The Estimation of Buttresses Volume and Classification of Leaf, Wood, and Lianas from Terrestrial Laser Scanning data

by

Tao Han

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

Forests account for one third of land area, and two third of global photosynthesis. The better we know about forests, the better decisions we can make on forest management and carbon cycle modeling. During the last decades, we see the development in remote sensing techniques for forest monitoring. For example, satellite images are efficient methods to monitor forest changes. However, those satellite images present forests in two dimensions, while they do not support rapid and robust assessment of accurate ground reference data. By contrast, Terrestrial Laser Scanning (TLS) can retrieve the three-dimensional vegetation structure with millimeter accuracy.

The main objectives of this PhD thesis are to: (1) present a new machine learning model to separate lianas and trees; and (2) develop a deep learning model to classify leaf and woody components; and (3) develop a non-destructive method to estimate buttress volume. Here, we use 3D point clouds collected by terrestrial laser scanning (TLS) to reach above objectives, where TLS can describe the 3D forest structure in millimeter-level details. This PhD research contributes to filling important knowledge gaps in contemporary scientific fields.

Chapter 2 describes the utilization of a new machine learning model, based on RandomForest (RF) and eXtreme Gradient Boosting (XGBoosting), to separate lianas and trees using TLS point clouds. In this chapter, we find that the XGBoosting algorithm achieves an overall accuracy of 0.88, with a recall of 0.66, higher than the RF algorithm (accuracyof 0.85 and recall of 0.56). In addition, we find the optimal radius search method is as accurate as the multiple radius search method, with F1 scores of 0.49 and 0.48, respectively.We also find that the RF algorithm shows a recall of 0.88 on the independent data. We conclude that the model in chapter 2 provides a new flexible approach to extract lianas from 3D point clouds, enabling more studies to evaluate lianas impact on tree and forest

structure.

Chapter 3 explores the utilization of a deep learning time-series approach to classify leaf and woody components from TLS point clouds. We found that the multivariable time series (MTS) method (accuracy of 0.96) outperformed the univariable time series (UTS) method (accuracy of 0.67 to 0.88) to classify leaf and woody components. Meanwhile, Residual Network (ResNet) spent much more time than Fully Convolutional Neural Network (FCN) and Long Short-Term Memory Fully Convolutional Neural Network (LSTM-FCN) in model development, while those three networks demonstrated similar performance on an independent dataset. Furthermore, we found that the Class Activation Map (CAM) of the proposed model can explain the black-box effect of deep learning. We demonstrated that deep learning algorithms coupled with TLS point clouds could accurately separateleaf and woody components, providing a good start point for future research to estimate forest structure parameters such as Leaf Area Index (LAI) and Wood Area Index (WAI).

A non-destructive method is developed to estimate buttress volume using 3D pointclouds in chapter 4. In this chapter, we found that the alpha shape algorithm (ASA) and slice triangulation (ST) performed better than allometric models for buttress volume estimation. Moreover, the ASA tended to work better than ST when the trees presentedmore and shallower horizontal buttresses. Concerning the allometric models, $D_{area130}$ was the most accurate predictor to estimate buttress volume, with a lower Akaike information criterion (AIC, -66.25) than DAB (-59.55) and Dconvex130 (30.56). At the same time, the DAB (RRMSE of 0.23) and $D_{area130}$ (RRMSE of 0.21) showed similar performance on validation data. Our results indicate that the ASA can help to correct the bias inthe present and past estimates of volume and biomass of large trees, which are keystone components to understanding biomass allocation and dynamics in tropical forests.

Preface

This thesis is original work by Tao Han. The workflow from chapter 2 to 4 was designed by myself, with the assistance of my supervisor Professor Arturo Sanchez-Azofeifa. Currently, Chapter 2 and chapter 3 are published in Remote Sensing.

Acknowledgement

” Life is a journey, not a destination. Learn to enjoy the ride”, said by Ralph Waldo Emerson. This PhD thesis is an exciting journey for me during the past four years, I get the chance to know many interesting people, who work in diverse range of scientific fields. Those people help me to gain a better understanding of what science should be, and howto keep balance between life and work.

