REMOTE SENSING TOOLS FOR DETECTING AND QUANTIFYING
LIANAS AND TREES AT THE TROPICAL DRY FOREST

by

Jose Antonio Guzmán Quesada

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Earth and Atmospheric Sciences
University of Alberta

ABSTRACT

Lianas are woody thick-stemmed climbers that use host trees to reach the forest canopy. Studies have shown a remarkable increase in liana abundance in the last two decades, while others have shown that liana abundance is associated with detrimental effects on forest dynamics. Liana abundance presents peaks in highly seasonal forests such as the Tropical Dry Forest (TDF); regions that are under threat for frequent droughts, fires, and anthropogenic pressures. Despite their abundance and relevance in these fragile ecosystems, there are no clear research priorities that help to conduct an efficient detection and monitoring of lianas. This dissertation aims to integrate new remote sensing perspectives to detect and quantify lianas and trees at the TDF. This was addressed using passive (Chapters 2 – 4) and active remote sensing (Chapter 5).

Using thermography, Chapters 2 explored the temporal variability of leaf temperature of lianas and trees at the canopy. Temperature observations were conducted in different seasons and ENSO years on lianas and trees infested and non-infested by lianas. The findings revealed that the presence of lianas on trees does not affect the temperature of exposed tree leaves; however, liana leaves tended to be warmer than tree leaves at noon. The results emphasize that lianas are an important biotic factor that can influence canopy temperature, and perhaps, its productivity.

Chapter 3 assessed the discrimination of liana and tree leaves using visible-near infrared (VIS-NIR) and longwave infrared (LWIR) spectra. This chapter compared the former contrasting spectral regions, four representations of leaf spectra, twenty-one algorithms of classification, and two contrasting life forms in the context of machine learning to explore the question of whether it is possible to discriminate between liana and tree leaves. The results revealed that both life forms

are more accurately discriminated using LWIR spectra (accuracy between 66% and 96%) compared with VIS-NIR spectra (accuracy between 50% and 69%). However, such accuracies of discrimination were achieved depending on the kind of spectral pre-processing and machine learning algorithm. The chapter's outcomes suggest the possibility to extend the discrimination between lianas and trees to airborne or satellite LWIR observations.

The prediction of leaf traits of lianas and trees using Partial Least-Square Regression (PLSR) models based on leaf reflectance or wavelet spectra is addressed in Chapter 4. This chapter revealed that the model performance differs between life forms or between reflectance/wavelet spectra models. Differences in model performance between life forms seemed to be the product of the intraspecific variability of leaf traits within these life forms. Likewise, it was shown that PLSR models based on wavelet spectra help to overcome current limitations of PLSR models based on reflectance spectra. The results showed that the variability of leaf traits between life forms influences predictive models. Thus, the variability of traits between plant groups may have an essential role in estimated errors associated with the mapping of leaf traits.

Using Terrestrial Laser Scanning, Chapter 5 evaluated the relationship between fractal geometry and tree-stands metrics on point clouds of trees. The chapter's results suggested that the intercept extracted from fractal geometry is an accurate and fast parameter that helps predict plant volume, crown coverage, or plant basal area at the tree or stand level. The fractal geometry also revealed that the fractal dimension is strongly associated with the presence/absence of leaves in the point cloud or the number of trees in the stands. Since this method is not susceptible to irregular structures, this method may potentially contribute to quantifying the volume of lianas or buttress roots of trees.

Chapter 6 provides future research directions that may help explain the drivers that lead the observed findings or the potential applicability of the results. Overall, this thesis highlighted the

need for new efficient and fast approaches that help assess the role and extent of lianas in the tropics. In the absence of a solid understanding of the presence and the effect of lianas in forest dynamics, future predictions of tropical forest productivity will remain speculative.

PREFACE

This dissertation is an original work by J. Antonio Guzmán Q. Chapter 2 of this thesis has been published as J.A. Guzmán, G.A. Sánchez-Azofeifa, and B. Rivard (2018) "Differences in leaf temperature between lianas and trees in the Neotropical canopy" *Forest*, vol. 9, issue 6, 307. I was responsible for the data processing, analysis as well as the manuscript composition. G.A. Sánchez-Azofeifa, the supervisory author, was involved with the data collection and the edits of the manuscript together with B. Rivard.

Chapter 3 has been published as J.A. Guzmán, B. Rivard, and G.A. Sánchez-Azofeifa (2018) "Discrimination of liana and tree leaves from a Neotropical Dry Forest using visible-near infrared and longwave infrared reflectance spectra" *Remote Sensing of Environment*, vol. 219, 135-144. I was responsible of data collection, analysis, and manuscript preparation. Both B. Rivard, and G.A. Sánchez-Azofeifa helped with the edits of the manuscript and their conceptualization.

Chapter 4 was also submitted for publication in *Remote Sensing of Environment* as J.A. Guzmán and G.A. Sánchez-Azofeifa "Prediction of leaf traits of lianas and trees via the integration of wavelet spectra in the visible-near infrared and thermal infrared domains". This chapter was accepted for publication at the moment of submitted the final thesis version to the Faculty of Graduate Studies and Research. For this chapter, I was responsible for data collection, analysis, and manuscript preparation. G.A. Sánchez-Azofeifa was the supervisory author and was involved the edits of the manuscript. The sample collection for Chapter 3 and 4 was performed under the scientific license ACG-PI-PC-033-2017 of the Costa Rican Government.

Chapter 5 was published as J.A. Guzmán, I. Sharp, F. Alencastro, and G.A. Sánchez-Azofeifa (2020) "On the relationship of fractal geometry and tree–stand metrics on point clouds derived from terrestrial laser scanning" *Methods in Ecology and Evolution,* 11: 1309– 1318. I was responsible of data collection, code development, analysis, and manuscript preparation. I. Sharp and F. Alencastro assisted with the data collection and preparation. G.A. Sánchez-Azofeifa, the supervisory author, was involved with the concept formation and edits of the manuscript.

Five other papers associated with some of the methods and the environment where this thesis took place were also published during Guzmán's PhD as a first author or co-author. These papers

are not part of the core dissertation of this thesis and are not included in this document. These papers can be found as

Guzmán, J.A., Laakso, K., López-Rodríguez, J.C., Rivard, B., Sánchez-Azofeifa, G.A. 2020. Using visible-near-infrared spectroscopy to classify lichens at a Neotropical Dry Forest. Ecological Indicators 111: 105999. Link

Guzmán, J.A., Sánchez-Azofeifa, G.A., Espírito-Santo, M.M. 2019. MODIS and PROBA-V NDVI Products Differ when Compared with Observations from Phenological Towers at Four Tropical Dry Forests in the Americas. Remote Sensing 11: 2316. Link

Rankine, C., Sánchez-Azofeifa, GA., Guzmán, JA., Espirito Santo, M., Sharp, I. 2017. Comparing MODIS and Near-Surface Vegetation Indexes for monitoring tropical dry forest phenology along a successional gradient using optical phenology towers. Environmental Research Letters. 12: 105007. Link

Sánchez-Azofeifa, GA, Guzmán, JA., Campos, CA, Castro, S., Garcia-Millan, V., Nightingale, J, Rankine, C. 2017. Twenty-first century remote sensing technologies are revolutionizing the study of tropical forests. Biotropica. 49: 604–619. Link

Sánchez-Azofeifa, GA., Guzmán, JA., Vega-Araya, M., Campos-Vargas, C., Durán, S., D'Souzaa, N., Gianoli, T., Portillo-Quintero, C., Sharp, I. 2017. Can Terrestrial Laser Scanner (TLS) and hemispherical photographs predict Tropical Dry Forest Succession with liana abundance? Biogeosciences. 14: 977–988. Link

To my little boy who radiates inspiration
and my wife who knows how to absorb my frustrations
and reflect understanding

ACKNOWLEDGEMENTS

I want to express my gratitude to everyone who has encouraged and supported my scientific path throughout these years. An immense thank you to my supervisor, Dr. Arturo Sánchez-Azofeifa. Working with Dr. Sánchez-Azofeifa since 2016 has allowed me to grow as a professional, inspiring my intellectual growth directly or indirectly. I sincerely appreciate all the support that he has provided me during this path. From fieldwork, traveling to congresses, teaching, deadlines, and even misunderstandings, Dr. Sanchez has made of this program a life experience that I will never forget. I will always have in mind a Costa Ricanism phrase that he told me at the beginning of the program '*si no llora, no mama*'; meaning that if you want to achieve something, you need to work hard and insist on getting it. This simple phrase, applied to different aspects of the program, changed my Ph.D. experience dramatically. Thank you.

I am deeply appreciative for my supervisor committee, Dr. Benoit Rivard and Dr. Janice Cooke, for sharing their thoughts and expertise in this life project. I am grateful for their discussions and comments during the different stages of this thesis from its beginning at the candidacy exam to its final presentation. I am especially grateful to Dr. Rivard who took his time to supervise me in the application of some of the methods used in this dissertation.

I am thankful to my CEOS lab mates, especially to Carlos Campos, Iain Sharp, and Sofia Calvo with whom I shared more than a month in the field eating sandwiches and several coffee cups. I appreciate the multiple feedbacks from my CEOS fellows through the lab meetings since that helped me develop several ideas. I would also like to express my thankfulness to the CEOS lab managers, Meimei Chong, Michael Hesketh, and Kati Laakso, for their constant help in organizing equipment and materials. CEOS mates, thanks for making this program a place of friendship and collegiality. I would also like to thank Santa Rosa National Park and its staff for allowing me to conduct research using their facilities.

Last but not least, the biggest thanks to my family, especially my parents, Julio and Rosibel, for their unending love and supporting my dreams. I appreciate them getting a loan that allowed me to come to Canada and lending me their car whenever I had fieldwork in Costa Rica. You have been with me every step of the way, even at a distance. To my wife, Elizabeth, who made this program a human experience. I deeply appreciate the long talks about different aspects of this thesis and her

selfless love to travel with me to the end of the world. To my little boy, Toñito, who with his smile makes my days during the pandemic locked up at home. Thanks both of you for always being there.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Appendices

ABBREVIATIONS

Abbreviation	Definition
AGB	Above-Ground Biomass
ANCOVA	Analyses of Covariance
AUS	Victoria, Australia
BA	Basal Area
BC	Bagged CART
BGLM	Bayesian Generalized Linear Model
CA	Crown Area
CAM	Eastern Cameroon
CAN	University of Alberta North Campus, Canada
CART	CART
CR	Costa Rica
CWT	Continuous Wavelet Transformation
DBH	Diameter at Breast Height
d_{MB}	Fractal dimension from fractal geometry
E	Size of a voxel edge
EnMAP	Environmental Mapping and Analysis Program
ENSO	El Niño-Southern Oscillation
EWT	Equivalent Water Thickness
FTIR	Fourier Transform Infrared Spectroscopy
GLM	Generalized Linear Model
GUY	Central Guyana
H	Tree Height
HISUI	Hyperspectral Imager Suite

IAI	Inter-American Institute for Global Change Research
IND	Mentaya River, Indonesia
intercept$_{MB}$	Intercept from fractal geometry
KNN	k-Nearest Neighbors
LDA	Linear Discriminant Analysis
LiDAR	Light Detection and Ranging
LMA	Leaf Mass Area
LWIR	Longwave infrared
MAP	Mean Annual Precipitation
MB	Minkowski–Bouligand
MLWIR	Mid- long- wave infrared
MULDA	Maximum Uncertainty Linear Discriminant Analysis
NN	Neural Network
NSERC	Natural Science and Engineering Research Council of Canada
P	Leaf photosynthesis
PCA	Principal Components Analysis
PER	South Western Amazon, Peru
PLSR	Partial Least Squares Regression
PNN	Parque Natural Metropolitano
P_{opt}	Maximum rate of photosynthesis at an optimum temperature
PRISMA	PRecursore IperSpettrale della Missione Applicativa
Q_{10}	Proportional increase in respiration with a 10 °C temperature rise
QDA	Quadratic Discriminant Analysis
QSM	Quantitative Structure Models
R	Leaf dark respiration
R^2	Coefficient of determination
R_{25}	Leaf respiration at 25 °C
RDA	Regularized Discriminant Analysis

RF	Random Forest
RLDA	Robust Linear Discriminant Analysis
RMDA	Robust Mixture Discriminant Analysis
RMSE	Root Mean Square Error
RMSRP	Root Mean Squared Error of Prediction
RQDA	Robust Quadratic Discriminant Analysis
RSE	Residual Standard Error
S	Size
SDA	Shrinkage Discriminant Analysis
SE	Standard error
SLA	Specific Leaf Area
SLDA	Stabilized Linear Discriminant Analysis
SMA	Standardize Major Axis regression
SMOTE	Synthetic Minority Over-Sampling Technique
SR	Single C5.0 Ruleset
SRNP–EMSS	Santa Rosa National Park – Environmental Monitoring Super Site
ST	Single C5.0 Tree
STRI	Smithsonian Tropical Research Institute
SVMLK	Support Vector Machines with Linear Kernel
SVMPK	Support Vector Machines with Polynomial Kernel
SVMRK	Support Vector Machines with Radial Basis Function Kernel
T_{air}	Air temperature
TBMs	Terrestrial Biosphere Models
T_d	Displayed leaf temperature
TDF	Tropical Dry Forest
T_{leaf}	Leaf temperature
T-LiDAR	Terrestrial-Light Detection and Ranging
TLS	Terrestrial Laser Scanning

T_{opt}	Optimum temperature
V	Tree Volume
VC	voxel-counting method
VIP	Variable Importance of Projection
VIS-NIR	Visible-near infrared
%RMSE	Percentage of Root Mean Square Error
α	Intercept of a linear mixed model
Ω	Difference in temperature between optimal temperature and the temperature in which photosynthesis drops to 37% of its value

CHAPTER 1

INTRODUCTION

1.1 Background

Climbing plants are a particular functional group that uses trees to grow, develop, and reproduce at the forest canopy. Pioneer work by Alwyn H. Gentry classified climbing plants into four main groups based on their climbing strategies, ecology, and morphology (Gentry, 1895, 1991): i) herbaceous epiphytes-hemiepiphytes, ii) woody hemiepiphyte, iii) vines, and iv) lianas. Among these climbing plants, lianas are described as woody thick-stemmed climbers that begin life as terrestrial seedlings to find later in their development host trees to reach the forest canopy (Figure 1–1) (Gentry, 1991). The ecological role of the interaction between lianas and trees is not well understood in comparison with other climbers (Stewart and Schnitzer, 2017). Despite growing around and over host trees, it is not entirely clear if the presence of trees affects lianas (i.e., lianas are competitors) or if lianas can survive without the presence of trees (i.e., lianas are obligate parasites) (Stewart and Schnitzer, 2017). Thus, hereinafter, I will refer to lianas as structural parasites of trees.

The presence of liana on trees can have a meaningful impact on forest dynamics and structure. During their growth and development, lianas tend to produce a higher proportion of their biomass in foliage while infesting their host (Putz, 1984; Wyka et al., 2013). The presence of liana on trees can influence their growth by the interception of light or influence tree mortality by increasing the mechanical strain on tree crowns (Rodríguez-Ronderos et al., 2016; Selaya and Anten, 2008; Visser et al., 2018). Moreover, liana-infested canopies tend to have a higher Woody Area Index, which in turn reduces canopy closure (Sánchez-Azofeifa et al., 2009b).

Lianas have a pivotal role in the biodiversity and structure of tropical ecosystems. Together with vines, lianas contribute up to 24% of the species richness of many tropical forests (Gentry, 1991). Lianas also comprise up to 24% of the woody stems in dry forests, while 18% in lowland and wet forests (Gentry, 1986, 1982). Likewise, it has been estimated that 5% of the above-ground total biomass of lowland rain forest are represented by this group of plants (DeWalt and Chave,

1

2004). The role of lianas in the ecosystems goes further than an element of biodiversity and structure; their abundance is strongly linked to the fauna presence and trophic chains' prevalence (Kilgore et al., 2010; Ødegaard, 2000). For instance, a higher abundance of lianas has been associated with richer communities of birds and ants as a result of the sheltering and nesting-eating environment created around them (Adams et al., 2017; Hilje et al., 2017; Schnitzer et al., 2020).
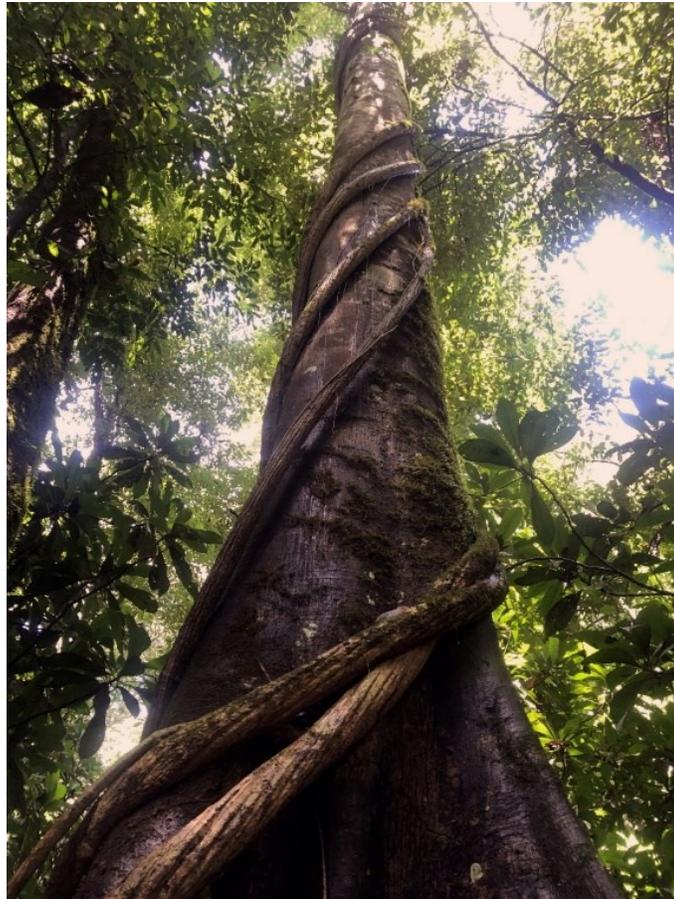
Figure 1–1. Photograph of a liana growing around a tree to reach the forest canopy. Photo credit: Marisol Luna (2020).

Despite being a group composed of several families, lianas are unevenly distributed worldwide; the majority of species are found to tropical environments (Gentry, 1991). Due to their distribution, lianas have been considered the key physiognomic feature that differentiates tropical from temperate forests (Croat, 1978). Overall, liana abundance and diversity vary with latitude and mean annual precipitation (MAP), showing a decline in woody vines at higher latitudes (Gallagher

and Leishman, 2012; Gentry, 1991) and high precipitation environments (Gentry, 1991; Parolari et al., 2020; Schnitzer and Bongers, 2011). According to Schnitzer (2005), the pantropical patterns of the liana abundance are related to their capacity to grow under water-stress conditions in comparison with trees; allowing them to present peaks of abundance in highly seasonal environments in contrast with trees.

Liana traits such as leaf size, seed mass, and Specific Leaf Area (SLA) also vary with latitude and MAP (Mello et al., 2020). For instance, lianas and other climbing plants tend to present larger leaf size and seed mass and lower SLA in regions close to the equator (Gallagher and Leishman, 2012). Likewise, liana traits also vary between environments with contrasting precipitation regimes. For example, lianas in wet environments tend to have higher leaf thickness and lower SLA compared to lianas in dry environments (Sánchez-Azofeifa et al., 2009a). Although liana traits seem to vary in latitudinal and precipitation gradients, the variation in lianas' species-specific attributes (i.e., tendrils, nectarines, floral type, seed type) tends to be spatially constrained (Meyer et al., 2020). Spatial limitations associated with the distribution of species-specific attributes of lianas are the product of the high diversity of the ecological strategies and evolutionary paths of this group (Meyer et al., 2020).

1.2 Lianas as a fingerprint of the global change

Since 2002, different studies have reported a remarkable increase in liana abundance in many temperate and tropical ecosystems (DeWalt et al., 2010; Londré and Schnitzer, 2006; Phillips et al., 2002; Schnitzer, 2015, 2005). Moreover, other studies have found that the presence of lianas is associated with detrimental effects related to a decrease in recruitment, growth, and survival of trees and forest biomass (Durán and Gianoli, 2013; Durán and Sánchez-Azofeifa, 2015; Estrada-Villegas et al., 2020; Lai et al., 2017; Martínez-Izquierdo et al., 2016; Schnitzer et al., 2014; Schnitzer and Carson, 2010; van der Heijden et al., 2015). For example, it has been estimated that lianas are responsible for 76% of forest net above-ground biomass reductions at a liana removal experiment in Panama (van der Heijden et al., 2015). Furthermore, lianas' detrimental effect on forest biomass seems to be higher in secondary forests (Lai et al., 2017), with this effect being more predominant as a function of the successional stage (Durán and Sánchez-Azofeifa, 2015; Estrada-Villegas et al., 2020). These increases in liana abundance and their detrimental effect suggest that lianas are an essential element of change in natural environments, impacting carbon

dynamics in tropical forests. As a result of this, researchers have cast doubts on the accuracy of models that do not include lianas to predict the current and future carbon cycle, highlighting the need to understand the liana's footprint in ecosystem dynamics (di Porcia e Brugnera et al., 2019; Meunier et al., 2020; Verbeeck and Kearsley, 2016).

Different hypotheses have emerged to explain the current increase in liana abundance (Schnitzer and Bongers 2011). For instance, studies suggest that increases in liana abundance might be associated with their response in leaf development and plant growth to high CO_2 concentrations in comparison with trees, which could be correlated with current increases of atmospheric CO_2 (Granados and Körner, 2002; Oki et al., 2014; Zotz et al., 2006). Other studies suggest that the efficient water transportation, deep soil water acquisition, and efficient leaf stomatal behavior of lianas may explain their proliferation in seasonal environments and their advantages over trees (Andrade et al., 2005; Chen et al., 2017, 2015; Ichihashi et al., 2017; Zhu and Cao, 2009). This hypothesis may also provide a plausible explanation of the extended leaf phenology of lianas over trees in a seasonal environment (Kalácska et al., 2005). However, the low rooting depth of lianas compared to trees does not seem to support part of the previous hypothesis (Smith-Martin et al., 2020). Another hypothesis, and probably the most accepted, suggests that the current trends in forest fragmentation and disturbance could be associated with increases in liana abundance due to a large group of abiotic and biotic factors that can be affected and that could benefit lianas over trees (Londré and Schnitzer, 2006; Magnago et al., 2017; van Melis et al., 2020). This hypothesis supports the decrease in liana abundance observed at the Congo Basin; product of the decimated elephant population that leads to reductions in forest disturbance in the region (Bongers et al., 2020). Schnitzer and Bongers (2011) clustered these hypotheses into four theme groups associated with putative mechanisms that explain increases in liana abundance: i) evapotranspirative demand, ii) natural disturbance, iii) fragmentation, and iv) elevated atmospheric CO_2. Regardless of the drivers that lead liana proliferation, these themes show common elements related to forest health and global change (i.e., atmospheric CO_2, water availability-droughts, forest disturbance) (Trumbore et al., 2015). Hence, lianas have been considered as a fingerprint of the impacts of the global change on tropical environments (Lewis et al., 2004).

1.3 The Tropical Dry Forest and liana abundance

As suggested before, liana abundance and diversity tend to increase toward seasonal environments with low MAP (Durán et al., 2015; Parolari et al., 2020; Schnitzer, 2005). A recent study conducted in Panama also suggests that liana proliferation and richness are favored by how rainfall is distributed throughout the year (Parolari et al., 2020), revealing that high dormancy environments –those with longer dry seasons– are prone to a high relative abundance and species richness of lianas. According to Parolari et al. (2020), the strong response of lianas to MAP and rainfall seasonality may suggest that future changes in rainfall regimes will determine the liana distribution and abundance in the tropics.

Across the tropics, the Tropical Dry Forest (TDF) is among the ecoregions with high rainfall seasonality and low MAP. Overall, the TDF has been described by Sánchez-Azofeifa et al. (2005) as a deciduous forest that presents a mean annual temperature above 25 °C, total annual precipitation between 700 and 2000 mm, and three or more dry months per year. For several decades, the TDF has been considered one of the most threatened ecosystems (Janzen, 1988) due to its attractive features for human settlement and development, which have led to the loss of wildlife habitats by deforestation and fragmentation (Portillo-Quintero and Sánchez-Azofeifa, 2010). The high liana abundance in these environments and the ongoing climatic changes associated with these (Stan et al., 2020) may imply that lianas are passive threats to the future structure and dynamics of these ecosystems.

Different research priorities have been proposed for the TDF, specifically to the neotropical dry forest. These research priorities are associated with creating linkages between social sciences and ecological studies for better science-policy nexus (Sánchez-Azofeifa et al., 2005) or the need to use improved sensor capabilities to better characterize biophysical variables or to map tree species (Sánchez-Azofeifa et al., 2003). Despite this, there are no clear research priorities that aim to integrate the study of lianas and trees at the TDF. Recently, da Cunha Vargas et al. (2020) attempted to evaluate the current state-of-the-art of ecological research of lianas and trees. Although da Cunha Vargas et al. (2020) identified the current need to better detect and map lianas, future perspectives do not go further than the current body of literature, perhaps because the emphasis of authors is more ecological with little or no knowledge of advance remote sensing topics. As such, it could be considered that there is no clear roadmap for future studies that aim to detect and quantify the extent of lianas in the TDF.

1.4 Remote sensing technologies applied to lianas and trees.

To better understand the implications of lianas and trees in a Tropical Dry Forest, new technologies need to be exploited that enhance the detection and monitoring of these life forms. In this regard, remote sensing technologies are revolutionizing the ways that studies of tropical forests are carried out (Sanchez-Azofeifa et al., 2017). It is likely, therefore, that new technologies and methods associated with this field provide feasible solutions to detect and quantify the extent and dynamics of lianas. Overall, the study of lianas and trees through passive or active remote sensing is not entirely new. Using passive sensors, for instance, several studies have addressed the optical properties of lianas from leaf, crown, and canopy observations (Asner and Martin, 2011; Avalos et al., 1999; Castro-Esau et al., 2004; Foster et al., 2008; Hesketh and Sánchez-Azofeifa, 2012; Kalacska et al., 2007; Li et al., 2018; Marvin et al., 2016; Oki et al., 2014; Sánchez-Azofeifa et al., 2009a; Sánchez-Azofeifa and Castro-Esau, 2006). Most of these studies were conducted using reflectance spectra encompassing the visible, near, and short-wave infrared regions (0.38 – 2.5 μm) where spectral signatures of lianas and trees are separable. Other studies based on passive sensors use RGB cameras or thermal infrared imagery to address the mapping of lianas and trees (Li et al., 2018; Waite et al., 2019; Yuan et al., 2019). Yuan et al. (2019) specifically, based part of their research on some of the main findings of Chapter 2 of this dissertation. On the other hand, using active sensors such as Terrestrial Laser Scanning (TLS), studies have addressed the potential impact of lianas on forest structure (Moorthy et al., 2018; Rodríguez-Ronderos et al., 2016; Sánchez-Azofeifa et al., 2017) or attempted to quantify the biomass of lianas and trees (Moorthy et al., 2020, 2019).

1.5 Motivations and scope of this dissertation.

In the face of a lacking roadmap for future studies, the unifying aim of this dissertation is to integrate new remote sensing perspectives for the detection and quantification of lianas and trees at the TDF. This aim is addressed by conducting multidisciplinary research that links fields such as remote sensing, modeling, and plant ecophysiology. This dissertation encompasses four research chapters that use passive (Chapter 2 – 4) and active (Chapter 5) remote sensing. Chapter 6, summarizes the main findings of this dissertation and provides future perspectives. The

motivation behind each chapter as well as the specific objectives and hypotheses are presented below.

Chapter 2 analyzes the temporal trends of intra- and inter-specific leaf temperature of full-sun exposed leaves of several liana species and four host trees. This is accomplished using thermography at the canopy of a seasonal forest in Panama. This chapter evaluates two hypotheses: i) the presence of lianas affects the leaf temperature of their host trees (intra-specific trends), and ii) leaves of lianas and their host trees exhibit differences in leaf temperature (inter-specific trends). These hypotheses are addressed by observing temporal variations of displayed leaf temperature (leaf temperature – ambient temperature) during the wet and dry seasons in contrasting El Niño (2015–2016) and La Niña years (2016–2017). To show the potential impact of the differences in leaf temperature between these life forms on carbon exchange, equations of respiration and photosynthesis response to temperature (Slot et al., 2013; Slot and Winter, 2017) are used to estimate rates of gas exchange. This chapter aims to highlight the fact that lianas are an important biotic factor at the canopy level, which may affect the forest temperature, and consequently, its productivity.

Chapter 3 assesses and compares the use of visible-near infrared (VIS-NIR, 0.45 – 0.95 μm) and longwave infrared (LWIR, 8 – 11 μm) spectra to discriminate leaves of lianas and trees. As mentioned in section 1.3, several studies have addressed the discrimination of lianas and trees using leaf optical properties in the visible, near, and short-wave infrared regions (Avalos et al., 1999; Castro-Esau et al., 2004; Hesketh and Sánchez-Azofeifa, 2012; Kalacska et al., 2007; Sánchez-Azofeifa et al., 2009a; Sánchez-Azofeifa and Castro-Esau, 2006). At these spectral regions, the signatures of lianas and trees are separable; however, its accuracy of discrimination depends on seasonal or phenological factors (Hesketh and Sánchez-Azofeifa, 2012). Since lianas and trees differ in the concentration of compounds such as lignin, cellulose, and hemicellulose (Asner and Martin, 2012, 2011) and these compounds can be detected at the LWIR region (Ribeiro da Luz and Crowley, 2007), this chapter hypothesizes that LWIR spectra provide better discrimination between these life forms than the VIS-NIR spectra. This chapter also addresses the effect of the spectral pre-processing associated with noise reduction and feature enhancement on the spectral discrimination of life forms. In combination with data-reduction techniques, twenty-one machine learning algorithms are evaluated to classify liana and tree leaves. This chapter aims

to highlight the potential of the LWIR spectral region for the accurate detection of liana and trees in tropical environments.

Chapter 4 evaluates the prediction of leaf traits of lianas and trees using leaf spectra. Most of the current studies attempt to predict a set of leaf traits from a broad group of plants without considering possible differences between groups. This lack of consideration of the group identity may introduce bias in the prediction of leaf traits, which may be large between groups of species that coexist in the same environment and present distinctive anatomical, physiological, biophysical, and biochemical features; as the case of lianas and trees (Asner and Martin, 2012; Sánchez-Azofeifa et al., 2009a; Slot and Winter, 2017). Although previous studies have already developed predictive models for some leaf traits using these life forms (Asner et al., 2011; Cheng et al., 2014, 2012), none of the current ones have addressed the potential impact of the life forms on leaf trait predictions. This chapter aims to analyze the effect of life forms on the prediction of leaf traits as well as to evaluate the integration of wavelet spectra with Partial Least Square Regression (PLSR) models to improve the prediction of leaf traits. Specifically, this chapter i) evaluates the ability of PLSR models based on reflectance and wavelet spectra to accurately predict leaf traits; ii) compares how these models identify spectral regions that could play an important role in predicting traits; and iii) compares how life forms may influence the prediction of traits and performance in these models. Three leaf traits are studied: Leaf Mass per Area (LMA), gravimetric Water Content (WC), and the Equivalent Water Thickness (EWT). Leaf spectra from the VIS-NIR (0.45 – 1.0 μm) and mid- long- wave (MLWIR, 2.55 – 11.0 μm) infrared spectral regions are used to predict leaf traits. The findings in this chapter aim to highlight the use of wavelet spectra to overcome limitations of PLSR models based on leaf reflectance for predicting traits. Likewise, this chapter attempt to highlight the need to explore the potential bias associated with plant groups to predict traits.

Chapter 5 examines the relationship of fractal geometry with tree and stand metrics. Unlike previous research chapters, this chapter is focused on active remote sensing using tree point clouds derived from Terrestrial Laser Scanning (TLS). Overall, methods associated with fractals have been widely used to determine bifurcation patterns of trees or to evaluate stressors effects on plant development. In a few instances, fractals have been used to predict tree or stand metrics. This chapter explores the use of fractal geometry applied to point clouds derived from TLS to predict tree and strand metrics. This is addressed using seven open databases of point clouds with and

without leaves. Four tree metrics are estimated on each point cloud: i) tree height, ii) diameter at breast height, iii) crown area, and iv) tree volume. In addition, parameters of fractal geometry are estimated on each point cloud using the voxel-counting method (Bunde and Havlin, 1994). The relationship between fractal geometry parameters and tree metrics is then explored. This chapter also explores to up-scale the observations at the tree-level to plot-level through the creation of artificial stands. Findings associated with this chapter highlights fractal geometry equations as a fast and accurate approach for predicting the selected metrics on irregular structures. Likewise, the method presented here may help to understand how plants or stands occupy their 3D space (Sánchez-Azofeifa et al., 2017). Even though Chapter 5 does not include lianas as the central group of study, the principles drawn from this reveal its potential application on irregular woody plants such as lianas. This could be considered the first step on characterizing the effect of lianas on tree architecture and structure. Thus, herein the significance of this chapter in this dissertation.

Finally, chapter 6 summarizes the main conclusions associated with the previous research-chapters. While the main findings of this dissertation are presented within separate thematic chapters, the common theme that weaves them together is remote sensing technologies applied to life forms. Therefore, this chapter attempts to offer general insight into future research directions to better understand lianas' role in the ecosystem using remote sensing technologies.

CHAPTER 2

DIFFERENCES IN LEAF TEMPERATURE BETWEEN LIANAS AND TREES IN THE NEOTROPICAL CANOPY

Abstract

Leaf temperature (T_{leaf}) influences photosynthesis and respiration. Currently, there is a growing interest in including lianas in productivity models due to their increasing abundance and their detrimental effects in the carbon stock of tropical ecosystems. Therefore, understanding the differences of T_{leaf} between lianas and trees is important for future predictions of productivity. Here, we determined the displayed leaf temperature ($T_d = T_{leaf} -$ air temperature) of several species of lianas and their host trees during El Niño - Southern Oscillation (ENSO) and non-ENSO years to evaluate if the presence of lianas affects the T_d of their host trees, and if leaves of lianas and their host trees exhibit differences in T_d. Our results suggest that close to midday, the presence of lianas does not affect the T_d of their host trees; however, lianas tend to have higher values of T_d than their hosts across seasons, in both ENSO and non-ENSO years. Although lianas and trees tend to have similar physiological-temperature responses, differences in T_d could lead to significant differences in rates of photosynthesis and respiration based on temperature response curves. Future models should thus consider differences in leaf temperature between these two life forms to achieve robust predictions of productivity.

Keywords

2.1 Introduction

Variations in leaf temperature (T_{leaf}) have been considered a factor that can affect the net primary productivity of the biosphere. Most Terrestrial Biosphere Models (TBMs) use kinetic constants of T_{leaf} dependence to evaluate the response of photosynthesis and respiration in a given

ecosystem (Rogers et al., 2017). However, the increasing effects of global warming and the high diversity of species and ecological strategies pose a challenge to determining with accuracy the T_{leaf} of different individuals and plant communities. In tree communities of temperate regions, Leuzinger & Körner (Leuzinger and Körner, 2007) have found that the presence or absence of certain tree species in the canopy can play a significant role in the control of the forest surface temperature. This control depends in large part on the spatial arrangement and leaf functional traits of such species, such as the stomatal conductance and the capacity of leaf cooling associated with water transpiration (Dai et al., 2004; Jones, 1999; Meinzer et al., 1997). Currently, there is no concrete evidence of how the spatial arrangement of species can affect the forest surface temperature of a tropical forest; nevertheless, it could be expected that the high diversity of species, life forms, and functional traits produce a highly dynamic surface temperature that could reduce our ability to predict different ecological processes.

Studies have shown several implications of the increasing temperature on ecological processes at different levels. For example, at the leaf level, increases in temperature above the photosynthesis optimum are associated with the decline of CO_2 assimilation rates (Slot and Winter, 2017; Vargas and Cordero, 2013), stomatal conductance (Slot et al., 2016), and increases in respiration (Slot et al., 2013); trends that can vary widely among species and life forms (Slot et al., 2014b, 2013; Slot and Winter, 2017). Likewise, at the ecosystem level in tropical forests, climate warming is associated with long-term increases in biomass (Lin et al., 2010) and the dominance of plants such as lianas (Durán et al., 2015).

Associated with this later life form, since 2002, studies have reported a notable increase in liana abundance in tropical and temperate environments (DeWalt and Chave, 2004; Londré and Schnitzer, 2006; Phillips et al., 2002; Schnitzer, 2015, 2005). Likewise, other studies have shown significant detrimental effects of lianas presence on the tree recruitment, growth, survival, and carbon stock (Durán and Gianoli, 2013; Martínez-Izquierdo et al., 2016; Schnitzer and Carson, 2010). Together, these trends have cast doubt on the accuracy of some TBMs that predict the carbon cycle (Verbeeck and Kearsley, 2016) highlighting the need to incorporate lianas as a future factor in such models for a better understanding of the ecosystem dynamics. Currently, most of the temperature–response studies that have compared the physiological performance of lianas and trees at the leaf-level suggest that there is no need to make a distinction between the physiological behavior of these life forms in future models (Slot et al., 2014b, 2014a, 2013; Slot and Winter,

2017). However, the higher interception of light by lianas (Rodríguez-Ronderos et al., 2016) and the greater competitive advantage of lianas in water use (Chen et al., 2015) suggests that their T_{leaf} may differ from that of their host trees, and might also affect T_{leaf} of their hosts; such differences in T_{leaf} between these life forms could reduce our ability to predict carbon fluxes in tropical forest canopies accurately.

In this study, we analyze the intra- and inter-specific trends in T_{leaf} for full-sun exposed leaves of several liana species and four host trees of a neotropical seasonal forest in Panama. We achieved this by using thermography and addressed two hypotheses at the canopy level: (i) the presence of lianas affects the leaf temperature of their host trees (intra-specific trends), and (ii) leaves of lianas and their host trees exhibit differences in leaf temperature (inter-specific trends). These hypotheses were addressed by observing temporal variations during the wet and dry seasons in contrasting El Niño (2015–2016) and La Niña years (2016–2017). Our hypotheses were tested using the displayed leaf temperature (T_{d}) as the difference of T_{leaf} obtained from the thermal images minus the ambient air temperature (T_{a}). The T_{d} was estimated as a proxy of T_{leaf} in order to reduce the effect of the variation of the surrounding T_{a} during each measurement and perform temporal comparisons.

We hypothesized that leaves of trees with lianas would experience higher values of T_{d} than leaves of trees without lianas. This is based on the negative effects of lianas on trees associated with the ability of lianas to reduce the water availability around their host trees (Ichihashi et al., 2017); a process that could affect the heat dissipation by transpiration of leaves of host trees (Lin et al., 2017). Likewise, we expect that leaves of lianas would show lower T_{d} in comparison with host tree leaves; due to their ability to grow in drought environments (Schnitzer, 2005) and their greater competitive advantage on the acquisition, regulation, and efficient use of water in comparison with trees (Cai et al., 2009; Chen et al., 2015; De Guzman et al., 2016; Zhu and Cao, 2009). In addition, we hypothesized that during years with little rainfall or few seasons with droughts (La Niña year or dry seasons), leaves of both life forms will show higher values of T_{d} due to the high evaporative demand of the surrounding environment (Bretfeld et al., 2018). To address our hypotheses, we used unpublished values of leaf emissivity for each life form to calculate the T_{leaf}. To show the impact of the differences in T_{d}, and consequently T_{leaf}, on carbon fluxes between these life forms, we calculated rates of respiration (R) and photosynthesis (P) using equations of temperature-response previously published by Slot et al. (2013) and Slot and Winter (2017), and

our estimations of T_{leaf}. Our results highlight the fact that lianas are an important biotic factor at the canopy level, that in turn can affect forest temperature. Therefore, their differential expression in T_{leaf} should be considered for future predictions of forest productivity; that could become a part of the challenge of including lianas in future global vegetation models (Schnitzer et al., 2016).

2.2 Materials and Methods

2.2.1 Study site

This study was conducted in Parque Natural Metropolitano (PNM, 8°59′39.95″ N, 79°32′34.68″ W, 150 m a.s.l.) that is located west of Panama City on the Pacific coast of the Republic of Panama. This site presents a tropical dry forest with a mean annual temperature of 26.5 °C and annual rainfall average of 1740 mm. In general, the region is characterized by two contrasting seasons: a wet season between May and December when most of the rainfall occurs, and a dry season between January and March (Figure A1–1). The PNM contains 265 ha of natural forest reserve with an old secondary forest of 80–150 years with tree heights of up to 40 m. This site has a 42 m standing crane with a 51 m long jib with a suspended cage that was used to access the top of the canopy. The crane covers approximately 8000 m^2 of forest in which 65 and 20 species of trees and lianas can be found, respectively (Avalos and Mulkey, 1999). Using this crane, we conducted four data collection campaigns: November 2015, February 2016, October 2016, and February 2017 (two in wet and two in dry seasons). The first two campaigns were conducted during a strong El Niño year, while the last two campaigns occurred during a starting La Niña year. A mosaic of the canopy at PNM captured on December 2015 can be observed at http://www.gigapan.com/gigapans/196831.

2.2.2 Species selection and field design

At the top of the canopy, we selected four of the most abundant tree species that were fully exposed to the sun: *Anacardium excelsum* (Bertero & Balb. ex Kunth), a late-successional species; *Annona spraguei* (Saff.), a mid-successional species; *Castilla elastica* (Liebm.), a mid-successional species; and *Luehea seemannii* (Triana & Planch), an early-successional species. We took between two and six fusion images (RGB and thermal) at the top of the crown for two or four individuals of each species with lianas and without lianas. The categorization of a tree with or without lianas was based on canopy observations without considering the possible presence of

lianas on trees that do not reach the top-crown. The fusion images used in this study were collected with a FLIR T400 thermal infrared camera (FLIR Systems AB, Danderyd, Sweden). Samples from *A. spraguei* were only collected during an El Niño year due to fact that our sampling trees died after this period. The fusion images were collected with a thermal camera that has a wavelength range between 7.3 and 13 µm, a standard calibration range from −20 to 650 °C, an image of 320 × 240 pixels, and a temperature error of 2% (e.g., 2% of 30 °C = ±0.6 °C). The thermal imagery was acquired between 10:00 a.m. and 12:00 p.m. (Figure A1–2). The distance between the camera and the target surface in the tree was estimated to be 1.8 m, leading to a nominal spatial resolution of ~2.64 mm per pixel.

2.2.3 Estimation of the leaf temperature

From each thermal image acquired with the FLIR T410, we estimated the T_{leaf} for the upper-middle region of the leaf blade from five leaves of trees and lianas (in trees with liana infestation). Selected leaves did not have apparent mechanical damage or evidence of herbivore attacks. The estimation of T_{leaf} from the thermal images was performed using the FLIR Tools 5.12 software (http://www.flir.com/instruments/display/?id=51975). To compute the T_{leaf}, we used the mean relative air humidity and air temperature (T_a) recorded every 15 min by a meteorological station located at the crane's structure. These meteorological station data sets were provided by the Physical Monitoring Program of the Smithsonian Tropical Research Institute (STRI). Likewise, we used the same value of T_a as a reflected temperature (or commonly known as background radiance) to compute the T_{leaf}; due to the fact that high emissivity and closed objects allow accurate temperature measurements in almost any background radiance conditions (Usamentiaga et al., 2014). In addition, to compute T_{leaf}, the emissivity of leaves for the four tree species and seven lianas species was estimated in February 2017 using the reference emissivity technique (López et al., 2012) (data unpublished). The calculation of the T_{leaf} for lianas was performed using the mean value of emissivity (0.983) estimated for this life form, while the calculation for leaves of trees was conducted using the mean value of emissivity determined for each species (*A. excelsum* = 0.976; *A. spraguei* = 0.977; *C. elastica* = 0.976; *L. seemannii* = 0.980). Following the estimation of T_{leaf}, we computed the displayed leaf temperature (T_d) for each leaf as the difference of T_{leaf} - T_a. For its nature, T_d can show positive and negative values, where positive values describe the hottest leaves and negative values the cooler leaves, according to the surrounding environment. As

mentioned in the introduction, this calculation was conducted as a parameter that can allow us to reduce the temporal variation of the ambient temperature between seasons and years. These values of T_d were used later to perform the statistical comparisons below.

2.2.4 Estimations of photosynthesis and leaf respiration

We calculated the leaf photosynthesis (P, μmol CO_2 m^{-2} s^{-1}) and leaf dark respiration (R, μmol CO_2 m^{-2} s^{-1}) to show that despite the similar physiological–temperature behavior of these life forms [7], differences in leaf temperature combined with the physiological performance can produce different estimations of productivity for these functional groups at the leaf level. These gas exchange traits were estimated using the equations and parameters published by Slot and Winter (2017) (See Equation (1)) and derived from Slot et al. (2013) (See Equation (2)) for the species of trees (excluding *A. spraguei*) and lianas of this study:

$$P = P_{\text{opt}} \times e^{-\left(\frac{T_{\text{leaf}} - T_{\text{opt}}}{\Omega}\right)^2} \tag{1}$$

$$R = R_{25} \times Q_{10}^{(T_{leaf} - 25)/10} \tag{2}$$

where P_{opt} is the maximum rate of photosynthesis at an optimum temperature (T_{opt}), Ω represents the difference in temperature between T_{opt} and the temperature in which P drops to 37% of its value at T_{opt}, R_{25} is the leaf respiration at 25 °C, Q_{10} is the proportional increase in R with a 10 °C temperature rise, and T_{leaf} is our actual measurements of leaf temperature using thermography. *A. spraguei* was excluded in this analysis due to the lack of data during the La Niña year. Because we did not identify lianas species, we estimated the gas exchange traits using the mean values reported by (Slot et al., 2013; Slot and Winter, 2017) as a functional group. Likewise, for the case of *A. excelsum,* we estimated P by the published values for Parque Nacional San Lorenzo, Panama. Using the selected leaves for our second hypothesis, we solved these equations assuming that the kinetic leaf temperature is equal to our estimation of T_{leaf}. This aims to simulate the productivity at the leaf level based on the surface canopy temperature that it is commonly used by most of the TBMs.

Although studies have reported a significant variation in photosynthesis and respiration performance of lianas and trees between seasons (Cai et al., 2009), we conducted these calculations

assuming a lack of seasonal variation in the physiological behavior. From the values of P and R extracted above, we calculated the normalized differences based on each species of host tree following:

$$\text{Normalized } P \text{ or } R = \frac{P \text{ or } R \text{ value} - P \text{ or } R \text{ mean}}{P \text{ or } R \text{ standar deviaion}} \qquad (3)$$

This latter calculation was implemented to reduce the temporal variation of leaf temperature which can produce erroneous interpretations from the comparisons of P and R across seasons and ENSO years. The resulting values from this normalization are unitless, and the magnitude of their variation can be compared between life forms, seasons, and ENSO years.

2.2.5 Data analysis

To address our hypotheses, we used linear mixed-effect models to compare the variability of the T_d according to the season, ENSO year, and: i) the presence of lianas on the T_d of trees, or ii) differences between leaf type (liana and tree) on the T_d. To test for the differences of T_d between leaf types, we only considered leaves of trees with lianas. On average, more than 240 estimations of T_d were used in each analysis. A detailed description of the sample size (number of tree individuals, thermal images, and T_d estimations) used in each analysis is shown in Tables A1–1 and A1–2. Due to the hierarchical nature of our design, linear mixed-effect models that combine fixed and random components (Zuur et al., 2009) were used to reduce the "random" factors that we cannot control in the field such as micro-climatic variations on leaves or the health of the individuals. Specifically, our model can be described by the following equation:

$$\begin{aligned} T_d = {}&\alpha + \beta_{1ij} + Season_{ij} + ENSO_{ij} + \beta_{1ij} \times Season_{ij} + \beta_{1ij} \times ENSO_{ij} \\ &+ Season_{ij} \times ENSO_{ij} + \beta_{1ij} \times Season_{ij} \times ENSO_{ij} + a_{i|j} + e_{ij} \end{aligned} \qquad (4)$$

where α represents the intercept, β_1 the presence of lianas for our first hypothesis and the leaf type for our second hypothesis, e the unexplained error, and a the random factor which is affected by each thermal image (i) nested within each individual (j). We considered each tree from each campaign of data collection as an independent sample. Likewise, we applied this same analysis to

compare the variations of the normalized differences of P and R using these parameters as response variables and β_1 as leaf type. These analyses were performed using the *nlme* package [36] under the open-source statistical software R version 3.3.1 (R Core Team, 2020). Box-Cox transformations were performed when the normality of the data was not reached. In all cases, we used the standard error (SE) as a descriptor of the sampling distribution around the mean.

