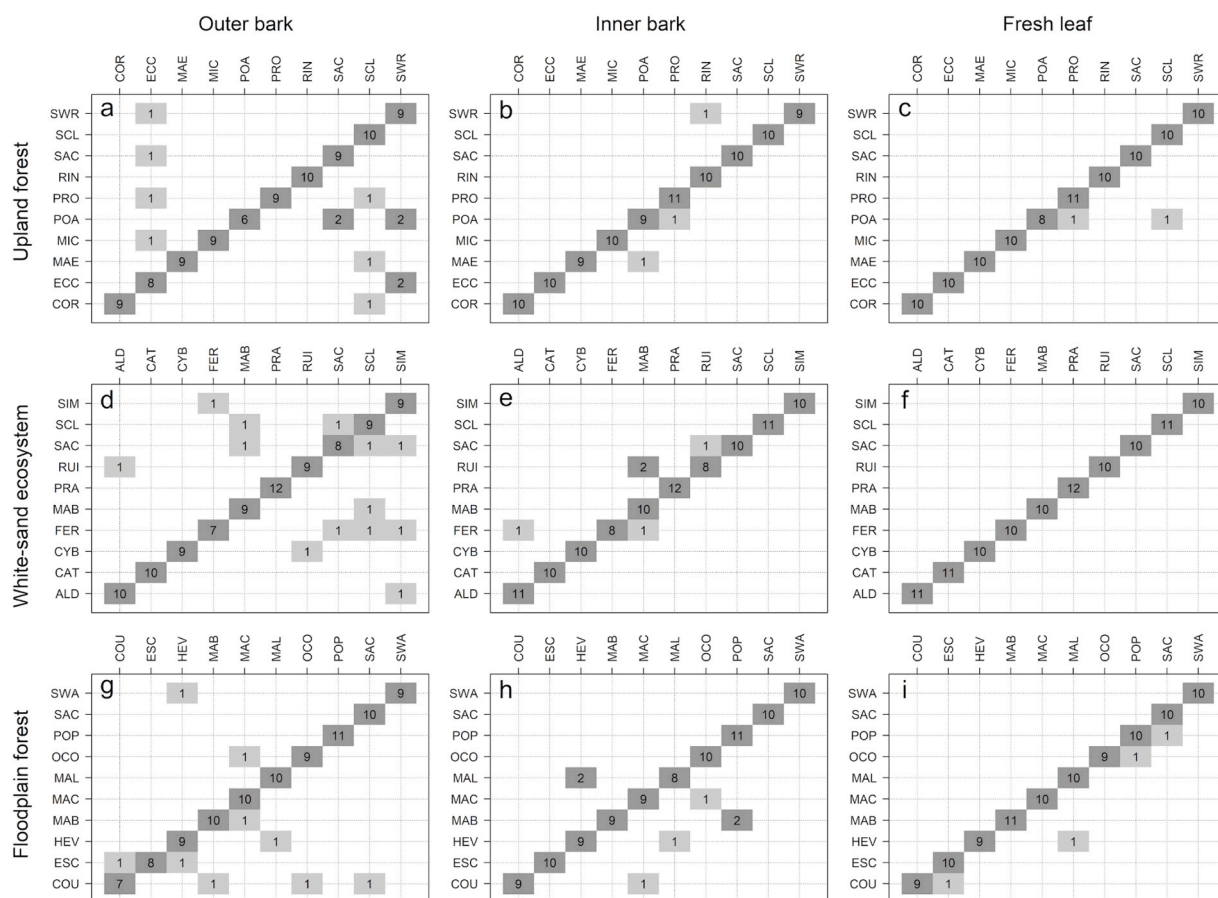
When combining all the individuals from the three ecosystems into a single general model, prediction accuracy remained high, ranging from 86 % to 98 % (Table 2). Fig. 5 illustrates the misclassifications. The outer bark model had the lowest accuracy, with 26 misclassified individuals from 308. In contrast, the inner bark and fresh leaf models performed similarly, achieving 97 % and 98 % accuracy in the LOO validation, respectively, with only nine misclassified individuals in the inner bark and seven in the fresh leaf model. External validation showed consistent prediction power for both fresh leaf (95 %) and inner bark (92 %), with only two

Table 2
Results of the linear discriminant analysis (LDA) for the three tissue types within each ecosystem; general model encompassing all ecosystems; general model built with the wavelengths selected by the stepwise method and external validation with data from other sites. Accuracy was assessed using the leave-one-out (LOO) and holdout (70/30 split) cross-validations. n represents the number of individual trees used.

Model	Validation	n	Outer bark	Inner bark	Fresh leaf
Upland forest	LOO	101	86 %	97 %	98 %
	70/30		78.5 %	95.5 %	97.2 %
White-sand ecosystem	LOO	105	82 %	95 %	100 %
	70/30		79.6 %	95.2 %	99.9 %
Floodplain forest	LOO	102	89 %	93 %	96 %
	70/30		88.3 %	92.2 %	92.4 %
General model	LOO	308	86 %	97 %	98 %
	70/30		86.5 %	96.2 %	97.7 %
General model with stepwise	LOO	308	71 %	91 %	98 %
General model x other sites	external validation	*	54 % *(26)	92 % *(26)	95 % *(43)

* the number of individuals for outer and inner bark was lower than for fresh leaf due to the flooded conditions in the floodplain forest ecosystem on the Xingu site at the time of collection, making it impossible to collect bark spectra.