First, I would like to express my sincerely gratitude to my PhD supervisor, Dr. Arturo Sanchez-Azofeifa. With his assistance on my research, I could complete all of my three projects in four years, at the University of Alberta. Not only Dr. Sanchez-Azofeifa gives me strong support for data collection in Costa Rica, but he provides me the opportunityto participate in academic meetings in Belgium and Italy. Besides, Dr. Arturo Sanchez-Azofeifa ensures me stay on track to complete my PhD on time. His valuable commentson my projects, is the main reason that I could publish my work in high quality journals. During the past four years, I really appreciate that he gives me the opportunity to undertakethis interesting research.

Second, I want to thank my Doctoral Supervisory Committee members, Dr. Arturo Sanchez-Azofeifa, Dr. Jeffrey Kavanaugh, Dr. Guillermo Castilla, Dr. Margriet Haagsma, and Dr. Benoit Rivard, for their guidance on my thesis proposal. I also want to acknowledgeChina Scholarship Council, the organization collaborates with the University of Alberta, funds me to take a PhD program in Canada for four years.

Third, I want to thank Dr. Michael Hesketh and Meimei Chong, who work at Center for Earth Observation Sciences (CEOS), department of Earth and Atmospheric Sciences (EAS), for all their help on equipment supply and technical support in the laboratory. I also thank my academic colleagues, Dr.Jing Chen, Dr. Sen Cao, Dr. Genping Zhao, Dr. Quan Wang, Dr. Lidong Zou, Dr Yaqian Long, Dr. Kati Laakaso, Dr. Chuanliang Sun, Dr. Kayla Stan, Dr. Sofia Calvo, Dr. Antonio Guzman, Dr. Yiming Yan, Dr. Xin Chen, Dr. Rui Jiang, Dr. Xu Yuan, Felipe Alencastro, Iain Sharp, Carlos Campos, Oscar Baron,Connor Bax, Steven Wagers, Patrick O'Brien, Yao Shi, He Yan, Cristina Ke, Amanda Lei, Jonathan Liu, Menglei Duan, Willow Liu, Nooshin Mashhadi, Mohammed A Yaseen, Ronny Alexander, and Meghan Grenke. Specifically, I want to thank Iain Sharp, Carlos Campos and Felipe Alencastro for their kind help in using terrestrial laser scanning to collect data, at the Santa Rosa National Park, Costa Rica.

Finally, I want to thank my girlfriend Xueyao Chen, for her support during the last three years. With her encouragement and understanding, I could overcome the difficulties whether in life or work. Furthermore, I want to extend my thanks to my families, especially my parents, for their constant love and support for 27 years.

Contents

List of Tables

List of Figures

List of Abbreviations and definitions

MODIS: Moderate Resolution Imaging Spectroradiometer

NDVI: Normalized Difference Vegetation Index

LiDAR: Light Detection and Ranging

GLAS: Geoscience Laser Altimeter System

ALS: airborne laser scanning

TLS: terrestrial laser scanning

TLS methods: time-of-flight (**TOF**) and phased-shift (**PS**) scanners

PR curve: Precision-Recall curve

AGB: above ground biomass

TP: terrestrial photogrammetry

ST: Slice Triangulation

ASA: alpha shape algorithm

DBH: diameter at breast height

RF: Random Forest, a learning classifier

XGBoosting: eXtreme Gradient Boosting, a learning classifier

Precision: is defined as the percentage of correctly classified points (**TP**, True positive) out of all classified points (TP and **FP**, False positive)

Recall*:* is the percentage of correctly classified points (TP) out of all true points (TP and **TN**, True negative)

FCN: Fully Convolutional Neural Network

LSTM-FCN: Long Short-Term Memory Fully Convolutional Neural Network

ResNet: Residual Network

UTS: univariable time series

MTS: multivariable time series

CAM: class activation map

TSC: time series classification

RNN: recurrent neural network

SVM: support vector machine

GAP: global average pooling layer

DAB: the diameter above the buttress

$D_{area130}$: the diameter calculated from non-convex area at the breast height

$D_{convex130}$: the diameter calculated from the convex-hull perimeter the breast height

Chapter1

Introduction

1.1 Introduction

Forests account for one-third of the land area, and two-thirds of global photosynthesis, and therefore forests is one of the most essential components in global ecosystems (Bergesenet al., 2019). Forest ecosystems not only offer habitats for a wide range of plants and animals, but also, they prevent floods, soil erosion, and mitigate climate change (Waringet al., 2007). Furthermore, the economic contributions of forests are well recognized, such as timber value, food, fuel, and bioproducts. In some developing countries, forests and trees may provide around twenty percent of income for rural households (Bergesen et al., 2019). Accurate and timely monitoring of forest and its biological parameters can provide essential information for forest inventory, wildlife protection, carbon cycle modelling, forestmanagement and conservation (Srinivasan et al., 2014).