2.3 Results

2.3.1 Leaf temperature of trees with and without lianas

The intra-specific comparison of T_d in each of our four tree species between leaves of trees with and without lianas suggests that the presence of lianas does not affect the T_d of host trees across seasons or ENSO years during our measurement times (Table 2–1). In general, tree species with and without lianas showed values of T_d in the range of −4.11 °C and 9.91 °C, with mean values close to 1.48 ± 0.09 °C (Figure 2–1). Regardless of the tree species, season, or ENSO year, trees without lianas showed values of T_d of 1.39 °C on average (±0.14), while trees with lianas showed values of T_d of 1.57 °C (±0.13). As such, trees with lianas were 12.94% hotter than trees without lianas based on the average value, but this difference was not significant. At the species level, the lowest values of T_d (−2.81 ± 0.10 °C) were observed for *L. seemannii* trees with lianas measured in the dry season of an El Niño year, while trees of *C. elastica* with lianas showed the highest values of T_d (7.10 ± 0.26 °C) in the wet season of the La Niña year. Across seasons and ENSO years, *A. excelsum* trees without lianas tended to have slightly higher values of T_d than trees of the same species with lianas. Conversely, during an El Niño year, trees without lianas of *A. spraguei* had marginally lower values of T_d than trees with lianas; however, both trends are not significant. Trees of *L. seemannii* and *A. excelsum* with and without lianas showed the most contrasting—albeit non-significant—trends between years, with leaves during the El Niño year showing lower values of T_d than during the La Niña year.
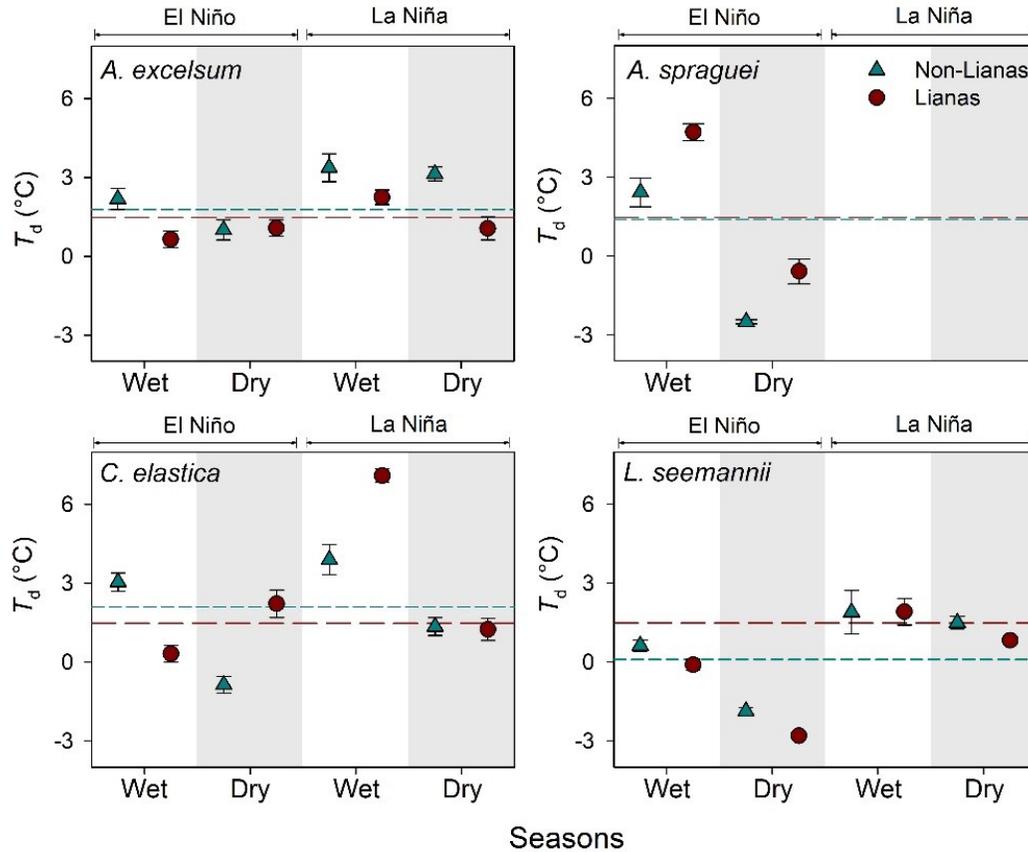
Figure 2–1. Displayed leaf temperature (T_d) of tropical tree leaves with and without lianas during the wet and dry season in contrasting ENSO years at the canopy of Parque Natural Metropolitano, Panama. Each point represents the mean (±SE). Short dashed lines represent the mean of displayed leaf temperature per tree, while long dashed lines represent the mean of all samples.

2.3.2 Leaf temperature of lianas and their host trees

The inter-specific comparisons of T_d suggest that liana leaves present higher values of T_d than leaves of their host trees (Table 2–2, Figure 2–2). In general, liana leaves showed T_d values in a range of −4.11 °C and 15.45 °C, with mean values close to 2.69 ± 0.17 °C; these values are 93.52% higher than those reported above for leaves of trees without lianas on mean values. This expression of the T_d between life forms is significantly affected by the season, where lianas tend to have higher values of T_d in the wet season in comparison with the dry season. The difference associated with each host tree suggests that for *L. seemannii*, leaves of lianas and their host have a significant interaction with ENSO, where leaves of both life forms during La Niña showed higher and

18

contrasting values of T_d with respect to an El Niño year. The effect of season, ENSO, and leaf type (tree or liana) was also observed in the host tree of *C. elastica* (Table 2–2).

Table 2–1. Effect of the liana presence, season, ENSO year, and their interaction on the displayed leaf temperature of tropical trees. Values represent *F-ratios* and values in parentheses describe the degree of freedoms; no statistical significance was found.

Factors	Species			
	A. excelsum	*A. spraguei*	*C. elastica*	*L. seemannii*
Presence	0.90	2.27	0.01	1.88
	(1, 14)	(1, 13)	(1, 13)	(1, 7)
Season	0.01	9.07	2.11	2.13
	(1, 14)	(1, 13)	(1, 13)	(1, 7)
ENSO	1.56	-	4.47	5.37
	(1, 14)		(1, 13)	(1, 7)
Presence*Season	0.01	0.02	0.60	0.01
	(1, 14)	(1, 13)	(1, 13)	(1, 7)
Presence*ENSO	0.02	-	0.09	0.01
	(1, 14)		(1, 13)	(1, 7)
Season*ENSO	0.02	-	1.14	2.45
	(1, 14)		(1, 13)	(1, 7)
Presence*Season*ENSO	0.25	-	3.55	0.12
	(1, 14)		(1, 13)	(1, 7)
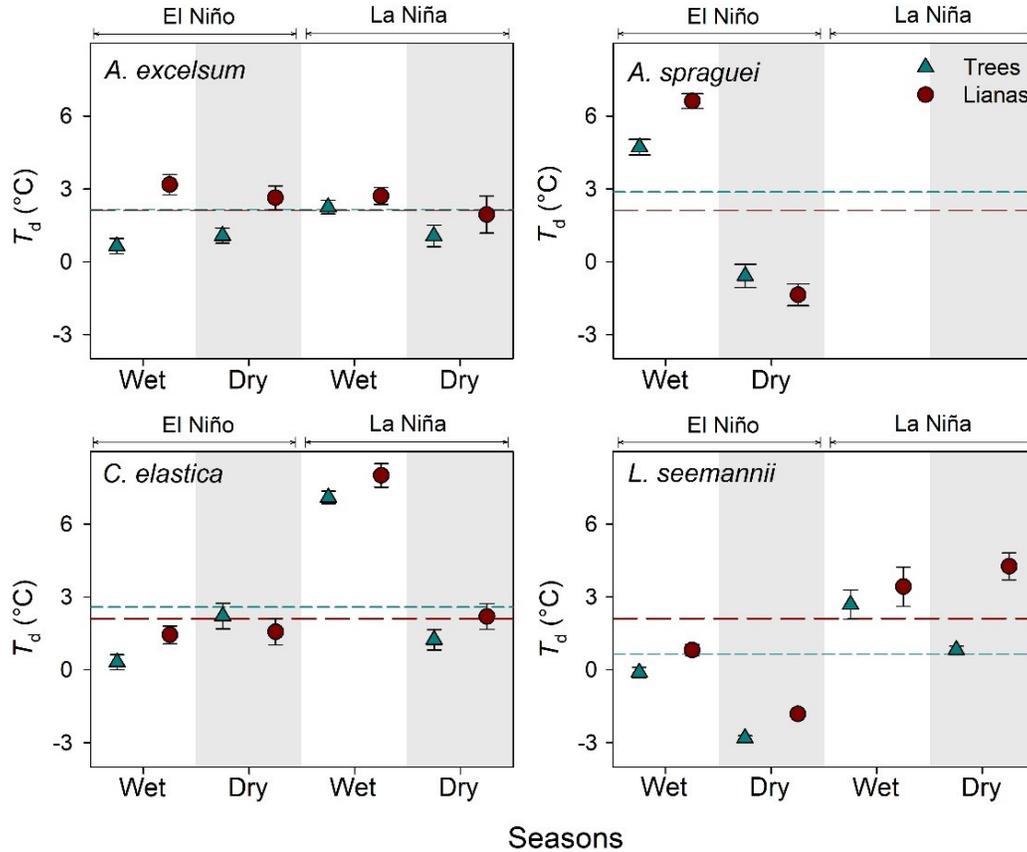
Figure 2–2. Displayed leaf temperature (T_d) of lianas and their host tree in four tree species during the wet and dry season in contrasting ENSO years at the canopy of Parque Natural Metropolitano, Panama. Each point represents the mean (±SE). Short dashed lines represent the mean of displayed leaf temperature per host tree, while long dashed lines represent the mean of all samples.

2.3.3 Comparisons of photosynthesis and leaf respiration between life forms

From the predictions of P and R using our estimations of T_{leaf}, we computed the normalized difference for each host tree in order to conduct a temporal comparison based on life forms, seasons, and ENSO years. Our results suggest that the differences in the expression of T_{leaf} in combination with the physiological performance of these life forms can produce different estimations of P and R between life forms in all host species, which can be affected (in some cases) by seasons and the ENSO (Table 2–3, Figure 2–3). These differences are more pronounced for P than R, for which leaves of lianas exhibit lower rates per unit leaf area than leaves of trees.

Table 2–2. Effect of the leaf type (tree or liana leaf), season, ENSO, and their interaction on the displayed temperature of leaves of/on host tropical trees. Values represent *F-ratios*, values in parentheses describe the degree of freedoms, and the asterisks represent the significance: * $p <$ 0.05; ** $p < 0.01$, *** $p < 0.001$.

Factors	Species			
	A. excelsum	*A. spraguei*	*C. elastica*	*L. seemannii*
Type	22.22 ***	12.96 ***	10.04 **	192.32 ***
	(1, 7)	(1, 2)	(1, 6)	(1, 4)
Season	0.02	9.06	0.40	1.12
	(1, 7)	(1, 2)	(1, 6)	(1, 4)
ENSO	0.15	-	2.46	1.87
	(1, 7)		(1, 6)	(1, 4)
Type*Season	13.60 ***	32.49 ***	5.92 *	16.22 ***
	(1, 7)	(1, 2)	(1, 6)	(1, 4)
Type*ENSO	0.69	-	2.40	7.32 **
	(1, 7)		(1, 6)	(1, 4)
Season*ENSO	0.01	-	2.46	2.40
	(1, 7)		(1, 6)	(1, 4)
Type*Season*ENSO	1.09	-	11.88 ***	0.44
	(1, 7)		(1, 6)	(1, 4)

Figure 2–3. Normalized differences of predictions of photosynthesis (*P*) and leaf dark respiration (*R*) of leaves of trees and lianas on four host species during the wet and dry season in contrasting ENSO years. Each point represents the mean (±SE). Short dashed lines represent the mean for trees, while long dashed lines represent the mean for lianas.

Table 2–3. Effect of the leaf type (tree or liana leaf), season, ENSO, and their interaction on the normalized difference of photosynthesis and leaf respiration of the values predicted. Values represent *F-ratios*, values in parentheses describe the degree of freedoms, and the asterisks represent the significance: * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$.

Factors	Species or host tree					
	Photosynthesis			Respiration		
	A. excelsum	*C. elastica*	*L. seemannii*	*A. excelsum*	*C. elastica*	*L. seemannii*
Type	1133.01 ***	11.42 ***	3432.71 ***	18.45 ***	420.99 ***	4.56 *
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)
Season	0.49	1.03	2.49	0.05	0.01	0.02
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)
ENSO	0.97	0.37	3.39	0.09	1.51	0.19
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)
Type*Season	12.33 ***	0.19	0.03	13.85 ***	5.43 *	0.15
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)
Type*ENSO	5.86 *	3.87 *	86.23 ***	0.57	0.99	39.76 ***
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)
Season*ENSO	1.69	0.41	0.34	0.17	1.29	2.71
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)
Type*Season*ENSO	4.97 *	0.01	10.61 **	0.87	20.99 ***	0.17
	(1, 7)	(1, 6)	(1, 4)	(1, 7)	(1, 6)	(1, 4)

2.4 Discussion

2.4.1 Leaf temperature of trees with and without lianas

We hypothesized that differences in T_d between trees with and without lianas due to the lianas presence could impact the availability of water –and therefore the heat dissipation by transpiration—of leaves of host trees for water competition. However, we found that the presence

of lianas on trees does not seem to affect the temperature of tree leaves. This unexpected observation could either be due to: (i) the lack of direct competition for water that may not reduce the surrounding water availability for trees (De Deurwaerder et al., 2018), or (ii) the possible compensation of more assignation of water to fully exposed leaves than leaves covered by lianas for transpiration cooling. It is important to note here that this result is based on four tree species which may not cover the whole spectrum of the liana-trees relationship. For example, Visser et al. (2018) suggest that liana infestation can have negative effects on tree population growth rates; however, this effect is more harmful to fast-growing species than slow-growing species. Therefore, we cannot exclude the possibility that lianas could have an effect on leaf temperature on trees of other life history groups, or even other sites. Likewise, it is important to note that in our study, we use the presence and absence of lianas on trees; that is why we consider that the density and extent of lianas on tree crowns may influence that host tree performance and should be examined in the future. On the other hand, the lack of a seasonal or ENSO effect on T_d could be related to the fact that trees from old secondary forests such as PNM tend to present similar sap flow during droughts and wet periods, as a result of favorable soil features and roots access to deeper soil water reservoirs (Bretfeld et al., 2018). These conditions may contribute to maintaining similar values of T_d by transpirational cooling regardless of the season or ENSO year.

2.4.2 Leaf temperature of trees and lianas

We hypothesized that leaves of lianas might have lower temperatures in comparison with their host tree due to their greater competitive advance on the acquisition, regulation, and efficient use of water as compared with trees. By contrast, our results suggest that liana leaves have considerably higher values of T_d than tree leaves. This result reinforces previous observations conducted on few leaves of three species of trees and two species of lianas using thermocouples (Sánchez-Azofeifa et al., 2011). Although the aim of this study was not to determine which biotic drivers can lead to differences in T_d between life forms, it is clear that several factors can influence T_{leaf}, such as leaf size (Leigh et al., 2017), leaf inclination (Medina et al., 1978), anatomical traits (Pérez-Estrada et al., 2000), or the presence of photo-protection pigments. Although we do not have direct measurements of leaf inclination associated with our T_d estimations, perhaps the leaf angle distribution of liana leaves on the canopy facing the sun more directly may contribute to the highest T_d. As Rey-Sánchez et al. (2016) pointed out, leaves that are facing the sun more directly

have been shown to have higher solar irradiation that strongly affects the T_{leaf}. On the other hand, the higher values of T_d of lianas could be related to the stomatal control. A recent study conducted in a botanical garden in China suggests that lianas tend to have earlier stomatal closure in comparison to trees to reduce the water transpiration close to noon (Chen et al., 2017). If this pattern occurs at the forest canopy of the PNM, it is expected that lianas would present higher values of temperature associated with the physiological regulation, which could be detected using thermography (Prytz et al., 2003). Moreover, the different T_d between life forms could be associated with the higher concentration of photo-protection pigments in trees in comparison with lianas (Sánchez-Azofeifa et al., 2009a). Currently, there is strong evidence that the presence of photo-protection pigments is associated with the efficiency of thermal dissipation of the excess of energy from light stress environments such as the canopies (Demmig-Adams, 1998; Demmig-Adams and Adams, 1992). Therefore, it could be expected that such differences in photo-protection pigments between life forms influence the efficiency of thermal dissipation and the T_d.

2.4.3 Effect of the temperature on photosynthesis and respiration of lianas and trees

Although lianas have a similar physiological–temperature response compared to trees (Slot et al., 2014a, 2013; Slot and Winter, 2017), higher displayed temperatures of lianas would have significant implications for future TBMs. Specifically, in a future scenario of modeling in which there is a lack of seasonal variation in the physiological behavior of leaves and the differences in leaf temperature between these life forms are not taken into consideration, predictions of productivity of lianas could be underestimated in comparison to trees, mainly regarding photosynthesis. Although our calculations of gas exchange are based on T_{leaf} at noon, differences in T_{leaf} through the day could have a major role in diurnal courses of productivity between life forms. A daily course of T_{leaf} of leaves of lianas and trees has been reported by Sánchez-Azofeifa et al. (2011), suggesting that liana leaves had higher overall temperatures than tree leaves. The Sánchez-Azofeifa et al. (2011) and our trends of T_{leaf} between life forms, together with their physiological sensitivity to the temperature, may lead to different expectations of productivity of lianas and trees. Therefore, future models should thus consider differences in leaf thermo-regulation between these life forms in order to achieve robust predictions of productivity.

2.5 Conclusions

Currently, there is a strong need to understand the variations of abiotic factors that can affect the forest productivity (Rogers et al., 2017). Like Slot and Winter (2017) pointed out, in the absence of a solid understanding of the abiotic controls such as temperature over physiological processes, future predictions of productivity will remain speculative. We have shown that the presence of lianas may not affect the leaf temperature of their host trees; however, lianas leaves tend to have higher values of temperature than their host trees. Our results highlight this difference in the expression of the temperature of lianas as an important biotic factor at the canopy level that can influence the forest temperature; therefore, their differential expression may have a significant weight in future predictions of forest productivity. We consider that future studies should explore the spatial variability of liana leaf temperature between forest strata such as in trees (Rey-Sánchez et al., 2016), and determine which drivers contribute to the differences in T_d. Likewise, the differential expression of leaf temperature between life forms should be considered for future studies in order to predict the productivity of ecosystems. In addition, studies should explore this differential expression of leaf temperature in other regions such as wet forest, where species present lower thermal optima of photosynthesis (Slot and Winter, 2017), in order to know if the magnitude and direction of our findings vary among environments.

CHAPTER 3

DISCRIMINATION OF LIANA AND TREE LEAVES FROM A NEOTROPICAL DRY FOREST USING VISIBLE-NEAR INFRARED AND LONGWAVE INFRARED REFLECTANCE SPECTRA

Abstract

Increases in liana abundance in tropical forests are pervasive threats to the current and future forest carbon stocks. Never before has the need been more evident for new approaches to detect the presence of liana in ecosystems, given their significance as fingerprints of global environmental change. In this study, we explore the use of longwave infrared reflectance (LWIR, 8-11µm) as a wavelength region for the classification of liana and tree leaves and compare classification results with those obtained using visible-near infrared reflectance data (VIS-NIR, 0.45-0.95 µm). Twenty sun leaves were collected from each of 14 liana species and 21 tree species located at the canopy or forest edge (*n*= 700) in Santa Rosa National Park, Costa Rica. LWIR and VIS-NIR reflectance measurements were performed on these leaves using a portable calibrated Fourier Transform Infrared Spectroscopy (FTIR) Agilent ExoScan 4100 and a UniSpec spectral analysis system, respectively. The VIS-NIR and LWIR data were first resampled. Then these two spectral libraries were pre-processed for noise reduction and spectral feature enhancement resulting in three datasets for each spectral region as follows: filtered only, filtered followed by extraction of the first derivative, and continuous wavelet transformation (CWT). Data reduction was then applied to these data sets using principal components analysis (PCA). The outputs obtained from the PCA were used to conduct the supervised classification of liana and tree leaves. In total, 21 classifiers were applied to datasets of training and testing to extract the classification accuracy and agreement for liana and tree leaves. The results suggest that the classification of leaves based on LWIR data can reach accuracy values between 66 and 96% and agreement values between 32 and 92%, regardless of the type of classifier. In contrast, the classification based on VIS-NIR data shows accuracy values between 50 and 70% and agreement values between 0.01 and 40%. The highest classification rates of liana and tree leaves were obtained from datasets pre-processed using the

CWT or from the extraction of the first derivative and classified using either random forest, *k*-nearest neighbor, or support vector machine with radial kernel. The results using the LWIR reflectance highlight the potential of this spectral region for the accurate detection of liana extent in tropical ecosystems. Future studies should consider this potential and test the regional monitoring of lianas.

Keywords

Life forms, leaf spectroscopy, spectral classification, longwave infrared, wavelet analysis.

3.1 Introduction

Lianas –woody vines– are a diverse and abundant group of plants with a pivotal role in the structure and dynamics of tropical forest (Rodríguez-Ronderos et al., 2016; Sánchez-Azofeifa et al., 2017; Schnitzer and Bongers, 2011). In general, more than ~24% of plant species richness in many tropical forests is represented by lianas (Gentry, 1991). This group of plants is considered a non-self-supporting structural parasite (See Stewart & Schnitzer 2017 discussion of their categorization) that uses host trees to reach the forest canopy (See Figure A2-1 for a graphic representation). Compared to trees, lianas tend to have a higher proportion of photosynthetic biomass per whole-plant biomass which contributes significantly to the interception of light and consequently to carbon storage (Durán et al., 2015; Rodríguez-Ronderos et al., 2016; van der Heijden et al., 2015; Wyka et al., 2013). Several studies have reported a notable increase in liana abundance in tropical and temperate environments (DeWalt et al., 2010; Londré and Schnitzer, 2006; Phillips et al., 2002; Schnitzer, 2015, 2005). Likewise, other studies have shown significant detrimental effects by lianas on tree recruitment, growth, survival, and carbon stock (Durán and Gianoli, 2013; Martínez-Izquierdo et al., 2016; Schnitzer and Carson, 2010). These trends bring into question the accuracy of productivity models and highlight the need to document the footprint of lianas and understand their role in the dynamics of ecosystems (Verbeeck and Kearsley, 2016).

To better understand the implications of lianas on ecosystems, new technologies need to be exploited that enhance the detection of this life form in the landscape. In this regard, unmanned aerial vehicles, airborne, and satellite technologies that remotely sense canopy reflectance properties may provide feasible solutions to monitor liana cover on trees over large scales. Currently, several studies have addressed the differences in optical properties of lianas and trees

from leaf and canopy observations (Asner and Martin, 2011; Avalos et al., 1999; Castro-Esau et al., 2004; Foster et al., 2008; Hesketh and Sánchez-Azofeifa, 2012; Kalacska et al., 2007; Marvin et al., 2016; Sánchez-Azofeifa et al., 2009a; Sánchez-Azofeifa and Castro-Esau, 2006). Most of these studies were conducted using reflectance spectra encompassing the visible, near, and short-wave infrared regions (VIS-NIR-SWIR, 380-2500 nm) where spectral signatures of lianas and trees are separable. In these regions, liana and tree leaves tend to differ in their reflectance near 535, 688, 985, and 2252 nm due to the lower concentration of carotenoids and chlorophyll, the lower leaf thickness, and the higher leaf water content of liana leaves (Kalacska et al., 2007; Sánchez-Azofeifa et al., 2009a). However as pointed out by Hesketh and Sánchez-Azofeifa (2012), seasonal or phenological effects on the reflectance of leaves for both life forms can have important implications for their automated classification across seasons. These effects can be associated with environmental drivers that may affect the allocation and removal of mobile compounds (i.e., pigments) and the water content of leaves.

The longwave infrared spectrum (LWIR, 8-14 µm) has not been explored for the separation of liana and tree leaves. LWIR investigations of trees and their leaves have shown that the spectra of leaves are controlled primarily by structural compounds at the leaf surface (Harrison et al., 2018; Ribeiro da Luz and Crowley, 2007). These compounds include cellulose, hemicellulose, cutin, silica, and terpenes imparting reflectance features that enhance the separability of species (Buitrago et al., 2018b; Harrison et al., 2018; Meerdink et al., 2016; Ribeiro da Luz, 2006; Ribeiro da Luz and Crowley, 2010, 2007; Ullah et al., 2012a). This spectral region may be characterized by lower temporal variability of leaf spectral signatures as suggested from limited data by the seminal work of Salisbury (1986) comparing LWIR signatures of green leaves from late autumn and senescent leaves. This body of work suggests that LWIR signatures may not be significantly affected by seasonal or phenological effects. If these results apply to a wide range of biomes, detection in the LWIR would present a unique remote sensing advantage. An additional advantage of working in this region for the discrimination of lianas and trees rests on the observation that these life forms present contrasting leaf concentrations of lignin, cellulose, and hemicellulose (Asner and Martin, 2012, 2011) that may be detected in LWIR spectra of plants.

This study assesses the use of the VIS-NIR and LWIR spectral region for the discrimination of liana and tree leaves. We hypothesize that LWIR reflectance spectra of leaves may provide better discrimination of the two life forms. Because the pre-processing of spectral data may affect the

spectral classification and its role in enhancing spectral features has been shown to improve spectral separability (Harrison et al., 2018; Rivard et al., 2008), this study also evaluates the impact of three common pre-processing methods in improving the classification of the two life forms. In addition, we evaluate 21 supervised-classifiers for the discrimination of liana and tree leaves. We do so to explore the quality of discrimination of some classifiers or "classifier families" in a field increasingly driven by data classification (Fernández-Delgado et al., 2014). We do not aim to identify the "best" classifier; our goal is to provide a broad perspective of considerations for future studies involving the classification of spectral libraries. This study highlights the LWIR region for future detection of these life forms in forests with the future aim to further understand the current role of lianas in ecosystem functioning.

3.2 Materials and Methods

3.2.1 Study site and sample collection

This study was conducted in the Santa Rosa National Park (SRNP, 10º48" N, 85º36" W) located on the Pacific coast of northwestern Costa Rica. The SRNP is in a tropical dry forest (Sánchez-Azofeifa et al., 2005). This site presents a wet-season extending from the middle of May to late November while a dry-season, during which most of the trees lose their leaves, encompasses the remaining months (Kalacska et al., 2004). This site has an air temperature that varies from 26 °C in the wet-season to 29 °C in the dry season, and has a mean annual precipitation of 1720 mm (Kalacska et al., 2004). The SRNP is composed of a mosaic of forest patches in different successional stages of natural regeneration and with different land-use histories associated with anthropogenic fires, deforestation, and land clearing for pasture and agriculture (Arroyo-Mora et al., 2005; Calvo-Alvarado et al., 2009; Sánchez-Azofeifa et al., 2017). The SRNP has 96 species of trees of different life history (Hilje et al., 2015) and approximately ~18 species of lianas that can reach the forest canopy. The abundance and height of trees in the canopy depends on the successional stage of the forest patches (Hilje et al., 2015; Li et al., 2017).

Fully exposed sun leaves of 14 species of lianas and 21 species of trees (Table 3–1) were collected during the wet-season of 2017 between the months of May and July. Leaves were sampled in the forest canopy or at the forest edge using an extension pruner. For each species, five healthy and mature leaves were collected from each of four individuals for a total of 700 leaves. The number of individuals and leaves selected were based on the variability of the LWIR

reflectance of tree leaves in prior research at our study area (Harrison et al., 2018) and the spatial distribution of individuals with low abundance. Leaves were then placed in moist paper and immediately stored in sealed plastic bags that were placed in a cooler and taken to the laboratory for collection of VIS-NIR and LWIR reflectance spectra within two hours. Despite the importance of the shortwave infrared region (SWIR, 1.1–2.6 μm) for the discrimination of liana and tree leaves (Kalacska et al., 2007), we do not conduct measurements in the SWIR region because we did not have a spectrometer with that spectral range at the time we were in the field.

3.2.2 Measurements of VIS-NIR reflectance spectra

The collection of leaf spectral reflectance measurements in the VIS-NIR followed previously defined protocols of data collection described by Castro-Esau et al. (2004) and Kalacska et al. (2007). Measurements were conducted using a UniSpec Spectral Analysis System (PP Systems, Amesbury, MA, USA). A leaf clip that holds a foreoptic provides illumination (7.0 W halogen bulb) to a leaf area of 4.15 mm^2. For each leaf, three measurements were taken at the middle leaf lamina avoiding the midrib and these were later averaged. Each measurement was the average of ~ 40-60 scans for the purpose of noise reduction. This data collected consists of reflectance for 256 bands spanning a spectral range of 306-1138 nm, each with a bandwidth of 3.3 nm.

3.2.3 Measurements of LWIR spectral reflectance

The collection of leaf spectral reflectance measurements in the LWIR was conducted using a portable Agilent 4100 ExoScan Fourier Transform Infra-Red (FTIR) spectrometer following a protocol previously described by Harrison et al. (2018). This spectrometer is equipped with an internal IR illumination source and we made use of a diffuse reflectance probe to which a diffuse infra-gold reference cap can be attached for background collection. The illuminated area for this probe has a diameter of 1.5 cm and a maximum depth of light penetration of 20-50 μm, depending on the medium. The FTIR probe was brought into contact with the leaf for the same leaf region as measured for collection of VIS-NIR spectra. Each reflectance measurement was the average of 150 scans for the purpose of noise reduction and was obtained following a measurement of the infra-gold background. The data collected consists of reflectance for 1799 bands spanning a spectral range from 2.5 to 15.4 μm (4000-650 cm^{-1}) with a resolution of 4 cm^{-1}.

Table 3–1. Liana and tree species sampled at the Santa Rosa National Park, Costa Rica.

Trees		Lianas	
Family	Species	Family	Species
Apocynaceae	*Stemmadenia obovata*	Apocynaceae	*Forsteronia sp.*
Bignoniaceae	*Crescentia alata*		*Forsteronia spicata*
Burseraceae	*Bursera simarouba*	Bignoniaceae	*Arrabidaea chica*
Dilleniaceae	*Curatella americana*		*Cydista aequinoctialis*
Euphorbiaceae	*Jatropha curcas*		*Cydista diversifolia*
	Sapium glandulosum		*Paulinia sp.*
Fabaceae/Caes	*Bauhinia ungulata*	Cucurbitaceae	*Cayaponia racemosa*
	Hymenaea courbaril	Dilleniaceae	*Tetracera volubilis*
Fabaceae/Pap	*Gliricidia sepium*	Malpighiaceae	*Heteropterys panamensis*
Fagaceae	*Quercus oleoides*		*Heteropterys sp.*
Hippocrateaceae	*Semialarium mexicanum*		*Hiraea reclinata*
Lauraceae	*Ocotea veraguensis*	Rhamnaceae	*Gouania polygama*
Malpighiaceae	*Byrsonima crassifolia*	Sapindaceae	*Serjania atrolineata*
Malvaceae	*Guazuma ulmifolia*		*Serjania schiedeana*
Meliaceae	*Cedrela odorata*		
	Trichilia americana		
Nyctaginaceae	*Pisonia aculeata*		
Sapindaceae	*Cochlospermum vitifolium*		
Simaroubaceae	*Simarouba glauca*		
Tiliaceae	*Luehea speciosa*		
Verbenaceae	*Rehdera trinervis*		

3.2.4 Spectral analysis

The VIS-IR and LWIR reflectance measurements were treated as independent datasets and processed in the following five steps: i) data resampling, ii) pre-processing, iii) data reduction, iv) training of classifiers, and v) testing of classifiers.

3.2.4.1 Resampling and pre-processing

The VIS-NIR and LWIR data were resampled to a bandwidth of 1 nm and 10 nm, respectively using the simple linear interpolation function '*resample*' in the *prospectr* package (Stevens and Ramirez-Lopez, 2013; version 0.1.3) of the R software version 3.4.1 (R Core Team, 2020). Each of the two datasets was then pre-processed for noise reduction and spectral feature enhancement resulting in three datasets for each spectral region as follows: filtered only, filtered followed by extraction of the first derivative, and continuous wavelet transformation (CWT). For filtering, a Savitzky-Golay smoothing filter was applied using a polynomial quadratic order and a 25-point window. The Savitzky-Golay filter was selected as it tends to preserve features of the initial spectrum such as maxima, minima that can be modified with other averaging filters (Ruffin et al., 2008). The first derivative was computed as a descriptor of spectral shifts. Filtering and derivatives were computed using the '*sgolayfilt*' function in the *signal* package of R (Ligges et al., 2014). A CWT was used as a multi-scale analysis for the extraction of absorption features (Rivard et al., 2008). In CWT each reflectance spectrum is represented as a sum of wavelets, each capturing spectral features of different scales. The CWT was computed based on a second-order derivative of Gaussian with a variance of 1 and a range of scales between 1 and 9. After the extraction of wavelets at these scales, wavelet spectra are summed between scales of 2 and 5. The features captured in wavelet spectra at these scales provide the most visible discrimination between life forms (Figure A2–2 and Figure A2–3). The CWT transformation was conducted using the '*wavCWT*' function in the *wmtsa* package of R (Constantine and Percival, 2017).

Ensuing analysis examining the effect of resampling on classification was conducted on the initial two datasets that were pre-processed and the three respective outputs for each spectral region for a total of 8 datasets. The spectral range for analysis was 450 - 950 nm (700 bands) for the VIS-NIR datasets and 8 - 11 µm (301 bands) for the LWIR datasets.

3.2.4.2 Data reduction and training datasets for classification

For each of the eight datasets a principal component analysis (PCA) was used as a data reduction technique to reduce the redundancy of some spectral bands, and at the same time highlight those uncorrelated bands that describe the maximum variability in reflectance signatures. The PCA applied was based on a covariance matrix. The number of components retained explained 95% of the total variance and these were later used as an input for the classification. Likewise, the eigenvectors of each band were used as descriptors of the loadings of bands. The PCA was computed using the '*pricomp*' function of R.

The scores resulting from the principal components analysis were randomly split in two for training and testing the classifiers. The data split was conducted based on species and not life forms. This procedure allows each species to have a similar representation of samples in both the training and testing datasets, which would not necessarily occur if the data were split based on life forms. This method of splitting was conducted using the '*createDataPartition*' function of the *caret* package in R (Kuhn, 2008). At the end of this procedure, datasets encompassed 140 samples of lianas and 210 samples of trees. This unbalanced distribution of the categorical predictors is statistically detrimental for the training classification process (Kuhn and Johnson, 2013). Therefore, the training datasets were then balanced using the Synthetic Minority Over-Sampling Technique (SMOTE) (Chawla et al., 2002). This technique over-samples the minor class (in our case lianas) and adds synthetic samples linking any/all of the *k* minority class nearest neighbors using a distance measure (Chawla et al., 2002). To apply this technique, five nearest neighbors and 200% of over- and under-samples were used. This procedure was conducted using the 'SMOTE' function of the *DMwR* package of R (Torgo, 2010).

3.2.4.3 Lianas and trees classification

Different algorithms were evaluated for the classification of life forms and to assess if LWIR data would provide better results than VIS-NIR data. Twenty-one classifiers were used (Table 2) and these were grouped according to their similarities (Figure A2–4). In general, these algorithms were selected because they require little or no tuning parameters for their computation; and because some have been widely used in studies that address the classification of lianas and trees. The classifiers were implemented using the *caret* package in R (Kuhn, 2008); a single computing environment that allows the training of each classifier, the tuning of parameters, and the testing of

predictions. During the training of each classifier, a 0.632 bootstrap technique based on 100 iterations was used to create and optimize the algorithm. This technique consists of producing a random sample of data taken with replacement, where the final bootstrap sample size is the same as the original dataset. As Kuhn and Johnson (2013) pointed out, during this process some of the original samples can be selected several times in the bootstrap, while others are not selected; commonly known as 'out-of-bag' samples. Therefore, for a given iteration the model will be optimized based on the prediction of out-of-bag samples. With this technique, it is expected that the probability of any given sample not being chosen after n samples is 0.368 and the number of chosen samples from the original samples is $0.632n$. Therefore, the model is optimized based on the remaining $0.368n$. Likewise, in certain classifiers there is a need to use tuning parameters that help to reduce its complexity and reinforce its repeatability (Ghosh et al., 2014); the tuning parameters used for those classifiers are described in Table 3–2.

Table 3–2. List of classifiers and corresponding tuning parameters used in this study.

Classifier	Abbreviation	Function	Library	Tuning parameters
Linear Discriminant Analysis	LDA	lda	MASS[1]	---
Quadratic Discriminant Analysis	QDA	qda	MASS[1]	---
Maximum Uncertainty Linear Discriminant Analysis	MULDA	Mlda	HiDimDA[2]	---
Stabilized Linear Discriminant Analysis	SLDA	slda	ipred[3]	---
Robust Linear Discriminant Analysis	RLDA	Linda	rrcov[4]	---
Robust Quadratic Discriminant Analysis	RQDA	QdaCov	rrcov[4]	---
Regularized Discriminant Analysis	RDA	rlda	klaR[5]	gamma = 0, lambda= 0

Robust Mixture Discriminant Analysis	RMDA	rmda	robustDA[6]	k = 2, model= VEV
Shrinkage Discriminant Analysis	SDA	sda	sda[7]	diagonal= F, lambda= 0
Generalized Linear Model	GLM	glm	---	---
Bayesian Generalized Linear Model	BGLM	bayesglm	arm[8]	---
Neural Network	NN	nnet	nnet[1]	size= 5, decay= 1×10^{-4}
Bagged CART	BC	treebag	e1071[9]	---
CART	CART	rpart1SE	rpart[10]	---
Random Forest	RF	rf	randomForest[11]	mtry=2
Single C5.0 Tree	ST	C5.0Tree	C50[12]	---
Single C5.0 Ruleset	SR	C5.0Rules	C50[12]	---
Support Vector Machines with Linear Kernel	SVMLK	svmLinear	kernlab[13]	Cost = 1
Support Vector Machines with Polynomial Kernel	SVMPK	svmPoly	kernlab[13]	Degree = 3, scale = 0.1, Cost = 1
Support Vector Machines with Radial Basis Function Kernel	SVMRK	svmRadial	kernlab[13]	Sigma= 0.045, Cost = 1
k-Nearest Neighbors	KNN	knn	---	*k* = 5

[1]Venables and Ripley (2002), [2]Duarte (2015), [3]Peters and Torsten (2017), [4]Todorov and Filzmoser (2009), [5]Weihs et al. (2005), [6]Bouveyron and Girard (2015), [7]Ahdesmaki et al. (2015), [8]Gelman and Su (2016), [9]Meyer et al. (2017), [10]Therneau et al. (2017), [11]Liaw and Wiener (2002), [12]Kuhn et al. (2015), [13]Karatzoglou et al. (2004).

The classifiers computed during the training were validated using the testing datasets. The validation process includes the computation of confusion matrixes to extract four descriptors: i) the accuracy describing the agreement between the observed and predicted classes; ii) the cohen's kappa coefficient representing the agreement between the accuracy observed and expected; iii) the

sensitivity that describes the *true positive* since it measures the probability in which a class (i.e., lianas or trees) of interest is predicted correctly for all samples of that class; and iv) the specificity that represents the 1- *false-positive* since it measures the probability in which non-event classes are predicted as non-event classes. These parameters were calculated using the '*confusionMatrix*' function of the *caret* package in R (Kuhn, 2008).

3. Results

3.1 Dataset evaluation

The eight datasets of average spectra per species are shown in Figure 3–1. Liana and tree leaf spectra are similar for all datasets of the VIS-NIR region but are less so for the datasets of the LWIR region. In the LWIR region, the degree of overlap of leaf and liana spectral signatures depends on the pre-processing method. From this perspective, the first derivative and CWT datasets provide a better discrimination of features (e.g., reflectance peaks and troughs) that distinguish leaves and lianas as can be seen between 8.5 and 8.9 µm in the CWT dataset.
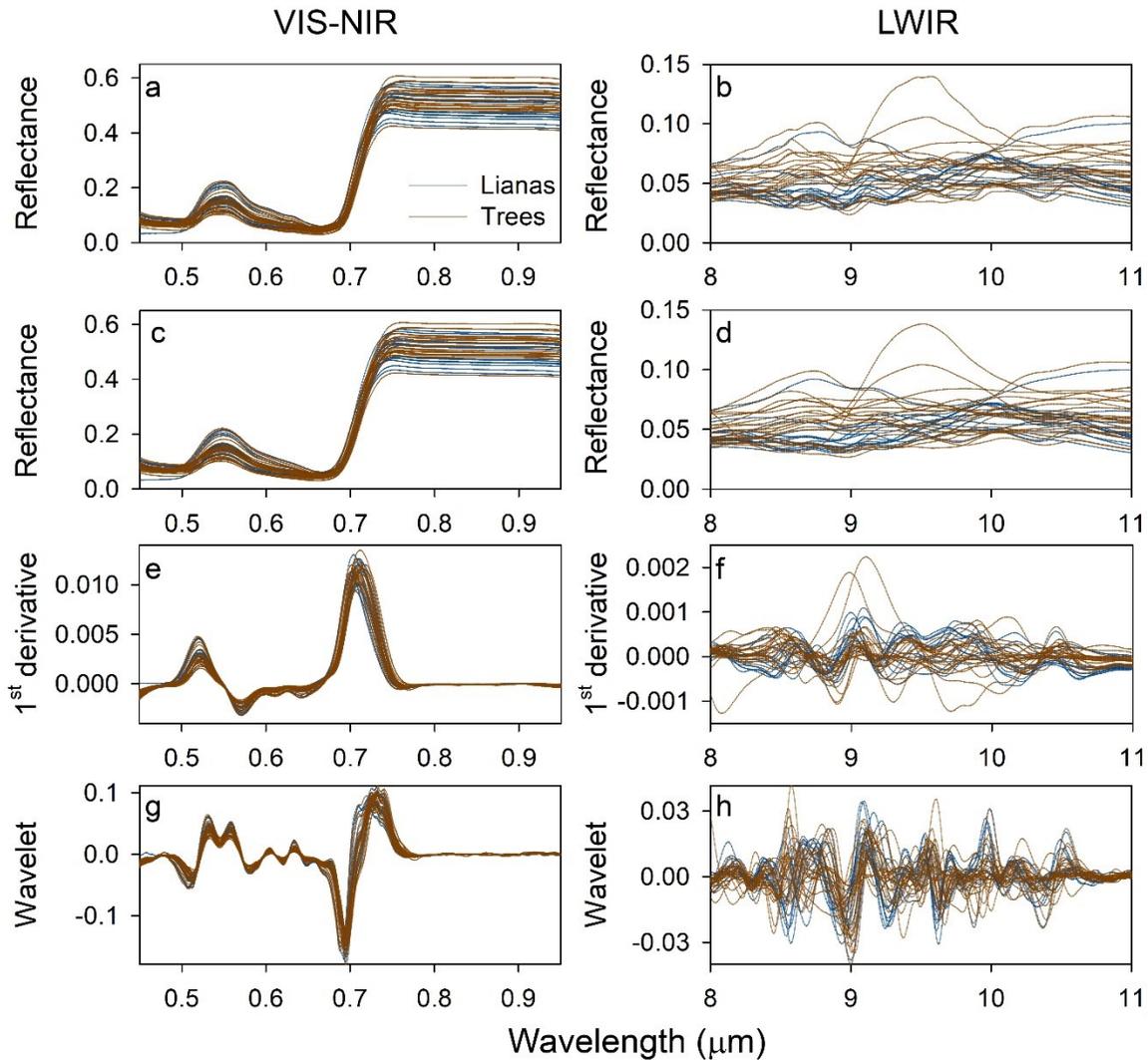
Figure 3–1. Visible-near infrared (VIS-NIR) and long wave infrared (LWIR) reflectance spectra of liana and tree leaves and their pre-processed equivalents. Unmodified raw spectra are shown in (a, b). Pre-processed spectra include filtered (c, d), first derivative (e, f), and continuous wavelet transformation (CWT) (g, h). Each spectrum represents the average for each species.

3.2 Principal component analysis

The number of components extracted from the PCA analysis that explains 95% of the total variance tends to differ according to the spectral region and the associated datasets (Table 3–3). Two components are required for the raw and filtered datasets in the VIS-NIR. That number increases to three for the first derivative and CWT datasets in the VIS-NIR and the raw and filtered datasets in the LWIR. In contrast, 8 and 17 components are required for the first derivative and

CWT datasets in the LWIR region. As expected, the percentage of the variance explained by the first component is higher than other components. However, the percentage of the variance explained by the first component as compared to that of the remaining components, is higher for the VIS-NIR datasets than the LWIR datasets. The scatterplot of the scores of the first two principal components for the lianas and the trees shows that scores for the VIS-NIR datasets tend to have greater overlap than the scores for the LWIR datasets (Figure 3–2). Likewise, the degree of overlap of the scores for these two classes in these scatterplots is impacted by the pre-processing, where scores for the CWT and first derivative datasets show less overlap than scores for the raw and filtered datasets.

The eigenvector loadings for the VIS-NIR datasets reveal that regions close to 0.55, 0.68, and 0.71 µm present an important contribution to the variance explained by the PCAs (Figure 3–3). This contribution is particularly apparent for the first derivative and CWT datasets (Figure 3–3e, g). On the other hand, the loadings of the eigenvectors of the LWIR datasets reveal several spectral regions that particularly contribute to the data variance. In general, the loadings of the first component for the raw and filtered LWIR datasets (Figure 3–3b, d) were of uniform and low value throughout the LWIR spectrum, but loadings extracted from the second and third component show a predominant contribution to the variance explained in regions close to 9.50 and 10.15 µm, respectively. In contrast, loadings extracted from the first derivative and CWT LWIR datasets (Figure 3–3f, h) shows numerous regions with a large contribution to the variance explained throughout the LWIR spectral region and throughout eigenvectors.

3.3 Classifiers

The classification of spectra of liana and tree leaves shows that the values of accuracy, kappa, sensitivity, and specificity are higher in classifications based on LWIR datasets than on VIS-NIR datasets (Figure 3–4). In general, the accuracy values based on LWIR datasets range from 0.66 to 0.96 (0.77 ± 0.08), while the accuracy values based on VIS-NIR datasets range from 0.50 to 0.69 (0.61 ± 0.04). The kappa values range from 0.33 to 0.92 (0.52 ± 0.16) for LWIR datasets, and from 0.01 to 0.37 (0.17 ± 0.08) for VIS-NIR datasets. In terms of sensitivity and specificity, values range from 0.70 ± 0.13 and 0.81 ± 0.09, respectively for LWIR datasets, and from 0.44 ± 0.12 and 0.73 ± 0.11 respectively for VIS-NIR datasets (Figure 3–4).

Table 3–3. Variation explained by components of the principal component analysis applied to the raw and pre-processed spectral libraries of the visible-near infrared (VIS-NIR) and longwave infrared (LWIR) leaf spectra.