Fig. 4. Confusion matrices resulting from LOO validation for each model tested for across different ecosystems. The diagonal represents correct predictions of species identification, while off the diagonal line are the incorrect predictions using the spectroscopic method. Species listed horizontally correspond to those identified by a taxonomist, while those listed vertically correspond to species "identified" using spectroscopy. Species abbreviations are as follows: Upland forest (COR = *Corythophora alta*, ECC = *Ecclinusa guianensis*, MAE = *Manilkara elata*, MIC = *Micrandropsis scleroxylon*, POA = *Pouteria anomala*, PRO = *Protium apiculatum*, RIN = *Rinorea guianensis*, SAC = *Sacoglottis guianensis*, SCL = *Scleronema micranthum*, and SWR = *Swartzia reticulata*); white-sand ecosystem (ALD = *Aldina heterophylla*, CAT = *Catostemma sclerophyllum*, CYB = *Cybianthus fulvopulverulentus*, FER = *Ferdinandusa chlorantha*, MAB = *Manilkara bidentata*, PRA = *Pradosia schomburgkiana*, RUI = *Ruizterania retusa*, SAC = *Sacoglottis guianensis*, SCL = *Scleronema micranthum*, and SIM = *Simaba guianensis*); floodplain forest (COU = *Couratari tenuicarpa*, ESC = *Eschweilera tenuifolia*, HEV = *Hevea spruceana*, MAB = *Manilkara bidentata,* MAC = *Macrolobium acaciifolium*, MAL = *Malouetia tamaquarina*, OCO = *Ocotea aciphylla*, POP = *Pouteria pachyphylla*, SAC = *Sacoglottis guianensis*, and SWA = *Swartzia acuminata*).
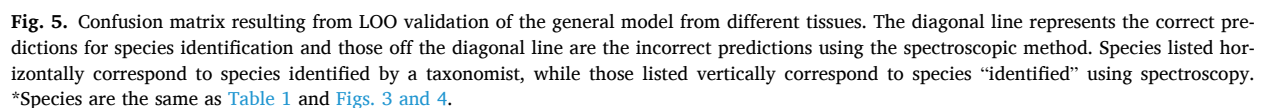
individuals incorrectly predicted in both tests (see confusion matrix in figure S3). As for the outer bark, the set with data from other sites did not show high efficiency, reaching only 54 % correct predictions.
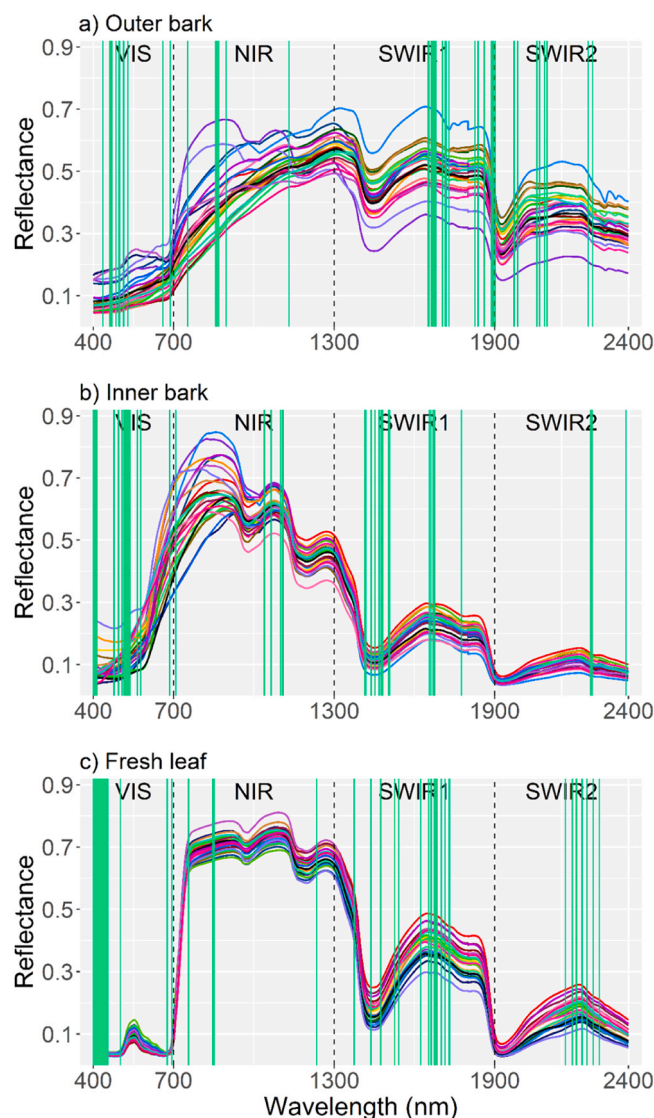
Using the stepwise technique, the most informative characteristics for species classification were identified for each tissue type using the general models (Fig. 6). The outer bark model (Fig. 6.a) selected the 52 most informative variables, with the SWIR I region (1300 nm–1900 nm) being the most representative. The inner bark model (Fig. 6.b) relied on 66 selected variables, primarily from the VIS and SWIR I regions. For fresh leaves (Fig. 6.c), 78 variables were selected, with 41 in the VIS region. Table S2 presents all selected variables for each of the three tissue models. When the general model was built using the wavelengths selected, the prediction power decreased for outer bark (71 %) and inner bark (91 %), while it remained consistent for fresh leaf (98 %).

4. Discussion

This study assessed the potential of spectral signatures from different tree tissues—outer bark, inner bark, and fresh leaves—for species identification across three distinct Amazonian ecosystems: upland, white-sand, and floodplain forest. Our results demonstrate that species discrimination was successfully achieved across all tested tissues and ecosystems. Fresh leaves exhibited the highest classification accuracy among the tissues analyzed, followed by inner and outer bark. Furthermore, we evaluated the feasibility of developing a generalized spectral model applicable across these three forest types. Despite the variations in spectral responses, the model achieved high accuracy in species recognition for the tested tissues. These findings highlight the promising potential of full-

Fig. 5. Confusion matrix resulting from LOO validation of the general model from different tissues. The diagonal line represents the correct predictions for species identification and those off the diagonal line are the incorrect predictions using the spectroscopic method. Species listed horizontally correspond to species identified by a taxonomist, while those listed vertically correspond to species "identified" using spectroscopy. *Species are the same as Table 1 and Figs. 3 and 4.

Fig. 6. Mean spectral reflectance curves of all species sampled at the ATTO site, with the wavelengths selected (vertical green lines) for the general model by tissue type using the stepwise selection technique. *Species are the same as Table 1 and Figs. 3, 4 and 5.

range spectroscopy as a powerful tool for enhancing biodiversity inventories and advancing field-based species identification in the Amazon.

The spectral models based on fresh leaves showed the highest effectiveness for species recognition in the field for all ecosystems. The spectral reflectance of the leaves is determined by their chemical adaptations to environmental conditions, including climate, nutrient availability, phenology, age of the individual, and lighting (Asner, 1998; Asner et al., 2014; Ponzoni, 2002), which influence their biochemical composition. The most accurate fresh leaf model was in the white-sand ecosystem, where the leaves exhibit scleromorphic traits, being thicker and enriched with secondary compounds to avoid herbivory, potentially enhancing their spectral signal. For the floodplain forest, the fresh leaf model showed slightly lower accuracies (92 % and 96 %) than those observed in the white-sand and upland. These results may have been influenced by leaf phenology, since in the field there is no such control, although only adult individuals and the preference for mature, healthy leaves were used. Our findings corroborate those of Prospere et al. (2014) and Liu et al. (2012) who successfully discriminated plant species *in situ* in wetlands with leaf spectra using different classification methods. However, collecting leaves requires climbing trees, an expensive and time-consuming field practice, especially considering the great heights of the trees in the Amazon. In addition to species discrimination in hyperdiverse environments, leaf spectra can also be robustly applied in differentiating between invasive and native species, which is important for the management of conservation areas (Mallmann et al., 2023).