Historically, field surveys and aerial photography were the two main approaches to monitor and manage forests. Field surveys provide an accurate and detailed way to measureforest until now, though it is very time-consuming, and expensive. By contrast, aerial photography is a cost-effective alternative to retrieve forest information, and it has been widely used in many forest studies and applications, such as forest management, extent and health (e.g., dead trees) mapping, etc. (Hussin et al., 2000; Liang et al., 2012).

With the advent of remote sensing techniques, airborne and spaceborne methods are becoming widely accepted as practical and efficient approaches for forest monitoring (Yu et al., 2013). Satellite

images, such as those from Sentinel 2, Landsat and the Moderate Resolution Imaging Spectroradiometer (MODIS), capture current and past forest information over large areas due to their wide spatial coverage and repeatedly temporal observations (Shimizu et al., 2019). For example, in the Mexican state of Michoacan, forest disturbance was monitored from 2000 to 2016 using MODIS NDVI time series (Gao et al., 2021). In another study, time-series biomass maps were generated using Landsat images between 1985 and 2010 for the west Carpathian Mountains, Poland (Main-Knorn et al., 2013). The former two are just two examples of probably thousands of studies on which different typesof remote sensing platforms and techniques have been used to study different questions associated to forest extent, health and even economic value. One key denominator of manyof these studies is that the accuracy of the different studies highly relies on the quality and quantity of the reference data and landscape heterogeneity. Also, those satellite images only present forests in two dimensions. Therefore, rapid and robust assessment of accurate ground reference data of three-dimensional (3D) forest structure at plot level is essential, which could support the utilization of satellite remote sensing products for forest monitoring (Liang et al., 2012).

LiDAR (Light Detection and Ranging) can obtain the 3D forest structure with a high accuracy, and therefore it is widely used in ecological and environmental studies (Holopainen et al., 2011). This technique uses an active mode to emit laser pulses to measure the distance between a given sensor and an object (Calders et al., 2013). LiDAR system can operate from spaceborne, airborne, and terrestrial platforms. Spaceborne LiDAR is an important tool to assess vegetation structure at larger scales. The Geoscience Laser Altimeter System (GLAS) is an example of a spaceborne LiDAR systems with potential for forestry applications. For instance, useful forest parameters, such as canopy height, crown depth, and a measure of canopy cover, were extracted using GLAS (Harding et al., 2005). In addition, Helmer et al. (2006) used the GLAS to estimate forest canopy heights in Amazon River Basin Forests, showing its capacity to quantify forest carbon pools forglobal inventories.

In terms of airborne systems, the Airborne Laser Scanning (ALS) is often used in forestry applications. Since the 1980s, many studies demonstrated the capacity of ALS to measure forest attributes, such as stand height, species, and ground elevation, etc. (Holmgren et al., 2004; Suratno et al., 2009). Wang et al. (2019) investigated the reliability of tree height obtained via ALS data and found that the estimation results were more reliable if the trees were taller. Similarly, Holmgren et al. (2004) used ALS data to discriminate Scot pine and Norway spruce on an individual tree level, the overall classification accuracy was 95% on all plots when compared with field measurements.

When it comes to terrestrial platforms or terrestrial laser scanning (TLS) systems,they can obtain the 3D forest structure with millimeter accuracy. For example, the spatial resolution of TLS ranges from 0.005 to 0.10 meters according to the model and type of the equipment, while that of ALS is between 0.1 and 1.0 meters (Yang et al., 2013). Although TLS has some restrictions on spatial coverage compared to spaceborne and airborne LiDAR, it has unprecedented potential for describing 3D details of forest structure at plot level, thus providing an efficient and effective approach to collect ground reference data (Béland et al., 2011; Hosoi et al., 2006; Liang et al., 2012).