Principal component	VIS-NIR				LWIR			
	Raw	SG	D	CWT	Raw	SG	D	CWT
PC1	0.79	0.80	0.71	0.80	0.70	0.71	0.40	0.401
PC2	0.19	0.18	0.23	0.13	0.17	0.17	0.29	0.222
PC3			0.03	0.04	0.09	0.08	0.11	0.098
PC4							0.06	0.055
PC5							0.04	0.048
PC6							0.03	0.035
PC7							0.02	0.022
PC8							0.01	0.015
PC9								0.011
PC10								0.009
PC11								0.008
PC12								0.007
PC13								0.006
PC14								0.005
PC15								0.004
PC16								0.003
PC17								0.003
Total variance explained	0.98	0.98	0.97	0.97	0.96	0.96	0.96	0.952
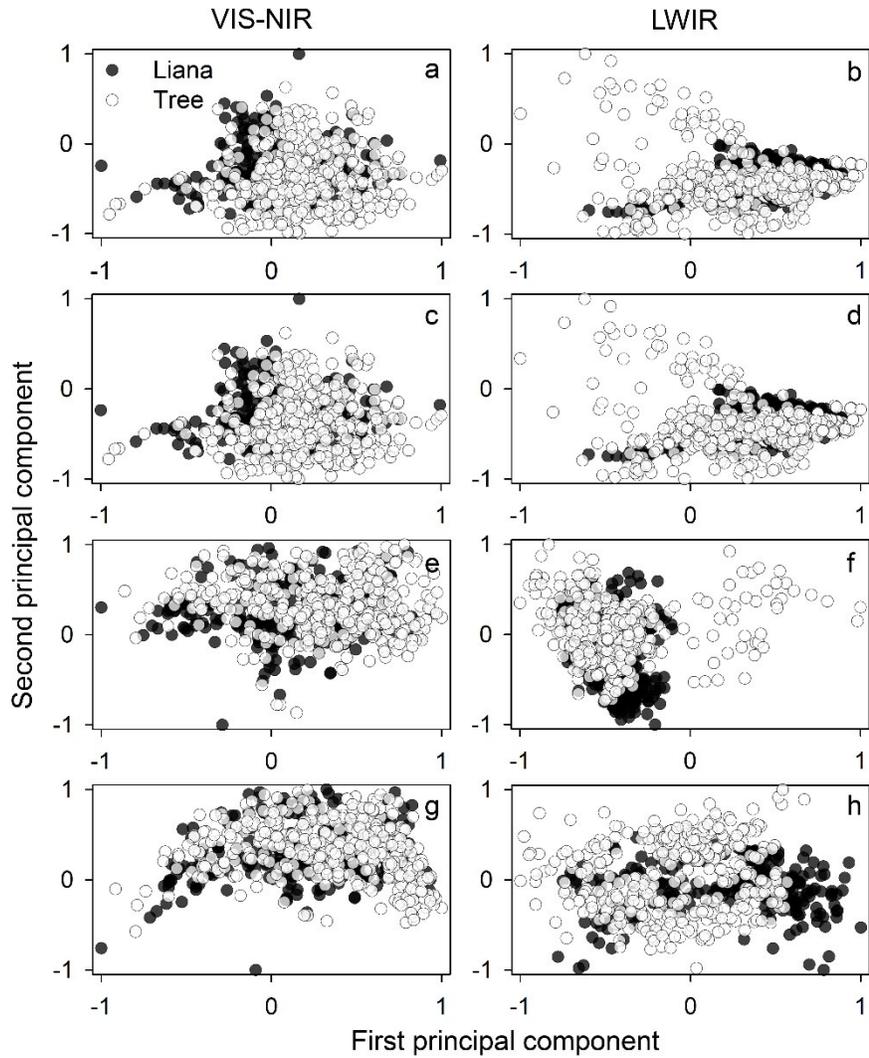
Figure 3–2. Scatterplots of the first two principal components scores obtained from the analysis of the spectral signatures of lianas and trees in the visible-near infrared (VIS-NIR) and longwave infrared (LWIR). Raw data (a, b), and pre-processed data filtered (c, d), first derivative (e, f), and continuous wavelet transformation (CWT) (g, h). Each point represents data from a single leaf spectrum.
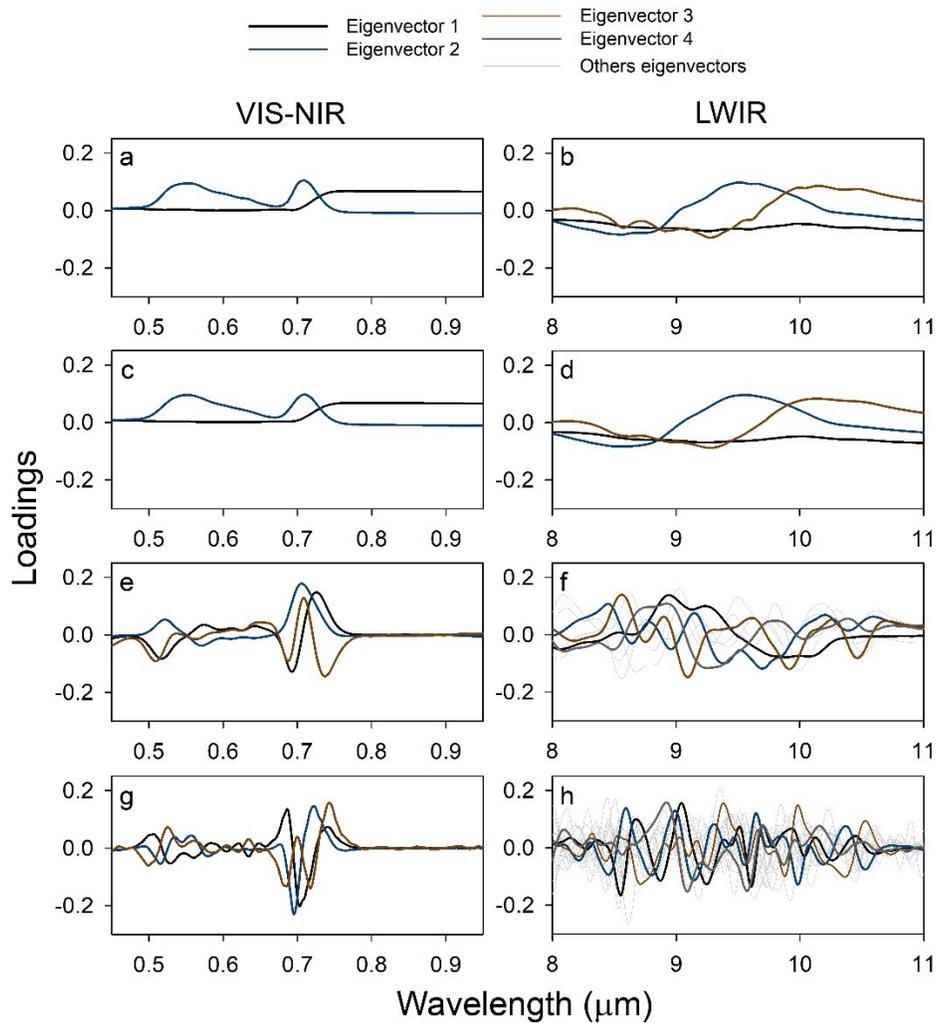
Figure 3–3. Eigenvectors loadings capturing 95% of the total variance explained by the principal components analysis conducted on the visible-near infrared (VIS-NIR) and longwave infrared (LWIR) region. Raw data (a and b), and pre-processed filtered (c and d), first derivative (e and f), and continuous wavelet transformation (g and h).

The classification results reveal that the first derivative and CWT datasets generally have higher values of accuracy and kappa than raw and filtered datasets (Figure 3–4). These findings are consistent for the VIS-NIR and LWIR for any given classifier and for the best performing classifiers.

In terms of the classifiers used, the results show that the selection of classifiers or families or classifiers can affect the overall classification of lianas and tree leaves using VIS-NIR or LWIR datasets. Classifiers associated with discriminant analysis or generalized linear models tend to have

lower values of discrimination than classifiers related to non-parametric models, decision trees or non-linear models. From this perspective, the results show that the three classifiers with highest accuracy and kappa values are RR, SVMRK, and KNN, while classifiers with lowest accuracy and kappa values are RMDA, QDA, and SLDA.
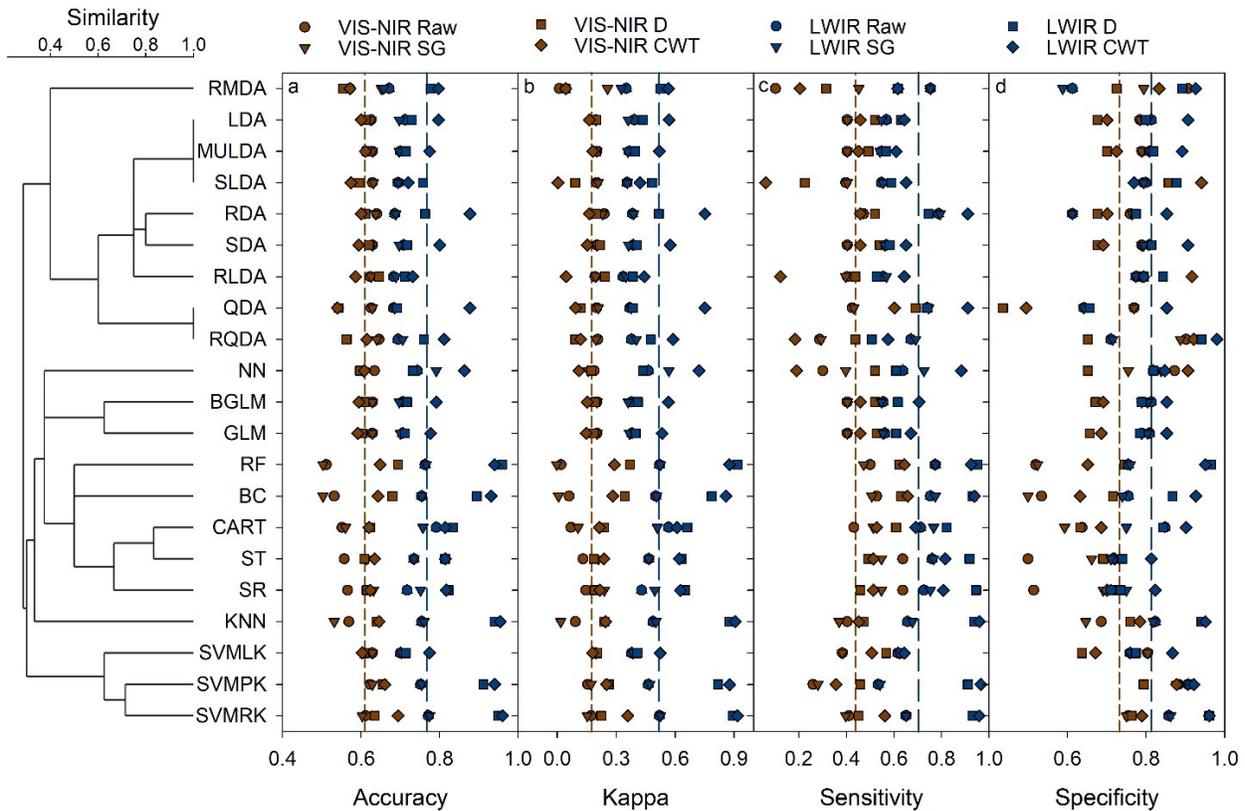
Figure 3–4. Descriptors of the classification of liana and tree leaf spectra based on the scores extracted from the principal component analysis of the visible-near infrared (VIS-NIR, brown symbols) and long wave infrared (LWIR, blue symbols) spectra. Descriptors of classification include accuracy, Kappa coefficient, sensitivity, and specificity. Raw data (circles), filtered (inverted triangles), first derivative (squares), and continuous wavelet transformation (CWT, lozenges). Vertical dotted and dashed lines for VIS-NIR (brown) and LWIR (blue) represent the average classification value across all methods. Refer to Table 2 for abbreviations for classifiers.

3.4 Discussion

The results of this study reveal the potential for improved discrimination of leaves of lianas and trees using spectral observations in the LWIR as compared to that collected in the VIS-NIR. This

potential can be described by the overall values of classifications using a wide range of classifiers computed on raw and pre-processed datasets and principal components analysis. For the two life forms of this study, the pre-processing method plays an essential role in the discrimination of leaves. Likewise, the choice of classifiers or families of classifiers affects the discrimination.

3.4.1 Influence of pre-processing on data reduction and classification

In general, the first derivative and CWT pre-processing methods are more effective for the classification of liana and tree leaves than the raw or filter spectral pre-processing methods. Previous studies have described the use of transformations, such as the first derivative and CWT methods, to highlight spectral features in the comparison of spectral libraries, to predict leaf functional traits, to detect vegetation damage, or to classify species (Cheng et al., 2012, 2010; Harrison et al., 2018; Rivard et al., 2008; Ullah et al., 2012b). This study compared these four pre-processing methods for the classification of leaves and revealed that they seem to be particularly crucial for classification due to the similarity of the reflectance spectra of these life forms. This similarity hinders the classification of liana and tree leaves in raw and filtered datasets because the classification of these data puts an emphasis on differences in the amplitude of reflectance rather than differences in spectral shape between these life forms. For classification purposes, the use of raw or filtered datasets could be promising for the discrimination of spectral libraries encompassing species with contrasting spectral characteristics. For example, in a similar study on nineteen species of herbaceous and woody plants, Buitrago et al. (2018) showed that it is possible to discriminate these species using few components extracted from a PCA applied to a raw spectra library (1.4 – 16.6 µm). However, for species with similar spectral characteristics, here that coexists together and share similar phylogenies, such as lianas and trees the latter approaches may not be suitable for classification. As shown in PCA applied on the first derivative or CWT LWIR datasets of this study, several subsets of the LWIR regions are highlighted that contribute to the discrimination of these life forms. The higher number of components needed to explain the data variance may be of concern for future classifications in regard to processing time, but should ultimately provide less redundant information that enhances the discrimination of these life forms. Future studies should consider how the spectral pre-processing methods could contribute to the prediction of leaf functional traits using leaf spectroscopy.

3.4.2 Effect of the classifier on life forms discrimination

To date the literature offers little insight on the impact of the selection of the classifier on the spectral discrimination of species, group of plants, or life forms. An exception is the study of Castro-Esau et al. (2004) that compared the effect of parametric and non-parametric classifiers on the discrimination of VIS-NIR leaf spectral libraries for lianas and trees. Our trends and their results suggest that non-parametric or non-linear classifiers perform better (i.e., higher accuracy and kappa) than parametric algorithms. In general, differences in the performance of classifiers for discrimination can be attributed to several reasons such as the assumptions of each algorithm, the selection of tuning parameters, data dimensionality, and the sample size (Belgiu and Drăguţ, 2016; James et al., 2013; Kuhn and Johnson, 2013; Shao and Lunetta, 2012). For example, in this study the family of discriminant analysis showed lower values of discrimination than other classifiers. Methods of discriminant analysis often assume that the observations have a Gaussian distribution (James et al., 2013), which may not be the case for our spectral data. In addition, these classifiers are associated with techniques of data reduction that maximize the differences between classes (Kuhn and Johnson, 2013), differences that could be diminished after the application of the PCA.

Our results also suggest that classification tree algorithms are part of the most promising classifiers to discriminate leaves of both life forms. Specifically, the high performance of the RF algorithm is also observed by Fernández-Delgado et al. (2014) in their comparison of classifiers applied to a diverse group datasets, suggesting that the RF classifier is part of the "best" family of algorithms for classification. However, it is important to highlight that this classifier could be susceptible to the selection of tuning parameters depending on the sample size and to the spectral variability of the samples (temporal and spatial). For example, in a comparative study of classifiers applied to the diagnosis of cancer, Statnikov et al. (2008) found that large RF datasets tend to be more sensitive to the selection of the tuning parameters than other classifiers such as SVMs. Though RF or classification tree algorithms are less computationally intensive than SVMs (Belgiu and Drăguţ, 2016; Cutler et al., 2007), their possible bias using large data sets should be considered in future studies of classification of spectral libraries or imaging spectroscopy data.

Our results reveal that NN and SVMs algorithms tend to present high values of discrimination of lianas and tree leaves. An advantage of using these classifiers is usually associated with limited requirements on training size benefiting the application to spectral libraries with a low number of spectral samples. However, their application tends to be computationally intensive and the

selection of tuning parameters is usually determined empirically that could be a concern for the reproducibility of studies. This last observation would also apply to other classifiers such as KNN, RF, SDA, RMDA, and RDA to name a few.

As the "No Free Lunch" theorem suggest (Wolpert, 1996; Wolpert and Macready, 1997), in the absence of clearly established knowledge, there is no classification algorithm that will always provide a better discrimination for a specific problem. Considering this, there is a need to report other parameters of classification to conduct an informative representation of the errors and cost associated with the discrimination of spectral data. This is mainly because studies traditionally report the accuracy and kappa values, but the first do not differentiate between the type of errors that have been created (Provost et al., 1998) and the second provide redundant or misleading information for decision making (Pontius and Millones, 2011). The use of parameters such as the sensitivity, specificity, indexes (i.e., Younden's Index) or Receiver Operating Characteristic curves (ROC) could be useful for the classification of spectral libraries as they can provide an overall view of the bias of discrimination that cannot be captured by only using accuracy or kappa values (Kuhn and Johnson, 2013). For example, our accuracy and kappa results for the VIS-NIR classification using SVMLK or NN may suggest that the pre-processing does not benefit the discrimination of leaves of life forms. However, using the values of sensitivity or specificity our results reveal that the pre-processing may benefit the classification of liana and tree leaves. Because of this, we recomment that future studies of classification of spectral libraries provide a multidimensional overview of their findings to support future strategies.

3.4.3 Future perspectives for the remotely sensed detection of life forms

The possibility to discriminate life forms based on leaf reflectance measurements in the VIS-NIR or LWIR region is an important step for future air- or space-borne mapping. Here, we show encouraging leaf-based results that could inspire future studies for the detection of life forms at large-scales. Currently, the detection of species and life forms using high spatial resolution hyperspectral airborne imagery has been applied with success (Marvin et al., 2016; Ribeiro da Luz and Crowley, 2010; Zhang et al., 2006). However, as pointed out in prior papers particularly in regards to the LWIR (Ribeiro da Luz and Crowley 2010, Harrison et al. 2018), the detection of species or life forms at large-scales requires considerations associated with spectral contrast, spatial resolution, image calibration, and improvements in atmospheric corrections. Preservation

of spectral contrast will be favoured by finer spatial resolution that will also minimize spectral mixing. In this regard surveys with pixels <5m and preferably closer to 1m will facilitate the spectral analysis with several pixels encompassing mature crowns though this level of detail will be at the expense of costs for regional coverage. In their 2010 study, Ribeiro da Luz and Crowley successfully used LWIR imagery at this scale of resolution to discriminate tree species in an arboretum in Virginia. As seen in their study we would expect that the detection of life forms in our investigation could be attempted in the future first on the basis of match filtering using spectral libraries of lianas and trees. Such libraries could also be used to label image based endmembers automatically or manually collected to identify the signature of the life forms. Regardless of the method used, future surveys should consider the synchronicity of the phenology between lianas (Kalácska et al., 2005) to insure maximum top of canopy detectability. Future or current airborne capabilities such as the Hyperspectral Thermal Emission Spectrometer (HyTES) that operates in the LWIR would be ideal tools to initiate this research and explore the detection of life forms in tropical ecosystems.

3.5 Conclusion

This study reveals the potential of the LWIR over the VIS-NIR reflectance to discriminate liana and tree leaves. This work expands previous LWIR studies by Ribeiro da Luz and Crowley (2010b), Ullah et al. (2012a), and Harrison et al. (2018) by using comprehensive methods of spectral pre-processing and techniques of classification applied to the discrimination of tree and liana leaves that can coexist in the same ecosystems. The overall performance of classification of liana and tree leaves is considerably improved using LWIR rather than VIS-NIR reflectance. However, the pre-processing method can affect this performance. Pre-processing methods such as CWT or the extraction of the first derivative tend to better highlight spectral features, and therefore, differentiate samples. Likewise, our results suggest that the most promising classification algorithms are the non-parametric or non-linear classifiers such as random forest, *k*-nearest neighbor, and support vector machine with radial kernel. The outcomes of our study highlight the potential of extending this kind of classification to the airborne or satellite level to further our understanding of the current dynamic of lianas in the ecosystem.

Acknowledgments

CHAPTER 4

PREDICTION OF LEAF TRAITS OF LIANAS AND TREES VIA THE INTEGRATION OF WAVELET SPECTRA IN THE VISIBLE-NEAR INFRARED AND THERMAL INFRARED DOMAINS

Abstract

Predicting leaf traits using models based on spectroscopic data can provide essential information to advance ecological research and future Earth system models. Most current models are based on Partial Least Squares Regression (PLSR) algorithms that attempt to predict a set of leaf traits of several plant groups using leaf spectra. However, PLSR models tend to be inconsistent in describing the importance of absorption features when used to predict leaf traits. Likewise, the effect of contrasting absorption features of different plant groups on the prediction and evaluation of PLSR models it is not well understood. Hence, this study focuses on using wavelet spectra to overcome current limitations of PLSR models and improve leaf traits predictions. Specifically, we explored the use of visible–near-infrared ($0.45 – 1.0$ μm) and mid- long-wave infrared spectra ($2.55 – 11$ μm) to predict three leaf traits of lianas and trees: Leaf Mass Area (LMA), Water Content (WC), and Equivalent Water Thickness (EWT). We also compare the effect of life forms on the prediction of traits by using sun leaves collected from 14 liana species and 21 tree species ($n= 700$) from a Neotropical Dry Forest. On each leaf, reflectance measurements were performed for both selected spectral regions; then, leaf traits were estimated from a leaf segment. Leaf reflectance was first resampled and then processed using continuous wavelet transformation (CWT) to derive the wavelet spectra. PLSR models linking the leaf traits and the reflectance or wavelet spectra were compared. Our results reveal that PLSR models based on wavelet spectra require fewer components to predict traits ($13 – 16$) than those based on reflectance ($25 – 29$). In addition, PLSR models' performance (e.g., R^2) of testing datasets tend to be higher for models based on wavelet spectra (LMA = 0.83; WC = 0.77; EWT = 0.68) than reflectance (LMA = 0.78; WC = 0.76; EWT = 0.49). Wavelet spectra models also seem to better characterize absorption features that drive the variability of leaf traits than models based on reflectance. However, life

forms play an essential role in model performance, where the prediction of lianas' traits presenting lower R^2 ($R^2 = 0.61 \pm 0.25$) than trees' traits ($R^2 = 0.69 \pm 0.15$) regardless of the type of spectra or leaf trait. Our findings highlight the use of wavelet spectra to overcome limitations of the PLSR models for predicting leaf traits and the need to explore potential bias associated with plant groups on the model evaluations.

Keywords

Leaf spectroscopy, leaf mass area, life forms, partial least-squares regression, thermal infrared, wavelet analysis.

4.1 Introduction

In the last few decades, there has been an increase in the number of studies that use spectroscopy techniques to predict leaf traits. Overall, these techniques are based on optical the interaction of light with a leaf that drives unique spectral signatures associated with its chemical and structural components (Curran, 1989). The close association between optical and biochemical/biophysical properties of leaves allows for quick and non-invasive characterization of a broad set of traits at different spatial scales. As a result, spectroscopy techniques are currently used to study the variation in leaf traits of leaves (Asner et al., 2011; Serbin et al., 2014), canopies (Asner and Martin, 2008; Wu et al., 2019), and biomes (Aguirre-Gutiérrez et al., 2021; Butler et al., 2017; Serbin et al., 2019; Wang et al., 2020).

Among leaf traits, Leaf Mass per Area (LMA, as the ratio of leaf dry mass to its leaf area), gravimetric Water Content (WC, as the percentage of the difference between leaf wet mass and dry mass divided by the wet mass), and the Equivalent Water Thickness (EWT, as the difference of leaf wet mass and dry mass divided by its area) are commonly quantified using leaf spectroscopy (Asner et al., 2011; Buitrago et al., 2018a; Cheng et al., 2014, 2012, 2011; Meerdink et al., 2016; Serbin et al., 2019, 2014; Streher et al., 2020; Ullah et al., 2012b; Wang et al., 2020). These traits tend to be considered functional due to their role on the growth, survival, and reproduction of species, which leads to a variety of plant strategies (Violle et al., 2007). The variation of leaf traits in ecosystems has been explored at the global and local scale, with findings indicating that at the global scale leaf traits variations tend to be modulated by climate through adjustments of leaf structure (Kuppler et al., 2020; Osnas et al., 2018; Wright et al., 2004); while

51

variations at the local scale tend to be modulated by the diversity and composition of communities (Messier et al., 2017; Swenson et al., 2020). The local variability of leaf traits could be an important element for the mapping of traits in tropical communities. Such ecosystems are home to a high diversity of species and life forms that coexist and present contrasting leaf traits and spectral features, such as the case of lianas and trees.

Lianas or woody vines are considered non-self-supporting structural parasites that use host trees to reach the forest canopy (see Stewart and Schnitzer, 2017 discussion of their categorization). At the canopy level, lianas develop a high proportion of leaf biomass per whole-plant biomass while infesting host trees (Rodríguez-Ronderos et al., 2016; Visser et al., 2018). In comparison with trees, liana leaves present lower LMA and higher WC (Asner and Martin, 2012; Ball et al., 2015; Cai et al., 2009; Mello et al., 2020; Sánchez-Azofeifa et al., 2009a; Slot et al., 2013). In addition, liana leaves present distinctive optical features in comparison with those of trees (Avalos et al., 1999; Castro-Esau et al., 2004; Guzmán et al., 2018; Hesketh and Sánchez-Azofeifa, 2012; Kalacska et al., 2007; Sánchez-Azofeifa et al., 2009a). Most of the previous studies that encompass the visible, near, and short-wave infrared regions (VIS-NIR-SWIR, 0.38 – 2.5 µm) suggest that liana leaves present higher reflectance close to 0.55 and 2.25 µm and lower reflectance in the near-infrared region (0.8 – 1.2 µm) compared with tree leaves. These differences in leaf reflectance appear due to the lower concentration of leaf pigments (e.g., carotenoids and chlorophyll), the higher water content, and the lower leaf thickness of liana leaves (Kalacska et al., 2007; Sánchez-Azofeifa et al., 2009a). Likewise, in the longwave infrared region (8 – 11 µm), liana and tree leaves present distinguish optical features close to 8.5, 8.9, and 9.2 µm (Guzmán et al., 2018). For mapping purposes, it has been documented that the presence of liana on tree crowns complicates the detection of tree species (Kalacska et al., 2007; Marvin et al., 2016; Sánchez-Azofeifa and Castro-Esau, 2006). Therefore, it is likely that the presence of lianas on trees hinders the mapping of traits at the canopy level due to its distinctive biophysical features, especially if models predicting leaf traits do not consider lianas as part of the canopy.

Understanding the spectroscopic mechanisms that drive the prediction of leaf traits and their potential errors associated with different plant groups (e.g., lianas vs. trees) is crucial for mapping traits and their future integration in Earth system models. As of today, only Asner et al. (2011) have explored the potential effect of life forms on leaf trait spectroscopy. Findings by Asner et al. (2011) for humid tropical forests suggest that different growth habits do not affect the

spectroscopy of LMA. However, lianas' spectral features in these ecosystems tend to resemble those of trees (Hesketh and Sánchez-Azofeifa, 2012; Sánchez-Azofeifa et al., 2009a). In tropical dry environments, in contrast, lianas present a higher relative abundance of individuals (Schnitzer and Bongers, 2011) and more contrasting optical features (Sánchez-Azofeifa et al., 2009a). Therefore, it is not well understood if the different growth habits with contrasting optical features in these dry environments influence leaf trait spectroscopy. Cheng et al. (2014, 2012, 2011) have investigated the use of spectroscopy to predict leaf traits of lianas and trees in dry environments, but the effect of growth habits on their predictive models were not explored. Sánchez-Azofeifa et al. (2009) and Ball et al. (2015) have investigated the association of spectral bands and the variability of leaf traits of lianas and trees in dry environments, but it was not investigated with the aim of creating predictive models.

In general, models to predict leaf traits tend to be classified as physically-based and data-driven methods (Féret et al., 2019). Physically-based methods, based on radiative transfer models such as PROSPECT (Jacquemoud and Baret, 1990), are consistent and well established. However, application to complex leaf morphologies and the search for the origin of potential errors associated with these methods tends to be challenging (Féret et al., 2019; Serbin et al., 2019). On the other hand, data-driven methods are based on the calibration of algorithms to estimate leaf traits. Among these, Partial Least Square Regressions (PLSR) introduced by Wold (1966) stands as a comprehensive multivariate algorithm since it reduces a large group of predictor variables (i.e., spectral information) to a few non-correlated components (Grossman et al., 1996). Despite its ability to handle the high collinearity of spectral data, PLSR models tend to be inconsistent in selecting spectral bands associated with absorption features that drive trait's prediction (Cheng et al., 2014; Streher et al., 2020). For this purpose, other algorithms that enhance the absorption features, such as continuous wavelet transformation (CWT), can accurately determine the association between leaf trait and spectral features (Cheng et al., 2014, 2012, 2011; Ullah et al., 2012b). However, the prediction of traits by CWT algorithms traditionally depends on a series of univariate methods applied to the wavelets (i.e., correlation scalogram) highly susceptible to multicollinearity when using hyperspectral data (Ullah et al., 2012b).

Hence, in this study, we integrate the ability of CWT to enhance absorption features and the competence of PLSR to handle high data collinearity as a data-driven method to predict leaf traits.

This study also assesses the use of the visible-near infrared (VIS-NIR, 0.45 – 1.0 µm) and mid-long-wave infrared spectral region (MLWIR, 2.55 – 11.0 µm) to predict leaf traits of lianas and trees from a Neotropical Dry Forest. The MLWIR region is used given its importance for predicting water-related leaf traits (Ullah et al., 2012) (i.e., WC and EWT), while the VIS and NIR region for their importance to describe leaf pigments and the leaf internal structure, respectively (Curran, 1989). Specifically, we (i) evaluate the ability of PLSR models based on reflectance and wavelet spectra to accurately predict leaf traits (i.e., LMA, WC, EWT); (ii) compare how these models identify spectral regions that could play an important role in predicting traits; and (iii) compare how life forms may influence the prediction of traits and performance in these models. Furthermore, this study highlights the integration of wavelet spectra with PLSR models as a method to improve the prediction of leaf traits, as well as the potential effect of life forms on the future mapping of leaf traits.

4.2 Materials and Methods

4.2.1 Study site and leaf spectra

This research used leaf spectra collected by Guzmán et al. (2018) and unpublished data of leaf traits measured from field samples collected at the same time. These samples were collected at the Santa Rosa National Park – Environmental Monitoring Super Site (SRNP–EMSS, 10º48" N, 85º36" W) located on the Pacific coast of northwestern Costa Rica. This site is a tropical dry forest that presents a wet-season during the middle of May to late November, and a dry-season encompassing the remaining months in which most of the trees lose their leaves (Kalacska et al., 2004; Sánchez-Azofeifa et al., 2005). The SRNP–EMSS has a mean annual precipitation of 1720 mm and an air temperature that varies from 26 °C in the wet-season to 29 °C in the dry season (Kalacska et al., 2004). The SRNP–EMSS is a mosaic of forest patches of with different ages, successional stages and land-use histories (Cao et al., 2015; Sánchez-Azofeifa et al., 2017). This site presents 96 species of trees with different life histories (Hilje et al., 2015), and approximately 18 species of lianas that reach the forest canopy (Adrian Guadamuz personal communication).

Leaf samples were collected during the 2017 wet-season. Samples consisted of sun leaves from four individuals of 14 species of lianas and 21 species of trees (Table S1). Five mature and healthy leaves were collected from each of four individuals, giving a total of 700 leaves. The number of individuals and leaf samples collected were determined based on the variability of the leaf spectra reported in previous studies in the region (Harrison et al., 2018), and the spatial distribution of plants with low abundance.

Spectral measurements were performed on each leaf using a UniSpec Spectral Analysis System (PP Systems, Amesbury, MA, USA) for the VIS-NIR region, and a portable Agilent 4100 ExoScan Fourier Transform Infra-Red (FTIR) spectrometer for the MLWIR region. Overall, the UniSpec Spectral Analysis System integrates a leaf clip that holds a foreoptic which provides illumination (7.0 W halogen bulb) to a leaf area of 4.15 mm^2. Three measurements were taken with the previous spectrometer at the middle leaf lamina of each leaf avoiding the midrib and these were later averaged. Each spectrum was the average of ~ 40-60 scans for the purpose of noise reduction. The data from VIS-NIR spectrometer consists of reflectance for 256 bands spanning a spectral range from 0.306 to 1.138 μm, each with a bandwidth of 3.3 nm. On the other hand, Agilent 4100 ExoScan Fourier Transform Infra-Red (FTIR) spectrometer is equipped with an internal IR illumination source and a diffuse reflectance probe. The illuminated area for this spectrometer has a diameter of 1.5 cm and a maximum depth of light penetration of 20-50 μm, depending on the medium. The reflectance probe was brought into contact with the leaf for the same leaf region as measured for collection of VIS-NIR spectra. Each reflectance measurement was the average of 150 scans for the purpose of noise reduction and was obtained following a measurement of the infra-gold background. The data collected consists of reflectance for 1799 bands spanning a spectral range from 2.5 to 15.4 μm (4000-650 cm^{-1}) with a resolution of 4 cm^{-1}. For more details on the instrument features, protocols of sample collection, and spectral measurements see (Harrison et al., 2018), Guzmán et al. (2018) and Foley et al. (2006).

4.2.2 Measurements of leaf traits

After the leaf spectral measurements, a leaf segment with known area (8.725 or 2.12 cm^2) was extracted per leaf using a leaf punch. Each segment was weighed, dried at ~55 °C for >72 h, and then weighed again to estimate three-leaf traits: i) Leaf Mass Area (LMA, g m^{-2}), Water Content (WC, %), and Equivalent Water Thickness (EWT, g m^2). The LMA was estimated as the ratio of

leaf dry mass and its area, WC as the percentage of the difference between wet mass and dry mass divided by the wet mass, and EWT as the difference between wet mass and dry mass divided by its area. The wet and dry mass estimations were conducted using a portable three-digit scale (0.001 g resolution) calibrated every 20 measurements. Leaf trait values per species are also described in Table A3–1.

4.2.3 Spectral processing

The MLWIR spectra were resampled to a bandwidth of 3.3 nm to match the spectral resolution of the VIS-NIR spectra and to ensure a uniform distribution of bands for further analysis. This was performed using a Gaussian function defined by the Full Width Half Maximum (FWHM) values to resample the high-resolution data to lower resolution bellow 4.2 µm and lower-resolution data to high-resolution above 4.2 µm. The FWHM is assumed to be equal to 3.3 nm over the spectral range. The spectral resampling was performed using the 'resample2' function in the *prospectr* package (Stevens and Ramirez-Lopez, 2020) of the R software version 4.0.4 (R Core Team, 2020).

The VIS-NIR spectra and the resampled MLWIR spectra were then transformed using a continuous wavelet transformation (CWT) (Grossmann and Morlet, 1984). Specifically, CWT was used as a method to isolate scales capturing features that may improve the prediction of leaf traits. The premise of this analysis is that a reflectance spectrum can be represented as a sum of wave-like functions (wavelets) of different scales (i.e., widths) (Torrence and Compo, 1998), each scale capturing different features (Rivard et al., 2008). For instance, lower scales (i.e., 2^1 or 2^2) can capture the small structure of the spectrum (e.g., random noise or small absorption features), while higher scales (i.e., 2^8 or 2^9) can capture large-scale structure (e.g., spectral continuum or broad absorption features) (Feng et al., 2018). For this transformation, the leaf reflectance from each spectrometer was decomposed in nine wavelets. The CWT was performed using a second-order Gaussian function derivative and applying a variance of 1 on each spectral region separately. Wavelets that best describe the spectral features were then summed, and these summed spectra, called wavelet from herein, were used in the ensuing models. Overall, different wavelet scales are sensitive to specific spectral features that drive the prediction of leaf traits (Cheng et al., 2011, 2014; Ullah et al., 2012); therefore, the selection of scales to be summed may have an influence on the predictive ability of PLSR models. Based on an exploratory analysis (Appendix A1), scales 2^1, 2^2, 2^3, 2^8, and 2^9 were summed to form the wavelet spectra form both spectral regions since

these scales tend to have a high performance for predicting the studied traits (Fig S1). The CWT was applied using the 'wavCWT' function in the *wmtsa* package of R (Constantine and Percival, 2017). The scales 2^8 and 2^9 also seems to be comparable with the scales selected by Ullah et al. (2012) for predicting water-related leaf traits.

Datasets of reflectance (i.e., VIS-NIR and resampled MLWIR spectra) and wavelet spectra from both spectral regions were then visually compared and used to build the Partial Least-Squares Regression (PLSR) models to predict leaf traits. Bands at the edges of the spectrometer ranges associated to low signal-to-noise ratios were deleted. The resulting spectral ranges used in the subsequent analysis were 0.45–1.0 μm (184 bands) for the VIS-NIR spectral region and 2.55–11 μm (2817 bands) for the MLWIR spectral region.

4.2.4 Prediction of traits using partial least-squares regression

We used Partial Least-Squares Regression (PLSR) to predict leaf traits using the *pls* package (Mevik and Wehrens, 2007) in R. In general, PLSR is used to predict chemometrics or leaf traits using spectra due to its ability to handle high collinearity of data in comparison with other statistical analysis tools such as standard stepwise linear regression (Grossman et al., 1996). For this analysis, we considered and compared the processed reflectance and wavelet spectra as predictor variables, and each of the log-transformed leaf traits as response variables. The PLSR analysis was performed following a similar approach as presented in Serbin et al. (2014) and Streher et al. (2020). For this study, our workflow is based on three main aspects: i) estimation of the optimal number of components, ii) evaluation of the variability of the predictor variables, and iii) evaluation of the model performance and prediction of traits. Initially, leaf samples ($n = 700$) were randomly split 60:40 ratio of samples for training and testing purposes. However, unlike Serbin et al. (2014) and Streher et al. (2020), the data were hierarchically split based on life forms and species samples, allowing each life form and species to have equal or similar representation in training models, respectively. Here our premise is that studies that want to evaluate and compare the prediction of leaf traits using PLSR between environments, life forms, or species need to integrate the potential intraspecific variability of each group (Osnas et al., 2018). Thus, PLSR models created under an unbalanced basis may be prone to more bias in those groups with less representation and higher intraspecific variability. In order to create training and testing datasets with a 60:40 ratio, trees samples were split using a 50:50 ratio (210 samples for training and

testing), while lianas samples using a 75:25 ratio (210 samples for training and 70 for testing). The ratios used to randomly split the samples represent the variability of leaf traits within life forms and between datasets (Figure A3–2). The data split for lianas and trees was conducted using the 'createDataPartition' function of the *caret* package of R (Kuhn, 2008).

4.2.4.1 Estimation of the optimal number of components

Selecting the optimal number of components by PLSR is an important step for predicting leaf traits; selecting less than the optimal number may lead to loss of information, while selecting more than the optimal number could lead to poor transferable models (Wiklund et al., 2007). Because of this, training datasets for each trait were initially used to estimate the optimal number of components that must be retained. The optimal number of components was estimated using the one-sigma method on 100 iterations of PLSR models created using 10-fold cross-validation and 50 components. In general, the one-sigma method returns the optimal number of components where the Root Mean Squared Error of Prediction (RMSRP) are within one standard error of the absolute minimum (Hastie et al., 2013). On each iteration, the same list of samples for cross-validation were applied to build both PLSR models based on reflectance or wavelet spectra. The mode of the optimal number of components for 100 iterations was then used for further analysis. The optimal number of components was selected using the function 'selectNcomp' of the *pls* package (Mevik and Wehrens, 2007).

4.2.4.2 Evaluation of the variability of the predictor variables

Once the optimal number of components per dataset and traits was estimated, we evaluated the variability of the predictor variables of the PLSR models. This was conducted by running 1000 iterative PLSR models. These iterative models use 70% of the training samples ($n = 294$) randomly selected as calibration each time. The data split for calibration during each iteration was also based on life forms allowing each life form to have a similar representation of samples for calibration purposes. Like the previous section, on each iteration the same list of samples selected for calibration was used to build both PLSR models based on reflectance or wavelet spectra. This procedure allows a paired comparison of the potential effect of the type of spectra on the model performance (Section 4.2.4.3). From each iteration, PLSR coefficients of the training models and the Variable Importance of Projection (VIP) were extracted using the optimal number of

components. The PLSR coefficients and the VIP describe regions of the spectrum that have an important role in predicting leaf traits (Mehmood et al., 2012). The PLSR coefficients were directly extracted from the 'plsr' function, while the VIP was estimated using the 'VIP' function of the *plsVarSel* package of R (Mehmood et al., 2012). The PLSR coefficients and the VIP from models based on reflectance and wavelet spectra were compared qualitatively. Likewise, peak detection analyses based on the local maxima were performed using the VIP average from the multiple iterations. The previous analyses were performed to show the potential of wavelet over reflectance spectra to describe spectral bands that might drive the prediction of leaf traits. For these analyses, the local threshold for peak detection was set as the half of the 95th percentile value of VIP per leaf trait and type of spectra.

4.2.4.3 Evaluation of the model performance and prediction of traits

Using the iterative models from the previous section and the optimal number of components, we evaluated the performance and variability of the PLSR models during the training and testing processes. The performance was evaluated during each iteration on the training and testing samples. Likewise, the model performance was evaluated during each iteration on samples from lianas and trees separately to evaluate the potential intraspecific bias of life form groups on the resulting models. From each iteration, observed-predicted relationships were performed using de-transformed values, then descriptors of the model performance such as the coefficient of determination (R^2), the Root Mean Square Error (RMSE), and the mean model bias were estimated. Likewise, we also estimated the RMSE value for each trait and dataset spectra as a percentage of the sample data range (%RMSE) following Feilhauer et al. (2010).

4.2.5 Data analysis

Initially, the life form effect on each leaf trait was compared using linear mixed models. For this, the life forms were considered as a fixed effect, and the species as a random factor to contemplate the potential effects of the species on the leaf trait variability. Then, the effect of the type of spectra (i.e., reflectance or wavelet) and the life form on the PLSR models were statistically compared. Specifically, Analyses of Covariance (ANCOVA) were performed to compare the effect of the type of spectra and the life form on the observed-predicted relationships per leaf trait on training and testing datasets. Here our premise is that if the observed-predicted relationships of

traits do not differ between life forms, a general model for lianas and trees could be used to study their traits. For these analyses, the iteration average of the predicted values was used as an independent continuous variable, while the observed values as response variable. In these analyses, we counted for the within-subjects error of the type of spectra since the samples remain unvaried during each iteration. In addition, paired *t-tests* were applied to compare the effect of the type of spectra on the performance parameters during the training and testing process. These comparisons were performed assuming that the estimated parameters are independent of the iterative procedure, similar to *t-test* bootstrapping procedures. Finally, paired *t-tests* were also used to compare the model performance between life form for each type of spectra during the training and testing process.

4.3 Results

4.3.1 Datasets evaluation

The comparisons using linear-mixed models reveal that liana and tree leaves tend to differ in their LMA and WC, but not in EWT (Table A3–2). Specifically, lianas leave presented lower LMA and higher WC than trees (Figure 4–1). Tree leaves showed higher coefficient of variation of leaf traits (LMA = 0.51; WC = 0.14; EWT = 0.35) than liana leaves (LMA = 0.40; WC = 0.11; EWT = 0.20). The linear-mixed models also reveal that leaf traits of tree species seem to be more variable than liana species when these are compared using the conditional variance-covariance matrixes (Figure A3–3). Despite this, species as a random effect presented a low contribution to the variance of the leaf traits between life forms (LMA = 0.04; WC = 0.002, EWT = 0.02) (Table A3–2).

Regarding leaf spectra, liana and tree leaves present characteristic leaf spectrum shape (reflectance or wavelet), which tend to be more contrasting between them in the MLWIR than the VIS-NIR region (Figure 4–2). In the VIS-NIR region, liana leaves tend to reflect more light than trees close to 0.55 μm (as described earlier by Sanchez-Azofeifa et al. 2009 in Panama), but tree leaves seem to reflect more light than lianas between 0.8 and 1.0 μm (Figure 4–2a). On the other hand, in the MLWIR region, liana leaves tend to present distinctive features (e.g., peaks and troughs) compared to tree leaves (Figure 4–2b, d). These distinctive features seem to be better differentiated using wavelet spectra in regions close to 2.70, 3.00, and 8.6 μm (Figure 4–2d).
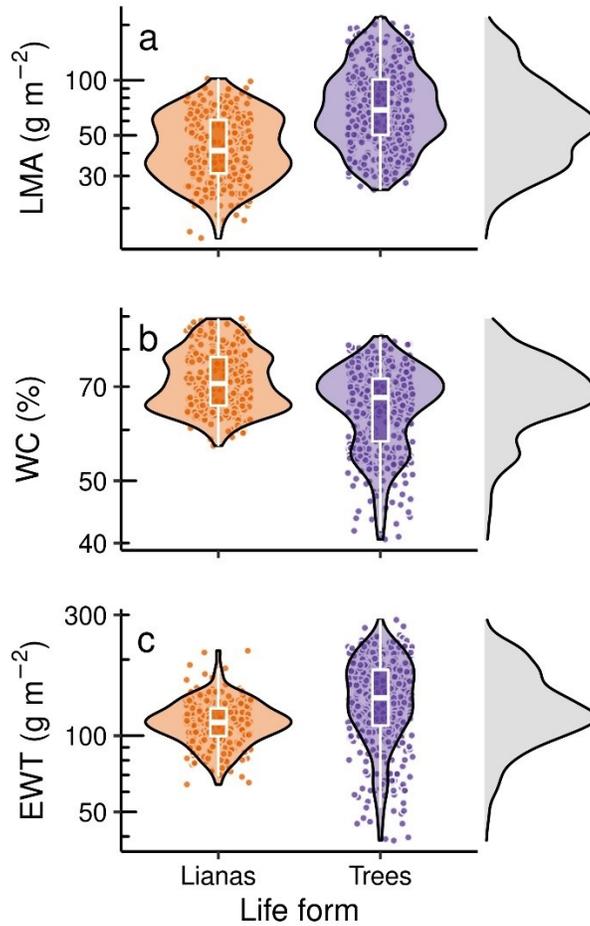
Figure 4–1. Violin plots comparing three functional leaf traits between lianas and trees at the Santa Rosa National Park – Environmental Monitoring Super Site, Costa Rica. a: Leaf Mass Area (LMA); b: Water Content (WC); c: Equivalent Water Thinness (EWT). Each point represents a sample, while the boxes the first, median, and third quartiles. The irregular polygons describe the kernel density distributions for each life form and the whole dataset outside of the plots.
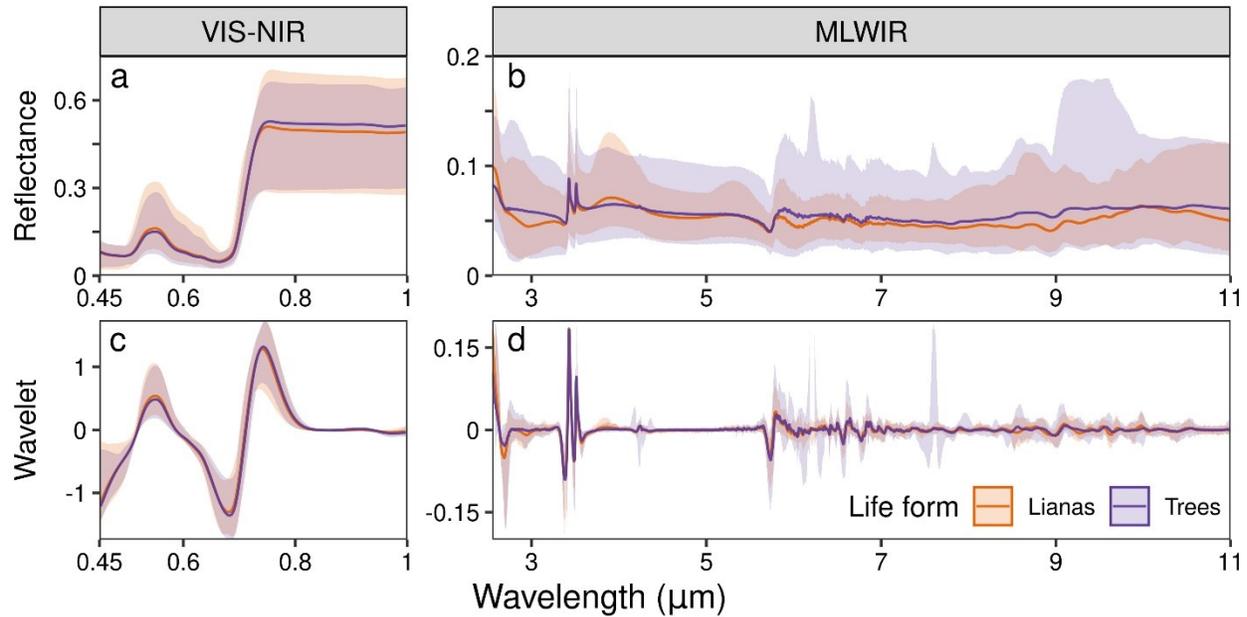
Figure 4–2. Visible-near infrared (VIS-NIR) and mid-long wave infrared (MLWIR) reflectance spectra (a, b) of liana and tree leaves and their wavelet using continuous wavelet transformation (c, d). Each line represents the average for life form, while the shade around each line the minimum and maximum value.