The inner bark was the second-best tissue for species identification, with the spectral model accuracy decreasing by less than 5 % compared to fresh leaves across all forest types. This slight reduction suggests that inner bark is a viable alternative for species

classification in forest inventories, balancing efficiency and practicality. Outer bark, while less effective, still achieved an accuracy of approximately 89 %, only 18 % lower than fresh leaves. Although this reduction is notable, it remains a reasonable outcome given the challenges associated with outer bark spectroscopy. Our findings align with Hadlich et al. (2018), who demonstrated similar results for upland forest trees in the Central Amazon. Our study expands these insights across different forest types, reinforcing the potential of inner and outer bark spectroscopy for species identification in diverse Amazonian ecosystems.

The lower accuracy of outer bark models can be attributed to their exposure to environmental factors, including dust and microorganisms (e.g. lichens), and substantial variation in bark thickness between species (Paine et al., 2010; Rosell, 2016). These effects were captured with the spectral variation within the tissue between species. They were pronounced in the white-sand ecosystem, probably due to the scleromorphic characteristics of these forests, where trees are functionally adapted to water scarcity, leading to distinct chemical and structural traits that affect spectral discrimination, such as bark thickness and texture. The high variability in bark morphology and texture contributed to more significant spectral variability. This variation in the outer bark, along with differences between species, also changes throughout the tree's ontogeny. However, further studies assessing the spectral responding of bark at different ontogenetic stages may provide valuable insights. The rough, cracked, fissured and fenestrated barks for example have the limitation of light scattering at the time of collection with the ASD Fieldspec 4 equipment in the field, leading to less consistent spectral signatures to the outer bark than the more uniform patterns observed in fresh leaves and inner bark. Because it is a tissue that is protected by the outer bark, the inner bark, however, presents the most conserved biochemical characteristics and has less intraspecific variation (Hadlich et al., 2018). In the field, with the removal of the outer bark, the surface becomes regular, facilitating the accuracy of the spectral curve exposed from inner bark. However, Juola et al. (2023) demonstrated that the use of *in situ* imaging spectroscopy of the outer bark of standing trees, together with information on the texture of the trunk bark, can significantly improve the accuracy of spectral models of ten boreal species, with an accuracy of between 93 % and 97.5 %. Thus the combination of spectroscopy with morphological data is a promising alternative for the development of accurate applications for identifying tree species and supporting inventory (Juola et al., 2022, 2023). On the other hand, for the floodplain forest, the model with the outer bark performed better between upland and white-sand, showing the lowest intraspecific variation and with a reflectance spectrum similar to that of upland. This is remarkable, considering that trunks experience long periods of flooding, which could accelerate the decomposition of their tissues, affecting predictions.

The general model's results were also promising, as accuracy remained high for fresh leaves, inner bark, and outer bark models. This suggests species identification in forest inventories does not require separate models for each forest type. By incorporating spectra from individuals across upland, white-sand, and floodplain forests, the general model captured a broader range of phenotypic variation, integrating more comprehensive physical and chemical information. And by testing whether this local model is efficient in predicting species from other locations, more than 300 km away from where the model was generated, we have shown that the potential is enormous, especially for fresh leaves and inner bark. Although few species and individuals were used for this test, this step is very important for making this technique applicable in different regions of the Amazon. Given the challenges of species delimitation in the Amazon, species recognition tests using field spectroscopy must account for the fact that locally trained models need to be tested on well-delimited species occurring in geographically distant regions. If the tested specimens represent undescribed species currently grouped under a known species, the model is likely to misclassify them, as the spectral variation may reflect biological diversity that has not yet been formally recognized. Two key factors must be carefully evaluated when testing the regionalization of the model: i. the spectral response of species in other ecosystems not tested in this study; ii. species with identification challenges may involve species complexes or distinct allopatric species that have yet to be formally identified or described. Using such models represents a significant step forward in improving the efficiency of species identification for biodiversity research and monitoring, ultimately contributing to more effective conservation strategies in the Amazon.

Identifying the most informative wavelength for species discrimination across different tissues is essential for optimizing spectral models, especially considering variations in the range of the electromagnetic spectrum captured by the equipment, spectral resolution, and the influence of chemical signals on classification accuracy. Although the selection of the most important variables did not improve the accuracy of the models, maintaining accuracy only for fresh leaf (98 %), we identified that the VIS (400–700 nm) and SWIR I (1300–1900 nm) regions were the most important for distinguishing species across all three tested tissues, with some variations among them. The fresh leaves models used 48 wavelengths in the VIS region, 18 in the SWIR I, and only 5 in the NIR region (700–1300 nm). Fresh leaf spectral responses are a function of leaf structure, water content, and the concentration of biochemicals (Curran et al., 1992). The discriminatory power in the VIS region was evidenced in the spectral range corresponding to the blue-green transition (400–499 nm) (Hennessy et al., 2020). In this region, the optical properties of the leaves remain stable due to the presence of biologically active pigments, such as chlorophyll and carotenoids (Gausman, 1983; Jones & Vaughan, 2010) that play a crucial role in this discrimination process. The NIR region, which reflects cell structure (Ponzoni *et al.*, 2015), had only 5 wavelengths selected by the spectral model. Additionally, specific wavelengths in the shortwave infrared (SWIR II) region (2.100–2300 nm) provide valuable information on cellulose and lignin content (Curran, 1989).

The spectral model of the outer bark selected 34 informative wavelengths in the SWIR I and II regions, of which 16 specific wavelengths belonged to lignin, hemicellulose, cellulose, aromatic compounds, and extractives (Jones et al., 2006; Schwanninger et al., 2011; Workman and Weyer, 2007) being the most representative region. For the inner bark, the spectral models selected 32 in the VIS region, and also detected 11 wavelengths belonging to lignin, cellulose, and water (Workman and Weyer, 2007), with VIS and SWIR I being the most important regions. When comparing the most informative wavelengths for species discrimination between outer and inner bark using the general model (upland, white-sand, and floodplain forests) to those presented by Hadlich et al. (2018) for upland forest, distinct differences emerged. In the present study, the most informative wavelengths for the outer bark were in SWIR I, whereas Hadlich et al. (2018) recognized VIS and NIR, as the most relevant. For inner bark, the selection of the most informative

regions in this study were VIS and SWIR I, while Hadlich et al. (2018) identified VIS and NIR as the key regions. The selection of the most informative wavelength between models can be influenced by differences in the dataset, such as species composition and environmental variation. Juola et al. (2022) found selected features within the VIS and NIR range for the outer bark. Therefore, the selection of variables that best discriminate between species must be refined because, in addition to the data set, this information can be influenced by different equipment with different resolutions and spectral range.