4.3.2 Prediction of traits using partial least-squares regression

4.3.2.1 Comparison of the optimal number of components

The extraction of the optimal number of components from the PLSR models tend to be affected by the type of spectra (Figure 4–3). Regardless of the leaf trait, the estimation of the optimal number of components using the wavelet spectra is considerably lower (between 13 and 16) than using the reflectance spectra (between 25 and 29) (Figure 4–3a, b, and c). Likewise, the estimation of the optimal number of components for WC and EWT in wavelet spectra models is less variable than reflectance models (Figure 4–3b and c). The RMSEP trends over the number of components for the studied leaf traits reveal that at the first components (< 20) the wavelet spectra tend to present lower RMSEP than reflectance spectra (Figure 4–3d, e, f); however, at a higher number of components (> 20) the reverse occurs.
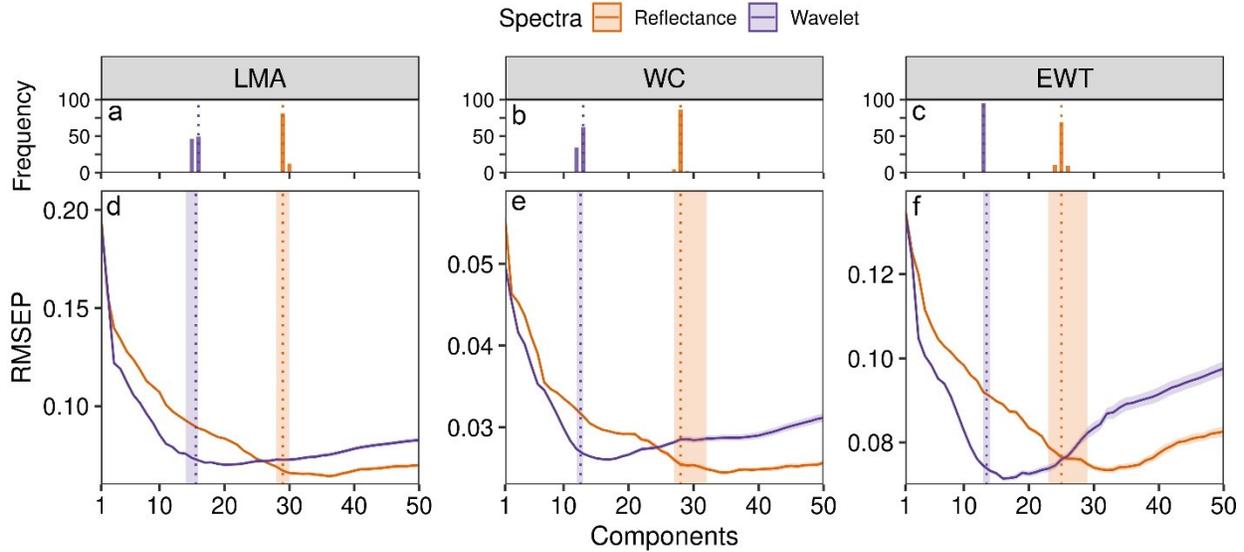
Figure 4–3. Comparison of the frequency of optimal number of components (a, b, c) and the Root Mean Squared Error of Prediction (RMSEP) (d, f, g) from PLSR models based on reflectance and wavelet spectra to predict leaf functional traits. Leaf Mass Area (LMA, $\log(g\,m^{-2})$), Water Content (WC, $\log(\%)$), and Equivalent Water Thinness (EWT, $\log(g\,m^{-2})$). For the RMSEP panels, each solid line represents the average of the 100 iterations, while the shade around each line its standard deviation. Vertical dotted lines represent the mode of the optimal number of components, while the vertical shade around these represents the minimum and maximum of all iterations.

4.3.2.2 Evaluation of the variability of the predictor variables

The extraction of the PLSR coefficients using the optimal number of components reveals to be more variable in models based on reflectance than wavelet spectra (Figure A3–4). Likewise, the VIP extracted from PLSR models suggests that the use of wavelet spectra enhances and reduces the predictor role of some spectral regions compared with reflectance spectra (Figure 4–4). For example, spectral regions close to 0.55, 0.71, 0.76, 3.43, and 3.5 µm tend to have an important role in predicting leaf traits for both types of spectra, but this importance is considerably higher for wavelet than reflectance spectra (Figure 4–4; Table 4–1). Moreover, regions close to 3.8, and 9.2 µm seems to have some degree of importance for predicting leaf traits using reflectance based on the VIP values, but not for wavelet spectra. The peak detection analyses applied to VIP from wavelet spectra models also identified regions associated with water absorption features (2.56, 2.70, and 6.20 µm) and lignin/cellulose (5.76, 6.58, and 7.60 µm) (Figure 4–4, Table 4–1); regions

which were not identified by models based on reflectance (expect for 2.60 μm on EWT-reflectance models).
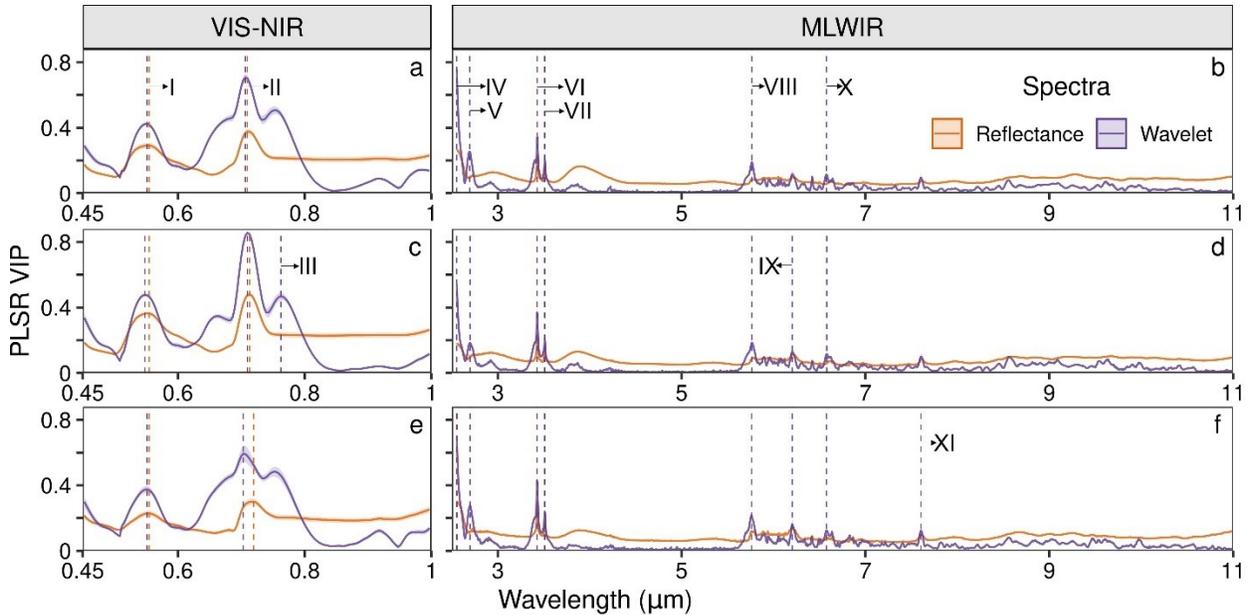
Figure 4–4. Variable importance of prediction (VIP) of PLSR models based on reflectance and wavelet spectra (CWT) to predict three-leaf functional traits: Leaf Mass Area (LMA) (a and b), Water Content (WC) (c and d), and Equivalent Water Thinness (EWT) (e and f). Each line represents the average of 1000 iterations, while the shade around each line represents its standard deviation. Vertical dashed lines represent the peaks of local maxima from the peak detection analyses, which are detailed in Table 1.

4.3.2.3 Evaluation of the model performance and prediction of traits

Overall, the PLSR model's iterations reveal the large variability of their performance during the testing and, to a lesser extent, during the training (Table 4–2). At the training level, the reflectance spectra perform slightly better to predict LMA and WC than the wavelet spectra based on: i) higher R^2, ii) lower RMSE, and iii) lower %RMSE (Table 4–2). However, at the testing level, wavelet PLSR models' performance is higher than those based on reflectance spectra for the three-leaf traits. The bias of the predicted values tends to be close to zero for models based on reflectance and wavelet spectra during the training and testing process. The ANCOVA comparisons reveal that the spectral processing seems to affect observed-predicted relationships of LMA and WC during training and EWT during testing processes (Figure A3–5a, e, and f; Table A3–3); however,

this effect is not visually clear for LMA and WC (Figure A3–5a, and e). Using testing datasets, LMA and EWT of trees from models based on reflectance tend to present lower slopes (Figure 4–5a and e) in comparison with models based on wavelet spectra (Figure 4–5b and f). This change in the slopes of the observed-predicted relationships between life forms and types of spectra does not appear in training datasets (Figure A3–6). Overall, the ANCOVA comparisons suggest that the life forms do not influence the observed-predicted relationships of leaf trait on training or testing datasets (Table A3–3).

In terms of the model performance between life forms, except for models based on LMA, liana samples tend to present lower R^2 than tree samples regardless of the type of spectra and dataset process (Figure 4–6a, b, and c; Figure A3–7a, b, and c). Surprisingly, the RMSE of observed-predicted relationships on liana samples is also lower than tree samples for the three-leaf traits (Figure 4–6g, h, i, and Figure A3–7g, h, and i). The bias of these iterative models for both life forms tends to be close to zero (Figure 4–6d, e, f, and Figure A3–7d, e, and f); however, the model's residuals reveal that higher values of leaf traits may be prone to more bias in comparison with lower values (Figure A3–8 and A3–9). This effect of the model residuals together with the intraspecific variability of leaf traits within life forms seems to drive the differences in the %RMSE (Figure 6j, k, l, and Figure S7j, k, and l). Overall, life forms with higher and less variable values of leaf traits tend to present higher %RMSE (i.e., WC in lianas, and LMA in trees). Likewise, lianas with low coefficient of variation of leaf traits tend to present low RMSE. Regardless the model performance, the spread of the performance parameters from multiple iterations is consistently higher in models based on reflectance than wavelet models (Figure 4–6 and Figure A3–7). This spread, which describe the precision of the models, can also be observed in Table 4–2; the standard deviations of performance parameters from wavelet models are lower than parameters from reflectance models in most cases (Table 4–2).

Table 4–1. Description of peaks of local maxima from the variable importance of prediction, their spectral location, leaf trait and PLSR model where were detected, and the potential features related to based on the existing literature.

Symbology	Location (μm)	Detected in	Features related to	Reference
I	0.55	All leaf traits and types of spectra	Photochemical pigments / Electron transition	Curran (1989)
II	0.71	All leaf traits and types of spectra	Inflection point between chlorophyll absorption features and volume scattering	Curran (1989)
III	0.76	WC-wavelet	Mesophyll structure / volume scattering	Curran (1989)
IV	2.56	EWT-reflectance LMA-wavelet WC-wavelet EWT-wavelet	Water (2.51 μm) / O–H stretching	Arshad et al. (2018) and Ullah et al. (2012)
V	2.68	All leaf traits using wavelet spectra	Water (2.91 μm) / O–H stretching	Arshad et al. (2018), Elvidge (1988), and Ullah et al. (2012)
VI	3.44	All leaf traits and types of spectra	Cellulose / CH_2 asymmetric stretching	Elvidge (1988)
VII	3.51	All leaf traits and types of spectra	Cellulose / CH_2 asymmetric stretching	Elvidge (1988)
VIII	5.76	All leaf traits using wavelet spectra	Cellulose or lignin (5.76 μm) / C=O stretching	Elvidge (1988) and Stewart et al. (1997)
IX	6.20	WC and EWT using wavelet spectra	Water / O–H stretching	Elvidge (1988) and Fabre et al. (2011)
X	6.58	All leaf traits using wavelet spectra	Lignin / Aromatic skeletal vibrations	Boeriu et al. (2004)
XI	7.60	EWT using wavelet spectra	Cellulose (7.60 μm) or lignin (7.54 μm) / CH_2 wagging	Boeriu et al. (2004) and Elvidge (1988)

Table 4–2. Comparison of the partial least-squares regression (PLSR) performance during the training and testing of models based on reflectance or wavelet spectra to predict three functional traits: leaf mass area (LMA), water content (WC), equivalent water thinness (EWT). Values represent the mean and standard deviation of 1000 iterative models. Paired *t-test* results and their significance are described below of each comparison.

Spectra	R^2 (10^{-2})		Bias (10^{-2})		RMSE		%RMSE	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
LMA (g m^{-2})								
Reflectance	94.12 ± 1.08	77.67 ± 4.88	-0.68 ± 0.56	-1.16 ± 0.82	8.80 ± 0.74	17.87 ± 1.71	4.87 ± 0.41	8.96 ± 0.86
Wavelet	93.55 ± 0.70	83.35 ± 1.89	-0.91 ± 0.58	0.62 ± 0.74	9.23 ± 0.48	15.47 ± 0.87	5.11 ± 0.27	7.76 ± 0.44
	15.74***	-37.11***	18.47***	-76.01***	-17.64***	44.17***	-17.64***	44.17***
WC (%)								
Reflectance	89.84 ± 1.07	76.24 ± 1.89	-0.11 ± 0.20	0.51 ± 0.29	2.98 ± 0.15	4.35 ± 0.17	6.11 ± 0.32	9.17 ± 0.36
Wavelet	86.16 ± 0.61	76.77 ± 1.29	-0.13 ± 0.20	0.52 ± 0.25	3.48 ± 0.07	4.31 ± 0.12	7.14 ± 0.16	9.07 ± 0.78
	105.61***	-8.08***	3.72***	-0.47	-102.54***	7.96***	-102.54***	7.97***
EWT (g m^{-2})								
Reflectance	80.93 ± 2.73	48.69 ± 21.92	-1.07 ± 0.66	0.60 ± 0.98	18.43 ± 1.09	32.93 ± 4.68	7.44 ± 0.44	14.19 ± 2.02
Wavelet	82.26 ± 1.92	67.70 ± 2.56	-1.09 ± 0.61	2.08 ± 0.79	17.79 ± 0.89	26.37 ± 1.05	7.18 ± 0.36	11.37 ± 0.45
	-14.67***	-27.58***	1.58**	-57.76***	17.86***	45.22***	17.86***	45.22***

p-value: * <0.05; ** <0.01, *** <0.001
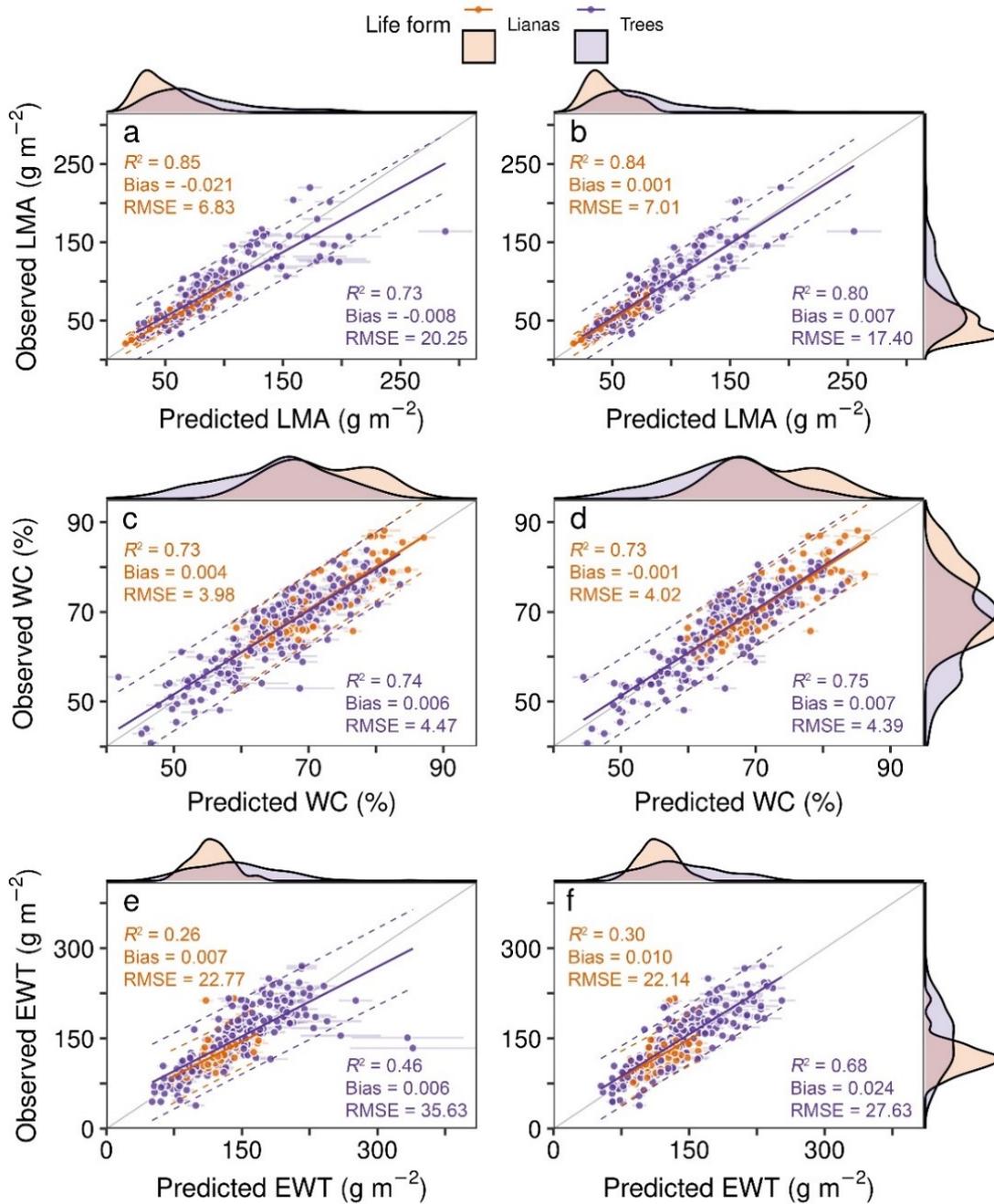
Figure 4–5. Observed and predicted leaf traits of lianas and trees from partial least-squares regression (PLSR) models using reflectance (a, c, e) and wavelet spectra (b, d, f). Points correspond to the testing dataset. Each point represents the average of 1000 iterations, while the error bars around each point their standard deviation. The solid grey lines indicate the 1:1 relationship. The solid lines represent the fitted regression while the dashed lines the 95% prediction intervals. Kernel density distributions are plotted next to each scatter plot. Statistics about their comparisons can be found in Table A3–3.
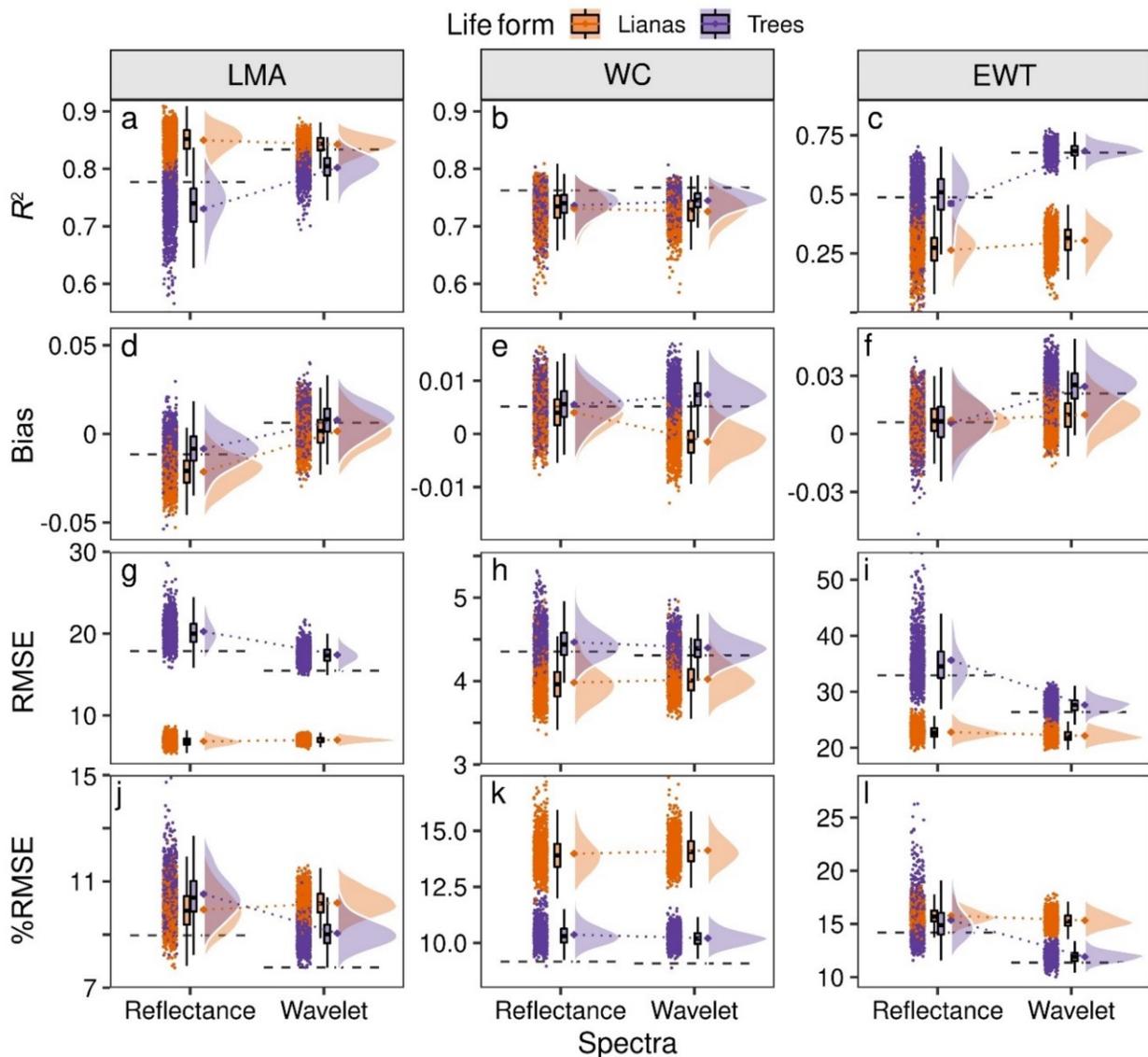
Figure 4–6. Raincloud plots comparing the performance between life forms of testing models of partial least-squares regression (PLSR) based on reflectance and continuous wavelet transformation (CWT) spectra to predict three functional traits: Leaf Mass Area (LMA, g m^{-2}), Water Content (WC, %), Equivalent Water Thinness (EWT, g m^{-2}). RMSE is the root mean square error, and %RMSE is the RMSE represented by its percentage. Each point represents a model iteration, the boxes the first, median, and third quartiles, and the irregular polygons describe the kernel density distributions. Dashed lines represent the mean regardless of the life form by type of spectra, while the dotted lines the mean trend of change within life forms between the type of spectra. Statistics of their comparison can be found in Table A4–4.

4.4 Discussion

Our results reveal the potential of wavelet spectra to overcome some limitations of PLSR models based on reflectance and to improve the prediction of leaf traits using spectroscopy. Despite this, our results also reveal the potential bias associated with the prediction of leaf traits when considering life forms, which has not been done until today. The following discussion highlights both previous aspects and provides insights into the future application of wavelet spectra in PLSR models.

4.4.1 Influence of the spectra processing on the PLSR models

Overall, our results reveal three main advantages regarding using wavelet vs. reflectance spectra in PLSR models to predict leaf traits: i) wavelet-based models requires fewer components to predict leaf traits, ii) wavelet-based models enhance the importance of spectral bands that drive the prediction of traits, and iii) wavelet-based models in training datasets improve the performance of predicting traits in accuracy and precision. Specifically, reductions in the optimal number of components in wavelet-based models are associated with increases in the variability of the spectral features highlighted by the CWT (Guzmán et al., 2018); leading to more parsimonious PLSR models that require fewer components to predict the same variability of a leaf trait. These reductions in the optimal number of components are directly translated into shorter processing time during the model's training and testing. Likewise, the low variability of the selected optimal number of components of wavelet-based models is translated into more robust and consistent models. Despite the broad spectral range and the higher number of spectral bands used in this study, the optimal number of components estimated for LMA and EWT from models based on wavelet spectra tend to be close (± 5) to those reported by other studies (Meerdink et al., 2016; Serbin et al., 2014; Streher et al., 2020; Wang et al., 2020).

In terms of the spectral band's importance (i.e., VIP), models based on wavelet spectra tend to better define the spectral bands that drive the association between spectral features and leaf traits. This is done by enhancing bands with a strong relationship between the absorption features variability and the leaf trait variability and diminish bands without any apparent contribution. Other studies have shown that PLSR models based on reflectance spectra are able to characterize the simultaneous contribution of absorption features of leaves to predict traits (Serbin et al., 2019,

2014; Streher et al., 2020); however, the role of spectral bands with importance without any apparent association between leaf optics and traits is considerable. In our study, for instance, reflectance-base models showed spectral bands close to 9.2 µm tend to have a similar importance for the prediction of the studied leaf traits than bands close to the red-edge region (0.71 and 0.76 µm). This occurs even though the former is a region with features that could be associated with the leaf surface constituents and not with the internal leaf structure or water content (Ribeiro da Luz, 2006; Ribeiro da Luz and Crowley, 2010, 2007). Although the red-edge regions does not present water absorption features either beneficial for predicting WC and EWT, their reflectance is controlled by the inflection point between chlorophyll absorption features (~ 0.68 µm) and volume scattering (> 0.71 µm) product of the leaf thickness and air spaces (Curran, 1989). It is well known that the leaf investment in the mesophyll strongly modulates the LMA, WC, and EWT through the mass and water storage (Niinemets, 1999; Poorter et al., 2009; Shipley et al., 2006). Therefore, based on the leaf economic spectrum (Osnas et al., 2018, 2013; Wright et al., 2004), it is expected that this region will present a meaningful association with the studied leaf traits. Likewise, it is likely that the contrasting spectral shapes in the VIS-NIR in conjunction with the differences observed in water content between life forms may contribute to an indirect association of the importance of this region to predict WC and EWT; a factor that was not expected. In terms of LMA, the volume scattering between 0.71 and 1.2 µm has been recognize as a region of importance for predicting LMA (Asner et al., 2011; Serbin et al., 2019; Streher et al., 2020). In our LMA wavelet-base model this importance is only observed in regions close to 0.75 µm due to the lacking of absorption features in the near-infrared region.

Water absorption features in the MLWIR also tend to be captured by PLSR models that use wavelet spectra. The VIP of models based on wavelet spectra presented regions close to 2.56, 2.70 and 6.20 µm that appear to exemplify the water absorption features of plant materials (Elvidge, 1988; Fabre et al., 2011; Ullah et al., 2012). These water absorption regions do not seem to be detected in reflectance-based PLSR models, highlighting its limitation in the description of spectral bands associated with leaf traits (Cheng et al., 2014). Other absorption features associated with lignin and cellulose tend to be highlighted by the peak detection algorithm in wavelet-based models. These features could be indirectly associated with the studied leaf traits given the contrasting concentration of lignin and cellulose documented for these life forms (Asner and Martin, 2012).

Regarding model performance, the spectral processing using CWT tend to improve the prediction of leaf traits in testing datasets. These improvements are observed at the accuracy of the models developed (high R^2 and low RMSE) and at the precision of the multiple iterations (low standard deviation). Despite this, the improvements in performance seem to be more subdued for WC, than LMA or EWT. This could be due to the nature of the WC values (i.e., values between 0 and 100) that may homogenize the predictive ability of PLSR models based on reflectance or wavelet spectra. The role of the distribution of values on PLSR models should be explored using other leaf traits based on percentage. In terms of R^2 and RMSE, the observed performances of LMA testing datasets are similar or higher to those reported by studies that use PLSR models and a spectral range close to 0.45 and 2.4 µm (Asner et al., 2011; Li et al., 2007; Serbin et al., 2019; Streher et al., 2020; Yang et al., 2016). Likewise, our models that use LMA testing datasets also seems to outperform those that integrate the thermal infrared domain (> 2.4 µm) in temperate tree species (Meerdink et al., 2016). In comparison with studies that use wavelet spectra and correlation scalograms on a variety of tree species (Cheng et al., 2014, 2011), the PLSR models also seem to outperform the prediction of LMA and WC. Despite this, other studies that integrate wavelet spectra or PLSR models to predict WC and include the shortwave infrared domains (1.2 – 2.4 µm) seems to outperform our predictions (Li et al., 2007; Meerdink et al., 2016; Ullah et al., 2012b). Overall, performance differences between studies need to be addressed with caution since there is an important variation in how the models were developed, and most of them tend to span different wavelength ranges, sensors, and plant groups.

4.4.2 Effects of the life forms on PLSR models

Our results reveal that the observed-predicted relationships of leaf traits appear to be unaffected by life forms on training and testing datasets. This seems to be in concordance with Asner et al. (2011), who suggest that the prediction of LMA using spectroscopy techniques is unaffected by growth habits. Despite this, our results suggest that life forms tend to affect the PLSR model performance. Except for LMA, liana EWT, and to a lesser extent liana WC, are poorly predicted in comparison with trees. Differences in performance occur even though the models were created under a balanced approach; a procedure that does not tend to be considered by studies which compare groups of plants (Asner et al., 2011; Serbin et al., 2019, 2014; Streher et al., 2020). These differences may suggest that the potential errors of predicting traits are susceptible to the plant

groups' identity that integrate the models rather than the overall model performance. Differences in performance between growth forms (i.e., graminoids, forbs, woody species) have also been observed by Streher et al. (2020), who developed models for each plant group. Overall, the life forms effect on the model performance appears to be the product of the intraspecific variability of traits within plant groups and the predictive ability of the model developed. Thus, life forms with higher values (i.e., LMA in trees and WC in lianas) seem to drive low-performance models in terms of R^2. On the other hand, life forms with low variability of leaf traits (i.e., lianas) seems to drive high-performance in terms of RMSE.

The effect of plant groups on the potential errors of predicting traits may play an essential role in the mapping of traits in tropical environments, where there is a high diversity of species with different growing habits coexisting together. Based on our results, for instance, biases at the local scale may differ spatially during the up-scaling of information when the forest canopy is infested or not by lianas. Ecosystems such as tropical dry forests could be more prone to this example due to the higher relative abundance of lianas in comparison with other tropical forests (Schnitzer and Bongers, 2011).

4.4.3 Future insights on the integration of wavelet spectra in PLSR models to predict traits.

In contrast to the traditional methods (e.g., Asner et al., 2011; Serbin et al., 2019, 2014; Streher et al., 2020), the integration of wavelet spectra with PLSR consistently improves model development. The potential of wavelet spectra to predict leaf traits has also been demonstrated by studies focused on band selection though correlation scalograms (Cheng et al., 2014, 2011; Ullah et al., 2012b). Based on the method presented, questions might remain regarding the selection and sum of the optimal number of scales to create the transformed spectra and predict different leaf traits. For instance, there is consistent evidence that suggests that specific wavelet scales are sensitive to specific spectral features (Cheng et al., 2014, 2011; Ullah et al., 2012b). Therefore, future studies should explore the potential effect of the selection and combination of wavelet scales with its integration with PLSR models to predict different leaf traits. This could be addressed qualitatively in other databases by looking at the variability of absorption features at different scales (e.g., using correlation scalograms) or quantitatively by creating and comparing models with different combinations of scales (e.g., such as Appendix A3–1).

Towards the upscaling of information, imagery in a given spectral-domain will need to be also processed using CWT. This step may require a higher computational demand in comparison with traditional methods based on reflectance spectra, but it could be tackled using parallel processing on each pixel. Future initiatives should be focused on benchmarks using different databases and spectral domains to address which method could be the most favorable to predict leaf traits. This previous aspect is crucial for the aim of mapping traits using current and future hyperspectral imagery from missions such as PRecursore IperSpettrale della Missione Applicativa (PRISMA), Hyperspectral Imager Suite (HISUI), and Environmental Mapping and Analysis Program (EnMAP). A suitable mapping of traits will benefit their integration in Earth system models (van Bodegom et al., 2014; Y. Yang et al., 2015).

4.5 Conclusions

Our study reveals the potential of wavelet spectra to overcome limitations and improve the prediction of LMA, WC, and EWT using PLSR models. This study also expands previous studies that enhance absorption features through the CWT (Cheng et al., 2014, 2011) and integrates the thermal infrared domain in PLSR models to predict leaf traits (Buitrago et al., 2018a; Meerdink et al., 2016). In comparison with reflectance spectra, the integration of wavelet spectra in PLSR leads to: i) more parsimonious and consistent models that require fewer components to predict the same variability of leaf traits, ii) models that tend to highlight better the spectral bands that drive the prediction of traits, and iii) increases in model performance based on accuracy (i.e., high R^2 and low RMSE) and precision (low standard deviation). Our study also reveals that the differences between life forms do not seem to affect the observed and predicted relationships derived from PLSR models. However, we observed that the model performance is affected by the intraspecific variability of the life forms, suggesting that the potential bias of prediction of leaf traits is susceptible to the identity of the plant groups (lianas vs. trees). Overall, the application and evaluation of methods such as the presented here may help advance the future mapping of leaf traits locally or globally as well as ecological studies based on leaf spectroscopy.

Data and code availability

The spectral data and leaf traits used in this study are available at the Tropi-Dry Dataverse (https://doi.org/10.7910/DVN/YBYO7W). Codes for data processing, analysis, and visualization are available at https://github.com/Antguz/PLSR_leaf-traits with the current version 0.2 archived at Zenodo (https://doi.org/10.5281/zenodo.4124182).

CHAPTER 5

ON THE RELATIONSHIP OF FRACTAL GEOMETRY AND TREE-STAND METRICS ON POINT CLOUDS DERIVED FROM TERRESTRIAL LASER SCANNING

Abstract

Fractals have been widely used to determine bifurcation patterns in trees or to analyze the homeostasis of the development of plants to different environments. In a few instances, fractals have been used to predict tree or stand metrics. Here we explore the use of fractal geometry based on the voxel-counting method (VC) to predict tree and stands metrics on point clouds derived from Terrestrial Laser Scanning. This was explored using 189 leaf-on and leaf-off point clouds from seven databases around the world. Four metrics were estimated at the tree level: height, diameter at breast height, crown area, and tree volume. At the stand level, artificial stands were created by adding trees to a given plot, and then the basal area, stand volume, and area coverage by crowns were estimated. The VC was applied to trees or stands creating voxels of different volumes (S) while counting the number of voxels (N) required to fill it. Log-log relationships between N and $1/S$ were used to estimate the fractal dimension (d_{MB}) and the intercept$_{MB}$. At the tree level, the intercept$_{MB}$ shows a stronger relationship with metrics for leaf-on ($r^2 = 0.26 - 0.90$) and leaf-off point clouds ($r^2 = 0.18 - 0.87$) than d_{MB} ($r^2 < 0.34$); however, d_{MB} seems to describe better the complexity embedded within leaf-on/leaf-off point clouds. The predictions by the intercept$_{MB}$ are affected by the presence/absence of leaves, but less affected by the random effects of the databases. At the stand level, both fractal geometry parameters (intercept$_{MB}$ and d_{MB}) tend to predict the variability of stand metrics ($r^2 = 0.61 - 0.98$). The estimation of tree and stand metrics based on fractal geometry equations can be considered a fast approach for predicting irregular structures. Using fractals on point clouds also allows us to understand the structural complexity of how trees or stands occupy their 3D-space. This complexity can be further used as a structural trait of trees or forest ecosystems. Fractal geometry equations can also help towards the development of large-scale biomass maps at different ecosystems.

Keywords

Fractal dimension, plant architecture, power-law, Terrestrial LiDAR, voxelization.

5.1 Introduction

Since its definition by Mandelbrot (1983), fractals have been considered as certain geometrical structures whose shape tend to be similar in their overall patterns when they are observed at different scales. Fractals applicability to the study of spatial and temporal patterns are so vast that, according to Halley et al. (2004), it has been implemented in every field of science. On plants specifically, the estimation of the fractal dimension is commonly used since it offers a relatively inexpensive tool to assess biological processes. For example, the fractal dimension has been applied to evaluate the bifurcation patterns of roots and branches (Fitter and Stickland, 1992), the homeostasis of development of plants during growth (Escós et al., 2003), as well as to estimate above-ground biomass (AGB) (van Noordwijk and Mulia, 2002; West et al., 1999). A reason for the popularity of applying fractals to plants is their strong relationship with the power-law (Halley et al., 2004), which is an accurate descriptor of the allometric scaling of plants (Niklas, 1994).

Despite their use, fractals have been rarely applied to large plants such as trees since it is difficult to quantify the geometrical structure of trees at large scales. However, the rapid development of new technologies such as Terrestrial Laser Scanners (TLS), also known as Terrestrial-Light Detection and Ranging (T-LiDAR) instruments, may help to overcome this limitation and help to study the fractal properties of trees and stands from different perspectives. Specifically, by employing point clouds derived from TLS it is possible to apply fractal geometry to determine how trees and other growth forms occupy their three-dimensional space. Currently, the fractal geometry has been applied to airborne LiDAR data as a method for the extraction of vegetation points (H. Yang et al., 2015), and it has been applied as a method to characterize the tree and stand structural complexity using TLS (Seidel, 2018; Seidel et al., 2019b; Suzuki, 2007). In only a few cases, however, has fractal geometry been used to characterize tree metrics (e.g., tree height, trunk diameter, etc.) (Seidel et al., 2019a), and never before has it been used to predict the volume of trees or stands using TLS derived point clouds.

The estimation of tree volume, and consequently the AGB, at tree or stand level is critical for the understanding of carbon dynamics in ecosystems (Houghton et al., 2009). Currently, there are three main methods based on TLS to retrieve tree volume and its AGB: allometric-based methods,

voxelization, and Quantitative Structure Models (QSM). Advantages and disadvantages of these methods have been discussed through their development and application (Calders et al., 2015; Drake et al., 2003; Greaves et al., 2015; Momo et al., 2017a). However, none of the current methods can describe how tree architecture is distributed in their space or how this architecture is related to their size.

Since the distribution of points in the space of a given scanned tree or forest stand is determined by its architecture, structure, and size, it is possible to use fractal geometry to evaluate how trees occupy their space, and therefore, predict their metrics (i.e., volume, crown area, diameter at breast height (DBH), and tree height). Here we explore the above idea and provide insights into the application of fractal geometry on tree point cloud as well as how it can be scaled to forest stands. Hence, we computed the fractal geometry on 189 point clouds of trees with and without leaves from seven databases distributed around the world. In addition, we also explore the effect that the presence/absence of leaves has on the parameters and relationships derived from fractal geometry. Point clouds *per se* are not an ideal fractal, but their properties associated with the distribution, aggregation, and resolution tend to be similar across different sets of scales, where methods of fractal geometry can be applied.

5.2. Materials and Methods

5.2.2 Databases of tree point clouds

This study was conducted using seven databases of point clouds collected using TLS at different locations: i) Central Guyana (GUY), ii) South Western Amazon, Peru (PER), iii) Mentaya River, Indonesia (IND), iv) Eastern Cameroon (CAM), v) Victoria, Australia (AUS), vi) Santa Rosa National Park – Environmental Monitoring Supersite, Costa Rica (CR), and vii) University of Alberta North Campus, Canada (CAN). The GUY, PER, and IND databases consist of 29 point clouds of leaf-on trees collected by Gonzalez de Tanago et al. (2017) and are available at http://lucid.wur.nl. These point clouds were created by establishing plots around each target tree for scanning (See Gonzalez de Tanago et al., 2017). The CAM database consists of 61 scanned trees collected by Momo et al. (2017b) and are available at Momo et al. (2017a). These point clouds were manually segmented to create leaf-off point clouds (See Momo et al. (2017b)). The AUS database consists of 65 leaf-on point clouds of trees from a native Eucalypt Open Forest collected by Calders et al. (2015) and are available at the TERN AusCover data archive

(http://dx.doi.org/10.4227/05/542B766D5D00D). These point clouds were collected using a semi-automated extraction of trees on scanned plots (See in Calders et al. (2015)). The CR and CAN database consists of 19 leaf-on and 15 leaf-off point clouds collected and processed by the authors of this manuscript. The CR database was generated by conducting 5 to 7 scans around each tree of interest, while the CAN database was performed using a manual extraction of trees on two scanning plots with more than 12 scans each. For more details on the description of data collection and processing of the CR and CAN databases see Appendix A4–1. The points clouds of the CR and CAN database are available at the Tropi-Dry dataverse repository (https://doi.org/10.7910/DVN/DYNAWT). In general, a description of the location, elevation, and average annual precipitation of these databases can be found in Table 5–1. In total, point clouds from 189 individual trees (76 leaf-off and 113 leaf-on) were used in this research. The overall database has 50 species, 46 genus, and 23 families (Table A4–1). Using these point clouds tree metrics were compared with their fractal geometry.

5.2.2 Tree metrics

Four tree metrics were computed directly from each point cloud. These metrics consist of descriptors of the three dimensions that are usually estimated in the field by direct measurements or harvest. Specifically, metrics associated with tree height (H, m), crown area (CA, m^2), Diameter at Breast Height (DBH, cm), and tree volume (V, m^3) were estimated. H was calculated as the difference between the maximum and minimum value of the z coordinate of the tree point cloud, while DBH was estimated as the diameter of an area of a circle (basal area, BA) from a convex hull object drawn using the x and y plane between a height (z) of 1.25 and 1.35 m. The CA was also calculated with a convex hull using the x and y plane of the point cloud and estimating their area. The above metrics were estimated using the 'tree_metrics' function from the *rTLS* package (Guzmán et al., 2020) for R (R Core Team, 2020) available at https://antguz.github.io/rTLS. The V was estimated using Quantitative Structural Models (QSM) following Raumonen et al. (2013). This method consists of segmenting tree point clouds into sections of stems and branches, then fitting cylinders to the perimeter of each segmented section to give shape and volume to the point cloud, therefore resulting in the construction of a QSM for the tree. The sum of the volume of the fitted cylinders was used as the V. The published databases of GUY, CAM, IND, PER, and AUS already have the QSMs fitted on their point clouds; these were used in this study.

Table 5–1. General descriptors of the databases used.

Site of the database	Acronym	Approximate location	Average annual precipitation (mm)	Elevation (m a.s.l.)	TLS scanner used	Leaves	n
Edmonton, Canada[1]	CAN	53° 31' 42.4" N 113° 31' 26.5" W	459	683	RIEGL VZ-400i	Off	15
Santa Rosa National Park – Environmental Monitoring Supersite, Costa Rica[1]	CR	10° 50' 16.84" N 85° 37' 6.62" W	1720	200-300	RIEGL VZ-400i	On	19
Central Guyana[2]	GUY	6° 2' 2.4" N 58° 41' 47.4" W	2195	117	RIEGL VZ-400	On	10
Eastern of Cameroon[3]	CAM	4° 02' 20.77" N 14° 55' 49.15" E	1500 − 2000	600 − 700	Leica C10	Off	61
Indonesia[2]	IND	2° 24' 36" S 113° 7' 48" E	2616	22	RIEGL VZ-400	On	10
Peru[2]	PER	12° 16' 12" S 69° 6' 0" W	2074	312	RIEGL VZ-400	On	9
Victoria, Australia[4]	AUS	37° 45' 32' S 145° 0' 57.6" E	500 − 800	150 − 600	RIEGL VZ-400	On	65

[1]Database by the authors; [2]Gonzalez de Tanago, Lau, Bartholomeus, Herold, Avitabile, Raumonen, Martius, Goodman, Manuri, et al. 2017; [3]Momo et al., 2017a; [4]Calders et al., 2015

5.2.3 Fractal geometry and voxel counting

The fractal geometry was evaluated for each tree point cloud to determine how the Euclidean distances between points, which represent the architectural pattern of the trees, changes with the scale at which it is measured. This was evaluated using the Minkowski–Bouligand method (MB) or commonly known as the box-counting method. In general, the box-counting method is a way of sampling complex patterns in a given object (usually an image or a 2D object) by breaking the object into smaller and smaller boxes while extracting the rate of change in the number of boxes needed to fill it (Bunde and Havlin, 1994). In our study, because point clouds are represented in three dimensions, voxels of a given volume were used instead of boxes. These voxels were created using a fixed grid cube to find the optimal coverage for the trees. In practice, a given point cloud can be covered using a single large voxel ($N_1 = 1$) of size S_1 (m^3), but as S is reduced in volume ($S_1 > S_2 ... > S_n$), the number of voxels required ($N > 1$) to cover it will increase (Figure 5–1a). Since N increases as a power function (Figure 5–1b), a positive d_{MB} and their intercept can be estimated using a linear model following:

$$\log N = d_{MB} \log\frac{1}{S} + intercept_{MB} \qquad (1)$$

where d_{MB} is the slope and the intercept$_{MB}$ is the intercept with the y-axis. This linear relationship was solved using standardize major axis (SMA) regressions applying the 'sma' function of the "smart" package in R (Warton et al., 2012). From this model, the coefficient of determination (r^2) was also estimated. The r^2 could be considered as a measurement of self-similarity as it describe how voxels show similar statistical properties at different scales (Mandelbrot, 1967). Since the tree dimension varies between individuals, the number of voxels required to cover a given tree also varies. Therefore, a large voxel (N_1, S_1) was created for each point cloud following:

$$E_1 = \arg max \{\Delta x, \Delta y, \Delta z\} + 0.001\ m \qquad (2)$$

where E_1 is the size of a voxel edge (m), and Δx, Δy, and Δz are the difference between the maximum and minimum coordinates of points in the point cloud. Using E_1, another eight different Es were defined between E_1 and the minimum distance (E_{min}, or commonly known as lower cut-off) of 0.01 m using a logarithmic sequence. In practice, when E tends to be equal or lower than

the minimum Euclidean distance between points, N is equal to the number of points in the cloud. Therefore, E_{min} needs to be specified to avoid a quantization error. S was calculated following E^3, and so, Eq. (2) was applied. Regardless of the tree dimension, the point cloud resolution, or the number of points in the cloud, the voxelization was performed adding voxels from the lower to the higher value of x, y, and z. Following these procedures, the resulting d_{MB} may have values between 0 and 1, where values close to 1 represents a tree that uniformly occupies their 3D space (such as a Menger sponge with the greatest surface-to-volume ratio), while values close to 0 represents a cylindrical tree (such as a pole-like object) or a point cloud with two points. On the other hand, the intercept$_{MB}$ may have positive and negative values, where high values tend to be associated with large objects that require several voxels to fill at different scales. In general, d_{MB} could potentially be affected by subsampling methods that impact the point density, but the intercept$_{MB}$ remains almost invariant despite these (Appendix S2, Figure S1). Voxels were created using the 'voxel_counting' function from the *rTLS* package (Guzmán et al., 2020).
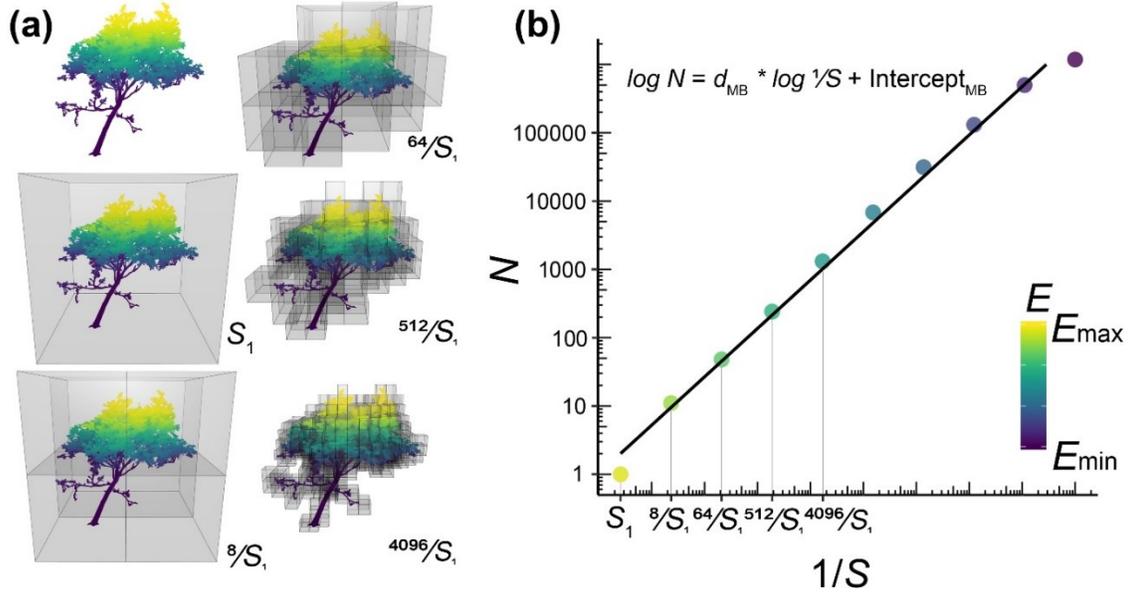
Figure 5–1. Schematic representation of the voxel counting method to calculate the fractal dimension (d_{MB}) and the intercept$_{MB}$. (a) A given point cloud can be covered by a large voxel ($N = 1$) of size S_1. However, a greater number of voxels ($N > 1$) will be required if $S < S_1$. (b) As N increases as a power function, the scaling factor d_{MB} can be solved using a log-log linear regression. S is estimated using the length of a voxel edge (E) as E^3.