Spectroscopy is a valuable technique for improving tree biodiversity inventories in the field. However, some equipment-related limitations can be refined. The ASD FieldSpec 4 is an expensive device, making it difficult to popularize this technique among forest companies and scientific institutions. Additionally, the equipment weighs around 8 kg and requires a computer connection and battery power for operation, which makes it challenging to use in dense forest environments where accessing trees can be difficult. However, to make this technique more widely accessible to the field, we expect smaller, more affordable, and high-precision spectrometers, such as a small device from the ASD FieldSpec 4 or other spectral companies. However, it is necessary to verify the prediction efficiency when such equipment with different spectral resolution and spectral range is employed. The development and adoption of such equipment would help to expand the use of spectroscopy, as it is necessary to assess the species discrimination power of entire plots, including rare species and species with identification problems (species clusters, undelimited species or hypotheses of new species). This would facilitate the implementation of quality control in species identification and improve biodiversity inventories in tropical forest monitoring.

CRediT authorship contribution statement

Caroline C. Vasconcelos: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation. **Florian Wittmann:** Writing – review & editing, Visualization, Validation, Resources, Funding acquisition, Data curation, Conceptualization. **Conde Maíra L. G.:** Writing – review & editing, Methodology, Data curation. **Caroline L. Mallmann:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis. **Layon O. Demarchi:** Writing – review & editing, Visualization, Resources, Methodology, Data curation. **Piedade Maria T. F.:** Writing – review & editing, Project administration, Funding acquisition, Data curation. **Gisele Biem Mori:** Writing – review & editing, Methodology, Investigation. **Flavia Machado Durgante:** Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jochen Schöngart:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Hilana Louise Hadlich:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Priscila Amaral de Sá:** Writing – review & editing, Methodology, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.gecco.2025.e03970.

Data Availability

Data will be made available on request.

References

Adeney, J.M., Christensen, N.L., Vicentini, A., Cohn-Haft, M., 2016. White-sand ecosystems in Amazonia. Biotropica. https://doi.org/10.1111/btp.12293.

Alvares, C.A., Stape, J.L., Sentelhas, P.C., De Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Z. 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507.

Andreae, M.O., Acevedo, O.C., Araùjo, A., Artaxo, P., Barbosa, C.G.G., Barbosa, H.M.J., Brito, J., Carbone, S., Chi, X., Cintra, B.B.L., Da Silva, N.F., Dias, N.L., Dias-Júnior, C.Q., Ditas, F., Ditz, R., Godoi, A.F.L., Godoi, R.H.M., Heimann, M., Hoffmann, T., Kesselmeier, J., Könemann, T., Krüger, M.L., Lavric, J.V., Manzi, A.O., Lopes, A.P., Martins, D.L., Mikhailov, E.F., Moran-Zuloaga, D., Nelson, B.W., Nölscher, A.C., Santos Nogueira, D., Piedade, M.T.F., Pöhlker, C., Pöschl, U., Quesada, C.A., Rizzo, L.V., Ro, C.U., Ruckteschler, N., Sá, L.D.A., De Oliveira Sá, M., Sales, C.B., Dos Santos, R.M.N., Saturno, J., Schöngart, J., Sörgel, M., De Souza, C.M., De Souza, R.A.F., Su, H., Targhetta, N., Tóta, J., Trebs, I., Trumbore, S., Van Eijck, A., Walter, D., Wang, Z., Weber, B., Williams, J., Winderlich, J., Wittmann, F., Wolff, S., Yáñez-Serrano, A.M., 2015. The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols. Atmos. Chem. Phys. 15, 10723–10776. https://doi.org/10.5194/acp-15-10723-2015.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sens. Environ. 64, 234–253. https://doi.org/10.1016/S0034-4257(98)00014-5.

Asner, G.P., Martin, R.E., Carranza-Jim Enez, L., Sinca, F., Tupayachi, R., Anderson, C.B., Martinez, P., 2014. Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. N. Phytol. 204, 127–139. https://doi.org/10.1111/nph.12895.

Baraloto, C., Ferreira, E., Rockwell, C., Walthier, F., 2007. Limitations and applications of parataxonomy for community forest management in Southwestern Amazonia. Ethnobot. Res. Appl. 5, 77–84. https://doi.org/10.17348/era.5.0.77-84.

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

Botelho, I., 2017. Uso da espectroscopia no infravermelho próximo (FT-NIR) como ferramenta na discriminação de espécies herborizadas de Burseraceae oriundas de diferentes locais da Amazônia Legal, 75.

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

Cardoso, D., Särkinen, T., Alexander, S., Amorim, A.M., Bittrich, V., Celis, M., Daly, D.C., Fiaschi, P., Funk, V.A., Giacomin, L.L., Goldenberg, R., Heiden, G., Iganci, J., Kelloff, C.L., Knapp, S., De Lima, H.C., Machado, A.F.P., Dos Santos, R.M., Mello-Silva, R., Michelangeli, F.A., Mitchell, J., Moonlight, P., De Moraes, P.L.R., Mori, S.A., Nunes, T.S., Pennington, T.D., Pirani, J.R., Prance, G.T., De Queiroz, L.P., Rapini, A., Riina, R., Rincon, C.A.V., Roque, N., Shimizu, G., Sobral, M., Stehmann, J.R., Stevens, W.D., Taylor, C.M., Trovó, M., Van Den Berg, C., Van Der Werff, H., Viana, P.L., Zartman, C.E., Forzza, R.C., 2017. Amazon plant diversity revealed by a taxonomically verified species list. Proc. Natl. Acad. Sci. USA 114, 10695–10700. https://doi.org/10.1073/PNAS.1706756114/-/DCSUPPLEMENTAL/PNAS.1706756114.SD02.XLSX.