5.2.4 Artificial forest stands

To prove that fractal geometry parameters can be used to predict the stand metrics, a series of artificial forest stands were created. These artificial stands were constructed using leaf-on point clouds described in Section 6.2.1. In general, the artificial stand's construction consists of randomly adding a given number of trees (point clouds) to a given plot area. Specifically, a randomly chosen tree is positioned in the plot and added to the empty stand based on a random set of $x\ y$ coordinates given to the basal trunk. Subsequently, a second and successive number of randomly chosen trees are also positioned in the plot using random $x\ y$ coordinates to their basal trunk, while ensuring that the crown area of these trees does not overlap by more than 5% with already existing crowns (Figure 5–2). In total, 1300 artificial forest stands were created using a plot area of a hectare, while the number of trees found in the plot ranged from 2 to 62.
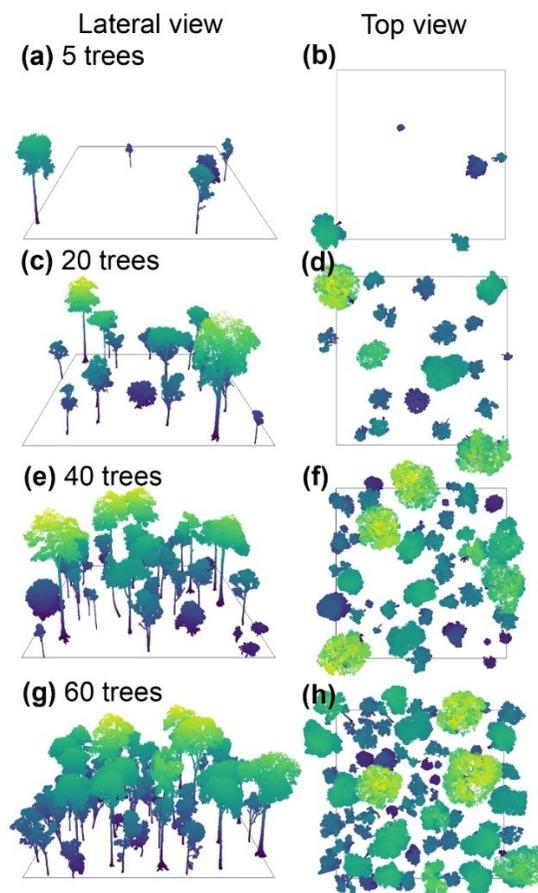
Figure 5–2. Schematic representation of artificial forest stands using the leaf-on point clouds from different databases. The plot area is one hectare.

5.2.5 Data analysis

Initially, the effect of the presence-absence of leaves on the fractal geometry parameters were addressed by comparing the d_{MB}, the intercept$_{MB}$, and r^2 between leaf-off and leaf-on point clouds using linear mixed models. For this, we considered presence-absence of leaves as a fixed effect, and the point cloud origin as a random factor in order to contemplate the potential effects of the different pre-processing and data acquisition methods for each database. Likewise, the relationships between the log-transformed tree metrics (*H*, *DBH*, *CA*, and *V*) with the fractal geometry parameters (d_{MB} and intercept$_{MB}$) and their interaction with the presence-absence of leaves were also examined using linear mixed models. For these models the fractal geometry parameters, the presence-absence of leaves, and their interaction were considered as fixed effects, while the database of each point cloud as a random effect. The premise around these analyses is that if the presence-absence of leaves do not interact with the fractal geometry parameters to predict tree metrics, a general equation can be used to predict each metric. However, the interaction of the presence-absence of leaves and the fractal parameters tend to affect the prediction of tree metrics as shown in Section 6.3.2. Therefore, linear mixed models were separately applied on leaf-on and leaf-off point clouds to test whether the fractal metrics can predict tree metrics. The fitting regressions were then evaluated using the marginal coefficient of determination (r^2) extracted directly from the model, the mean bias, and the residual standard error (RSE). In this specific case, the bias and RSE were estimated as:

$$Bias\ (P, j) = \frac{P_{est}(j) - P_{obs}(j)}{P_{obs}\ (j)} \tag{3}$$

$$RSE\ (P, j) = \sqrt{\frac{1}{N_j - 1} \sum_{i \in j} (P_{est}\ (i, j) - P_{obs}\ (i, j))^2} \tag{4}$$

where *P* is each estimated (est) and observed (obs) tree metric, and *N* is the number of point clouds (*j*) per leaf-on or leaf-off dataset (*i*). Linear mixed models were applied using the 'lmer' function from the *lme4* package (Bates et al., 2015).

Similar as above, the relationship between log-transformed stand metrics (i.e., V_S, BA_S, and CA_S) and the fractal geometry parameters were examined using simple linear regressions. The fitted model of these relationships was also evaluated using the r^2, mean bias, and RSE. These regressions

were fitted using the 'lm' function of R. Finally, observed and predicted values of tree and stand metrics based on intercept$_{MB}$ equations were compared using analyses of covariance. These comparisons were performed in order to know if a general equation can be applied regardless of the level of observations (tree or stands). These analyses were performed using the 'aov' function of R and the unbalance of samples were addressed using type III error of the 'Anova' function of the *car* package (Fox and Weisberg, 2019).

5.3 Results

5.3.1 Tree metrics

The scanned trees showed considerable variation in their metrics between and within databases (Table 5–2). Regardless of the presence or absence of leaves, trees presented a range in *DBH* between 0.10 and 2.17 m with mean values close to 0.53 ± 0.34 m. Likewise, the range variation of *H* was between 5.55 and 53.62 m with mean values of 24.67 ± 12.14 m. Trees with the highest *DBH* values do not necessarily tend to be the tallest. For example, the highest *DBH* from the CR database was found in *Ficus sp.* (2.17 m) which is associated with an *H* close to 10.50 m, while the tallest tree *Cylicodiscus gabunensis* (*H* = 53.62 m) from the CAM database is associated with a *DBH* of 1.61 m. Moreover, the *CA* varies in a range between 1.12 and 739.72 m^2 with a mean value of 131.82 ± 142.05 m^2. The highest *CA* is associated with the tree with the highest *DBH* (*Ficus sp.*). In the same way, trees presented a higher variation in the estimations of QSM volume with a range between 0.07 and 98.01 m^3 and mean values of 8.10 ± 14.99 m^3.

5.3.2 Fractal geometry and its relationship with tree metrics.

Among all the SMA regressions between *N* and *S* the results show r^2 values higher than 0.95. Leaf-off point clouds seem to present higher and less variable values of r^2 in comparison with leaf-on (Figure 5–3c). However, the linear mixed model reveals that the r^2 values are not affected by the presence-absence of leaves in the point clouds (Table A4–2). The databases considered as random factors in this model explain more than 80% of the r^2 variance. On the other hand, the intercept$_{MB}$ presented a range between 1.27 and 3.39 with a mean of 2.50 ± 0.42, while d_{MB} showed values between 0.45 and 0.73 with a mean of 0.59 ± 0.06. Leaf-on point clouds tend to have higher values of d_{MB} and intercept$_{MB}$ than leaf-off (Figure 5–3a, b); however, these differences are not significant based on the comparisons of the mixed models (Table A4–2). Likewise, the databases

considered as random factors explain more than 84 and 24% of the variance of d_{MB} and the intercept$_{MB}$ in these models (Table A4–2, Figure A4–2). Values per database of the fractal geometry parameters can be found in Table A4–3.

Table 5–2. Mean and standard deviation of the tree metrics per database and presence or absence of leaves. The acronyms of the databases are described in Table 1.

Database	Tree height (m)	DBH (cm)	Crown area (m^2)	QSM volume (m^3)
CAN	11.50 ± 4.04	44.05 ± 31.26	142.59 ± 85.37	3.20 ± 3.30
CR	12.20 ± 4.51	62.80 ± 48.26	155.47 ± 169.85	11.69 ± 21.64
GUY	33.07 ± 2.10	73.66 ± 12.65	238.24 ± 78.31	12.90 ± 4.35
CAM	33.73 ± 12.42	58.17 ± 40.75	142.96 ± 177.47	12.35 ± 20.56
IND	29.09 ± 5.61	58.39 ± 19.16	108.23 ± 70.27	7.46 ± 6.66
PER	40.46 ± 6.95	90.00 ± 23.53	364.12 ± 139.82	21.19 ± 13.62
AUS	18.66 ± 4.19	37.39 ± 15.14	67.03 ± 45.44	1.73 ± 12.29
Leaf-on	21.53 ± 9.21	50.91 ± 29.82	124.41 ± 126.29	6.45 ± 11.55
Leaf-off	29.34 ± 14.34	55.38 ± 39.28	142.88 ± 162.96	10.54 ± 18.80
All databases	24.67 ± 12.14	52.71 ± 33.92	131.84 ± 142.05	8.10 ± 14.99

Results also showed that the parameters extracted from the fractal geometry can predict tree metrics to some extent (Table A4–4); however, it seems that the interaction of the presence-absence of leaves with the fractal geometry parameters tends to affect the prediction of all tree metrics according to the mixed model (Table A4–4). Overall, the variance explained by the random effect of the mixed models is higher in the prediction of metrics by d_{MB} (82 ± 5%) than intercept$_{MB}$ (51 ± 33%) (Table A4–4, Figure A4–3). Due to the above interaction, linear mixed models were performed separately on leaf-off and leaf-on point clouds. These models also revealed that the fractal geometry parameters are reliable predictors of tree metrics (Figure 5–4, Table 5–3). However, the ability to predict tree metrics is more evident when using the intercept$_{MB}$ (Figure 5–4b, d, f, h) rather than d_{MB} (Figure 5–4a, c, e, g). The former is based on the high values of r^2, and lower values of mean bias and RME observed in the relationships with intercept$_{MB}$ rather than d_{MB} (Table 5–3). Likewise, the variance explained by the databases in these mixed models are lower in relationships based on the intercept$_{MB}$ (56 ± 31%) (Table A4–6) than the d_{MB} (86 ± 9%) (Table

A4–5). In addition, the conditional variance-covariance matrixes of the databases seem to be less variable in models based on the intercept$_{MB}$ than the d_{MB} (Figure A4–4). Regardless of the fractal geometry parameter employed, it seems that the prediction of tree metrics tends to be stronger on relationships using leaf-on point clouds than leaf-off point clouds. This is based on the high values of r^2, and the lower values of mean bias and RME observed in the relationships with leaf-on than leaf-off (Table 5–3).
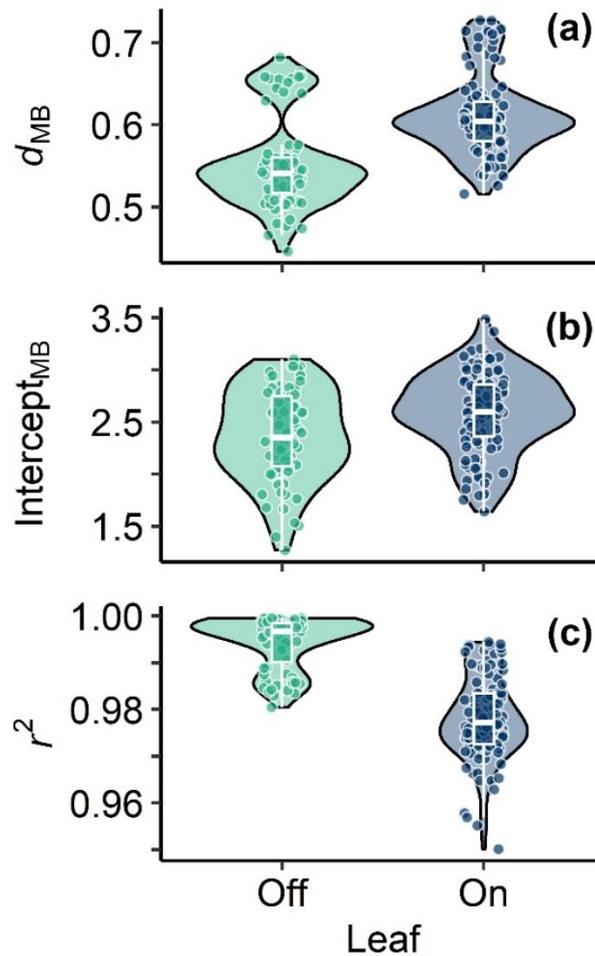
Figure 5–3. Violin plot comparing parameters extracted from fractal geometry using the voxel-counting method on tree point clouds with and without leaves. d_{MB} represents the fractal dimension and r^2 the coefficient of determination.
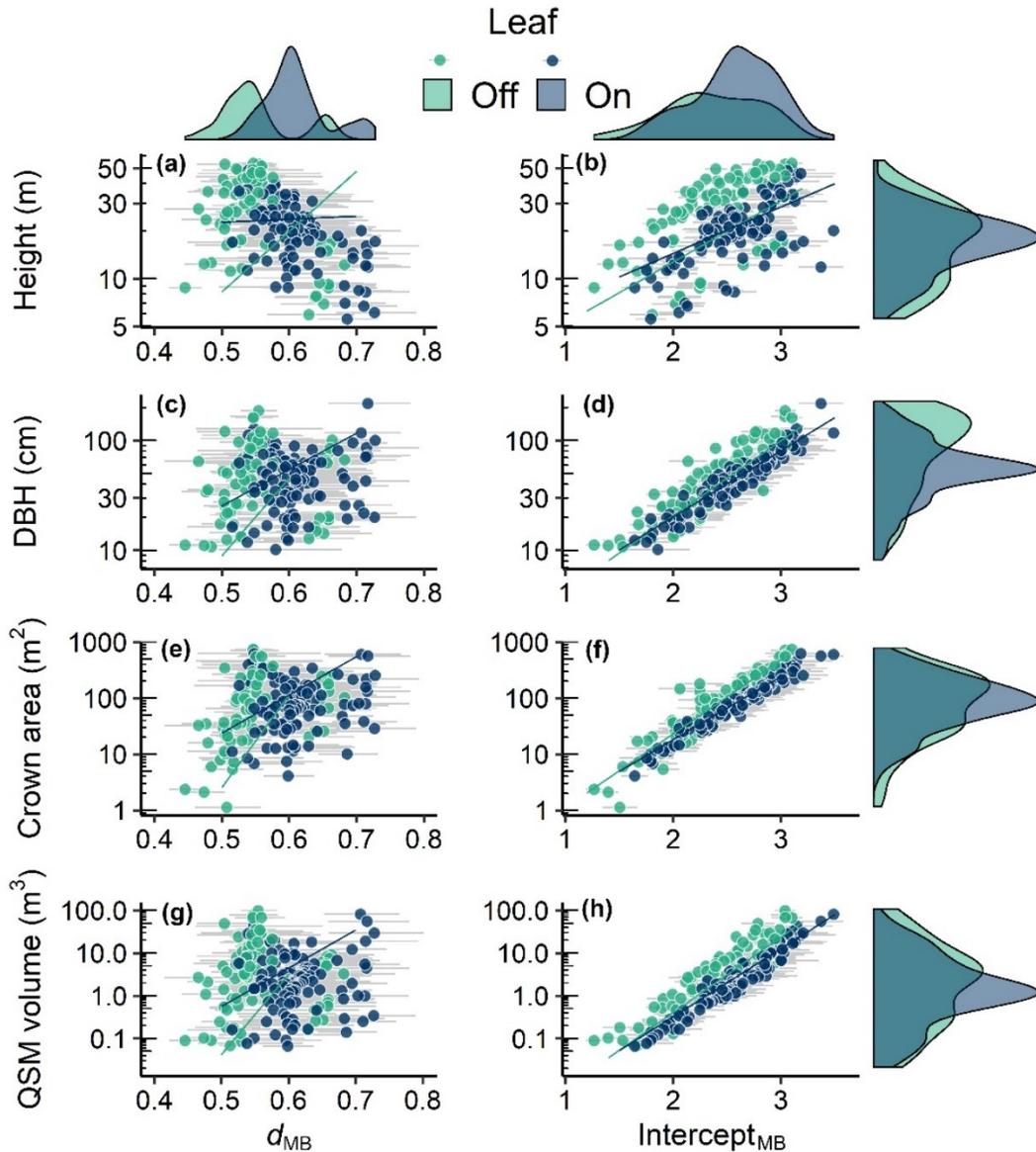
Figure 5–4. Relationship of tree metrics and the fractal geometry parameters (d_{MB}: fractal dimension and intercept$_{MB}$) extracted from the voxel-counting method applied on tree point clouds with and without leaves. Density plots were plotted next to each scatterplot to describe the distribution values. Gray line ranges next to each point represent the upper and lower error. Regressions statistics can be found in Table 5–3.

Table 5–3. Relationship between tree metrics and fractal geometry parameters (fractal dimension (d_{MB}) and intercept$_{MB}$) using linear mixed model. The r^2 represent the marginal coefficient of determination and the asterisks next to each value the significance of the linear parameters. Tree metrics were first log-transformed before fitting. More details of the model statistics can be found in Table A4–5 and A4–6.

Fractal geometry	Tree metrics	Leaf-on					Leaf-off				
		Intercept	Slope	r^2	Bias	RME	Intercept	Slope	r^2	Bias	RME
d_{MB}	H	1.26***	0.19	0.00	0.01	0.11	-0.99	3.81***	0.09	0.02	0.16
	DBH	-0.25	3.31***	0.17	0.02	0.19	-2.60**	7.09***	0.21	0.03	0.26
	CA	-2.03*	6.81***	0.18	0.05	0.32	-7.90**	16.60***	0.32	0.42	0.43
	V	-4.73***	8.96***	0.16	0.15	0.47	-10.75**	18.75***	0.21	-0.67	0.64
Intercept$_{MB}$	H	0.57***	0.30***	0.26	0.00	0.06	0.33	0.39***	0.18	0.00	0.08
	DBH	0.10	0.60***	0.89	0.00	0.07	-0.05	0.68***	0.70	0.01	0.12
	CA	-1.00***	1.13***	0.83	0.01	0.09	-1.27***	1.32***	0.87	0.18	0.21
	V	-3.67***	1.59***	0.90	-0.50	0.12	-3.92**	1.77***	0.74	-0.03	0.24

p-value: * <0.05; ** <0.01, *** <0.001

5.3.3 Fractal geometry and its relationship with stand metrics.

Linear regressions between fractal geometry parameters and stand metrics reveal strong relationships ($r^2 > 0.61$), which are more evident using the intercept$_{MB}$ ($r^2 > 0.95$) than d_{MB} (r^2 between 0.61 and 0.67) (Table 5–4). In both cases, d_{MB} and the intercept$_{MB}$ seem to increase with increases in the number of trees that compose a given stand (Figure 5–5). The comparisons of the observed and predicted metrics between trees and stands reveal that the slopes of these levels of evaluations tend to be different between them ($p < 0.001$) (Table A4–7), where tree relationships present higher slopes than stand relationships (Figure 5–6).
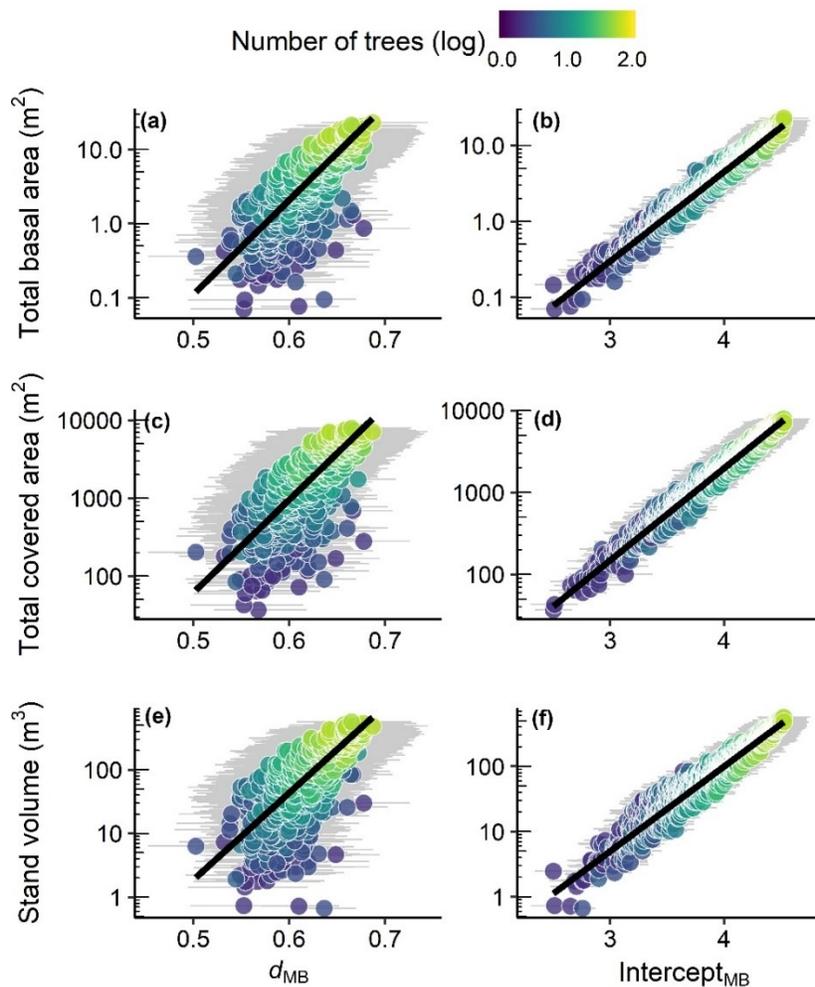
Figure 5–5. Relationship between stand metrics and parameters extracted from the fractal geometry using the voxel counting method. d_{MB} represents the fractal dimension. The color gradient represents the number of trees that compose each artificial stand. Gray line ranges next to each point represent the upper and lower error.

Table 5–4. Results from the relationship of the stand metrics and parameters derived from fractal geometry (fractal dimension (d_{MB}) and intercept$_{MB}$). Stand metrics were first log-transformed before fitting. All the regression parameters were significant at a *p-value* <0.001.

Fractal geometry	Stand metrics	Intercept	Slope	r^2	Bias	RSE
d_{MB}	Total basal area	-7.33 ± 0.15	12.75 ± 0.25	0.67	-0.67	0.27
	Total covered area	-4.18 ± 0.15	11.93 ± 0.25	0.65	0.01	0.27
	Stand volume	-6.56 ± 0.19	13.66 ± 0.30	0.61	0.04	0.33
Intercept$_{MB}$	Total basal area	-4.05 ± 0.02	1.18 ± 0.01	0.98	0.12	0.07
	Total covered area	-1.20 ± 0.02	1.12 ± 0.01	0.98	0.01	0.07
	Stand volume	-3.20 ± 0.03	1.29 ± 0.01	0.95	0.01	0.12

5.4 Discussion

We propose here a novel method to estimate tree-stand metrics from point clouds based on equations derived from fractal geometry. Overall, our results reveal that fractal geometry parameters (d_{HB} and intercept) tend to be, to some extent, reliable predictors of metrics as well as the complexity of point clouds. Since the significance and performance of d_{MB} and the intercept$_{MB}$ differ, the below discussion is focused on highlighting each of these in a forestry and ecological context, as well as on providing future directions in the application of fractal geometry to forest communities.

5.4.1 Relationship of the fractal dimension (d_{MB}) and tree-stand metrics.

Despite the results suggesting that the d_{MB} tends to present a weak relationship with tree or stand metrics, at the tree level it may differentiate the presence of leaves in point clouds, while at the stand level it tends to be associated with the number of trees in a given stand. In general, d_{MB} has been recognized as a descriptor of the structural complexity of point clouds at the tree or stand level (Seidel, 2018; Seidel et al., 2019a, 2019b). Since dense forest stands and leaf-on point clouds tend to present high d_{MB}, these could be considered as objects with high structural complexity. Here, we did not find evidence that the structural complexity described by d_{MB} is affected by tree size or the vertical and horizontal extent of trees as suggested by Seidel, Ehbrecht, Dorji, et al. (2019). Herein, we consider that d_{MB} is mainly affected by how trees occupy their 3D space. For

example, the presence of leaves in trees produces more scattered point clouds than leaf-off trees; increasing the point dispersion in their 3D space and, therefore, their complexity. More scattered point clouds, such as leaf-on, may also reduce the self-similarity between voxels at different scales explaining the lower values of r^2. At the artificial stand level, the random addition of trees to a plot increases the point dispersion; therefore its complexity. In natural stands, Seidel, Ehbrecht, et al. (2019) described that increases in the complexity related to the number of trees tend to be associated with the most complex-structured tree in a given stand. However, other stand features not discussed by these authors could also drive the variability of d_{MB} such as the distance between trees, canopy closure, or the vertical distribution of vegetation.
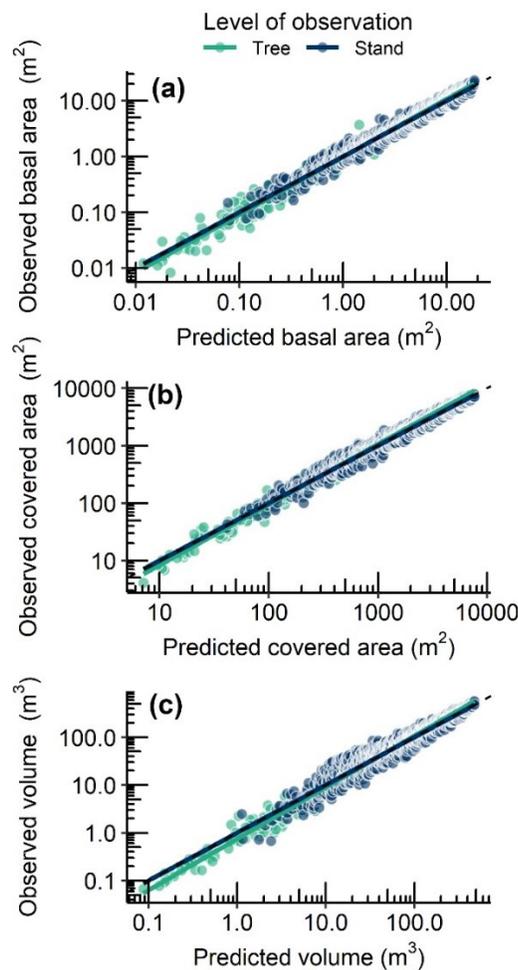
Figure 5–6. Relationship between the observed and predicted metrics from equations based on the intercept$_{MB}$ at two levels of observation (trees and artificial stands). All the predicted values were affected by the level of observation (Table A4–7).

5.4.2 Relationship of the intercept$_{MB}$ derived from fractal geometry and tree-stand metrics.

The intercept$_{MB}$ derived from fractal geometry is the stronger predictor of tree and stand metrics. Although its use is not common in the literature in comparison with d_{MB}, the intercept$_{MB}$ seems to vary concomitantly with the size of trees and stands. Likewise, its relationship with tree metrics tends to be less affected by the random effect of the databases, suggesting that it is a reliable and unbiased parameter that can be used on point clouds processed by different methods. Despite this, the presence or absence of leaves may affect the prediction by the intercept$_{MB}$, and surprisingly, our results suggest that predictions of metrics using leaf-on point clouds are stronger than using leaf-off point clouds. However, it is expected that less scattered point clouds such as leaf-off will lead to better estimations of tree metrics. The latter could be the result of different pre-processing methods and the nature of the leaf-off databases leading to weaker estimations (e.g., scanning trees, and then manual leaf segmentation on point clouds in CAM databases vs. stand scanning of trees without leaves to segment in CAN databases). These potential differences between databases are partially observed in the random effect of the models, which tend to be higher on leaf-off point clouds than leaf-on.

On the other hand, based on the high r^2 and lower RME, it seems that the intercept$_{MB}$ tends to predict stand metrics better when compared with tree metrics. However, the comparison of observed and expected metrics seems to differ between these two levels of evaluation. It is expected that a high number of iterations of artificial stands may lead to confident estimation of the regression parameters, an element that is not present in relationships based on samples trees. Despite this, it is important to highlight that the use of artificial stands may lead to a cumulative sum of errors in the estimation of metrics when adding trees to a given plot due to the un-natural overlap of tree crowns. Therefore, future validation of these method is required using point clouds of natural stands.

5.4.3 Future directions of the application of fractal geometry to point clouds.

The estimation of tree and stand metrics based on fractal geometry equations reveals to be an accurate approach able to predict the variability of structures at different scales. This method can be applied to predict metrics of irregular surfaces since it does not use geometrical structures to fit tree or stand elements (e.g., fitting cylinders on branches). At the tree level, this method could be beneficial to predict metrics on trees with buttress or lianas where branches, stems, and buttress

roots do not tend to be perfectly cylindrical. At the stand level, on the other hand, this method does not require *a priori* tree segmentation; therefore, it avoids errors and the processing time associated with the separation of trees for post-processing.

Using fractals on point clouds also allows us to understand the structural complexity of how trees or stands occupy their 3D-space. This complexity can be further used as an architectural trait towards the plant structural economic spectrum (Verbeeck et al., 2019) or as a possible descriptor of forest ecosystems. At the ecosystem level, for example, the d_{MB} could be used to characterize the level of liana infestation in plots since these tend to reduce the degree of organization of natural spaces, which are typically utilized by trees (Sánchez-Azofeifa et al., 2017). Since the accuracy of prediction of metrics from fractal geometry equations depends on the number of point clouds available, this method will benefit from an increase on the number of tree-stands point clouds repositories around the world. Likewise, the accuracy and validation of fractal geometry equations could be enhanced using field data (e.g., harvest volume or biomass). With the generation of more databases, this method may support processes of calibration and validation of volume and AGB of stands, and therefore, help to develop large-scale biomass maps.

Data accessibility

The TLS data collected by the authors are available at Tropi-Dry dataverse: https://doi.org/10.7910/DVN/DYNAWT. The code used to estimate some tree metrics, apply the voxel-counting method, and create the artificial stands is available at https://github.com/Antguz/rTLS as an R package with the current version 0.1.2 archived at Zenodo (Guzmán et al., 2020; https://doi.org/10.5281/zenodo.3525573).

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Summary of key findings

This dissertation explored new perspectives on the integration of remote technologies that helps to detect and quantify differences between lianas and trees at two Tropical Dry Forests. This was achieved by using passive (Chapters 2 – 4), and active remote sensing (Chapter 5). The following paragraphs present the main results of my chapters as well as their implications.

The temporal variability of leaf temperature (T_{leaf}) of lianas and trees using thermography was addressed in Chapter 2. This chapter expands previous observations using thermocouples for both life forms in a seminal work by Sánchez-Azofeifa et al. (2011) by integrating leaf and ambient temperature from different seasons and ENSO years. A key finding of this chapter reveals that lianas' presence on trees did not affect the temperature of exposed tree leaves; however, liana leaves tended to be warmer than tree leaves at noon. The findings emphasize that lianas are an important biotic factor that can influence canopy temperature, concluding that lianas' presence or absence on trees may have a significant weight on the canopy energy balance, and perhaps, its productivity.

The discrimination of liana and tree leaves using visible-near infrared (VIS-NIR) and longwave infrared (LWIR) spectra were assessed and compared in Chapter 3. In this chapter, both spectral regions (VIS-NIR and LWIR), four representations of leaf spectra, twenty-one algorithms of classification, and two contrasting life forms were compared in the context of machine learning to explore the proper practices to discriminate liana and tree leaves. The main finding of this chapter suggested that both life forms are more accurately discriminated using LWIR spectra (accuracy between 66% and 96%) compared with VIS-NIR spectra (accuracy between 50% and 69%). However, such accuracies of discrimination depended on the kind of spectral representation and machine learning algorithm. The chapter's outcomes suggest the possibility to extend the discrimination between lianas and trees to the airborne or satellite level that uses LWIR sensors.

Chapter 4 addressed the prediction of leaf traits of lianas and trees using Partial Least-Square Regression (PLSR) models based on leaf reflectance or wavelet spectra. Findings from this chapter indicate that observed/predicted relationships between life forms; however the model performance did, mainly in testing datasets. Differences in model performance between life forms seemed to be the product of the intraspecific variability of leaf traits within these life forms. The model comparisons also demonstrated that wavelet spectra help to overcome current limitations of PLSR models based on reflectance spectra, expanding previous observations on the application of wavelet spectra to predict traits by Cheng et al. (2014 and 2011) and Ullah et al. (2012). Compared with models based on reflectance, the integration of wavelet spectra with PLSR models helped to i) reduce the variability and number of optimal components, ii) enhance the spectral bands that drive the prediction of leaf traits, and iii) improve the model performance in terms of accuracy (high R^2 and low RMSE) and precision (low standard deviation). This chapter highlights that the variability of leaf traits between life forms played an important role in evaluating models of leaf traits. Therefore, the intraspecific variability of traits within plant groups, such as these life forms, needs to be considered for the mapping of leaf traits.

The relationship between fractal geometry and tree-stands metrics on point clouds of trees derived from Terrestrial Laser Scanning (TLS) was evaluated in Chapter 5. The findings showed that the intercept extracted from the voxel-counting method (Bunde and Havlin, 1994) (i.e., the relationship between the number of voxels and the inverse of voxel size) is an accurate parameter for predicting metrics at the tree or stand level. Likewise, results revealed that the fractal dimension is strongly associated with the presence/absence of leaves in the point cloud and the number of trees in the stands. Since this method did not depend on predefined cylinder geometries that resemble homogeneous trunks or branches, it may contribute to quantifying the volume of lianas or buttress roots of trees.

6.2 Future directions for the study of life forms using remote sensing

As any doctoral dissertation, several questions may result from previous research chapters. This section aims to provide future directions for some of these emerging inquiries in the context of the dissertation goals. Some of these future directions presented in this section may or may not be detailed previously within each research chapter. Likewise, some of these future directions are associated with the mechanisms that may help to explain or expand some of the observed findings.

6.2.1 Chapter 2: thermography for the study of leaf temperature between life forms

Based on my findings and the current body of literature, it seems clear that lianas and trees present contrasting T_{leaf} at the canopy. In this regard, future studies should address potential drivers that lead to the observed variations as well as their potential impacts on the energy balance at the canopy level as well as forest productivity. As mentioned in section 2.4.2, factors such as leaf inclination, leaf size, and water evapotranspiration may influence T_{leaf}. These factors need to be measured directly on the leaves; therefore, crane facilities like the one described in this chapter could facilitate exploring the variables mentioned above.

Future studies should also expand T_{leaf} measurements to diurnal observations and encompass other forest types in order to know the potential variability of T_{leaf} throughout the day-night and between environments. For instance, lianas present higher evapotranspiration rates and efficiency of water use in comparison with trees (Cai et al., 2009; Chen et al., 2015; De Guzman et al., 2016; Zhu and Cao, 2009). Contrary to the observed results, it could be expected that lianas with higher evapotranspiration rates, and therefore higher transpiration cooling, present lower temperatures compared with trees. As mentioned above, several factors may contribute to the contrasting findings in T_{leaf}. However, based on the chapter observations and the documented strong stomatal closure of lianas close to noon (Chen et al., 2017), I consider that these life forms may present changes in the T_{leaf} behavior within a diurnal cycle. In this sense, it is likely that lianas present coolest leaves in early morning due to their higher evapotranspiration rates and become warmer at noon due to their strong stomatal closure compared with trees. These fluctuations of diurnal T_{leaf} cycles between life forms seem to appear in Sánchez-Azofeifa et al. (2011); however, the authors did not discuss these since their focus was to assess the use of wireless sensing networks for environmental monitoring.

6.2.3 Chapter 3: LWIR spectra for the study of life forms.

The comparisons between spectral regions demonstrate the ability of LWIR spectra over VIS-NIR to accurately classify leaves of lianas and trees in a Tropical Dry Forest. As mentioned in section 3.4.3, these results could inspire future studies to detect these life forms using air- or space-borne technologies. Although the mapping of lianas and trees requires LWIR imagery with high spectral contrast and spatial resolution, current sensor capabilities such as the Hyperspectral

Thermal Emission Spectrometer (HyTES) may help to explore this. For instance, LWIR spectral libraries of lianas and trees, such as those presented in Chapter 2, could be used as a base for endmember extraction when the aim is to identify life forms using HyTES.

Other questions and perspectives not explored in this chapter may arise from the current findings and the literature. For instance, Hesketh and Sánchez-Azofeifa (2012) suggest that the accuracy of lianas and trees discrimination at shorter wavelengths depends on phenological events. In this context, it is not entirely clear how these phenological events influence life forms' discrimination using LWIR spectra. Although Salisbury (1986) suggests that the LWIR leaf spectra of temperate species are characterized by lower temporal variability, a recent study reveals that the earlier stages of leaf development highly influence the LWIR spectra (Richardson et al., 2020). Findings by Richardson et al. (2020) also reveal that slight variations at early stages of leaf development can impact how much longwave energy is emitted by leaves at a given temperature (i.e.,leaf emissivity (ε)); thus, impacting the radiometric T_{leaf} measurements. Richardson's findings together with the results of this chapter may also imply that lianas and trees differ in their ability to emit longwave energy. As such, future studies should explore how phenological events impact the discrimination of life forms using LWIR spectra and how differences in longwave reflectance ($1 - \varepsilon$) impact the expresion of T_{leaf} in these life forms. The latter could also be a complement for future directions associated with Chapter 2.

6.2.4 Chapter 4: Prediction of leaf traits of plant groups.

The intraspecific variability of leaf traits of lianas and trees seems to play an important role in evaluating models that predict leaf traits. In a broad context, the chapter findings may imply the need to consider the differences among plant groups to evaluate leaf spectra's ability to predict leaf traits; especially in studies that deal with contrasting plant groups (e.g., trees vs. grasses). For instance, this approach could be considered necessary by initiatives that attempt to create multibiome models aimed to predict a set of leaf traits (e.g., Serbin et al. (2019) and Wang et al. (2020)). Future studies should also explore whether models based on different leaf traits (e.g., chlorophyll, nitrogen, phosphorus, non-structural carbohydrates, etc.) are also susceptible to the observed variability of model performance between life forms or plant groups. To scale up the observations for the mapping of traits, future studies should explore how the presence of lianas on

trees impacts the prediction of canopy traits. For instance, which are potential biases for predicting canopy traits in forests infested or not by lianas.

The integration of summed-wavelet spectra with PLSR models leads to improved prediction of leaf traits compared with PLSR models based on reflectance spectra. As mentioned in section 4.4.3, questions may remain regarding the number and selection of scales used in the Continuous Wavelet Transformation (CWT) in order to create the summed-wavelet spectra (i.e., the spectral input for PLSR) to predict other leaf traits. For example, Cheng et al. (2014) and Cheng et al. (2011) suggest that different wavelet scales are sensitive to specific spectral features that drive the prediction of leaf traits. As mentioned in Appendix A3–3, this implies that the predictive ability of PLSR models to predict a given leaf trait could be impacted by the selection of scales. Therefore, future studies should explore the potential effect of the selection and combination of wavelet scales with its integration with PLSR models on the prediction of leaf traits. This could be addressed qualitatively in other databases by looking to the variability of absorption features at different scales (e.g., using correlation scalograms) or quantitative by creating and comparing models with different combinations of scales (e.g., such as Appendix A3–3). In a changing world interested in predicting leaf traits for future Earth System models, methods such as presented in Chapter 4 that compare different approaches to predict leaf traits may help guide proper routes for mapping leaf traits. Therefore, future initiatives should develop benchmarks using different databases, spectral domains, and plant groups to address which predictive model could be the most favorable to predict leaf traits.

6.2.5 Chapter 5: fractal geometry to quantify life forms and their complexity.

The ability of fractal geometry equations to accurately predict tree and stand metrics was demonstrated in Chapter 5. The method presented here, associated with the intercept of fractal geometry, may inspire future studies to quantify metrics on irregular structures such as lianas or trees with buttress roots. In the context of liana research, this could be used as the first step in characterizing the effect of lianas on metrics designed to quantify tree architecture. For instance, on segmented point clouds of lianas and trees, it could be possible to apply the presented fractal geometry equations to quickly quantify the volume for both life forms. Likewise, using segmented point clouds of tree species infested and non-infested by lianas, it could be possible to estimate the fractal dimension (d_{MB}) in order to compare how tree species modify the arrangement of their

crowns in the face of mechanical constraints by lianas. On the other hand, the d_{MB} as a descriptor of the structural complexity could advance our understanding of how trees or lianas occupy their 3D space at the individual or plot level. For example, since lianas tend to reduce the degree of organization of natural spaces (Sánchez-Azofeifa et al., 2017), the d_{MB} could be used to describe the presence or degree of liana infestation in a plot. This may have the potential to be applied not just to point clouds derived from TLS, but also to detailed point clouds derived from unmanned aerial vehicles or airborne sensors. Furthermore, the structural complexity inferred by the d_{MB} could be used as a descriptor of forest ecosystems (e.g., phenology –presence/absence of leaves–) or as a plant trait towards the plant structural economic spectrum (Shenkin et al., 2020; Verbeeck et al., 2019).

As mentioned in section 5.4.3, the presented equations rely on the accuracy of estimation of tree metrics on point clouds. Therefore, these results could benefit or be improved with field data (e.g., harvest volume or biomass). In the near future, this method may support processes of calibration and validation of volume and above-ground biomass of trees and stands, especially in this evolving field with an increasing number of point clouds available to the public and the development of new LiDAR sensors every year. In the absence of a full understanding of the presence and the effect of lianas in forest dynamics, future predictions of tropical forest productivity will remain speculative.

REFERENCES

Adams, B.J., Schnitzer, S.A., Yanoviak, S.P., 2017. Trees as islands: canopy ant species richness increases with the size of liana-free trees in a Neotropical forest. Ecography 40, 1067–1075.

Aguirre-Gutiérrez, J., Rifai, S., Shenkin, A., Oliveras, I., Bentley, L.P., Svátek, M., Girardin, C.A.J., Both, S., Riutta, T., Berenguer, E., Kissling, W.D., Bauman, D., Raab, N., Moore, S., Farfan-Rios, W., Figueiredo, A.E.S., Reis, S.M., Ndong, J.E., Ondo, F.E., N'ssi Bengone, N., Mihindou, V., Moraes de Seixas, M.M., Adu-Bredu, S., Abernethy, K., Asner, G.P., Barlow, J., Burslem, D.F.R.P., Coomes, D.A., Cernusak, L.A., Dargie, G.C., Enquist, B.J., Ewers, R.M., Ferreira, J., Jeffery, K.J., Joly, C.A., Lewis, S.L., Marimon-Junior, B.H., Martin, R.E., Morandi, P.S., Phillips, O.L., Quesada, C.A., Salinas, N., Schwantes Marimon, B., Silman, M., Teh, Y.A., White, L.J.T., Malhi, Y., 2021. Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. Remote Sensing of Environment 252, 112122.

Ahdesmaki, M., Zuber, V., Gibb, S., Strimmer, K., 2015. sda: shrinkage discriminant analysis and CAT score variable selection. Available at: https://cran.r-project.org/web/packages/sda/index.html , Accessed date: 8 June 2018.

Andrade, J.L., Meinzer, F.C., Goldstein, G., Schnitzer, S.A., 2005. Water uptake and transport in lianas and co-occurring trees of a seasonally dry tropical forest. Trees 19, 282–289.

Arroyo-Mora, J.P., Sánchez-Azofeifa, G.A., Rivard, B., Calvo, J.C., Janzen, D.H., 2005. Dynamics in landscape structure and composition for the Chorotega region, Costa Rica from 1960 to 2000. Agriculture, Ecosystems & Environment 106, 27–39.

Arshad, M., Ullah, S., Khurshid, K., Ali, A., 2018. Estimation of leaf water content from mid- and thermal-infrared spectra by coupling genetic algorithm and partial least squares regression. Journal of Applied Remote Sensing 12, 22203–22210.

Asner, G.P., Martin, R.E., 2012. Contrasting leaf chemical traits in tropical lianas and trees: implications for future forest composition. Ecology Letters 15, 1001–1007.

Asner, G.P., Martin, R.E., 2011. Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. New Phytologist 189, 999–1012.

Asner, G.P., Martin, R.E., 2008. Spectral and chemical analysis of tropical forests: Scaling from

leaf to canopy levels. Remote Sensing of Environment 112, 3958–3970.

Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright, S.J., Lugo, A.E., 2011. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. Ecological Applications 21, 85–98.

Avalos, G., Mulkey, S.S., 1999. Seasonal changes in liana cover in the upper canopy of a neotropical dry forest. Biotropica 31, 186–192.

Avalos, G., Mulkey, S.S., Kitajima, K., 1999. Leaf optical properties of trees and lianas in the outer canopy of a tropical dry forest. Biotropica 31, 517–520.

Ball, A., Sanchez-Azofeifa, A., Portillo-Quintero, C., Rivard, B., Castro-Contreras, S., Fernandes, G., 2015. Patterns of leaf biochemical and structural properties of cerrado life forms: implications for remote sensing. PLOS ONE 10, e0117659.

Bates, D., Mächler, M., Molker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67, 1–48.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114, 24–31.

Boeriu, C.G., Bravo, D., Gosselink, R.J.A., van Dam, J.E.G., 2004. Characterisation of structure-dependent functional properties of lignin with infrared spectroscopy. Industrial Crops and Products 20, 205–218.

Bongers, F., Ewango, C.E.N., van der Sande, M.T., Poorter, L., 2020. Liana species decline in Congo basin contrasts with global patterns. Ecology 101, e03004.

Bouveyron, C., Girard, S., 2015. robustDA: Robust Mixture Discriminant Analysis. Available at: https://cran.r-project.org/web/packages/robustDA/index.html ,Accessed date: 8 June 2018

Bretfeld, M., Ewers, B.E., Hall, J.S., 2018. Plant water use responses along secondary forest succession during the 2015–2016 El Niño drought in Panama. New Phytologist 219, 885–899.

Buitrago, M.F., Groen, T.A., Hecker, C.A., Skidmore, A.K., 2018a. Spectroscopic determination of leaf traits using infrared spectra. International Journal of Applied Earth Observation and Geoinformation 69, 237–250.

Buitrago, M.F., Skidmore, A.K., Groen, T.A., Hecker, C.A., 2018b. Connecting infrared spectra with plant traits to identify species. ISPRS Journal of Photogrammetry and Remote Sensing 139, 183–200.

Bunde, A., Havlin, S., 1994. A Brief Introduction to Fractal Geometry, in: Bunde, A., Havlin, S.

(Eds.), Fractals in Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–26.

Butler, E.E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K.R., Fazayeli, F., Banerjee, A., Atkin, O.K., Kattge, J., Amiaud, B., Blonder, B., Boenisch, G., Bond-Lamberty, B., Brown, K.A., Byun, C., Campetella, G., Cerabolini, B.E.L., Cornelissen, J.H.C., Craine, J.M., Craven, D., de Vries, F.T., Díaz, S., Domingues, T.F., Forey, E., González-Melo, A., Gross, N., Han, W., Hattingh, W.N., Hickler, T., Jansen, S., Kramer, K., Kraft, N.J.B., Kurokawa, H., Laughlin, D.C., Meir, P., Minden, V., Niinemets, Ü., Onoda, Y., Peñuelas, J., Read, Q., Sack, L., Schamp, B., Soudzilovskaia, N.A., Spasojevic, M.J., Sosinski, E., Thornton, P.E., Valladares, F., van Bodegom, P.M., Williams, M., Wirth, C., Reich, P.B., 2017. Mapping local and global variability in plant trait distributions. Proceedings of the National Academy of Sciences 114, E10937 LP-E10946.

Cai, Z.-Q., Schnitzer, S.A., Bongers, F., 2009. Seasonal differences in leaf-level physiology give lianas a competitive advantage over trees in a tropical seasonal forest. Oecologia 161, 25–33.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., Kaasalainen, M., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods in Ecology and Evolution 6, 198–208.

Calvo-Alvarado, J., McLennan, B., Sánchez-Azofeifa, A., Garvin, T., 2009. Deforestation and forest restoration in Guanacaste, Costa Rica: Putting conservation policies in context. Forest Ecology and Management 258, 931–940.

Cao, S., Yu, Q., Sanchez-Azofeifa, A., Feng, J., Rivard, B., Gu, Z., 2015. Mapping tropical dry forest succession using multiple criteria spectral mixture analysis. ISPRS Journal of Photogrammetry and Remote Sensing 109, 17–29.

Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Caelli, T., 2004. Discrimination of lianas and trees with leaf-level hyperspectral data. Remote Sensing of Environment 90, 353–372.

Chawla, N. V, Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research 16, 321–357.

Chen, Y.-J., Cao, K.-F., Schnitzer, S.A., Fan, Z.-X., Zhang, J.-L., Bongers, F., 2015. Water-use advantage for lianas over trees in tropical seasonal forests. New Phytologist 205, 128–136.

Chen, Y.-J., Schnitzer, S.A., Zhang, Y.-J., Fan, Z.-X., Goldstein, G., Tomlinson, K.W., Lin, H., Zhang, J.-L., Cao, K.-F., 2017. Physiological regulation and efficient xylem water transport

regulate diurnal water and carbon balances of tropical lianas. Functional Ecology 31, 306–317.