Cavender-Bares, J., White, D.M., Ahlstrand, N.I., Austin, M.W., Bastianelli, D., Bazan, S., Boughalmi, K., Cardinal-McTeague, W., Chacón-Madrigal, E., Couvreur, T.L. P, Davis, C., Durgante, F.M., Grace, O.M., Guzmán Q, J.A., Hansen, K., Hernández-Leal, M.S., Hopkins, M.J.G., Jackson, R., Kothari, S., Lee, A.K., Léveillé-Bourret, É., Pinto-Ledezma, J., Quinteros Casaverde, N.L., Meireles, J.E., Neto-Bradley, B., Nichodemus, C.O., Ree, R.H., Schmull, M., Soltis, D.E., Soltis, P.S., Tuomisto, H., Ustin, S.L., Vasconcelos, C.C., 2025. Next-generation specimen digitization: capturing reflectance spectra from the world's herbaria for modeling plant biology across time, space, and taxa. New Phytologist. https://doi.org/10.1111/nph.70645.

Chauvel, A., Lucas, Y., Boulet, R., 1987. On the genesis of the soil mantle of the region of Manaus, Central Amazonia, Brazil. Experientia 43, 234–241. https://doi.org/10.1007/BF01945546.

Clark, M.L., Roberts, D.A., 2012. Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. Remote Sens. 4, 1820–1855. https://doi.org/10.3390/rs4061820.

Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sens. Environ. https://doi.org/10.1016/0034-4257(89)90069-2.

Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., Peterson, D.L., 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. Remote Sens. Environ. 39, 153–166. https://doi.org/10.1016/0034-4257(92)90133-5.

Cysneiros, V.C., Mendonça Júnior, J.O., Lanza, T.R., Moraes, J.C.R., Samor, O.J.M., 2018. Espécies madeireiras da Amazônia: riqueza, nomes populares e suas peculiaridades. Pesqui. Florest. Bras. 38, 1. https://doi.org/10.4336/2018.pfb.38e201801567.

da Silva, C.E., do Nascimento, C.S., de Freitas, J.A., de Araújo, R.D., do Nascimento, C., Higuchi, N., 2024. Application of infrared spectroscopy in the identification of fallen trees from the Amazon Rainforest (Myristicaceae. For. Sci. 70, 102–112. https://doi.org/10.1093/forsci/fxae001.

Damasco, G., Daly, D.C., Vicentini, A., Fine, P.V.A., 2019. Reestablishment of Protium cordatum (Burseraceae) based on integrative taxonomy. Taxon 68, 34–46. https://doi.org/10.1002/tax.12022.

Demarchi, L.O., Klein, V.P., Aguiar, D.P.P., Marinho, L.C., Ferreira, M.J., Lopes, A., da Cruz, J., Quaresma, A.C., Schöngart, J., Wittmann, F., Piedade, M.T.F., 2022. The specialized white-sand flora of the Uatumã Sustainable Development Reserve, central Amazon, Brazil. Check List 18, 187–217. https://doi.org/10.15560/18.1.187.

Draper, F.C., Baker, T.R., Baraloto, C., Chave, J., Costa, F., Martin, R.E., Pennington, R.T., Vicentini, A., Asner, G.P., 2020. Quantifying tropical plant diversity requires an integrated technological approach. Trends Ecol. Evol. 35, 1100–1109. https://doi.org/10.1016/J.TREE.2020.08.003.

Durgante, F.M., Higuchi, N., Almeida, A., Vicentini, A., 2013. Species spectral signature: discriminating closely related plant species in the Amazon with near-infrared leaf-spectroscopy. For. Ecol. Manag. 291, 240–248. https://doi.org/10.1016/j.foreco.2012.10.045.

Esau, K., 1974. Anat. Das. Plantas Com. Sement.

Ferri, M., 1999. Botânica: morfologia interna das plantas: anatomia - Mário Guimarães Ferri - Google Livros [WWW Document]. URL ⟨https://books.google.com.br/books?hl=pt-BR&lr=id=CBSM694l6HsC&oi=fnd&pg=PA3&dq=Ferri,+M.G.,+1999.+Bot⟩⟨ànica:+morfologia+interna+das+plantas:+anatomia.+NBL+Editora&ots=G3BUj0R_vu&sig=L_jW0tcCglUA1c0m_nBi8mRPPZc&redir_esc=y#v=onepage&q&f=false⟩ (accessed 5.9.25).

Foley, W.J., McIlwee, A., Lawler, I., Aragones, L., Woolnough, A.P., Berding, N., 1998. Ecological applications of near infrared reflectance spectroscopy - a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. Oecologia. https://doi.org/10.1007/s004420050591.

Gaui, T.D., Costa, F.R.C., Coelho de Souza, F., Amaral, M.R.M., de Carvalho, D.C., Reis, F.Q., Higuchi, N., 2019. Long-term effect of selective logging on floristic composition: a 25 year experiment in the Brazilian Amazon. For. Ecol. Manag. 440, 258–266. https://doi.org/10.1016/j.foreco.2019.02.033.

Gausman, H.W., 1983. Visible light reflectance, transmittance, and absorptance of differently pigmented cotton leaves. Remote Sens. Environ. 13, 233–238. https://doi.org/10.1016/0034-4257(83)90041-X.

Gomes, A.C.S., Andrade, A., Barreto-Silva, J.S., Brenes-Arguedas, T., López, D.C., de Freitas, C.C., Lang, C., de Oliveira, A.A., Pérez, A.J., Perez, R., da Silva, J.B., Silveira, A.M.F., Vaz, M.C., Vendrami, J., Vicentini, A., 2013. Local plant species delimitation in a highly diverse Amazonian forest: do we all see the same species? J. Veg. Sci. 24, 70–79. https://doi.org/10.1111/j.1654-1103.2012.01441.x.

Goodwin, Z.A., Harris, D.J., Filer, D., Wood, J.R.I., Scotland, R.W., 2015. Widespread mistaken identity in tropical plant collections. Curr. Biol. https://doi.org/10.1016/j.cub.2015.10.002.

Hadlich, H.L., Durgante, F.M., dos Santos, J., Higuchi, N., Chambers, J.Q., Vicentini, A., 2018. Recognizing Amazonian tree species in the field using bark tissues spectra. For. Ecol. Manag. 427, 296–304. https://doi.org/10.1016/j.foreco.2018.06.002.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L., 2009. Análise multivariada de dados. Bookman editora.

Hennessy, A., Clarke, K., Lewis, M., 2020. Hyperspectral classification of plants: a review of waveband selection generalisability. Remote Sens. https://doi.org/10.3390/RS12010113.