Cheng, T., Rivard, B., Sánchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sensing of Environment 115, 659–670.

Cheng, T., Rivard, B., Sánchez-Azofeifa, A.G., Féret, J.-B., Jacquemoud, S., Ustin, S.L., 2014. Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous wavelet analysis. ISPRS Journal of Photogrammetry and Remote Sensing 87, 28–38.

Cheng, T., Rivard, B., Sánchez-Azofeifa, A.G., Féret, J.-B., Jacquemoud, S., Ustin, S.L., 2012. Predicting leaf gravimetric water content from foliar reflectance across a range of plant species using continuous wavelet analysis. Journal of Plant Physiology 169, 1134–1142.

Cheng, T., Rivard, B., Sánchez-Azofeifa, G.A., Feng, J., Calvo-Polanco, M., 2010. Continuous wavelet analysis for the detection of green attack damage due to mountain pine beetle infestation. Remote Sensing of Environment 114, 899–910.

Constantine, W., Percival, D., 2017. wmtsa: Wavelet Methods for Time Series Analysis. Available at: https://cran.r-project.org/web/packages/wmtsa/index.html, Accesseddate: 8 June 2018.

Croat, T., 1978. Flora of Barro Colorado Island. Stanford University Press, Stanford, California, US.

Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sensing of Environment 30, 271–278.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in Ecology. Ecology 88, 2783–2792.

da Cunha Vargas, B., Grombone-Guaratini, M.T., Morellato, L.P.C., 2020. Lianas research in the Neotropics: overview, interaction with trees, and future perspectives. Trees Accepted on 22 October 2020.

Dai, Y., Dickinson, R.E., Wang, Y.-P., 2004. A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. Journal of Climate 17, 2281–2299.

De Deurwaerder, H., Hervé-Fernández, P., Stahl, C., Burban, B., Petronelli, P., Hoffman, B., Bonal, D., Boeckx, P., Verbeeck, H., 2018. Liana and tree below-ground water competition—evidence for water resource partitioning during the dry season. Tree Physiology 38, 1071–1083.

De Guzman, M.E., Santiago, L.S., Schnitzer, S.A., Álvarez-Cansino, L., 2016. Trade-offs between water transport capacity and drought resistance in neotropical canopy liana and tree species. Tree Physiology 37, 1404–1414.

Demmig-Adams, B., 1998. Survey of thermal energy dissipation and pigment composition in sun and shade leaves. Plant and Cell Physiology 39, 474–482.

Demmig-Adams, B., Adams, W.W., 1992. Photoprotection and other responses of plants to high light stress. Annual Review of Plant Physiology and Plant Molecular Biology 43, 599–626.

DeWalt, S.J., Chave, J., 2004. Structure and biomass of four lowland Neotropical forests. Biotropica 36, 7–19.

DeWalt, S.J., Schnitzer, S.A., Chave, J., Bongers, F., Burnham, R.J., Cai, Z., Chuyong, G., Clark, D.B., Ewango, C.E.N., Gerwing, J.J., Gortaire, E., Hart, T., Ibarra-Manríquez, G., Ickes, K., Kenfack, D., Macía, M.J., Makana, J.-R., Martínez-Ramos, M., Mascaro, J., Moses, S., Muller-Landau, H.C., Parren, M.P.E., Parthasarathy, N., Pérez-Salicrup, D.R., Putz, F.E., Romero-Saltos, H., Thomas, D., 2010. Annual rainfall and seasonality predict pan-tropical patterns of liana density and basal area. Biotropica 42, 309–317.

di Porcia e Brugnera, M., Meunier, F., Longo, M., Moorthy, S.M., De Deurwaerder, H., Schnitzer, S.A., Bonal, D., Faybishenko, B., Verbeeck, H., 2019. Modeling the impact of liana infestation on the demography and carbon cycle of tropical forests. Global Change Biology 25, 3767–3780.

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B., Hofton, M., 2003. Above-ground biomass estimation in closed canopy Neotropical forests using LiDAR remote sensing: factors affecting the generality of relationships. Global Ecology and Biogeography 12, 147–159.

Duarte, P., 2015. Package 'HiDimDA.' Available at: https://cran.r-project.org/web/packages/HiDimDA/index.html, Accessed date: 8 June 2018.

Durán, S.M., Gianoli, E., 2013. Carbon stocks in tropical forests decrease with liana density. Biology Letters 9, 20130301.

Durán, S.M., Sánchez-Azofeifa, G.A., 2015. Liana Effects on Carbon Storage and Uptake in Mature and Secondary Tropical Forests, in: Parthasarathy, N. (Ed.), Biodiversity of Lianas. Springer International Publishing, Cham, pp. 43–55.

Durán, S.M., Sánchez-Azofeifa, G.A., Rios, R.S., Gianoli, E., 2015. The relative importance of

climate, stand variables and liana abundance for carbon storage in tropical forests. Global Ecology and Biogeography 24, 939–949.

Elvidge, C.D., 1988. Thermal infrared reflectance of dry plant materials: 2.5–20.0 μm. Remote Sensing of Environment 26, 265–285.

Escós, J., Alados, C.L., Emlen, J.M., 2003. The impact of grazing on plant fractal architecture and fitness of a Mediterranean shrub Anthyllis cytisoides L. Functional Ecology 11, 66–78.

Estrada-Villegas, S., Hall, J.S., van Breugel, M., Schnitzer, S.A., 2020. Lianas reduce biomass accumulation in early successional tropical forests. Ecology 101, e02989.

Fabre, S., Lesaignoux, A., Olioso, A., Briottet, X., 2011. Influence of Water Content on Spectral Reflectance of Leaves in the 3–15 μm Domain. IEEE Geoscience and Remote Sensing Letters 8, 143–147.

Feilhauer, H., Asner, G.P., Martin, R.E., Schmidtlein, S., 2010. Brightness-normalized Partial Least Squares Regression for hyperspectral data. Journal of Quantitative Spectroscopy and Radiative Transfer 111, 1947–1957.

Feng, J., Rogge, D., Rivard, B., 2018. Comparison of lithological mapping results from airborne hyperspectral VNIR-SWIR, LWIR and combined data. International Journal of Applied Earth Observation and Geoinformation 64, 340–353.

Féret, J.-B., le Maire, G., Jay, S., Berveiller, D., Bendoula, R., Hmimina, G., Cheraiet, A., Oliveira, J.C., Ponzoni, F.J., Solanki, T., de Boissieu, F., Chave, J., Nouvellon, Y., Porcar-Castell, A., Proisy, C., Soudani, K., Gastellu-Etchegorry, J.-P., Lefèvre-Fonollosa, M.-J., 2019. Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: Potential and limitations of physical modeling and machine learning. Remote Sensing of Environment 231, 110959.

Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D., 2014. Do we need hundreds of classifiers to solve real world classification problems? Journal of Machine Learning Research 15, 3133–3181.

Fitter, A.H., Stickland, T.R., 1992. Fractal characterization of root system architecture. Functional Ecology 6, 632–635.

Foley, S., Rivard, B., Sanchez-Azofeifa, G.A., Calvo, J., 2006. Foliar spectral properties following leaf clipping and implications for handling techniques. Remote Sensing of Environment 103, 265–275.

Foster, J.R., Townsend, P.A., Zganjar, C.E., 2008. Spatial and temporal patterns of gap dominance by low-canopy lianas detected using EO-1 Hyperion and Landsat Thematic Mapper. Remote Sensing of Environment 112, 2104–2117.

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, Third. ed. Sage, Thousand Oaks, CA.

Gallagher, R. V, Leishman, M.R., 2012. A global analysis of trait variation and evolution in climbing plants. Journal of Biogeography 39, 1757–1771.

Gelman, A., Su, Y.-S., 2016. arm: data analysis using regression and multilevel/hierarchical models. Available at: https://cran.r-project.org/web/packages/arm/index.html, Accessed date: 8 June 2018.

Gentry, A., 1986. Species richness and floristic composition of Choco region plant communities. Caldasia 15, 71–91.

Gentry, A., 1895. An ecotaxonomic survey of Panamanian lianas, in: D'Arcy, W., Correa, M. (Eds.), Historia Natural de Panam. Missouri Botanical Garden, Saint Louis, US, pp. 29–42.

Gentry, A.H., 1991. The distribution and evolution of climbing plants, in: Putz, F.E., Mooney, H.A. (Eds.), The Biology of Vines. Cambridge University Press, Cambridge, UK, pp. 3–49.

Gentry, A.H., 1982. Patterns of neotropical plant species diversity, in: Hecht, M.K., Wallace, B., Prance, G.T. (Eds.), Evolutionary Biology. Springer US, Boston, MA, pp. 1–84.

Ghosh, A., Fassnacht, F.E., Joshi, P.K., Koch, B., 2014. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. International Journal of Applied Earth Observation and Geoinformation 26, 49–63.

Gonzalez de Tanago, J., Lau, A., Bartholomeus, H., Herold, M., Avitabile, V., Raumonen, P., Martius, C., Goodman, R.C., Disney, M., Manuri, S., Burt, A., Calders, K., 2018. Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. Methods in Ecology and Evolution 9, 223–234.

Granados, J., Körner, C., 2002. In deep shade, elevated CO_2 increases the vigor of tropical climbing plants. Global Change Biology 8, 1109–1117.

Greaves, H.E., Vierling, L.A., Eitel, J.U.H., Boelman, N.T., Magney, T.S., Prager, C.M., Griffin, K.L., 2015. Estimating aboveground biomass and leaf area of low-stature Arctic shrubs with terrestrial LiDAR. Remote Sensing of Environment 164, 26–35.

Grossman, Y.L., Ustin, S.L., Jacquemoud, S., Sanderson, E.W., Schmuck, G., Verdebout, J., 1996. Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. Remote Sensing of Environment 56, 182–193.

Grossmann, A., Morlet, J., 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM Journal on Mathematical Analysis 15, 723–736.

Guzmán, J., Hernandez, R., Sanchez-Azofeifa, G., 2020. rTLS: Tools to process point clouds derived from Terrestrial Laser Scanning. rTLS: Tools to process point clouds derived from Terrestrial Laser Scanning. Zenodo, https://doi.org/10.5281/zenodo.3525573

Guzmán, J.A., Rivard, B., Sánchez-Azofeifa, G.A., 2018. Discrimination of liana and tree leaves from a Neotropical Dry Forest using visible-near infrared and longwave infrared reflectance spectra. Remote Sensing of Environment 219, 135–144.

Halley, J.M., Hartley, S., Kallimanis, A.S., Kunin, W.E., Lennon, J.J., Sgardelis, S.P., 2004. Uses and abuses of fractal methodology in ecology. Ecology Letters 7, 254–271.

Harrison, D., Rivard, B., Sánchez-Azofeifa, A., 2018. Classification of tree species based on longwave hyperspectral data from leaves, a case study for a tropical dry forest. International Journal of Applied Earth Observation and Geoinformation 66, 93–105.

Hastie, T., Friedman, J., Tibshirani, R., 2013. The Elements of Statistical Learning: data mining, inference, and prediction, 10th print. ed. Springer Series in Statistics, Berlin, Germany.

Hesketh, M., Sánchez-Azofeifa, G.A., 2012. The effect of seasonal spectral variation on species classification in the Panamanian tropical forest. Remote Sensing of Environment 118, 73–82.

Hilje, B., Calvo-Alvarado, J., Jiménez-Rodríguez, C., Sánchez-Azofeifa, A., 2015. Tree species composition, breeding systems, and pollination and dispersal syndromes in three forest successional stages in a tropical dry forest in Mesoamerica. Tropical Conservation Science 8, 76–94.

Hilje, B., Stack, S., Sánchez-Azofeifa, A., 2017. Lianas abundance is positively related with the avian acoustic community in Tropical Dry Forests. Forests.

Houghton, R.A., Hall, F., Goetz, S.J., 2009. Importance of biomass in the global carbon cycle. Journal of Geophysical Research: Biogeosciences 114, G00E03.

Ichihashi, R., Chiu, C.-W., Komatsu, H., Kume, T., Shinohara, Y., Tateishi, M., Tsuruta, K., Otsuki, K., 2017. Contribution of lianas to community-level canopy transpiration in a warm-temperate forest. Functional Ecology 31, 1690–1699.

Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment 34, 75–91.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning: with Applications in R. Springer New York, New York, NY.

Janzen, D.H., 1988. Tropical dry forests. The most endangered major tropical ecosystem, in: Wilson, E.O. (Ed.), Biodiversity. National Academy Press, Washington, DC, USA, pp. 130–137.

Jones, H.G., 1999. Use of thermography for quantitative studies of spatial and temporal variation of stomatal conductance over leaf surfaces. Plant, Cell & Environment 22, 1043–1055.

Kalacska, M., Bohlman, S., Sánchez-Azofeifa, G.A., Castro-Esau, K., Caelli, T., 2007. Hyperspectral discrimination of tropical dry forest lianas and trees: comparative data reduction approaches at the leaf and canopy levels. Remote Sensing of Environment 109, 406–415.

Kalácska, M., Calvo-Alvarado, J.C., Sánchez-Azofeifa, G.A., 2005. Calibration and assessment of seasonal changes in leaf area index of a tropical dry forest in different stages of succession. Tree Physiology 25, 733–744.

Kalacska, M., Sanchez-Azofeifa, G.A., Calvo-Alvarado, J.C., Quesada, M., Rivard, B., Janzen, D.H., 2004. Species composition, similarity and diversity in three successional stages of a seasonally dry tropical forest. Forest Ecology and Management 200, 227–247.

Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. kernlab - An S4 Package for Kernel Methods in R. Journal of Statistical Software 1, 1–20.

Kilgore, A., Lambert, T.D., Adler, G.H., 2010. Lianas influence fruit and seed use by rodents in a tropical forest. Tropical Ecology 51, 265–271.

Kuhn, M., 2008. Building predictive models in R using the caret package. Journal of Statistical Software 28, 1–26. Available at: https://cran.r-project.org/web/packages/C50/index.html, Accessed date: 8 June 2018.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY.

Kuhn, M., Weston, S., Coulter, N., Culp, M., 2015. C50: C5.0 decision trees and rule-based models.

Kuppler, J., Albert, C.H., Ames, G.M., Armbruster, W.S., Boenisch, G., Boucher, F.C., Campbell, D.R., Carneiro, L.T., Chacón-Madrigal, E., Enquist, B.J., Fonseca, C.R., Gómez, J.M., Guisan, A., Higuchi, P., Karger, D.N., Kattge, J., Kleyer, M., Kraft, N.J.B., Larue-Kontić, A.-

A.C., Lázaro, A., Lechleitner, M., Loughnan, D., Minden, V., Niinemets, Ü., Overbeck, G.E., Parachnowitsch, A.L., Perfectti, F., Pillar, V.D., Schellenberger Costa, D., Sletvold, N., Stang, M., Alves-dos-Santos, I., Streit, H., Wright, J., Zych, M., Junker, R.R., 2020. Global gradients in intraspecific variation in vegetative and floral traits are partially associated with climate and species richness. Global Ecology and Biogeography 29, 992–1007.

Lai, H.R., Hall, J.S., Turner, B.L., van Breugel, M., 2017. Liana effects on biomass dynamics strengthen during secondary forest succession. Ecology 98, 1062–1070.

Leigh, A., Sevanto, S., Close, J.D., Nicotra, A.B., 2017. The influence of leaf size and shape on leaf thermal dynamics: does theory hold up under natural conditions? Plant, Cell & Environment 40, 237–248.

Leuzinger, S., Körner, C., 2007. Tree species diversity affects canopy leaf temperatures in a mature temperate forest. Agricultural and Forest Meteorology 146, 29–37.

Lewis, S.L., Malhi, Y., Phillips, O.L., 2004. Fingerprinting the impacts of global change on tropical forests. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 359, 437–462.

Li, L., Ustin, S.L., Riano, D., 2007. Retrieval of Fresh Leaf Fuel Moisture Content Using Genetic Algorithm Partial Least Squares (GA-PLS) Modeling. IEEE Geoscience and Remote Sensing Letters 4, 216–220.

Li, W., Campos-Vargas, C., Marzahn, P., Sanchez-Azofeifa, A., 2018. On the estimation of tree mortality and liana infestation using a deep self-encoding network. International Journal of Applied Earth Observation and Geoinformation 73, 1–13.

Li, W., Cao, S., Campos-Vargas, C., Sanchez-Azofeifa, A., 2017. Identifying tropical dry forests extent and succession via the use of machine learning techniques. International Journal of Applied Earth Observation and Geoinformation 63, 196–205.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2, 18–22.

Ligges, U., Short, T., Kienzle, P., Schnackenber, S., Billinghurst, D., Borchers, H., Carezia, A., Dupuis, P., Eaton, J., Farhi, E., Hable, K., Hornik, K., Krey, S., Lash, B., Leisch, F., Mersmann, O., Neis, P., Ruohio, J., Smith, J., Stewart, D., Weingessel, A., 2014. signal: signal processing. Available at: https://cran.r-project.org/web/packages/signal/index.html, Accessed date: 8 June 2018.

Lin, D., Xia, J., Wan, S., 2010. Climate warming and biomass accumulation of terrestrial plants: a

meta-analysis. New Phytologist 188, 187–198.

Lin, H., Chen, Y., Zhang, H., Fu, P., Fan, Z., 2017. Stronger cooling effects of transpiration and leaf physical traits of plants from a hot dry habitat than from a hot wet habitat. Functional Ecology 31, 2202–2211.

Londré, R., Schnitzer, S., 2006. The distribution of lianas and their change in abundance in temperate forests over the past 45 years. Ecology 87, 2973–2978.

López, A., Molina-Aiz, F.D., Valera, D.L., Peña, A., 2012. Determining the emissivity of the leaves of nine horticultural crops by means of infrared thermography. Scientia Horticulturae 137, 49–58.

Magnago, L.F.S., Magrach, A., Barlow, J., Schaefer, C.E.G.R., Laurance, W.F., Martins, S.V., Edwards, D.P., 2017. Do fragment size and edge effects predict carbon stocks in trees and lianas in tropical forests? Functional Ecology 31, 542–552.

Mandelbrot, B.B., 1983. The fractal geometry of nature. Freeman, New York.

Mandelbrot, B.B., 1967. How long is the coast of britain? Statistical self-similarity and fractional dimension. Science 156, 636–639.

Martínez-Izquierdo, L., García, M.M., Powers, J.S., Schnitzer, S.A., 2016. Lianas suppress seedling growth and survival of 14 tree species in a Panamanian tropical forest. Ecology 97, 215–224.

Marvin, D.C., Asner, G.P., Schnitzer, S.A., 2016. Liana canopy cover mapped throughout a tropical forest with high-fidelity imaging spectroscopy. Remote Sensing of Environment 176, 98–106.

Medina, E., Sobrado, M., Herrera, R., 1978. Significance of leaf orientation for leaf temperature in an amazonian sclerophyll vegetation. Radiation and Environmental Biophysics 15, 131–140.

Meerdink, S.K., Roberts, D.A., King, J.Y., Roth, K.L., Dennison, P.E., Amaral, C.H., Hook, S.J., 2016. Linking seasonal foliar traits to VSWIR-TIR spectroscopy across California ecosystems. Remote Sensing of Environment 186, 322–338.

Mehmood, T., Liland, K.H., Snipen, L., Sæbø, S., 2012. A review of variable selection methods in Partial Least Squares Regression. Chemometrics and Intelligent Laboratory Systems 118, 62–69.

Meinzer, F.C., Andrade, J.L., Goldstein, G., Holbrook, N.M., Cavelier, J., Jackson, P., 1997. Control of transpiration from the upper canopy of a tropical forest: the role of stomatal, boundary layer and hydraulic architecture components. Plant, Cell & Environment 20, 1242–

1252.

Mello, F.N.A., Estrada-Villegas, S., DeFilippis, D.M., Schnitzer, S.A., 2020. Can functional traits explain plant coexistence? A case study with tropical lianas and trees. Diversity 12, 397.

Messier, J., McGill, B.J., Enquist, B.J., Lechowicz, M.J., 2017. Trait variation and integration across scales: is the leaf economic spectrum present at local scales? Ecography 40, 685–697.

Meunier, F., Verbeeck, H., Cowdery, B., Schnitzer, S.A., Smith-Martin, C.M., Powers, J., Xu, X., Slot, M., De Deurwaerder, H.P.T., Detto, M., Bonal, D., Longo, M., Santiago, L.S., Dietze, M., 2020. Unraveling the relative role of light and water competition between lianas and trees in tropical forests. Journal of Ecology First published: 29 October 2020.

Mevik, B.-H., Wehrens, R., 2007. The pls Package: Principal Component and Partial Least Squares Regression in R. Journal of Statistical Software 18, 1–24.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2017. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. Available at: https://cran.r-project.org/web/packages/e1071/index.html, Accessed date: 8 June 2018.

Meyer, L., Kissling, W.D., Lohmann, L.G., Hortal, J., Diniz-Filho, J.A.F., 2020. Deconstructing species richness–environment relationships in Neotropical lianas. Journal of Biogeography 47, 2168– 2180.

Momo, S.T., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., de Coligny, F., Kamdem, N.G., Libalah, M., Mofack, G.I.I., Le Moguédec, G., Pélissier, R., Barbier, N., 2017a. Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach. Methods in Ecology and Evolution 9, 905–916.

Momo, S.T., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., de Coligny, F., Kamdem, N.G., Libalah, M., Mofack, G.I.I., Le Moguédec, G., Pélissier, R., Barbier, N., 2017b. Data from: Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: a comparison with traditional destructive approach. Dryad Digital Repository. https://doi.org/10.5061/dryad.10hq7

Moorthy, S.M., Bao, Y., Calders, K., Schnitzer, S.A., Verbeeck, H., 2019. Semi-automatic extraction of liana stems from terrestrial LiDAR point clouds of tropical rainforests. ISPRS Journal of Photogrammetry and Remote Sensing 154, 114–126.

Moorthy, S.M., Calders, K., Di Porcia e Brugnera, M., Schnitzer, S.A., Verbeeck, H., 2018. Terrestrial Laser Scanning to Detect Liana Impact on Forest Structure. Remote Sensing 10, 810.

Moorthy, S.M., Raumonen, P., Van den Bulcke, J., Calders, K., Verbeeck, H., 2020. Terrestrial laser scanning for non-destructive estimates of liana stem biomass. Forest Ecology and Management 456, 117751.

Niinemets, Ü., 1999. Research review. Components of leaf dry mass per area – thickness and density – alter leaf photosynthetic capacity in reverse directions in woody plants. New Phytologist 144, 35–47.

Niklas, K.J., 1994. Plant allometry: the scaling of form and process. University of Chicago Press, Chicago.

Ødegaard, F., 2000. The relative importance of trees versus lianas as hosts for phytophagous beetles (Coleoptera) in tropical forests. Journal of Biogeography 27, 283–296.

Oki, Y., Sanchez-Azofeifa, G., Portillo-Quintero, C., Yamarte-Loreto, P., Fernandes, G.W., 2014. Carbon dioxide enrichment effects in the spectral signature of lianas and tree species from Tropical Dry Forests, in: Sanchez-Azofeifa, A., Powers, J.S., Fernandes, G.W., Quesada, M. (Eds.), Tropical Dry Forests in the Americas: Ecology, Conservation, and Management. Taylor & Francis, Boca Raton, FL, USA, pp. 367–373.

Osnas, J.L.D., Katabuchi, M., Kitajima, K., Wright, S.J., Reich, P.B., Van Bael, S.A., Kraft, N.J.B., Samaniego, M.J., Pacala, S.W., Lichstein, J.W., 2018. Divergent drivers of leaf trait variation within species, among species, and among functional groups. Proceedings of the National Academy of Sciences 115, 5480 LP – 5485.

Osnas, J.L.D., Lichstein, J.W., Reich, P.B., Pacala, S.W., 2013. Global Leaf Trait Relationships: Mass, Area, and the Leaf Economics Spectrum. Science 340, 741 LP – 744.

Parolari, A.J., Paul, K., Griffing, A., Condit, R., Perez, R., Aguilar, S., Schnitzer, S.A., 2020. Liana abundance and diversity increase with rainfall seasonality along a precipitation gradient in Panama. Ecography 43, 25–33.

Pérez-Estrada, L.B., Cano-Santana, Z., Oyama, K., 2000. Variation in leaf trichomes of Wigandia urens: environmental factors and physiological consequences. Tree Physiology 20, 629–632.

Peters, A., Torsten, H., 2017. ipred: Improved Predictors. Available at: https://cran.r-project.org/web/packages/ipred/index.html, Accessed date: 8 June 2018.

Phillips, O.L., Vasquez Martinez, R., Arroyo, L., Baker, T.R., Killeen, T., Lewis, S.L., Malhi, Y., Monteagudo Mendoza, A., Neill, D., Nunez Vargas, P., Alexiades, M., Ceron, C., Di Fiore, A., Erwin, T., Jardim, A., Palacios, W., Saldias, M., Vinceti, B., 2002. Increasing dominance of large lianas in Amazonian forests. Nature 418, 770–774.

Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing 32, 4407–4429.

Poorter, H., Niinemets, Ü., Poorter, L., Wright, I.J., Villar, R., 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. New Phytologist 182, 565–588.

Portillo-Quintero, C.A., Sánchez-Azofeifa, G.A., 2010. Extent and conservation of tropical dry forests in the Americas. Biological Conservation 143, 144–155.

Provost, F., Fawcett, T., Kohavi, R., 1998. The case against accuracy estimation for comparing induction algorithms, in: 98 Proceedings of the Fifteenth International Conference on Machine Learning. pp. 445–453.

Prytz, G., Futsaether, C.M., Johnsson, A., 2003. Thermography studies of the spatial and temporal variability in stomatal conductance of Avena leaves during stable and oscillatory transpiration. New Phytologist 158, 249–258.

Putz, F.E., 1984. The natural history of lianas on Barro Colorado Island, Panama. Ecology 65, 1713–1724.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R foundation for Statistical Computing. R foundation for Statistical Computing. Available at: http://www.R-project.org/.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P., 2013. Fast automatic precision tree models from terrestrial laser scanner data. Remote Sensing 5, 491–520.

Rey-Sánchez, A.C., Slot, M., Posada, J.M., Kitajima, K., 2016. Spatial and seasonal variation in leaf temperature within the canopy of a tropical forest. Climate Research 71, 75–89.

Ribeiro da Luz, B., 2006. Attenuated total reflectance spectroscopy of plant leaves: a tool for ecological and botanical studies. New Phytologist 172, 305–318.

Ribeiro da Luz, B., Crowley, J.K., 2010. Identification of plant species by using high spatial and spectral resolution thermal infrared (8.0–13.5 μm) imagery. Remote Sensing of Environment

114, 404–413.

Ribeiro da Luz, B., Crowley, J.K., 2007. Spectral reflectance and emissivity features of broad leaf plants: Prospects for remote sensing in the thermal infrared (8.0–14.0 μm). Remote Sensing of Environment 109, 393–405.

Richardson, A.D., Aubrecht, D.M., Basler, D., Hufkens, K., Muir, C.D., Hanssen, L., 2020. Developmental changes in the reflectance spectra of temperate deciduous tree leaves, and implications for thermal emissivity and leaf temperature. New Phytologist First published: 04 September 2020.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sensing of Environment 112, 2850–2862.

Rodríguez-Ronderos, M.E., Bohrer, G., Sanchez-Azofeifa, A., Powers, J.S., Schnitzer, S.A., 2016. Contribution of lianas to plant area index and canopy structure in a Panamanian forest. Ecology 97, 3271–3277.

Rogers, A., Medlyn, B.E., Dukes, J.S., Bonan, G., von Caemmerer, S., Dietze, M.C., Kattge, J., Leakey, A.D.B., Mercado, L.M., Niinemets, Ü., Prentice, I.C., Serbin, S.P., Sitch, S., Way, D.A., Zaehle, S., 2017. A roadmap for improving the representation of photosynthesis in Earth system models. New Phytologist 213, 22–42.

Ruffin, C., King, R.L., Younan, N.H., 2008. A combined derivative spectroscopy and Savitzky-Golay filtering method for the analysis of hyperspectral data. GIScience & Remote Sensing 45, 1–15.

Salisbury, J.W., 1986. Preliminary measurements of leaf spectral reflectance in the 8-14 μm region. International Journal of Remote Sensing 7, 1879–1886.

Sanchez-Azofeifa, A., Guzmán, J.A., Campos, C.A., Castro, S., Garcia-Millan, V., Nightingale, J., Rankine, C., 2017. Twenty-first century remote sensing technologies are revolutionizing the study of tropical forests. Biotropica 49, 604–619.

Sánchez-Azofeifa, G.A., Castro-Esau, K., 2006. Canopy observations on the hyperspectral properties of a community of tropical dry forest lianas and their host trees. International Journal of Remote Sensing 27, 2101–2109.

Sánchez-Azofeifa, G.A., Castro, K., Wright, S.J., Gamon, J., Kalacska, M., Rivard, B., Schnitzer, S.A., Feng, J.L., 2009a. Differences in leaf traits, leaf internal structure, and spectral

reflectance between two communities of lianas and trees: Implications for remote sensing in tropical environments. Remote Sensing of Environment 113, 2076–2088.

Sánchez-Azofeifa, G.A., Castro, K.L., Rivard, B., Kalascka, M.R., Harriss, R.C., 2003. Remote sensing research priorities in tropical dry forest environments. Biotropica 35, 134–142.

Sánchez-Azofeifa, G.A., Guzmán-Quesada, J.A., Vega-Araya, M., Campos-Vargas, C., Durán, S.M., D'Souza, N., Gianoli, T., Portillo-Quintero, C., Sharp, I., 2017. Can terrestrial laser scanners (TLSs) and hemispherical photographs predict tropical dry forest succession with liana abundance? Biogeosciences 14, 977–988.

Sánchez-Azofeifa, G.A., Kalácska, M., Espírito-Santo, M.M. do, Fernandes, G.W., Schnitzer, S.A., 2009b. Tropical dry forest succession and the contribution of lianas to wood area index (WAI). Forest Ecology and Management 258, 941–948.

Sánchez-Azofeifa, G.A., Quesada, M., Rodríguez, J.P., Nassar, J.M., Stoner, K.E., Castillo, A., Garvin, T., Zent, E.L., Calvo-Alvarado, J.C., Kalacska, M.E.R., Fajardo, L., Gamon, J.A., Cuevas-Reyes, P., 2005. Research priorities for neotropical dry forests. Biotropica 37, 477–485.

Sánchez-Azofeifa, G.A., Rankine, C., do Espirito Santo, M.M., Fatland, R., Garcia, M., 2011. Wireless sensing networks for environmental monitoring: two case studies from tropical forests, in: IEEE Seventh International Conference on EScience. pp. 70–76.

Sánchez-Azofeifa, G.A., Castro-Esau, K., 2006. Canopy observations on the hyperspectral properties of a community of tropical dry forest lianas and their host trees. International Journal of Remote Sensing 27, 2101–2109.

Schnitzer, S., 2015. Increasing liana abundance in neotropical forests: causes and consequences, in: Ecology of Lianas. John Wiley & Sons, Ltd, pp. 451–464.

Schnitzer, S., 2005. A mechanistic explanation for global patterns of liana abundance and distribution. The American Naturalist 166, 262–276.

Schnitzer, S., Bongers, F., 2011. Increasing liana abundance and biomass in tropical forests: emerging patterns and putative mechanisms. Ecology Letters 14, 397–406.

Schnitzer, S., van der Heijden, G., Mascaro, J., Carson, W.P., 2014. Lianas in gaps reduce carbon accumulation in a tropical forest. Ecology 95, 3008–3017.

Schnitzer, S.A., Carson, W.P., 2010. Lianas suppress tree regeneration and diversity in treefall gaps. Ecology Letters 13, 849–857.

Schnitzer, S.A., Michel, N.L., Powers, J.S., Robinson, W.D., 2020. Lianas maintain insectivorous bird abundance and diversity in a neotropical forest. Ecology 101, e03176.

Schnitzer, S.A., van der Heijden, G.M.F., Powers, J.S., 2016. Reply to Verbeeck and Kearsley: Addressing the challenges of including lianas in global vegetation models. Proceedings of the National Academy of Sciences 113, E5–E6.

Seidel, D., 2018. A holistic approach to determine tree structural complexity based on laser scanning data and fractal analysis. Ecology and Evolution 8, 128–134.

Seidel, D., Annighöfer, P., Stiers, M., Zemp, C.D., Burkardt, K., Ehbrecht, M., Willim, K., Kreft, H., Hölscher, D., Ammer, C., 2019a. How a measure of tree structural complexity relates to architectural benefit-to-cost ratio, light availability, and growth of trees. Ecology and Evolution 9, 7134–7142.

Seidel, D., Ehbrecht, M., Annighöfer, P., Ammer, C., 2019b. From tree to stand-level structural complexity — Which properties make a forest stand complex? Agricultural and Forest Meteorology 278, 107699.

Seidel, D., Ehbrecht, M., Dorji, Y., Jambay, J., Ammer, C., Annighöfer, P., 2019c. Identifying architectural characteristics that determine tree structural complexity. Trees 33, 911–919.

Selaya, N.G., Anten, N.P.R., 2008. Differences in biomass allocation, light interception and mechanical stability between lianas and trees in early secondary tropical forest. Functional Ecology 22, 30–39.

Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C., Townsend, P.A., 2014. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. Ecological Applications 24, 1651–1669.

Serbin, S.P., Wu, J., Ely, K.S., Kruger, E.L., Townsend, P.A., Meng, R., Wolfe, B.T., Chlus, A., Wang, Z., Rogers, A., 2019. From the Arctic to the tropics: multibiome prediction of leaf mass per area using leaf reflectance. New Phytologist 224, 1557–1568.

Shao, Y., Lunetta, R.S., 2012. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS Journal of Photogrammetry and Remote Sensing 70, 78–87.

Shenkin, A., Bentley, L.P., Oliveras, I., Salinas, N., Adu-Bredu, S., Marimon-Junior, B.H., Marimon, B.S., Peprah, T., Choque, E.L., Trujillo Rodriguez, L., Clemente Arenas, E.R., Adonteng, C., Seidu, J., Passos, F.B., Reis, S.M., Blonder, B., Silman, M., Enquist, B.J.,

Asner, G.P., Malhi, Y., 2020. The Influence of Ecosystem and Phylogeny on Tropical Tree Crown Size and Shape. Frontiers in Forests and Global Change 3, 109.

Shipley, B., Lechowicz, M., Wright, I., Reich, P., 2006. Fundamental trade-offs generating the worldwide leaf economics spectrum. Ecology 87, 535–541.

Slot, M., Garcia, M.N., Winter, K., 2016. Temperature response of CO2 exchange in three tropical tree species. Functional Plant Biology 43, 468–478.

Slot, M., Rey-Sánchez, C., Gerber, S., Lichstein, J.W., Winter, K., Kitajima, K., 2014a. Thermal acclimation of leaf respiration of tropical trees and lianas: response to experimental canopy warming, and consequences for tropical forest carbon balance. Global Change Biology 20, 2915–2926.

Slot, M., Rey-Sánchez, C., Winter, K., Kitajima, K., 2014b. Trait-based scaling of temperature-dependent foliar respiration in a species-rich tropical forest canopy. Functional Ecology 28, 1074–1086.

Slot, M., Winter, K., 2017. In situ temperature response of photosynthesis of 42 tree and liana species in the canopy of two Panamanian lowland tropical forests with contrasting rainfall regimes. New Phytologist 214, 1103–1117.

Slot, M., Wright, S.J., Kitajima, K., 2013. Foliar respiration and its temperature sensitivity in trees and lianas: in situ measurements in the upper canopy of a tropical forest. Tree Physiology 33, 505–515.

Smith-Martin, C.M., Xu, X., Medvigy, D., Schnitzer, S.A., Powers, J.S., 2020. Allometric scaling laws linking biomass and rooting depth vary across ontogeny and functional groups in tropical dry forest lianas and trees. New Phytologist 226, 714–726.

Stan, K., Sanchez-Azofeifa, A., Calvo-Rodriguez, S., Castro-Magnani, M., Chen, J., Ludwig, R., Zou, L., 2020. Climate change scenarios and projected impacts for forest productivity in Guanacaste Province (Costa Rica): lessons for tropical forest regions. Regional Environmental Change 20, 14.

Statnikov, A., Wang, L., Aliferis, C.F., 2008. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinformatics 9, 319.

Stevens, A., Ramirez-Lopez, L., 2020. An introduction to the prospectr package. Availableat: https://cran.r-project.org/web/packages/prospectr/index.html, Accessed date: 8 June 2018.

Stewart, D., Yahiaoui, N., McDougall, G.J., Myton, K., Marque, C., Boudet, A.M., Haigh, J., 1997. Fourier-transform infrared and Raman spectroscopic evidence for the incorporation of cinnamaldehydes into the lignin of transgenic tobacco (Nicotiana tabacum L.) plants with reduced expression of cinnamyl alcohol dehydrogenase. Planta 201, 311–318.

Stewart, T.E., Schnitzer, S.A., 2017. Blurred lines between competition and parasitism. Biotropica 49, 433–438.

Streher, A.S., Torres, R. da S., Morellato, L.P.C., Silva, T.S.F., 2020. Accuracy and limitations for spectroscopic prediction of leaf traits in seasonally dry tropical environments. Remote Sensing of Environment 244, 111828.

Suzuki, M.T., 2007. A three dimensional box counting method for measuring fractal dimensions of 3D models, in: Proceedings of the Eleventh IASTED International Conference on Internet and Multimedia Systems and Applications, IMSA '07. ACTA Press, Anaheim, CA, USA, pp. 42–47.

Swenson, N.G., Hulshof, C.M., Katabuchi, M., Enquist, B.J., 2020. Long-term shifts in the functional composition and diversity of a tropical dry forest: a 30-yr study. Ecological Monographs n/a, e01408.

Therneau, T., Atkinson, B., Ripley, B., 2017. rpart: recursive partitioning and regression trees. Available at: https://cran.r-project.org/web/packages/rpart/index.html, Accessed date: 8 June 2018.

Todorov, V., Filzmoser, P., 2009. An object-oriented framework for robust multivariate analysis. Journal of Statistical Software 32, 1–47.

Torgo, L., 2010. Data Mining with R, learning with case studies. Chapman and Hall/CRC.

Trumbore, S., Brando, P., Hartmann, H., 2015. Forest health and global change. Science 349, 814–818.

Ullah, S., Schlerf, M., Skidmore, A.K., Hecker, C., 2012a. Identifying plant species using mid-wave infrared (2.5–6μm) and thermal infrared (8–14μm) emissivity spectra. Remote Sensing of Environment 118, 95–102.

Ullah, S., Skidmore, A.K., Naeem, M., Schlerf, M., 2012b. An accurate retrieval of leaf water content from mid to thermal infrared spectra using continuous wavelet analysis. Science of The Total Environment 437, 145–152.

Usamentiaga, R., Venegas, P., Guerediaga, J., Vega, L., Molleda, J., Bulnes, F.G., 2014. Infrared

thermography for temperature measurement and non-destructive testing. Sensors 14, 12305–12348.

van Bodegom, P.M., Douma, J.C., Verheijen, L.M., 2014. A fully traits-based approach to modeling global vegetation distribution. Proceedings of the National Academy of Sciences 111, 13733 LP – 13738.

van der Heijden, G., Powers, J., Schnitzer, S., 2015. Lianas reduce carbon accumulation and storage in tropical forests. Proceedings of the National Academy of Sciences 112, 13267–13271.

van Melis, J., Camargo, M.G.G., Carvalho, P.G., Morellato, L.P.C., Grombone-Guaratini, M.T., 2020. Contrasting edge effect on lianas and trees in a cerrado savanna remnant. Austral Ecology First published: 03 November 2020.

van Noordwijk, M., Mulia, R., 2002. Functional branch analysis as tool for fractal scaling above- and belowground trees for their additive and non-additive properties. Ecological Modelling 149, 41–51.

Vargas, G., Cordero, R.A., 2013. Photosynthetic responses to temperature of two tropical rainforest tree species from Costa Rica. Trees 27, 1261–1270.

Venables, W., Ripley, B., 2002. Modern Applied Statistics with S, Fourth. ed. Springer, New York, USA.

Verbeeck, H., Bauters, M., Jackson, T., Shenkin, A., Disney, M., Calders, K., 2019. Time for a plant structural economics spectrum. Frontiers in Forests and Global Change 2, 43.

Verbeeck, H., Kearsley, E., 2016. The importance of including lianas in global vegetation models. Proceedings of the National Academy of Sciences 113, E4–E4.

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007. Let the concept of trait be functional! Oikos 116, 882–892.

Visser, M.D., Schnitzer, S.A., Muller-Landau, H.C., Jongejans, E., de Kroon, H., Comita, L.S., Hubbell, S.P., Wright, S.J., 2018. Tree species vary widely in their tolerance for liana infestation: A case study of differential host response to generalist parasites. Journal of Ecology 106, 781–794.

Waite, C.E., van der Heijden, G.M.F., Field, R., Boyd, D.S., 2019. A view from above: Unmanned aerial vehicles (UAVs) provide a new tool for assessing liana infestation in tropical forest canopies. Journal of Applied Ecology 56, 902–912.

Wang, Z., Chlus, A., Geygan, R., Ye, Z., Zheng, T., Singh, A., Couture, J.J., Cavender-Bares, J.,

Kruger, E.L., Townsend, P.A., 2020. Foliar functional traits from imaging spectroscopy across biomes in eastern North America. New Phytologist 228, 494–511.

Warton, D.I., Duursma, R.A., Falster, D.S., Taskinen, S., 2012. smatr 3– an R package for estimation and inference about allometric lines. Methods in Ecology and Evolution 3, 257–259.

Weihs, C., Ligges, U., Luebke, K., Raabe, N., 2005. klaR analyzing german business cycles, in: Baier, D., Decker, R., Schmidt-Thieme, L. (Eds.), Data Analysis and Decision Support. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 335–343.

West, G.B., Brown, J.H., Enquist, B.J., 1999. A general model for the structure and allometry of plant vascular systems. Nature 400, 664.

Wiklund, S., Nilsson, D., Eriksson, L., Sjöström, M., Wold, S., Faber, K., 2007. A randomization test for PLS component selection. Journal of Chemometrics 21, 427–439.

Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., Gonzalez de Tanago, J., Bartholomeus, H., Brede, B., Herold, M., 2017. Data acquisition considerations for Terrestrial Laser Scanning of forest plots. Remote Sensing of Environment 196, 140–153.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares, in: Krishnaiaah, P. (Ed.), Multivariate Analysis. Academic Press, New York, USA, pp. 391–420.

Wolpert, D.H., 1996. The lack of a priori distinctions between learning algorithms. Neural Computation 8, 1341–1390.

Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation 1, 67–82.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., Villar, R., 2004. The worldwide leaf economics spectrum. Nature 428, 821–827.

Wu, J., Rogers, A., Albert, L.P., Ely, K., Prohaska, N., Wolfe, B.T., Oliveira Jr, R.C., Saleska, S.R., Serbin, S.P., 2019. Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment and leaf age in lowland moist tropical forests.

New Phytologist 224, 663–674.

Wyka, T.P., Oleksyn, J., Karolewski, P., Schnitzer, S.A., 2013. Phenotypic correlates of the lianescent growth form: a review. Annals of Botany 112, 1667–1681.

Yang, H., Chen, W., Qian, T., Shen, D., Wang, J., 2015. The extraction of vegetation points from LiDAR using 3D fractal dimension analyses. Remote Sensing 7, 10815–10831.

Yang, X., Tang, J., Mustard, J.F., Wu, J., Zhao, K., Serbin, S., Lee, J.-E., 2016. Seasonal variability of multiple leaf traits captured by leaf spectroscopy at two temperate deciduous forests. Remote Sensing of Environment 179, 1–12.

Yang, Y., Zhu, Q., Peng, C., Wang, H., Chen, H., 2015. From plant functional types to plant functional traits: A new paradigm in modelling global vegetation dynamics. Progress in Physical Geography: Earth and Environment 39, 514–535.

Yuan, X., Laakso, K., Marzahn, P., Sanchez-Azofeifa, G.A., 2019. Canopy temperature differences between liana-infested and non-liana infested areas in a Neotropical Dry Forest. Forests.

Zhang, J., Rivard, B., Sánchez-Azofeifa, A., Castro-Esau, K., 2006. Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. Remote Sensing of Environment 105, 129–141.

Zhu, S.-D., Cao, K.-F., 2009. Hydraulic properties and photosynthetic rates in co-occurring lianas and trees in a seasonal tropical rainforest in southwestern China. Plant Ecology 204, 295–304.

Zotz, G., Cueni, N., Körner, C., 2006. In situ growth stimulation of a temperate zone liana (Hedera helix) in elevated CO2. Functional Ecology 20, 763–769.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R. Springer Science and Business Media, New York, NY, USA.

APPENDICES

Appendix 1. Supporting information for Chapter 2.

Table A1–1. Number of trees (NT), thermal infrared images collected (TIR), and leaves selected (LS) to address the first hypothesis during the wet and dry season in contrasting ENSO years at the canopy of Parque Natural Metropolitano, Panama. NL: host trees without lianas; L: host trees with lianas.

ENSO	Season	Condition	A. excelsum			A. spraguei			C. elastica			L. seemannii		
			NT	TIR	LS	NT	TIR	LS	NT	TIR	LS	NT	TIR	LS
El Niño	Wet	NL	4	14	64	2	3	15	4	10	54	2	6	30
El Niño	Wet	L	4	17	58	2	8	39	4	16	74	2	4	10
El Niño	Dry	NL	2	8	40	2	4	20	2	5	25	2	6	30
El Niño	Dry	L	2	6	30	2	7	34	2	6	30	2	7	34
La Niña	Wet	NL	2	5	25	---	---	---	2	5	25	2	3	19
La Niña	Wet	L	2	6	25	---	---	---	2	7	35	2	5	16
La Niña	Dry	NL	3	9	45	---	---	---	3	11	54	3	8	45
La Niña	Dry	L	3	6	30	---	---	---	2	7	34	2	5	25

Table A1–2. Number of host trees (HT), thermal infrared images collected (TIR), and leaves selected (LS) to address the second hypothesis during the wet and dry season in contrasting ENSO years at the canopy of Parque Natural Metropolitano, Panama.