Hopkins, M.J.G., 2007. Modelling the known and unknown plant biodiversity of the Amazon Basin. J. Biogeogr. 1400–1411. https://doi.org/10.1111/j.1365-2699.2007.01737.x.

Hopkins, M.J.G., 2019. Are we close to knowing the plant diversity of the Amazon? An. Acad. Bras. Cienc. 91. https://doi.org/10.1590/0001-3765201920190396.

Householder, J.E., Schöngart, J., Piedade, M.T.F., Junk, W.J., ter Steege, H., Montero, J.C., de Assis, R.L., de Aguiar, D.P.P., Pombo, M.M., Quaresma, A.C., Demarchi, L.O., Parolin, P., Lopes, A., Feitoza, G.V., Durgante, F.M., Albuquerque, B.W., Chu, A., Enßlin, D., Fabian, T., Fettweiß, K., Hirsch, M., Hombach, M., Hubbuch, A., Hutter, B., Jäger, T., Kober-Moritz, R., Lindner, M.K.R., Maier, F., Nowak, J., Petridis, Z., Schierling, L., Snjaric, E., Egger, G., Schneider, E., Damm, C., Wittmann, F., 2021. Modeling the ecological responses of tree species to the flood pulse of the Amazon Negro River Floodplains. Front. Ecol. Evol. 9. https://doi.org/10.3389/fevo.2021.628606.

Jacquemoud, S., Ustin, S., 2019. Leaf optical properties. Leaf Opt. Prop. https://doi.org/10.1017/9781108686457.

Jones, P.D., Schimleck, L.R., Peter, G.F., Daniels, R.F., Clark, A., 2006. Nondestructive estimation of wood chemical composition of sections of radial wood strips by diffuse reflectance near infrared spectroscopy. Wood Sci. Technol. 40, 709–720. https://doi.org/10.1007/s00226-006-0085-6.

Jones, H.G., Vaughan, R.A., 2010. Remote sensing of vegetation: principles, techniques, and applications. Oxford University Press, USA.

Junk, W.J., Piedade, M.T.F., Lourival, R., Wittmann, F., Kandus, P., Lacerda, L.D., Bozelli, R.L., Esteves, F.A., Nunes da Cunha, C., Maltchik, L., Schöngart, J., Schaeffer-Novelli, Y., Agostinho, A.A., 2014. Brazilian wetlands: their definition, delineation, and classification for research, sustainable management, and protection. Aquat. Conserv. Mar. Freshw. Ecosyst. 24, 5–22. https://doi.org/10.1002/aqc.2386.

Junk, W.J., Piedade, M.T.F., Schöngart, J., Cohn-Haft, M., Adeney, J.M., Wittmann, F., 2011. A classification of major naturally-occurring amazonian lowland wetlands. Wetlands 31, 623–640. https://doi.org/10.1007/s13157-011-0190-7.

Junk, W.J., Wittmann, F., Schöngart, J., Piedade, M.T.F., 2015. A classification of the major habitats of Amazonian black-water river floodplains and a comparison with their white-water counterparts. Wetl. Ecol. Manag 23, 677–693. https://doi.org/10.1007/s11273-015-9412-8.

Juola, J., Hovi, A., Rautiainen, M., 2022. A spectral analysis of stem bark for boreal and temperate tree species. Ecol. Evol. 12, 1–14. https://doi.org/10.1002/ece3.8718.

Juola, J., Hovi, A., Rautiainen, M., 2023. Classification of tree species based on hyperspectral reflectance images of stem bark. Eur. J. Remote Sens. 56. https://doi.org/10.1080/22797254.2022.2161420.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI Int. Jt. Conf. Artif. Intell. 1137–1143.

Lang, C., Costa, F.R.C., Camargo, J.L.C., Durgante, F.M., Vicentini, A., 2015. Near infrared spectroscopy facilitates rapid identification of both young and mature Amazonian tree species. PLoS One 10. https://doi.org/10.1371/journal.pone.0134521.

Liu, X.H., Sun, Y., Wu, Y., 2012. Reduction of hyperspectral dimensions and construction of discriminating models for identifying wetland plant species. Guang Pu Xue Yu Guang Pu Fen. Xi/Spectrosc. Spectr. Anal. 32, 459–464. https://doi.org/10.3964/j.issn.1000-0593(2012)02-0459-06.

Lobo, G. de S., Wittmann, F., Piedade, M.T.F., 2019. Response of black-water floodplain (igapó) forests to flood pulse regulation in a dammed Amazonian river. For. Ecol. Manag. 434, 110–118. https://doi.org/10.1016/j.foreco.2018.12.001.

Mallmann, C.L., Pereira Filho, W., Dreyer, J.B.B., Tabaldi, L.A., Durgante, F.M., 2023. Leaf-level field spectroscopy to discriminate invasive species (Psidium guajava L. and Hovenia dulcis Thunb.) from native tree species in the Southern Brazilian Atlantic Forest. Remote Sens 15. https://doi.org/10.3390/rs15030791.

Meireles, J.E., Cavender-Bares, J., Townsend, P.A., Ustin, S., Gamon, J.A., Schweiger, A.K., Schaepman, M.E., Asner, G.P., Martin, R.E., Singh, A., Schrodt, F., Chlus, A., O'Meara, B.C., 2020. Leaf reflectance spectra capture the evolutionary history of seed plants. N. Phytol. 228, 485–493. https://doi.org/10.1111/nph.16771.

Mori, G.B., 2019. Estrutura funcional das comunidades arbóreas de florestas alagáveis na Amazônia Central. Thesis.

Nelson, B.W., Ferreira, C.A.C., Da Silva, M.F., Kawasaki, M.L., 1990. Endemism centres, refugia and botanical collection density in Brazilian Amazonia. Nature 345, 714–716. https://doi.org/10.1038/345714a0.

Oliveira-Filho, A.T., Dexter, K.G., Pennington, R.T., Simon, M.F., Bueno, M.L., Neves, D.M., 2021. On the floristic identity of Amazonian vegetation types. Biotropica 53, 767–777. https://doi.org/10.1111/btp.12932.

Paine, C.E.T., Stahl, C., Courtois, E.A., Patiño, S., Sarmiento, C., Baraloto, C., 2010. Functional explanations for variation in bark thickness in tropical rain forest trees. Wiley Online Libr. Paine C. Stahl. EA Court. S Patiño C. Sarmiento C. BaralotoFunctional Ecol. 2010•Wiley Online Libr. 24, 1202–1210. https://doi.org/10.1111/J.1365-2435.2010.01736.X.