ENSO	Season	Leaf type	Species											
			A. excelsum			*A. spraguei*			*C. elastica*			*L. seemannii*		
			HT	TIR	LS	NT	TIR	LS	NT	TIR	LS	NT	TIR	LS
El Niño	Wet	Tree	4	17	58	2	8	39	4	16	74	2	4	10
El Niño	Wet	Liana		25	119		9	45		14	69		5	25
El Niño	Dry	Tree	2	6	30	2	7	34	2	6	30	2	7	34
El Niño	Dry	Liana		7	35		6	29		6	30		7	35
La Niña	Wet	Tree	2	6	25	---	---	---	2	7	35	2	5	19
La Niña	Wet	Liana		6	30	---	---	---		7	35		5	26
La Niña	Dry	Tree	3	6	30	---	---	---	2	7	34	2	5	25
La Niña	Dry	Liana		5	25	---	---	---		7	33		5	25
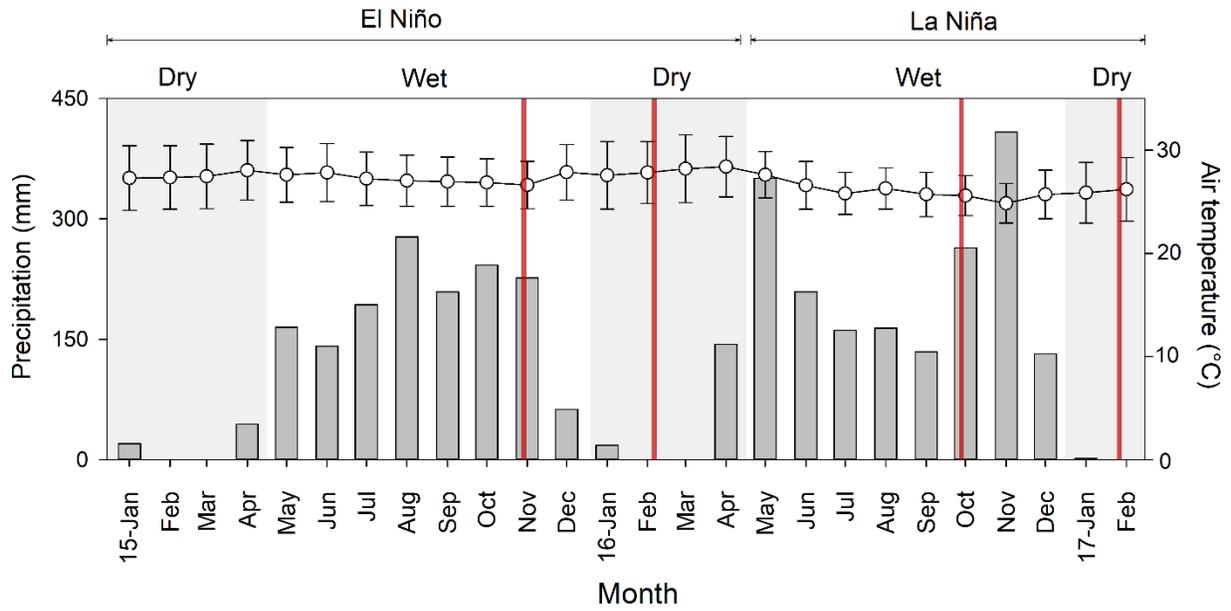
Figure A1–1. Monthly total precipitation and average monthly air temperature during wet and dry seasons in contrasting ENSO years at Parque Natural Metropolitano, Panama. Vertical red lines represent the campaigns of data collection.
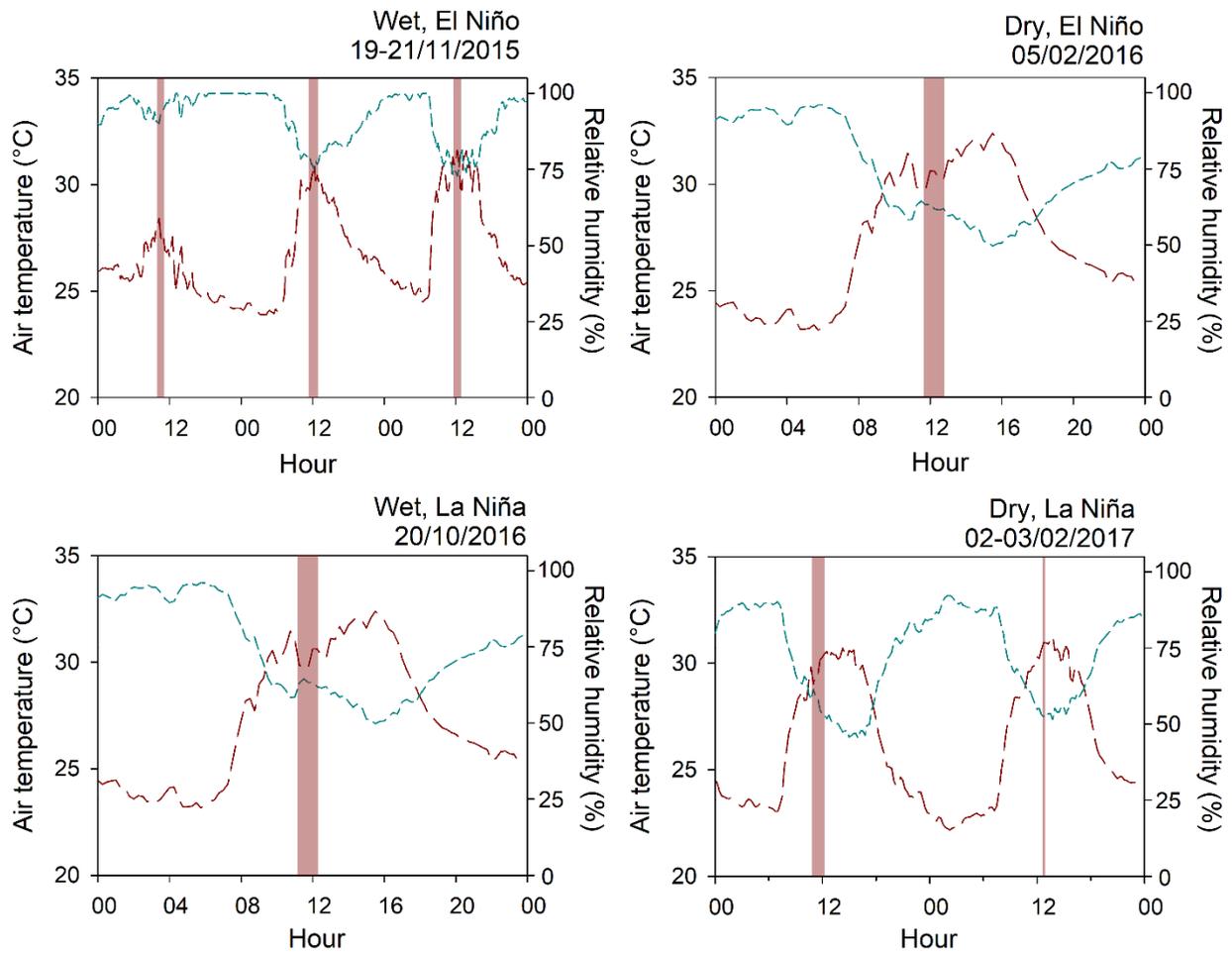
Figure A1–2. Mean values of temperature (red lines) and relative humidity (blue lines) during the days of data collection according to the meteorological station on the crane at Parque Natural Metropolitano, Panama. Vertical red bars represent the time of data collection.

Appendix 2. Supporting information for Chapter 3.
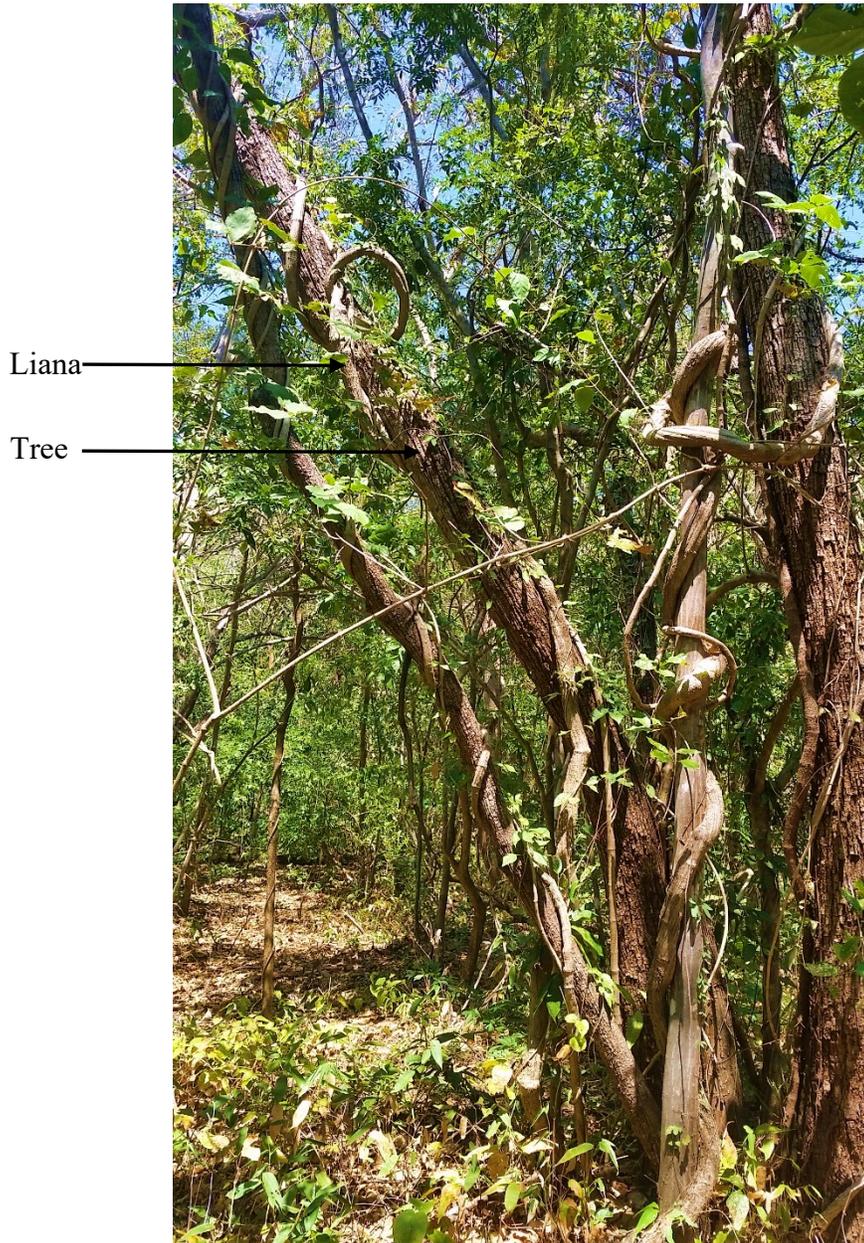
Liana

Tree

Figure A2–1. Picture of a liana growing around a tree at the Santa Rosa National Park, Costa Rica.
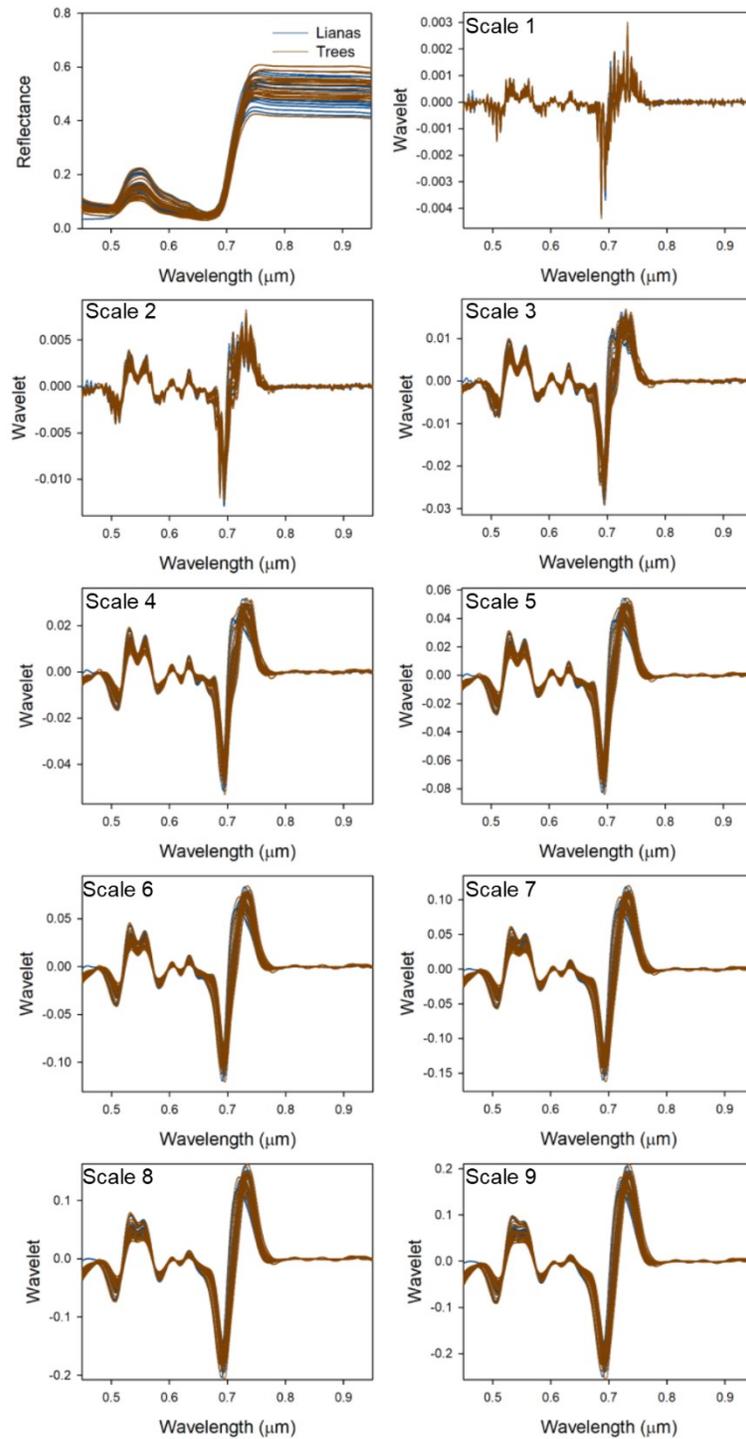
Figure A2–2. Wavelet spectra at scales 1 to 9 extracted from the continuous wavelet transformation applied to the VIS-NIR spectral libraries. Each line represents the wavelet transformed averaged of each species. For reference the upper left frame shows the reflectance spectra from which the wavelets were derived.
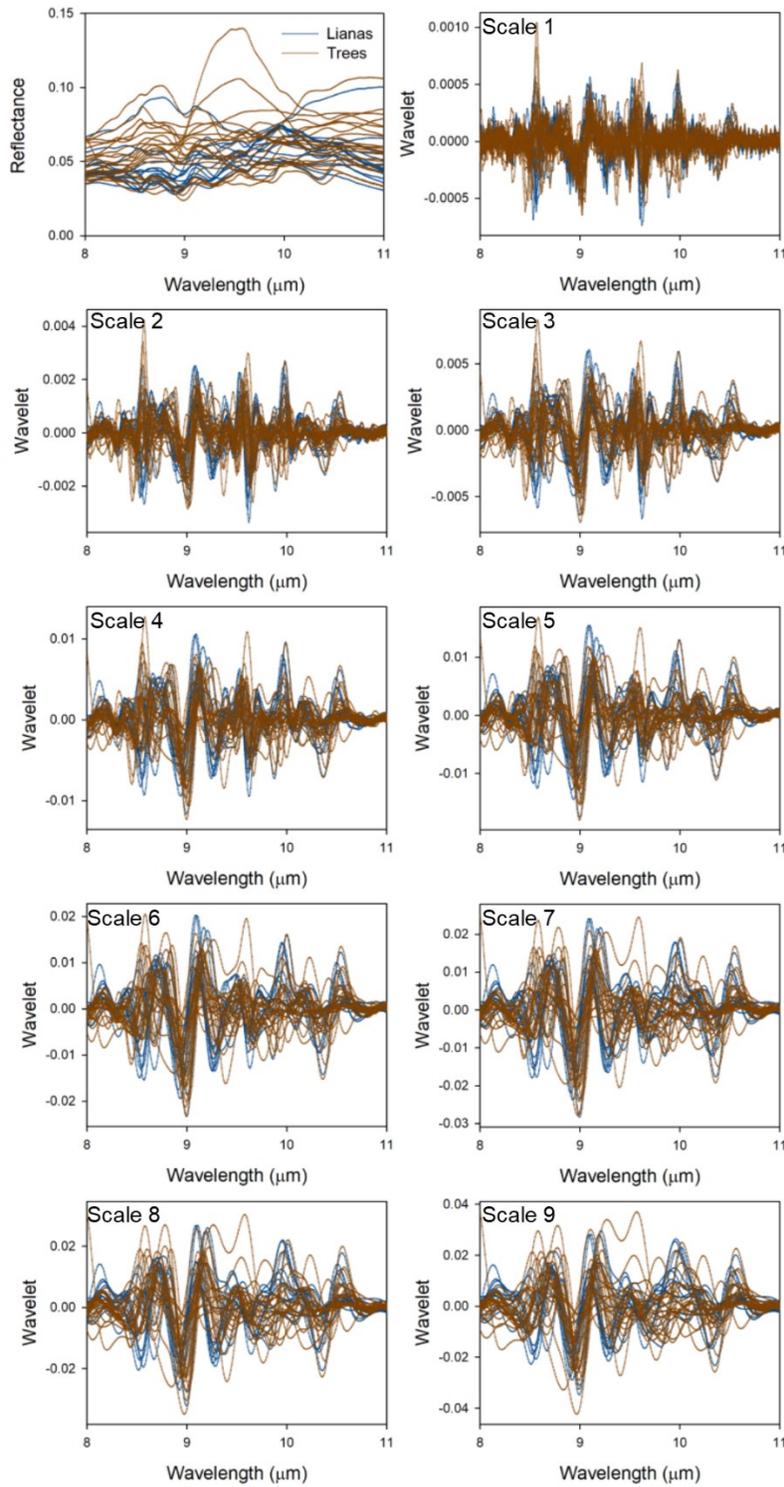
Figure A2–3. Wavelet spectra at scales 1 to 9 extracted from the continuous wavelet transformation applied to the LWIR spectral libraries. Each line represents the wavelet transformed averaged for each species. For reference the upper left frame shows the reflectance spectra from which the wavelets were derived.
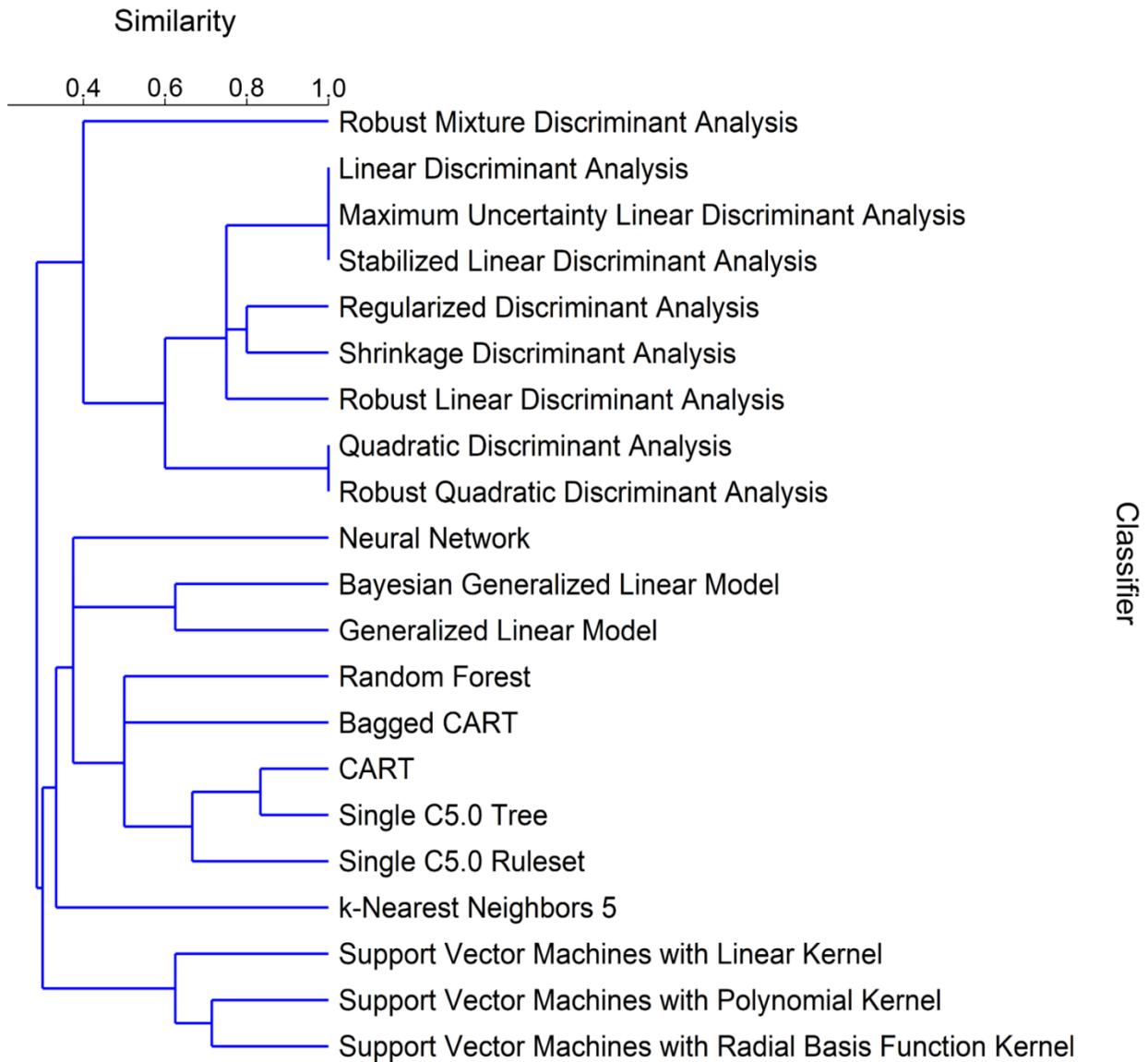
Figure A2–4. Cluster of Jacquard similarity of the classifiers used to discriminate liana and tree leaves based on the reflectance in the visible-near infrared or longwave infrared reflectance region. Data source: http://topepo.github.io/caret/models-clustered-by-tag-similarity.html

Appendix 3. Supporting information for Chapter 4.

Appendix A3–1. Optimal combination of summed wavelets for predicting leaf traits.

From the application of the continuous wavelet transformation to reflectance spectra, it is possible to isolate scales capturing different spectral features (Rivard et al., 2008). Wavelet scales that best describe the spectral features can be summed, and these summed spectra could be used in the ensuing models. However, different wavelet scales are sensitive to specific spectral features that drive the prediction of leaf traits (Cheng et al., 2014, 2011; Ullah et al., 2012b); therefore, the selection of scales to be summed may have an influence on the predictive ability of PLSR models. This appendix explores 372 potential combinations of summed wavelets to predict leaf traits. The extraction of wavelets was performed following Section 4.2.3. The resulting spectra include the potential combination of wavelets using a total of 2, 3, 4, and 5 scales. Using summed wavelet spectra of the samples selected for training purposes (Section 4.2.4), PLSR models were built for each leaf trait (i.e., LMA, WC, and EWT). The wavelet spectra used in the PLSR encompass the spectral range described at the end of Section 4.2.3. Overall, for a given summed wavelet spectra and leaf trait, a PLSR model was created to first estimate the optimal number of components following Section 4.2.4.1. Once the optimal number of components was estimated, descriptors of model performance (R^2 and RMSE; Section 4.2.4.3) were computed from the same PLSR model. The selection of the optinam summed wavelet combination that best predict the leaf traits was based on the model's performance (low RMSE and high R^2). Thus, this exploratory analysis reveals that the sum of scales 1, 2, 3, 8, and 9 consistently leads to high-performance models for predicting the three-leaf traits (Figure A3–1). This sum of scales was selected as the optimum wavelet spectra for the study.

References

Cheng, T., Rivard, B., Sánchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sensing of Environment 115, 659–670.

Cheng, T., Rivard, B., Sánchez-Azofeifa, A.G., Féret, J.-B., Jacquemoud, S., Ustin, S.L., 2014. Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using

continuous wavelet analysis. ISPRS Journal of Photogrammetry and Remote Sensing 87, 28–38.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sensing of Environment 112, 2850–2862.

Ullah, S., Skidmore, A.K., Naeem, M., Schlerf, M., 2012. An accurate retrieval of leaf water content from mid to thermal infrared spectra using continuous wavelet analysis. Science of The Total Environment 437, 145–152.

Table A3–1. Liana and tree species, and their measured leaf traits collected at the Santa Rosa National Park – Environmental Monitoring Super Site, Costa Rica. Leaf Mass Area (LMA), Water Content (WC), and Equivalent Water Thickness (EWT). The values represent the mean and standard deviation of 20 leaves.

Family	Species	Leaf traits		
		LMA (g m^{-2})	WC (%)	EWT (g m^{-2})
Trees				
Apocynaceae	*Stemmadenia obovate*	33.24 ± 3.7	77.99 ± 2.73	119.54 ± 20.51
Bignoniaceae	*Crescentia alata*	95.03 ± 10.56	68.92 ± 1.44	211.04 ± 24.83
Burseraceae	*Bursera simarouba*	66.65 ± 6.52	68.79 ± 1.65	146.93 ± 12.8
Dilleniaceae	*Curatella Americana*	145.73 ± 15.42	55.82 ± 1.92	184.7 ± 24.43
Euphorbiaceae	*Jatropha curcas*	56.91 ± 13.01	69.03 ± 4.51	125.1 ± 5.95
Euphorbiaceae	*Sapium glandulosum*	103.27 ± 35.4	57.49 ± 4.45	142.81 ± 53.74
Fabaceae/Caesalpinoideae	*Bauhinia ungulate*	54.33 ± 7.53	60.53 ± 3.76	84.41 ± 16.24
Fabaceae/Caesalpinoideae	*Hymenaea courbaril*	56.08 ± 14.32	50.76 ± 3.1	57.49 ± 13.73
Fabaceae/Papilionoideae	*Gliricidia sepium*	55.29 ± 7.23	74.98 ± 1.76	165.82 ± 19.35
Fagaceae	*Quercus oleoides*	173.19 ± 24.42	46.71 ± 6.03	152.68 ± 28.48
Hippocrateaceae	*Semialarium mexicanum*	76.3 ± 19.01	69.23 ± 4.44	168.54 ± 17.42
Lauraceae	*Ocotea veraguensis*	100.57 ± 7.93	52.44 ± 2.01	110.77 ± 6.64
Malpighiaceae	*Byrsonima crassifolia*	120.97 ± 18.7	65.49 ± 3.24	228.42 ± 20.59
Malvaceae	*Guazuma ulmifolia*	72.49 ± 11.71	67.6 ± 5.28	154.78 ± 37.94
Meliaceae	*Cedrela odorata*	49.39 ± 7.95	72.88 ± 4.13	132.97 ± 13.37
Meliaceae	*Trichilia americana*	51.12 ± 5.48	78.14 ± 2.36	183.21 ± 14.58
Nyctaginaceae	*Pisonia aculeata*	37.4 ± 3.7	72.72 ± 1.28	99.67 ± 9.17
Sapindaceae	*Cochlospermum vitifolium*	73.75 ± 6.73	64.81 ± 2.58	136.56 ± 17.69
Simaroubaceae	*Simarouba glauca*	147.39 ± 15.41	57.42 ± 3.25	199.76 ± 27.28
Tiliaceae	*Luehea speciose*	31.92 ± 3.51	68.83 ± 1.87	70.72 ± 9.1
Verbenaceae	*Rehdera trinervis*	96.5 ± 12.69	69.06 ± 2.59	215.36 ± 21.21
Lianas				
Apocynaceae	*Forsteronia sp.*	58.85 ± 6.89	67.02 ± 3.22	119.54 ± 9.19
Apocynaceae	*Forsteronia spicata*	31.86 ± 6.34	73.4 ± 3.99	87.28 ± 4.56
Bignoniaceae	*Arrabidaea chica*	31.9 ± 4.76	77.14 ± 2.23	107.23 ± 7.44
Bignoniaceae	*Cydista aequinoctialis*	30.03 ± 4.42	79 ± 3.14	113.52 ± 12.93
Bignoniaceae	*Cydista diversifolia*	34.67 ± 4.64	69.31 ± 2.19	78.22 ± 8.59

Bignoniaceae	*Paulinia sp.*	55.87 ± 6.4	65.55 ± 3.18	106.13 ± 6.59
Cucurbitaceae	*Cayaponia racemose*	22.58 ± 4.69	86.4 ± 1.61	143.32 ± 26.28
Dilleniaceae	*Tetracera volubilis*	31.29 ± 8.11	81.26 ± 2.67	135.7 ± 31.59
Malpighiaceae	*Heteropterys panamensis*	67.91 ± 7.93	63.75 ± 1.81	118.97 ± 7.36
Malpighiaceae	*Heteropterys sp.*	47.32 ± 7.39	69.16 ± 3.09	105.94 ± 12.8
Malpighiaceae	*Hiraea reclinate*	74.56 ± 11.61	62.53 ± 2.18	124.07 ± 15.85
Rhamnaceae	*Gouania polygama*	38.91 ± 9.96	71.14 ± 4.41	94.04 ± 12.01
Sapindaceae	*Serjania atrolineata*	75.99 ± 10.28	64.32 ± 3.13	136.56 ± 10.81
Sapindaceae	*Serjania schiedeana*	42.69 ± 7.2	75.31 ± 3.82	130.14 ± 10.39

Table A3–2. Results from the linear mixed models comparing the leaf traits between lianas and trees. Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Values between parentheses represent confidence intervals, and the asterisks next to each value the significance of the linear parameters. σ^2 describes the variance of the model, τ_{00} the variance of the databases as a random effect, and ICC the interclass correlation coefficients. The conditional variance-covariance matrixes of the random effects can be found in Figure A3–2.

	Leaf trait (dependent variable)		
	LMA	WC	EWT
Intercept	1.63***	1.85***	2.05 **
	(1.52 – 1.73)	(1.82 – 1.88)	(1.98 – 2.12)
Life form	0.23**	-0.04*	0.09
	(0.09 – 0.36)	(-0.08 – 0.01)	(-0.01 – 0.18)
Random effects			
σ^2	0.01	0.00	0.00
τ_{00} database	0.04	0.00	0.02
ICC	0.87	0.86	0.81
N species	35	35	35
Observations	700	700	700
Marginal R^2	0.22	0.12	0.08
Conditional R^2	0.90	0.88	0.82

p-value: * <0.05; ** <0.01, *** <0.001

Table A3–3. Results of the analysis of covariance comparing the effect of the predicted leaf traits (log scale), life forms, spectra processing, and their interaction on the observed traits (log scale). Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). The models consider the random effect of the samples on the spectral processing.

Process	Factor	Leaf trait		
		LMA	WC	EWT
Training	Predicted traits $_{(df = 1/414)}$	10847.80***	5721.90***	2430.31***
	Life form $_{(df = 1/414)}$	1.48	1.12	0.47
	Predicted * Life form $_{(df = 1/414)}$	0.37	0.61	2.85
	Predicted * Spectra $_{(df = 1/414)}$	21.34***	69.82***	1.62
	Predicted * Life form * Spectra $_{(df = 1/414)}$	0.50	0.26	0.12
Testing	Predicted traits $_{(df = 1/274)}$	2427.45***	1216.27***	757.219
	Life form $_{(df = 1/274)}$	3.68	0.17	1.10
	Predicted * Life form $_{(df = 1/274)}$	0.07	0.04	0.36
	Predicted * Spectra $_{(df = 1/274)}$	0.20	0.03	5.33*
	Predicted * Life form * Spectra $_{(df = 1/274)}$	1.57	0.24	1.30

p-value: *** <0.001

Table A3–4. Statistical comparisons of the parameters from partial least-squares regression (PLSR) between lianas and trees from models based on reflectance and wavelet spectra to predict three functional traits: Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Values represent the paired *t-test*, all of them with 999 of the degree of freedom.

Process	Parameter	Spectra	Trait		
			LMA	WC	EWT
Training	R^2	Reflectance	18.35***	-40.76***	-114.73***
		CWT	-2.63**	-46.19***	-130.00***
	Bias	Reflectance	5.90***	12.17***	8.24***
		CWT	34.36***	-20.08***	44.43***
	RMSE	Reflectance	-180.47***	-34.84***	-130.29***
		CWT	-248.46***	-98.38***	-137.6***
	%RMSE	Reflectance	-55.38***	39.13***	75.56***
		CWT	-49.12***	41.95***	79.07***
Testing	R^2	Reflectance	58.83***	-3.67***	-22.48***
		CWT	47.42***	-17.20***	-172.71***
	Bias	Reflectance	-31.77***	-8.86***	2.76**
		CWT	-14.58***	-62.31***	-34.56***
	RMSE	Reflectance	-208.53***	-42.15***	-68.92***
		CWT	-315.28***	-43.72***	-108.89***
	%RMSE	Reflectance	-14.94***	105.25***	4.85***
		CWT	52.89***	149.36***	124.12***

** *p-value* < 0.01; *** *p-value* < 0.001
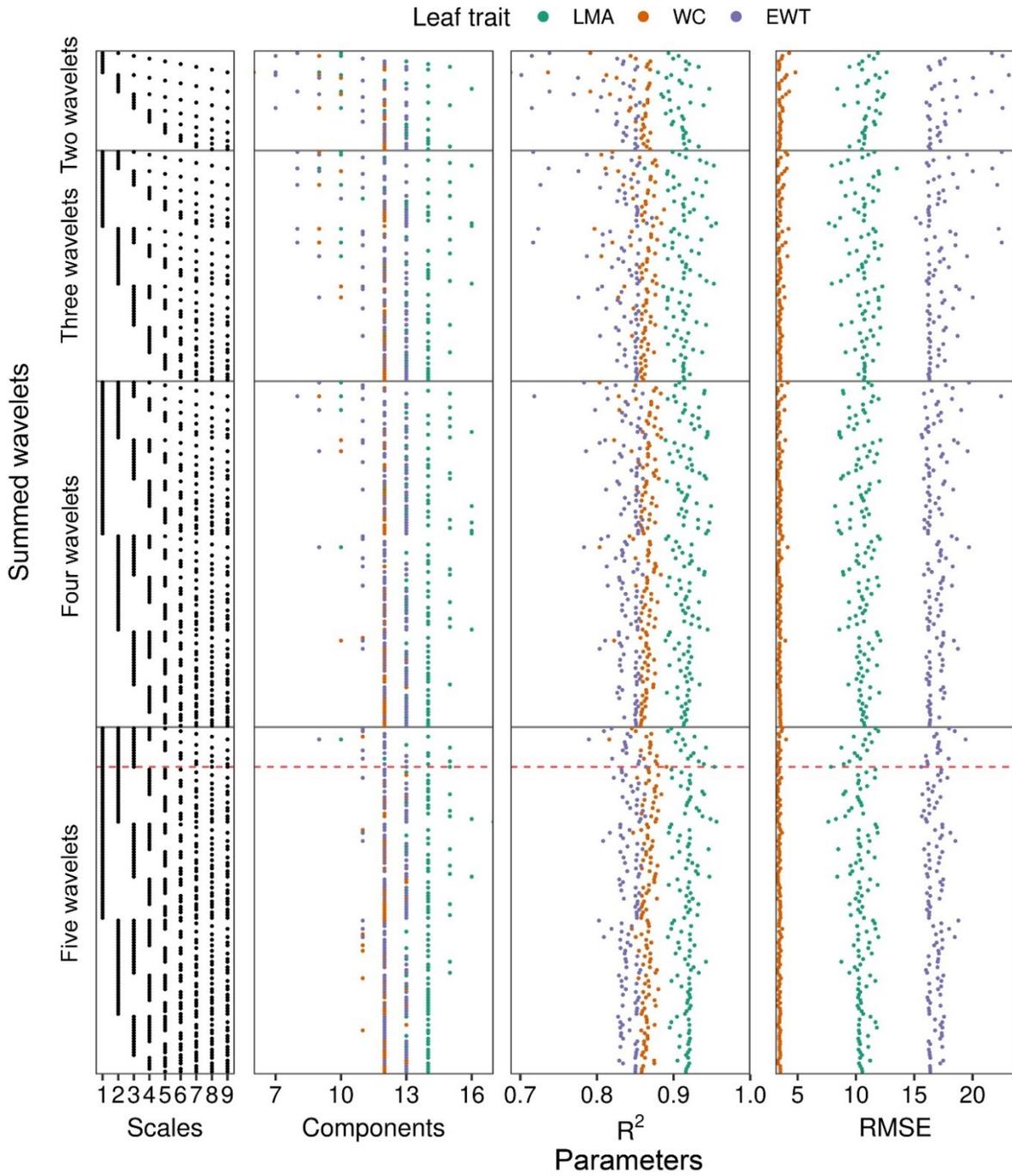
Figure A3–1. Summed wavelet combinations and their performance to predict three-leaf traits: Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). The panel of scales represents the potential wavelet combination in a horizontal plane. The performance of a given combination of scales is described by the optimal number of components, the coefficient of determination (R^2), and the root mean square error (RMSE). The red dashed line represents the combination of wavelets used for the main manuscript (scales 1, 2, 3, 8, and 9).
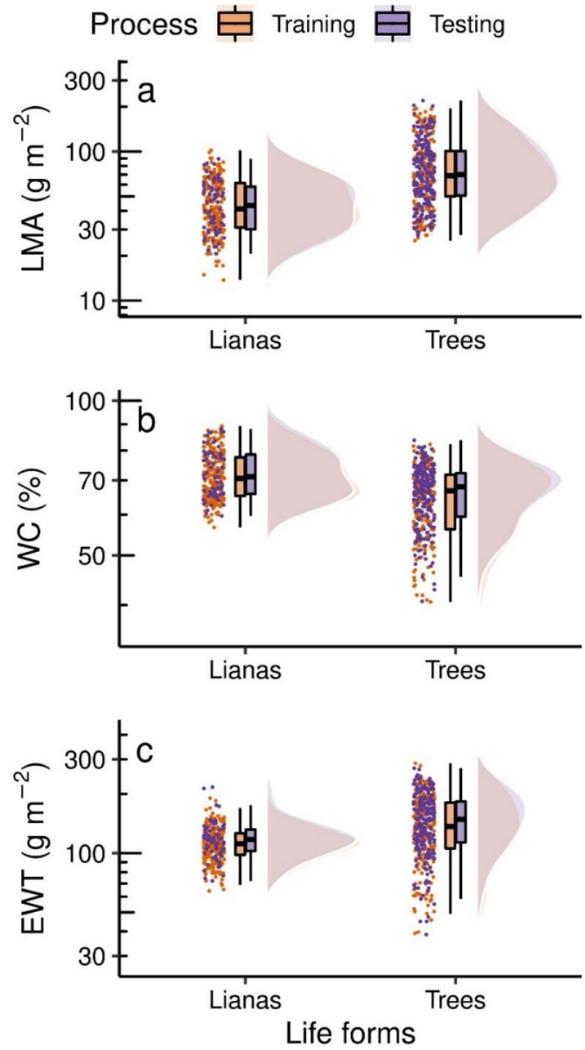
Figure A3–2. Comparison of the leaf traits of lianas and trees used for training and testing purposes. Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT).
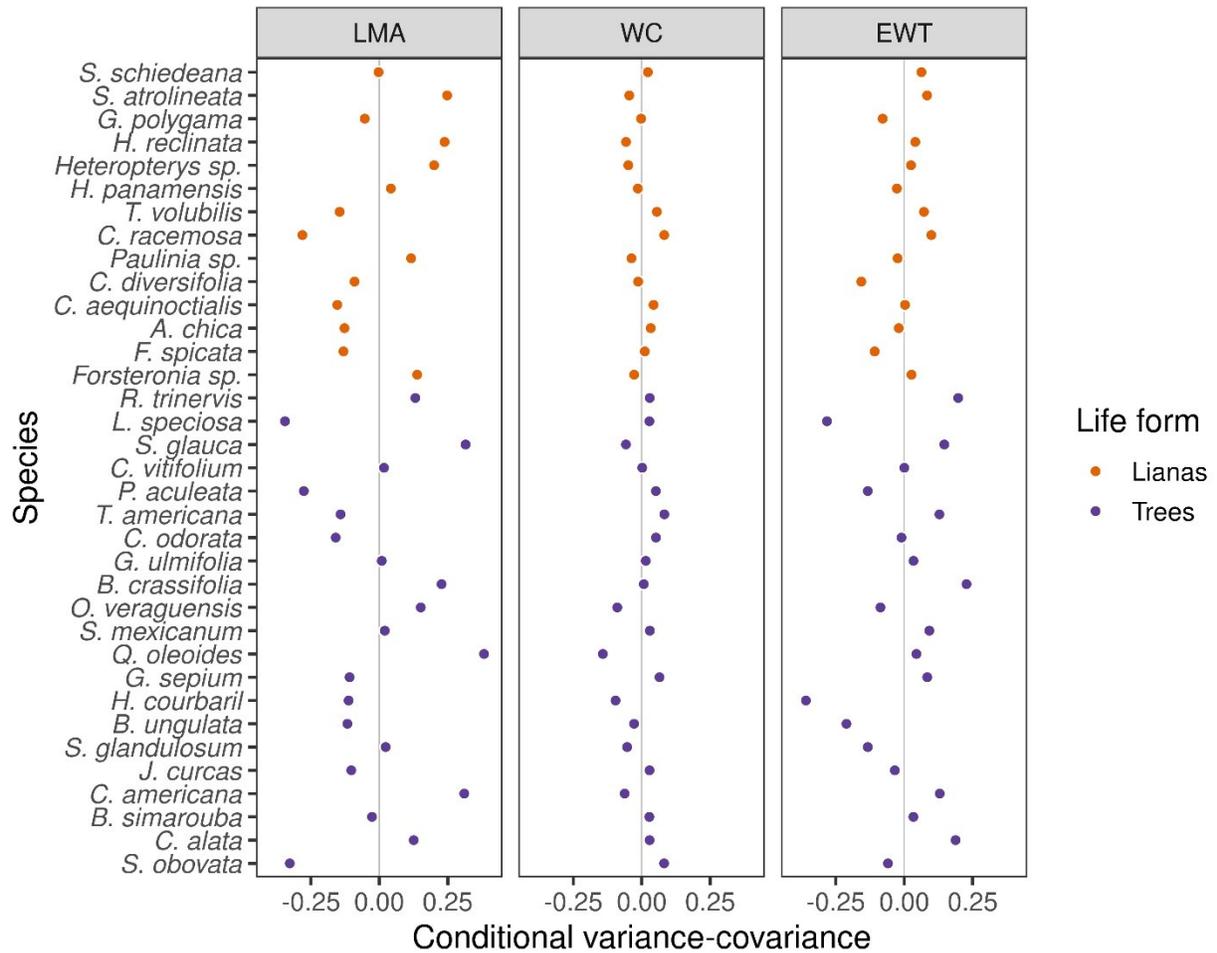
Figure A3–3. Conditional variance-covariance matrixes of the random effect of the species on models comparing the leaf traits between lianas and trees. Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). More statistics of the models can be found in Table A3–2.
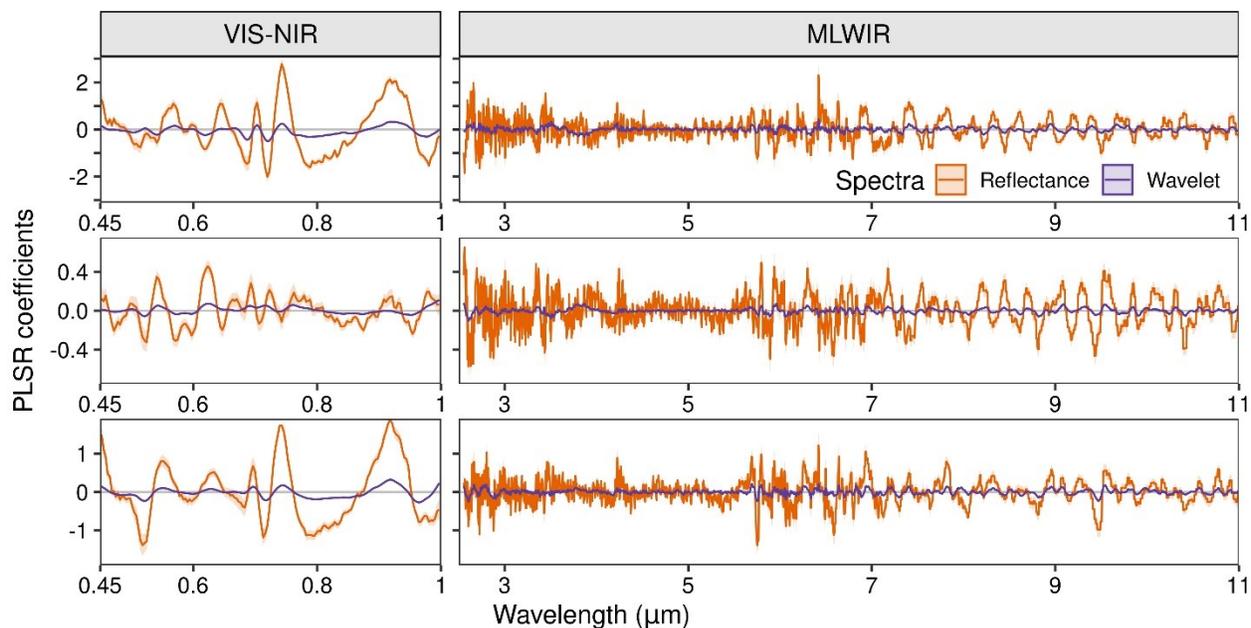
Figure A3–4. PLSR coefficients of the models performed on processed reflectance and continuous wavelet transformation (CWT) spectra for the three functional traits: Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Each line represents the average of 1000 iterations, while the shade around each line represents the standard deviation.

Figure A3–5. Observed and predicted leaf traits of lianas and trees from partial least-squares regression (PLSR) models based on reflectance (a, c, e) and wavelet spectra (b, d, f). Points correspond to the training dataset. Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Each point represents the average of 1000 iterations, while the error bars around each point represent the standard deviation. The solid grey lines indicate the 1:1 relationship. The solid lines represent the fitted regression, while the dashed lines the 95% prediction intervals. More statistics their comparisons can be found in Table A3–3, respectively.