Pasquini, C., 2003. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. J. Braz. Chem. Soc. https://doi.org/10.1590/S0103-50532003000200006.

Pasquini, C., 2018. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. Anal. Chim. Acta 1026, 8–36. https://doi.org/10.1016/j.aca.2018.04.004.

Pastore, T.C.M., Braga, J.W.B., Coradin, V.T.R., Magalhães, W.L.E., Okino, E.Y.A., Camargos, J.A.A., De Muñiz, G.I.B., Bressan, O.A., Davrieux, F., 2011. Near infrared spectroscopy (NIRS) as a potential tool for monitoring trade of similar woods. Discrim. true Mahog. cedar andiroba curupixá. degruyter. Com. 65, 73–80. https://doi.org/10.1515/HF.2011.010/HTML.

Pigozzo, R., 2011. Espectroscopia de infravermelho-próximo em madeiras neotropicais: aplicação na identificação e predição de propriedades físicas.

Ponzoni, F., 2002. Sensoriamento remoto no estudo da vegetação: diagnosticando a mata atlântica [WWW Document].

Prata, E.M.B., Sass, C., Rodrigues, D.P., Domingos, F.M.C.B., Specht, C.D., Damasco, G., Ribas, C.C., Fine, P.V.A., Vicentini, A., 2018. Towards integrative taxonomy in Neotropical botany: Disentangling the Pagamea guianensis species complex (Rubiaceae). Bot. J. Linn. Soc. 188, 213–231. https://doi.org/10.1093/botlinnean/boy051.

Procópio, L.C., Secco, R.D.S., 2008. A importância da identificação botânica nos inventários florestais: O exemplo do "tauari" (Couratari spp. e Cariniana spp. - Lecythidaceae) em duas áreas manejadas no estado do Pará. Acta Amaz 38, 31–42. https://doi.org/10.1590/S0044-59672008000100005.

Prospere, K., McLaren, K., Wilson, B., 2014. undefined, 2014. Plant Species Discrimination in a Tropical Wetland Using In Situ Hyperspectral Data, 2014•mdpi. com mdpi. comK Prosper. K. McLaren B WilsonRemote Sens. 6, 8494–8523. https://doi.org/10.3390/rs6098494.

R Core Team. (2024). R: A Language and Environment for Statistical Computing (4.4.1). R Foundation for Statistical Computing. https://www.r-project.org/.

Rosell, J.A., 2016. Bark thickness across the angiosperms: more than just fire. N. Phytol. 211, 90–102. https://doi.org/10.1111/nph.13889.

Schöngart, J., Wittmann, F., Faria de Resende, A., Assahira, C., de Sousa Lobo, G., Rocha Duarte Neves, J., da Rocha, M., Biem Mori, G., Costa Quaresma, A., Oreste Demarchi, L., Weiss Albuquerque, B., Oliveira Feitosa, Y., da Silva Costa, G., Vieira Feitoza, G., Machado Durgante, F., Lopes, A., Trumbore, S.E., Sanna Freire Silva, T., ter Steege, H., Val, A.L., Junk, W.J., Piedade, M.T.F., 2021. The shadow of the Balbina dam: a synthesis of over 35 years of downstream impacts on floodplain forests in Central Amazonia. Aquat. Conserv. Mar. Freshw. Ecosyst. 31, 1117–1135. https://doi.org/10.1002/aqc.3526.

Schwanninger, M., Rodrigues, J.C., Fackler, K., 2011. A review of band assignments in near infrared spectra of wood and wood components. J. Infrared Spectrosc. https://doi.org/10.1255/jnirs.955.

Stropp, J., Richard,Ladle, J., Emilio, T., Lessa, Thainá, Joaquín Hortal, 2022. Taxonomic uncertainty and the challenge of estimating global species richness. https://doi.org/10.1111/jbi.14463.

Targhetta, N., Kesselmeier, J., Wittmann, F., 2015. Effects of the hydroedaphic gradient on tree species composition and aboveground wood biomass of oligotrophic forest ecosystems in the central Amazon basin. Folia Geobot. 50, 185–205. https://doi.org/10.1007/s12224-015-9225-9.

ter Steege, H., Mota de Oliveira, S., Pitman, N.C.A., Sabatier, D., Antonelli, A., Guevara Andino, J.E., Aymard, G.A., Salomão, R.P., 2019. Towards a dynamic list of Amazonian tree species. Sci. Rep. 9, 1–5. https://doi.org/10.1038/s41598-019-40101-y.

ter Steege, H., Pitman, N.C.A., Sabatier, D., Baraloto, C., Salomão, R.P., Guevara, J.E., Phillips, O.L., Castilho, C.V., Magnusson, W.E., Molino, J.F., Monteagudo, A., Vargas, P.N., Montero, J.C., Feldpausch, T.R., Coronado, E.N.H., Killeen, T.J., Mostacedo, B., Vasquez, R., Assis, R.L., Terborgh, J., Wittmann, F., Andrade, A., Laurance, W.F., Laurance, S.G.W., Marimon, B.S., Marimon, B.H., Vieira, I.C.G., Amaral, I.L., Brienen, R., Castellanos, H., López, D.C., Duivenvoorden, J.F., Mogollón, H.F., Matos, F.D.D.A., Dávila, N., García-Villacorta, R., Diaz, P.R.S., Costa, F., Emilio, T., Levis, C., Schietti, J., Souza, P., Alonso, A., Dallmeier, F., Montoya, A.J.D., Piedade, M.T.F., Araujo-Murakami, A., Arroyo, L., Gribel, R., Fine, P.V.A., Peres, C.A., Toledo, M., Aymard, C., Baker, G.A., Cerón, T.R., Engel, C., Henkel, J., Maas, T.W., Petronelli, P., Stropp, J., Zartman, J., Daly, C.E., Neill, D., Silveira, D., Paredes, M., Chave, M.R., Lima Filho, J., Jørgensen, D.D. A., Fuentes, P.M., Schöngart, J., Valverde, J., Di Fiore, F.C., Jimenez, A., Mora, E.M., Phillips, M.C.P., Rivas, J.F., Van Andel, G., Von Hildebrand, T.R., Hoffman, P., Zent, B., Malhi, E.L., Prieto, Y., Rudas, A., Ruschell, A., Silva, A.R., Vos, N., Zent, V., Oliveira, S., Schutz, A.A., Gonzales, A.C., Nascimento, T., Ramirez-Angulo, M.T., Sierra, H., Tirado, R., Medina, M., Van Der Heijden, M.N.U., Vela, G., Torre, C.I.A., Vriesendorp, E.V., Wang, C., Young, O., Baider, K.R.,

Balslev, C., Ferreira, H., Mesones, C., Torres-Lezama, I., Giraldo, A., Zagt, L.E.U., Alexiades, R., Hernandez, M.N., Huamantupa-Chuquimaco, L., Milliken, I., Cuenca, W., Pauletto, W.P., Sandoval, D., Gamarra, E.V., Dexter, L.V., Feeley, K.G., Lopez-Gonzalez, K., Silman, M.R, G., 2013. Hyperdominance in the Amazonian tree flora. Science 342 (80). https://doi.org/10.1126/science.1243092.

ter Steege, H., prado, P.I., de Lima, F., pos, R.A., de Souza Coelho, L, E., de Andrade Lima Filho, D., Salomão, R.P., Leão Amaral, I., Dionízia de Almeida Matos, F., Castilho, C.V., phillips, O.L., Ernesto Guevara, J., de Jesus Veiga carim, M., Cárdenas López, D., Magnusson, W.E., Wittmann, F., Pires Martins, M., Sabatier, D., Victória Irume, M., Renan da Silva Guimarães, J., Molino, J.-F., Bánki, O.S., Teresa Fernandez piedade, M., pitman, A., Ferreira Ramos, N.C., Monteagudo Mendoza, J., Martins Venticinque, A., Garcia Luize, E., Núñez Vargas, B., Percy, Sanna Freire Silva, T., Márcia Moraes de Leão Novo, E., Farias costa Reis, neidiane, terborgh, J., Gilberto Manzatto, A., Regina Casula, K., Honorio coronado, E.N., Carlos Montero, J., Duque, A., costa, C., castaño Arboleda, F.R., Schöngart, N., Eugene Zartman, J., Killeen, C., Marimon, T.J., Hur Marimon-Junior, B.S., Vasquez, B., Mostacedo, R., Demarchi, B., Feldpausch, L.O., Engel, T.R., Petronelli, J., Baraloto, P., Assis, C., Castellanos, R.L., Fragomeni Simon, H., Brilhante de Medeiros, M., Quaresma, A, M., W Laurance, S.G., Rincón, L.M., Andrade, A., Sousa, T.R., Luís camargo, J., Schietti, J., Laurance, W.F., Lima de Queiroz, H., Eduardo Mendonça nascimento, H., Aparecida Lopes, M., de Sousa farias, E., Leonardo Lima Magalhães, J., Brienen, R., Aymard, C., David Cardenas Revilla, G.A., Célia Guimarães Vieira, J., Barçante Ladvocat cintra, I., Stevenson, B., Oliveira feitosa, P.R., Duivenvoorden, Y., Mogollón, J.F., Araujo-Murakami, H.F., Valle ferreira, A., Rafael Lozada, L., comiskey, J., Julio de Toledo, J.A., Damasco, J., Dávila, G., Lopes, N., García-Villacorta, A., Draper, R., Vicentini, F., Cornejo Valverde, A., Lloyd, F., Gomes, J.F., neill, V.H., Alonso, D., Dallmeier, A., francisco, Coelho de Souza, F., Gribel, R., Arroyo, L., Antunes Carvalho, F., Praia Portela de Aguiar, D., Dantas do Amaral, D., Petratti pansonato, M., Feeley, K.J., Berenguer, E., fine, A., Carneiro Guedes, P.V., Barlow, M., ferreira, J., Villa, J., Cristina Peñuela Mora, B., Jimenez, M., Carlos Licona, E.M., cerón, J., thomas, C., Maas, R., Silveira, P., Henkel, M., Stropp, T.W., Ríos Paredes, J., Dexter, M., Daly, K.G., Baker, D., Huamantupa-Chuquimaco, T.R., Milliken, I., Pennington, W., Sebastián Tello, T., Luis Marcelo Pena, J., Peres, J., Klitgaard, C.A., Fuentes, E., Silman, A., Di fiore, M.R., von Hildebrand, A., Chave, P., van Andel, J., Richard Hilário, T.R., Fernando Phillips, R., Rivas-torres, J., Costa Noronha, G., Prieto, J., Gonzales, A., de Sá Carpanedo, T., Pepe Gallardo Gonzales, R., Zárate Gómez, G., de Jesus Rodrigues, R., Zent, D., Ruschel, A.R, E.L., 2020. Biased-corrected richness estimates for the Amazonian tree flora. In: nature.comH Ter Steege, PI Prado, RAF Lima, E Pos, L Souza Coelho, D Andrade Lima FilhoScientific reports, 10, p. 10130. https://doi.org/10.1038/s41598-020-66686-3, 2020●nature.com.

ter Steege, H., Vaessen, R.W., Cárdenas-López, D., Sabatier, D., Antonelli, A., De Oliveira, Mota, Pitman, S., Jørgensen, N.C.A., Salomão, R.P, P.M., 2016. The discovery of the Amazonian tree flora with an updated checklist of all known tree taxa, 2016●nature.com Nat. comH Ter. Steege RW Vaessen D. C. árdenasLópez D. Sabatier A Antonelli SM OliveiraScientific Rep.. https://doi.org/10.1038/srep29549.

Tsuchikawa, S., 2007. A review of recent near infrared research for wood and paper. Appl. Spectrosc. Rev. https://doi.org/10.1080/05704920601036707.

Vasconcelos, C. da C., Ferraz, I.D.K., Adrianzén, M.U., Camargo, J.L.C., Terra-Araujo, M.H., 2021. Chromolucuma brevipedicellata (Sapotaceae, Chrysophylloideae), a new tree species from central Amazonia, Brazil. Brittonia 73, 211–219. https://doi.org/10.1007/s12228-021-09663-y.

Vasconelos, C.C., Adriazén, M.U., Camargo, J.L.C., Terra-Araujo, M.H., 2020. Pouteria kossmanniae (Sapotaceae): a new species from Central Amazonia, Brazil. Phytotaxa 447, 265–275. https://doi.org/10.11646/PHYTOTAXA.447.4.4.

Williams, Titus, 1988. Assessment of sampling stability in ecological applications of discriminant analysis. Ecol. Soc. Am. 69, 1275–1285.

Workman, J., Weyer, L., 2007. Practical Guide to Interpretive Near-Infrared Spectroscopy, Practical Guide to Interpretive Near-Infrared Spectroscopy. CRC Press. https://doi.org/10.1201/9781420018318.