Figure A3–6. Observed and predicted leaf traits of lianas and trees from partial least-squares regression (PLSR) models based on reflectance (a, c, e) and wavelet spectra (b, d, f). Points correspond to the training dataset. Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Each point represents the mean of 1000 iterations, while the error bars around each point represent the standard deviation. The solid grey lines indicate the 1:1 relationship. The solid lines represent the fitted regression, while the dashed lines the 95% prediction intervals. Kernel density distributions are plotted next to each scatter plot. More statistics about their comparisons can be found in Table A3–3.
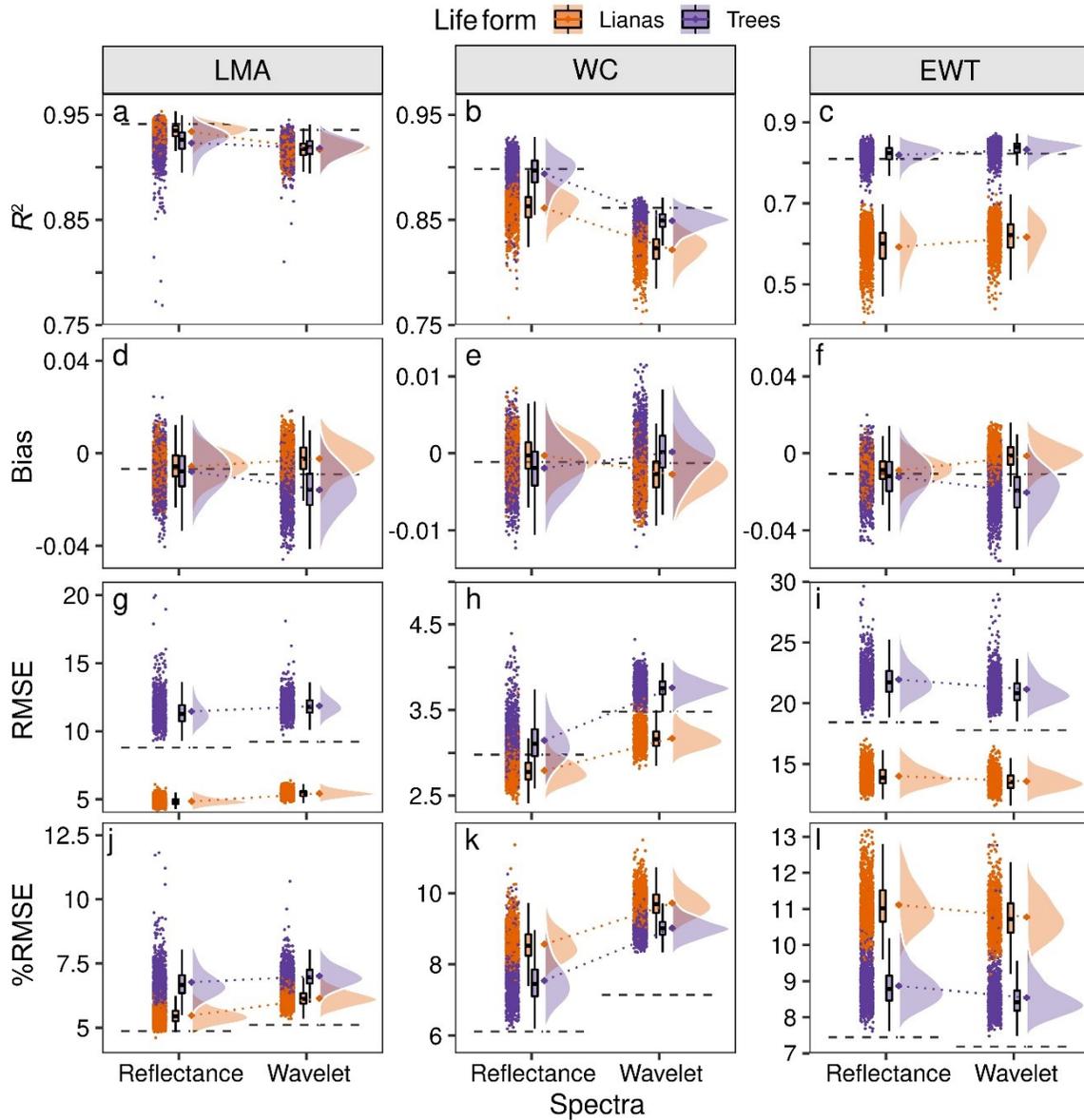
Figure A3–7. Raincloud plots comparing the performance between life forms of training models of partial least-squares regression (PLSR) based on reflectance and continuous wavelet transformation (CWT) spectra to predict three functional traits: Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). RMSE is the root mean square error, and %RMSE is the RMSE represented by its percentage. Each point represents a model iteration, the boxes the first, median, and third quartiles, and the irregular polygons describe the kernel density distributions. Dashed lines represent the mean regardless of the life form, while the dotted lines the mean trend of change within life forms between the type of spectra. Statistics of their comparison can be found in Table A3–4.
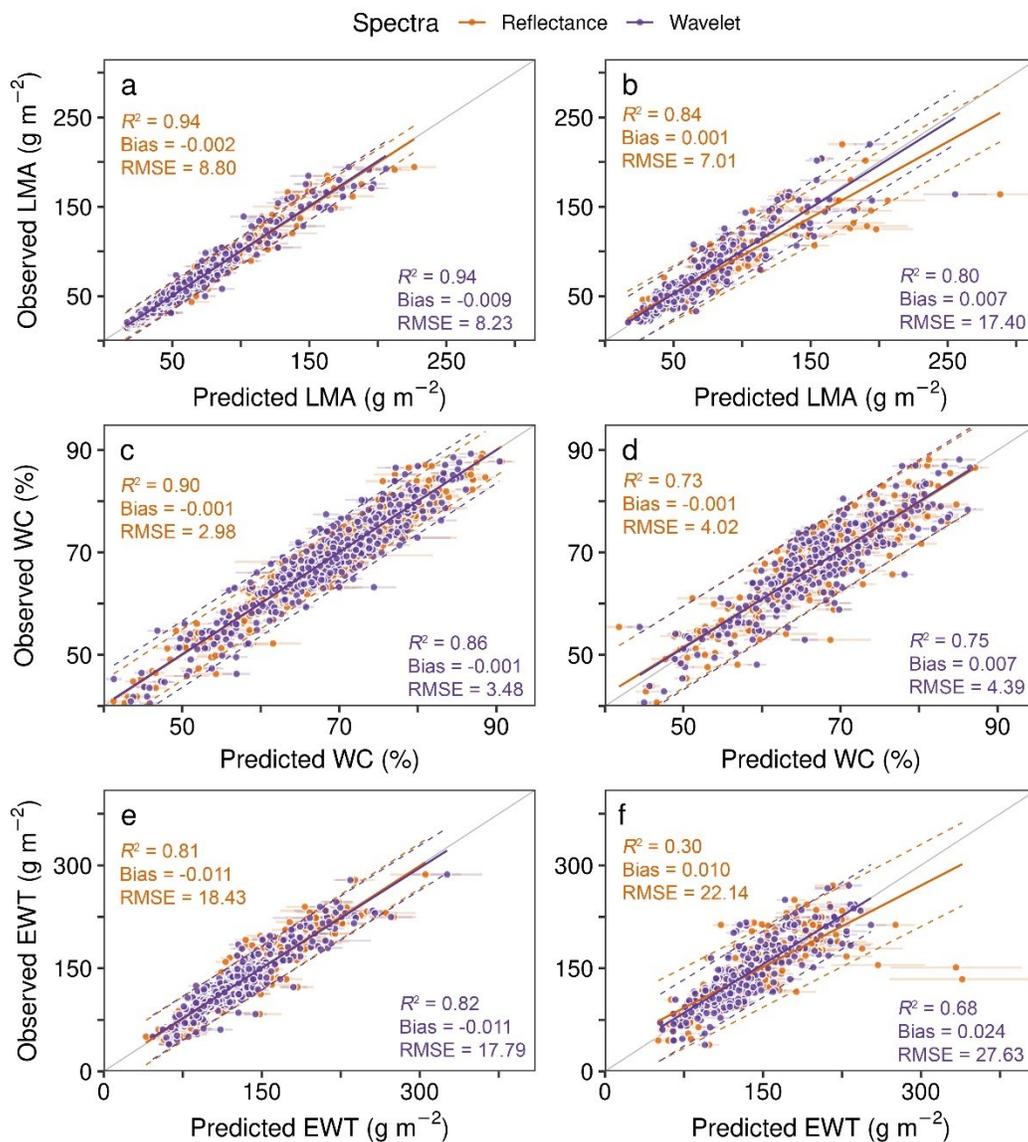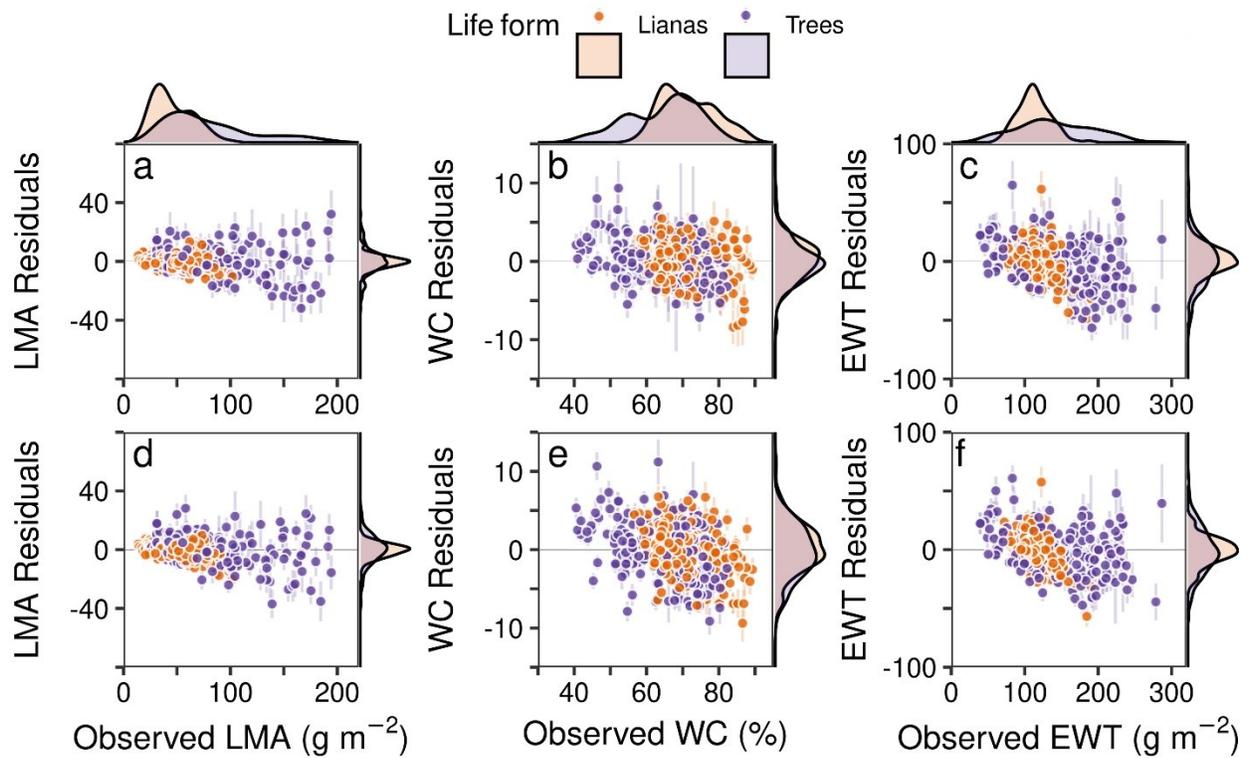
Figure A3–8. Residuals of the predicted leaf traits of lianas and trees from partial least-squares regression (PLSR) models based on reflectance (a, b, c) and wavelet spectra (d, e, f). Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Points correspond to the training dataset. Each point represents the mean of 1000 iterations, while the error bars around each point represent the standard deviation.
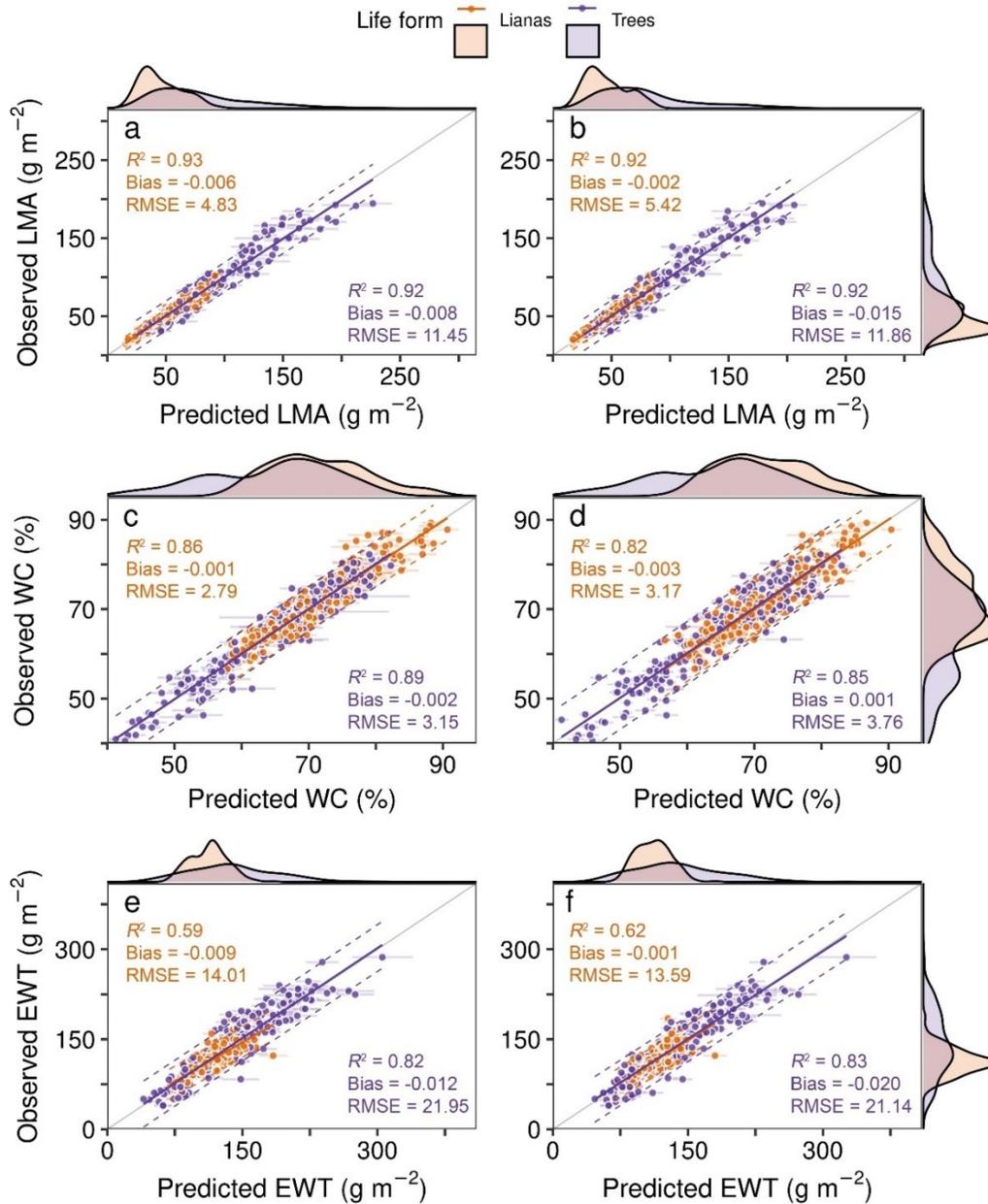
Figure A3–9. Residuals of the predicted leaf traits of lianas and trees from partial least-squares regression (PLSR) models based on reflectance (a, b, c) and wavelet spectra (d, e, f). Leaf Mass area (LMA), Water Content (WC), Equivalent Water Thinness (EWT). Points correspond to the testing dataset. Each point represents the mean of 1000 iterations, while the error bars around each point represent the standard deviation.

Appendix 4. Supporting information for Chapter 5.

Appendix A4–1 Data collection and pre-processing for Santa Rosa National Park – Environmental Monitoring Supersite and University of Alberta North Campus databases.
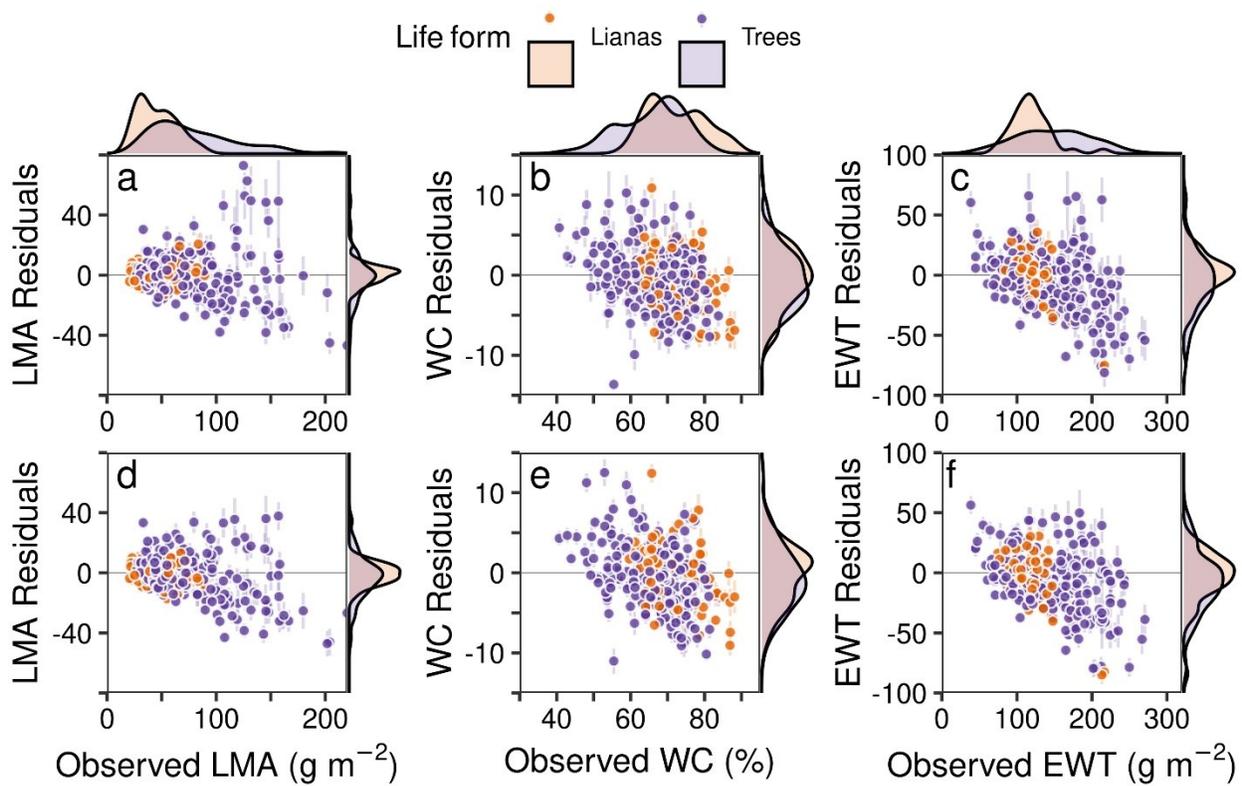

A4–1.1 Sites description

A4–1.1.2 Santa Rosa National Park – Environmental Monitoring Supersite database

This database was collected at the Santa Rosa National Park – Environmental Monitoring Supersite, Costa Rica (10°48″ N, 85°36″ W) (CR), where trees were scanned during the wet-season of 2017 between May and July. This site is located northwest of Costa Rica and presents one of the last remnants of Tropical Dry Forest in the country (Sánchez-Azofeifa et al., 2005). The wet season extends from May to late November while a dry-season, during which most of the trees lose their leaves, encompasses the remaining months (Kalacska et al., 2004). The air temperature of the site varies from 26 °C in the wet-season to 29 °C in the dry season, and the region presents a mean annual precipitation of 1720 mm (Kalacska et al., 2004). In general, SRNP is composed of a mosaic of forest patches in different successional stages of natural regeneration and with different land-use histories associated with anthropogenic fires, deforestation, and land clearing for pasture and agriculture (Arroyo-Mora et al., 2005a; Calvo-Alvarado et al., 2009; Sánchez-Azofeifa et al., 2017). In this park, it is possible to find 96 species of trees of different life history (Hilje et al., 2015) where the abundance and the height canopy species depends on the successional stage of the forest patches (Hilje et al., 2015; Li et al., 2017).

Nineteen trees of 13 species in 9 families were scanned (Table A4–1). The selection of trees was based on the degree of isolation for a subsequent quick segmentation and their full-sun exposure. Some of the species selected present a high cultural (i.e., *E. cyclocarpum, Q. oleoides*), agronomic (i.e., *P. guajava*), and forestry (e.g., *C. odorata, S. macrophylla*) importance.

A4–1.1.3 University of Alberta North Campus databases

This database was collected at the University of Alberta North Campus (53°31″ N, 113°31″ W) located in the city of Edmonton, Canada (CAN); the most northerly city in North America. This site presents a continental climate with a relatively low annual precipitation of 480 mm. The average minimum temperature during winter is close to -14.8 °C, while during summer to 23.1 °C.

On-campus, 18 trees of *Ulmus americana* located in two lots of 9 trees were scanned in October of 2017. The selection of *U. americana* trees was based on the degree of isolation for a subsequent quick segmentation and the absence of previous pruning.

A4–1.2 Tree scanning and processing

The 3D scanning for these databases was performed using a Riegl VZ-400i (RIEGL, Horn, Austria), which is a time-of-flight LiDAR sensor with multiple returns. This sensor has a beam divergence of 0.35 mrad or (5 mm accuracy at 100 m range) with a scanning frequency of 1200 kHz, allowing to reach 500,000 measurements s^{-1}. The Riegl VZ-400i was mounted to a lightweight carbon fiber tripod, which hosted an external Silver-Nikel 15V battery allowing for scanning multiple trees in a single day.

The CR scanning of trees was performed following Wilkes et al., (2017) recommendations. Specifically, the trees were scanned in a radial pattern at cardinal coordinates approximately 5 – 15m away from the target tree, depending on its height. This scanning design minimizes the tree occlusion, creates overlapping fields of view between scan positions, and allows the full 3D rendering of trees. Around each target tree, six cylinder targets were placed on 2 m poles outside the bounds of the scan positions to act as retro-reflective targets for future scan co-registration. Likewise, an additional 5 cm circular targets were placed within the bounds of the scan position to achieve a fine-registration of scan positions. On the other hand, the CAN scanning was set up following a 30 x 50 m plot pattern in the two lots of trees. Six cylinder targets were placed on 2 m poles outside the bounds of the plot as retro-reflective targets for future scan co-registration. On each lot, 9 to 12 scan positions were used to cover the distribution of the trees. The scan was performed using the scanner in a horizontal and vertical position to get a full hemispherical view of the tree canopy.

For each scan project (i.e., tree or plot), a coarse co-registration, fine registration, and filtering were performed using RIEGL's RiSCAN PRO® software. Specifically, the coarse co-registration was conducted using the retro-reflective targets distributed throughout the plot, where scan positions would coordinate and orient to one another using these common tie points. Afterward, a fine registration was performed using a multi-station adjustment approach. This approach was conducted to correct errors in translation and rotation of scan positions modifying the orientation and position of each scan relative to a locked or anchored scan. Using this adjustment, the accuracy

of registration between scans was lower than ~ 0.005 m. Finally, creating a common coordinate system for all the scan positions, a transformation matrix was applied to each scan to create point clouds. On the resulting point clouds, a filter was applied to remove the backscattered pulses. In general, backscattered pulses can be considered "noisy" due to these points tend to have a different pulse shape than that of the outgoing pulse (Pfennigbauer and Ullrich, 2010). A change in the pulse shape typically occurs when the outgoing pulse hits the edge of an object, such as a leaf or a piece of bark, and only a portion of the outgoing pulse reflected (Vaccari, et al., 2013). The filter was applied to that reflected pulse that tends to be different in ±2 standard deviation of the outgoing pulse. Then, the resulting point clouds were manually segmented, removing all the points that do not belong to the target trees. On the remaining points, a statistical outlier removal (SOR) filter was applied using a *k*-nearest neighbor of 10 and a standard deviation of 2. After the application of the filter, point clouds were subsample using a resolution of 0.01 m. The manual segmentation, the application of the SOR, and the resampling were performed on CloudCompare (2017).

A4–1.3 References

Arroyo-Mora, J. P., Sánchez-Azofeifa, G. A., Kalacska, M. E. R., Rivard, B., Calvo-Alvarado, J. C., & Janzen, D. H. (2005). Secondary forest detection in a neotropical dry lorest landscape using Landsat 7 ETM+ and IKONOS imagery. Biotropica, 37(4), 497–507.

Calvo-Alvarado, J., McLennan, B., Sánchez-Azofeifa, A., & Garvin, T. (2009). Deforestation and forest restoration in Guanacaste, Costa Rica: Putting conservation policies in context. Forest Ecology and Management, 258(6), 931–940.

CloudCompare. (2017). Retrieved from http://www.cloudcompare.org

Hilje, B., Calvo-Alvarado, J., Jiménez-Rodríguez, C., & Sánchez-Azofeifa, A. (2015). Tree species composition, breeding systems, and pollination and dispersal syndromes in three forest successional stages in a tropical dry forest in Mesoamerica. Tropical Conservation Science, 8(1), 76–94.

Kalacska, M., Sanchez-Azofeifa, G. A., Calvo-Alvarado, J. C., Quesada, M., Rivard, B., & Janzen, D. H. (2004). Species composition, similarity and diversity in three successional stages of a seasonally dry tropical forest. Forest Ecology and Management, 200(1–3), 227–247.

Li, W., Cao, S., Campos-Vargas, C., & Sanchez-Azofeifa, A. (2017). Identifying tropical dry forests extent and succession via the use of machine learning techniques. International Journal of Applied Earth Observation and Geoinformation, 63, 196–205.

Pfennigbauer, M., & Ullrich, A. (2010). Improving quality of laser scanning data acquisition through calibrated amplitude and pulse deviation measurement. In M. D. Turner & G. W. Kamerman (Eds.), Laser Radar Technology and Applications XV (p. 7684).

Sánchez-Azofeifa, G. A., Guzmán-Quesada, J. A., Vega-Araya, M., Campos-Vargas, C., Durán, S. M., D'Souza, N., … Sharp, I. (2017). Can terrestrial laser scanners (TLSs) and hemispherical photographs predict tropical dry forest succession with liana abundance? Biogeosciences, 14(4), 977–988.

Sánchez-Azofeifa, G. A., Quesada, M., Rodríguez, J. P., Nassar, J. M., Stoner, K. E., Castillo, A., … Cuevas-Reyes, P. (2005). Research priorities for neotropical dry forests. Biotropica, 37(4), 477–485.

Vaccari, S., van Leeuwen, M., Calders, K., Coops, N. C., & Herold, M. (2013). Bias in lidar-based canopy gap fraction estimates. Remote Sensing Letters, 4(4), 391–399.

Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., Gonzalez de Tanago, J., … Herold, M. (2017). Data acquisition considerations for Terrestrial Laser Scanning of forest plots. Remote Sensing of Environment, 196, 140–153.

Appendix A4–2. The impact of the point cloud resolution on the estimations of fractal geometry.

We evaluate the effect of the point cloud resolution on the estimations of fractal geometry using the point clouds of the CAN database. For this goal, after the application of the SOR filter described in Appendix A4–1 we subsample the point clouds using two methods: spatial grid and random point removal. For the first methods, three distances of the spatial grid were used: i) 0.001 m, ii) 0.005 m, and iii) 0.01 m (current resolution used in the manuscript). For the second method, 50 and 15% of the points were removed based on the total number of points (100%). Between both methods, points clouds using 0.001 m of resolution are equivalent to point clouds with 100% of points. Using these subsampling methods, the resulting point density of the clouds is reduced each step (Figure A4–1a, b). On these point clouds, we then apply the voxel-counting method described in section 2.3 to derive the potential effect of the point density on d_{MB}, the intercept$_{MB}$, and r^2.

Our methods reveal that d_{MB} and r^2 are partially affected by the point density, where a high reduction in the point density leads to reductions in d_{MB} and r^2 (Figure A4–1). Despite this, in both subsampling methods the intercept$_{MB}$ tends to remain almost invariant to the changes in point density. Therefore, reductions in the point density seem to affect the way of how points are distributed in the space, but not how these could coverage or describe a given surface.

Table A4–1. The number of samples (*n*) and tree species of the point clouds used grouped by database and family. The acronyms of the databases are described in Table 6–1.

Database	Family	Species	*n*
CAN	Ulmaceae	*Ulmus Americana*	15
CR	Bignoniaceae	*Crescentia alata*	2
	Dilleniaceae	*Curatella americana*	1
	Fabaceae	*Ateleia herbert-smithii*	2
		Enterolobium cyclocarpum	1
		Gliricidia sepium	1
		Quercus oleoides	1
	Meliaceae	*Cedrela odorata*	2
		Swietenia macrophylla	2
		Trichilia Americana	1
	Moraceae	*Ficus sp.*	1
	Myrtaceae	*Psidium guajava*	2
	Simaroubaceae	*Simarouba glauca*	1
	Sterculiaceae	*Guazuma ulmifolia*	2
GUY	Fabaceae	*Eperua falcata*	1
		Eperua grandiflora	7
		Ormosia coutinhoi	1
		Pithecellobium jupunba	1
CAM	Annonaceae	*Annickia chlorantha*	3
	Combretaceae	*Terminalia superba*	9
	Euphorbiaceae	*Macaranga barteri*	2
	Fabaceae	*Baphia leptobotrys*	3
		Cylicodiscus gabunensis	5
		Erythrophleum suaveolens	5
		Pentaclethra macrophylla	1
		Pterocarpus soyauxii	6
	Lecythidaceae	*Petersianthus macrocarpus*	6
	Malvaceae	*Duboscia macrocarpa*	2
		Eribroma oblongum	4
		Mansonia altissima	3
		Triplochiton scleroxylon	6
	Meliaceae	*Entandrophragma cylindricum*	2
	Myristicaceae	*Pycnanthus angolensis*	4
IND	Chrysobalanaceae	*Parastemon urophyllus*	1
	Dipterocarpaceae	*Shorea sp.*	2
		Shorea teysmanniana	1
	Ebenaceae	*Diospyros evena*	1
	Hypericaceae	*Cratoxylum glaucum*	1
	Meliaceae	*Aglaia rubiginosa*	1

	Tetrameristaceae	*Tetramerista glabra*	3
PER	Burseraceae	*Dacryodes peruviana*	1
	Combretaceae	*Buchenavia macrophylla*	1
	Elaeocarpaceae	*Sloanea eichleri*	1
	Fabaceae	*Pseudopiptadenia suaveolens*	1
	Lauraceae	*Nectandra longifolia*	1
	Lecythidaceae	*Couratari macrocarpa*	2
	Malvaceae	*Pterygota amazonica*	2
AUS	Myrtaceae	*Eucalyptus leucoxylon*	
		Eucalyptus microcarpa	65
		Eucalyptus tricarpa	

Table A4–2. Results from the linear mixed models comparing the effect of the presence-absence of leaves on the estimations of fractal geometry parameters. Values between parentheses represent confidence intervals, and the asterisks next to each value the significance of the linear parameters. σ^2 describes the variance of the model, τ_{00} the variance of the databases as a random effect, and ICC the interclass correlation coefficients. The conditional variance-covariance matrixes of the random effects can be found in Figure A4–2.

	Fractal geometry parameter (dependent variable)		
	d_{MB}	Intercept$_{MB}$	$r^2{}_{MB}$
Constant	-0.56***	2.26***	-0.01
	(-0.72 – -0.40)	(1.59 – 2.93)	(-0.02 – 0.00)
Leaf on/off	0.03	0.70	-0.01
	(-0.16 – 0.22)	(-0.11 – 1.50)	(-0.02 – 0.00)
Random effects			
σ^2	0.00	0.62	0.00
τ_{00} database	0.01	0.21	0.00
ICC	0.85	0.25	0.80
N databases	7	7	7
Observations	189	189	189
Marginal R^2	0.014	0.12	0.289
Conditional R^2	0.85	0.35	0.85

p-value: * <0.05; ** <0.01, *** <0.001

Table A4–3. Mean and standard deviation of the parameters extracted from the fractal geometry using the voxel counting method. Acronyms of the databases are described in Table 6–1. d_{MB} represents the fractal dimension, and r^2 the coefficient of determination.

Database	d_{MB}	Intercept$_{MB}$	r^2 (x10^{-2})
CAN	0.65 ± 0.01	2.47 ± 0.41	98.55 ± 0.17
CR	0.70 ± 0.02	2.65 ± 0.48	99.08 ± 0.30
GUY	0.60 ± 0.02	2.99 ± 0.09	97.36 ± 0.24
CAM	0.53 ± 0.03	2.34 ± 0.44	99.65 ± 0.35
IND	0.58 ± 0.02	2.64 ± 0.24	98.35 ± 0.55
PER	0.55 ± 0.01	2.99 ± 0.17	96.41 ± 1.01
AUS	0.60 ± 0.02	2.44 ± 0.32	97.63 ± 0.58
Leaf-on	0.61 ± 0.05	2.58 ± 0.38	97.82 ± 0.89
Leaf-off	0.55 ± 0.06	2.37 ± 0.44	99.43 ± 0.55
All databases	0.59 ± 0.06	2.50 ± 0.42	98.47 ± 1.11

Table A4–4. Linear mixed model statistics of the relationship between tree metrics with fractal geometry parameters, the presence-absence of leaves, and their interaction. Values between parentheses represent confidence intervals, and the asterisks next to each value the significance of the linear parameters. σ^2 describes the variance of the model, τ_{00} the variance of the databases as a random effect, and ICC the interclass correlation coefficients. The conditional variance-covariance matrixes of the random effects can be found in Figure A4–3.

Predictors	d_{MB}				Intercept$_{MB}$			
	H	DBH	Crown area	QSM volume	H	DBH	Crown area	QSM volume
Intercept	-0.93*	-2.47***	-7.74***	-10.51***	0.33	-0.04	-1.26***	-3.92***
	(-1.80 – -0.06)	(-3.79 – -1.16)	(-10.03 – -5.45)	(-13.84 – -7.18)	(0.00 – 0.66)	(-0.21 – 0.12)	(-1.47 – -1.04)	(-4.31 – -3.53)
d_{MB} or Intercept$_{MB}$	3.71***	6.89***	16.34***	18.35***	0.39***	0.68***	1.31***	1.77***
	(2.50 – 4.93)	(4.89 – 8.88)	(12.94 – 19.73)	(13.39 – 23.30)	(0.35 – 0.43)	(0.63 – 0.73)	(1.23 – 1.39)	(1.67 – 1.86)
Leaf on/off	2.14***	2.15*	5.63***	5.66*	0.24	0.15	0.21	0.25
	(1.00 – 3.28)	(0.41 – 3.89)	(2.61 – 8.65)	(1.27 – 10.06)	(-0.16 – 0.64)	(-0.09 – 0.38)	(-0.12 – 0.54)	(-0.27 – 0.77)
Interaction	-3.44***	-3.46*	-9.39***	-9.19**	-0.09***	-0.08*	-0.17**	-0.17*
	(-5.06 – -1.81)	(-6.14 – -0.78)	(-13.95 – -4.83)	(-15.85 – -2.54)	(-0.15 – -0.04)	(-0.15 – -0.01)	(-0.29 – -0.05)	(-0.32 – -0.03)
Random effects								
σ^2	0.02	0.05	0.15	0.31	0.01	0.01	0.02	0.03
τ_{00} database	0.13	0.17	0.63	1.31	0.05	0.01	0.00	0.05
ICC	0.88	0.77	0.81	0.81	0.91	0.39	0.14	0.60
N database	7	7	7	7	7	7	7	7
Observations	189	189	189	189	189	189	189	189
Marginal R^2	0.18	0.36	0.43	0.34	0.27	0.82	0.90	0.85
Conditional R^2	0.90	0.85	0.89	0.87	0.93	0.89	0.91	0.84

p-value: * <0.05; ** <0.01, *** <0.001

Table A4–5. Linear mixed model statistics of the relationship between tree metrics with the fractal dimension (d_{MB}), the presence-absence of leaves, and their interaction. Values between parentheses represent confidence intervals, and the asterisks next to each value the significance of the linear parameters. σ^2 describes the variance of the model, τ_{00} the variance of the databases as a random effect, and ICC the interclass correlation coefficients. The conditional variance-covariance matrixes of the databases can be found in Figure A4–4.

Predictors	Leaf-off				Leaf-on			
	H	DBH	Crown area	QSM volume	*H*	DBH	Crown area	QSM volume
Intercept	-0.99	-2.60**	-7.90**	-10.75**	1.26***	-0.25	-2.03*	-4.73***
	(-2.25 – 0.27)	(-4.36 – -0.83)	(-10.89 – -4.91)	(-15.16 – -6.34)	(0.66 – 1.86)	(-1.21 – 0.71)	(-3.74 – -0.31)	(-7.19 – -2.27)
d_{MB}	3.81***	7.09***	16.60***	18.75***	0.19	3.31***	6.81***	8.96***
	(2.36 – 5.26)	(4.64 – 9.54)	(12.57 – 20.64)	(12.77 – 24.72)	(-0.74 – 1.13)	(1.78 – 4.84)	(4.11 – 9.50)	(5.09 – 12.81)
Random effects								
σ^2	0.03	0.07	0.19	0.42	0.01	0.04	0.11	0.23
τ_{00} database	0.44	0.53	1.70	3.64	0.05	0.08	0.36	0.74
ICC	0.94	0.88	0.90	0.90	0.80	0.69	0.76	0.76
N database	2	2	2	2	5	5	5	5
Observations	76	76	76	76	113	113	113	113
Marginal R^2	0.09	0.21	0.32	0.21	0.01	0.17	0.18	0.16
Conditional R^2	0.95	0.91	0.93	0.92	0.80	0.74	0.81	0.80

p-value: * <0.05; ** <0.01, *** <0.001

Table A4–6. Linear mixed model statistics of the relationship between tree metrics with the intercept of fractal geometry (intercept$_{MB}$), the presence-absence of leaves, and their interaction. Values between parentheses represent confidence intervals, and the asterisks next to each value the significance of the linear parameters. σ^2 describes the variance of the models, τ_{00} the variance of the databases as a random effect, and ICC the interclass correlation coefficients. The conditional variance-covariance matrixes of the databases can be found in Figure A4–4.

Predictors	Leaf-off				Leaf-on			
	H	DBH	Crown area	QSM volume	*H*	DBH	Crown area	QSM volume
Intercept	0.33	-0.05	-1.27***	-3.92**	0.57***	0.10	-1.00***	-3.67***
	(-0.18 – 0.83)	(-0.31 – 0.21)	(-1.57 – -0.98)	(-4.54 – -3.31)	(0.38 – 0.76)	(-0.02 – -0.22)	(-1.16 – -0.84)	(-3.91 – -3.43)
Intercept$_{MB}$	0.39***	0.68***	1.32***	1.77***	0.30 ***	0.60***	1.13***	1.59***
	(0.35 – 0.43)	(0.62 – 0.75)	(1.20 – 1.43)	(1.64 – 1.89)	(0.26 – 0.33)	(0.56 – 0.65)	(1.08 – 1.18)	(1.52 – 1.66)
Random effects								
σ^2	0.01	0.01	0.05	0.06	0.00	0.01	0.01	0.02
τ_{00} database	0.13	0.02	0.00	0.15	0.03	0.00	0.01	0.03
ICC	0.95	0.61	0.04	0.71	0.88	0.20	0.46	0.62
N database	2	2	2	2	5	5	5	5
Observations	76	76	76	76	113	113	113	113
Marginal R^2	0.18	0.70	0.87	0.74	0.26	0.89	0.83	0.90
Conditional R^2	0.96	0.88	0.88	0.93	0.91	0.91	0.96	0.96

p-value: * <0.05; ** <0.01, *** <0.001

Table A4–7. Results of the analysis of covariance comparing the effect of the predicted tree metrics, the data type (tree or stand), and their interaction of the observed tree metrics.

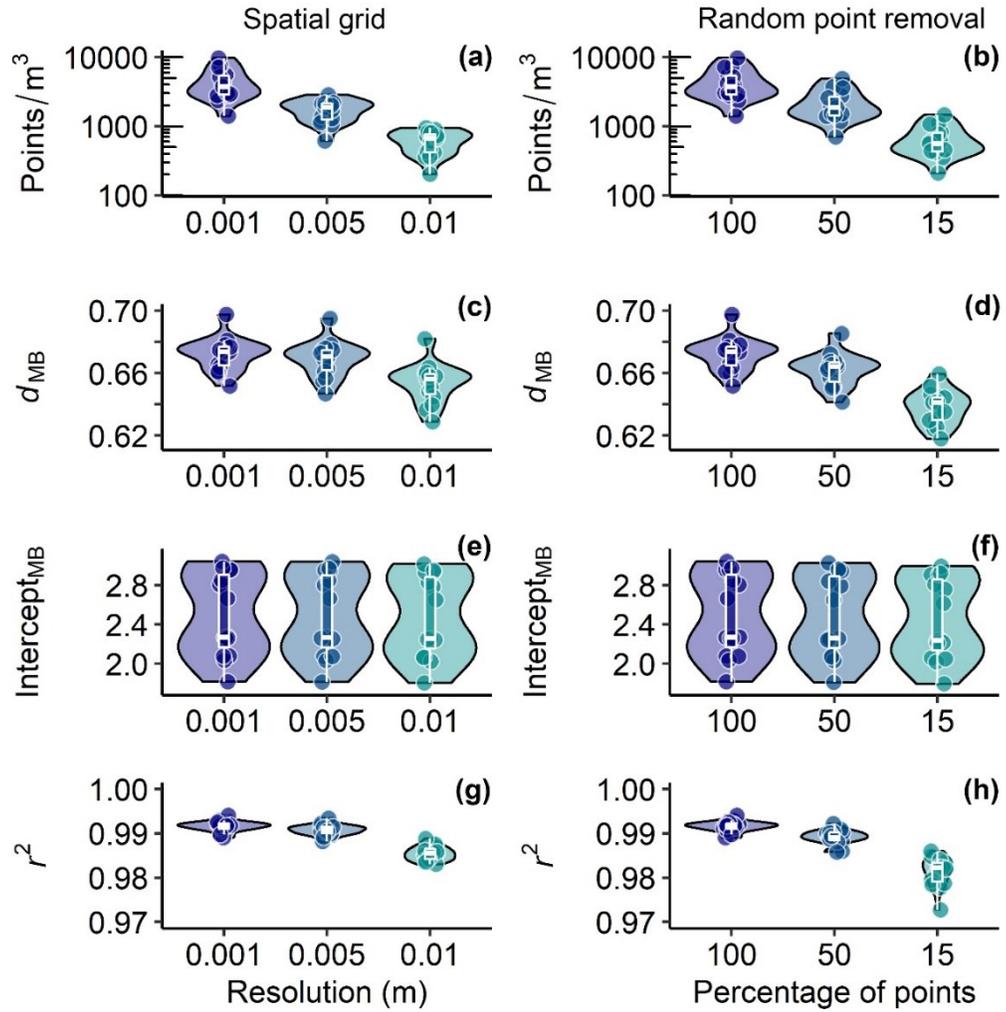| Factors	Observed values					
	Basal area		Covered area		Volume	
	F-ratio	*p-value*	*F-ratio*	*p-value*	*F-ratio*	*p-value*
Predicted value	41617.77	<0.001	54263.13	<0.001	21605.44	<0.001
Tree / Stand	0.17	0.68	15.16	<0.001	39.20	<0.001
Interaction	4.40	0.04	10.41	0.002	13.08	<0.001

Figure A4–1. Effect of the subsampling methods (spatial grid and random point removal) on the point density of point clouds and their estimation of fractal geometry parameters. The procedure to create this figure is described in Appendix A4–2.
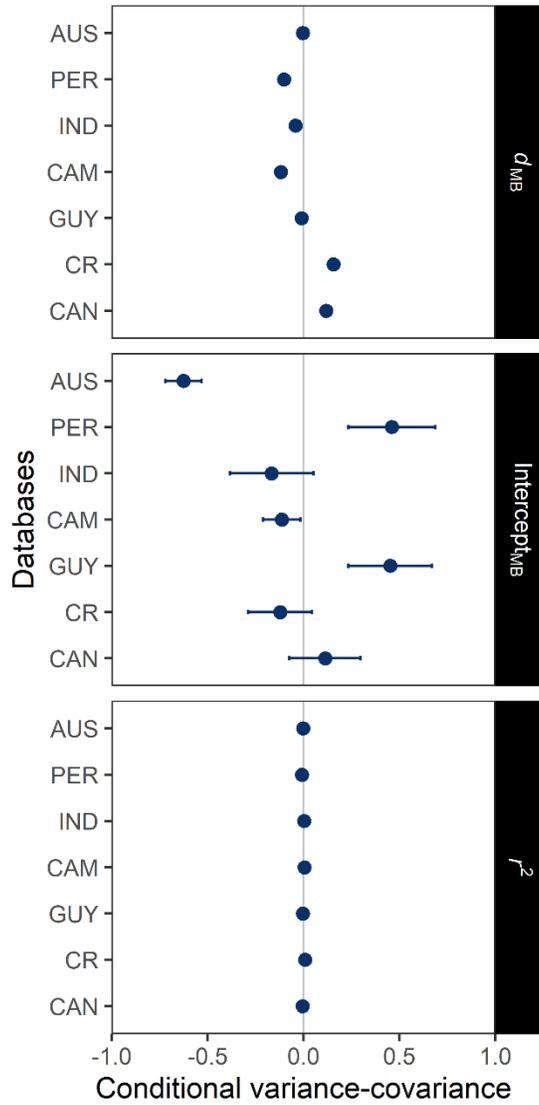
Figure A4–2. Conditional variance-covariance matrixes of the random effect of the databases on models comparing the effect of the presence/absence of leaves on fractal geometry parameters. The error bars represent confidence intervals. More statistics of the models can be found in Table A4–3.
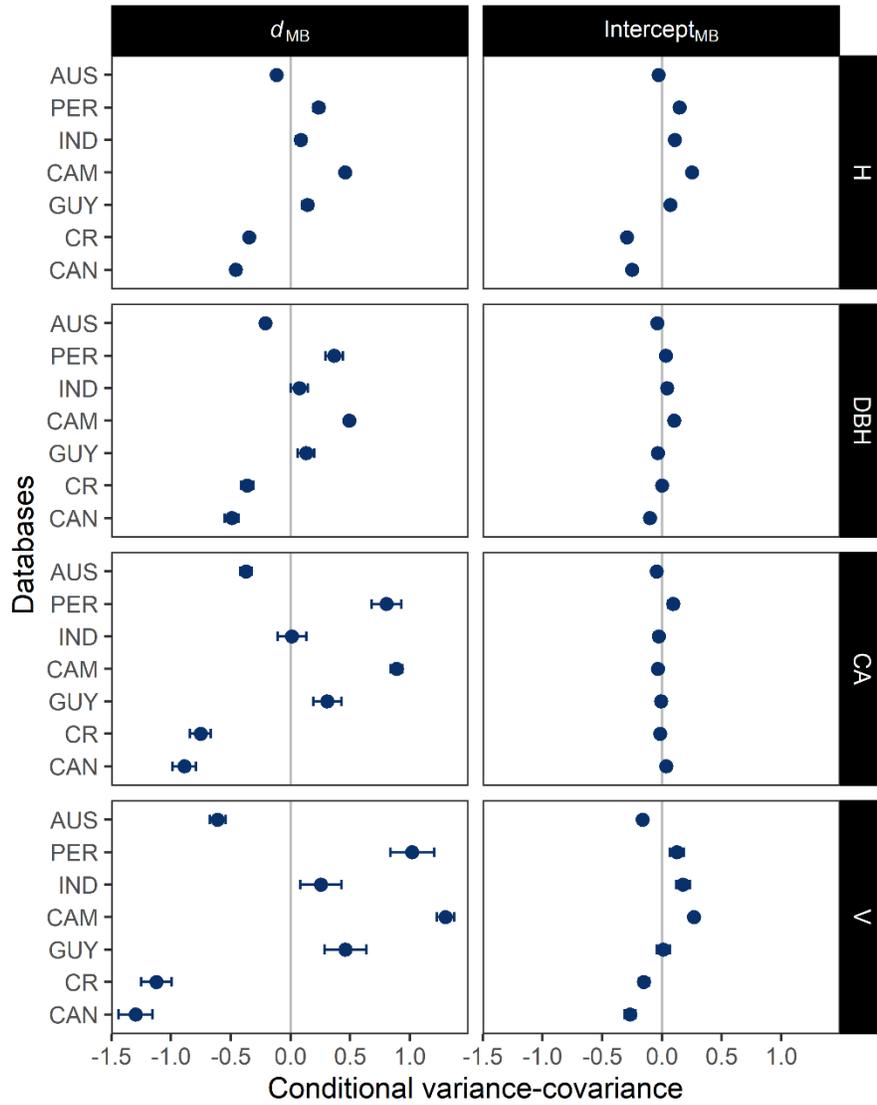
Figure A4–3. Conditional variance-covariance matrixes of the random effect of the databases on models that predict tree metrics using fractal geometry parameters (d_{MB} and Intercept$_{MB}$). The error bars represent confidence intervals. Tree height (*H*), diameter at the breast height (DBH), crown area (CA), and QSM volume (V). Statistics of the models can be found in Table A4–4.
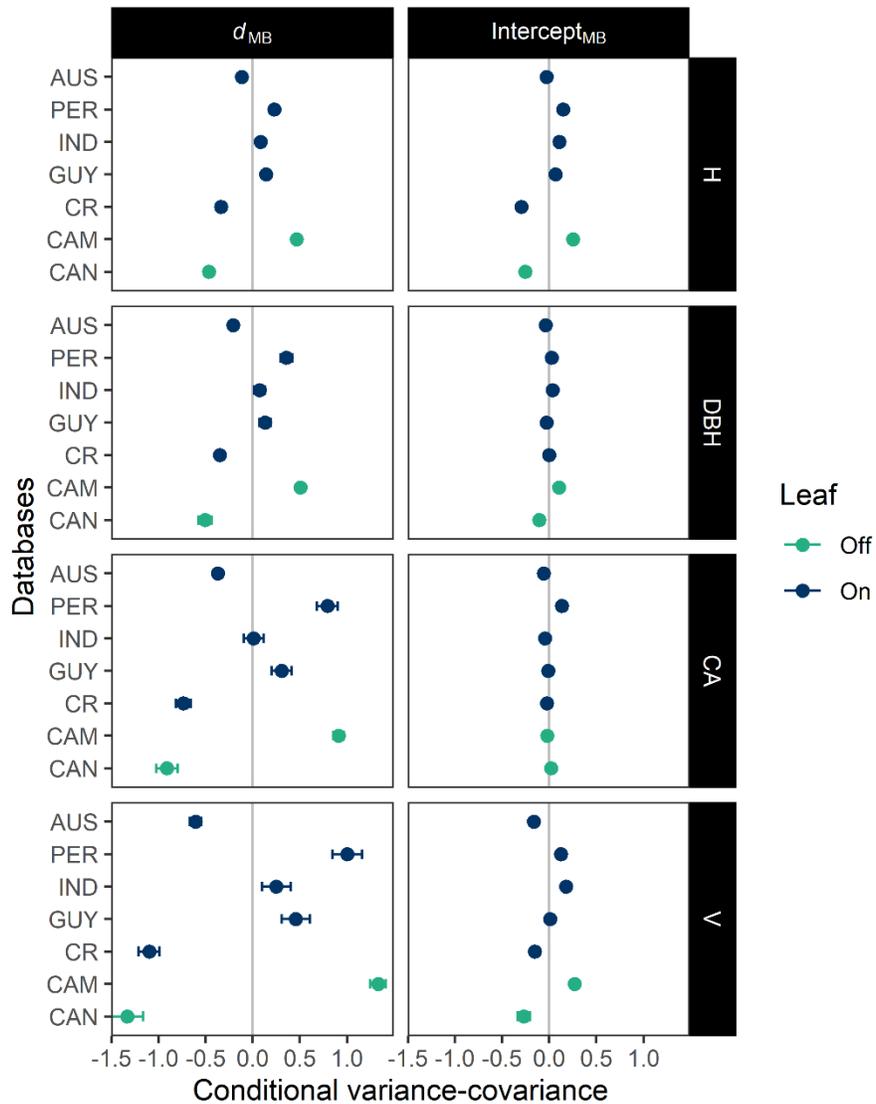
Figure A4–4. Conditional variance-covariance matrixes of the random effect (databases) of linear mixed models that predict tree metrics using fractal geometry parameters (d_{MB} and $Intercept_{MB}$) on tree point clouds with and without leaves. Colors denote different models. The error bars represent confidence intervals. Tree height (H), diameter at the breast height (DBH), crown area (CA), and QSM volume (V). More statistics of the models can be found in Table A4–5 and A4–6.