There are two types of TLS methods: time-of-flight (TOF) and phased-shift (PS) scanners, which are widely used on forest applications (Calders et al., 2020; Liang et al., 2016). The PS scanners continuously emit a laser beam into multiple phases and determinethe distance between sensor and target by comparing the phase shifts of the returned laser energy, while the TOF TLS measures that distance, using the time of the return of each laser signal emitted by the sensor (Newnham et al., 2015). TOF TLS is often regarded as the better option for vegetation studies, as it can provide rapid assessment of forest structure and cost-effective forest structure metrics (Liang et al., 2018). TOF TLS has two return systems: discrete return and waveform. The discrete return system records only the position of one contact between the laser beam and target, while the waveform records the whole of the return signal reflected from the target, providing more information about vegetation structure (Calders et al., 2020).

During the last two decades, TLS has seen many successful applications on forest measurement and management (Calders et al., 2015; Disney et al., 2018; Hackenberg et al., 2015; Holopainen et al., 2011; Liang et al., 2012). For instance, Kankare et al. (2014) used TLS point clouds to measure tree attributes including, tree height, DBH, diameter at 6 m and the lower living and dead branch height. The relative root mean squared errors (RMSE) of those parameters were 7.1%, 5.9%, 8.9%, 9.6%, and 42.9% when comparing with field measurements. Although the highest errors of dead branch heights were due to the occlusion effect in the point cloud, the accuracy could increase if the understory vegetation can be cleaned. Also, Moorthy et al. (2018) explored the utility of TLS data to detect and quantify changes in forest structure after lianas removal in a tropical forest. Local structure changes such as vertical plant profiles and canopy height change can be successfully detected after lianas removal. Furthermore, Calders et al. (2015) used TLS data to estimate above ground biomass (AGB), and compared this estimate with reference data from field measurements and AGB derived from allomeric models. Compared with reference data, the AGB estimates derived from TLS showed a total

AGB overestimation of 9.68%, while that from allomeric models showed an underestimation of 36.57%-29.85%.

In addition to TLS, terrestrial photogrammetry (TP) could also generate dense 3D point clouds for field-based forest studies (Bauwens et al., 2017; Cushman et al., 2021). For example, Bauwens et al. (2017) investigated the potential of TP for measuring and modelling 43 irregular stems. After clearing the understory vegetation around the focal tree, the photogrammetric process showed a higher success rate (80%) for trunk shape reconstruction. TP is a useful, cost-effective method for 3D modeling of trunks, though it can be affected by lower light and obstructive vegetation conditions in the forest understory,especially in structurally dense tropical forests (Cushman et al., 2021).

Though TLS has achieved many advances in forest inventory, there are still some knowledge gaps that remain. Specifically, elements such as volume estimation of irregular trees in tropical forests, foliar and woody materials discrimination, the quantification of lianas, stem quality assessment and the monitoring of change on biomass are some of the many challenges still exist in contemporary fields (Kankare et al., 2014; Liang et al., 2016; Moorthy et al., 2018; Zhu et al., 2018).

Lianas are woody climbers with relatively thick stem that use trees as structure support to reach the forest canopy (Ingwell et al., 2010; Martínez-Izquierdo et al., 2016). Although lianas account for a small proportion (around 5%) of the above-ground biomass in tropical forests, they have a large percentage of leaf productivity and canopy cover (Moorthy et al., 2018; Rodríguez-Ronderos et al., 2016; Schnitzer, 2005). Apart from competing with trees for above and below ground resources, lianas also occupy gaps and illuminated areas on the upper part of the canopy more efficiently (Letcher and Chazdon, 2009). Therefore, an increase in lianas abundance will suppress tree regeneration, promote tree mortality, and decrease tree growth, thereby affecting the whole forest carbon sequestration (Ingwell et al., 2010; Martínez-Izquierdo et al., 2016).

Forest canopy structure refers to the size, shape, orientation, position, and connectedness, of tree components, including stem, branches, and leaves, in a three-dimensional space. Also, the biochemical and structural properties of forest canopy are often used as metrics to understand relations between the atmosphere and the land surface (Béland et al., 2014). Meanwhile, the amount of green leaves in canopy controls key ecological processes such as photosynthesis activity, gas exchange, and light interception, while the woody part contributes to the total volume of trees (Béland et al., 2014; Tao et al., 2015). Accurate separation of foliage and woody materials from the canopy has the potential to

improve our understanding and management of tropical forests from a structural point of view.

Buttresses are large, wide roots on all sides of a shallowly rooted tree (Mehedi et al., 2012). Buttresses are considered mechanical structures that prevent the tree from falling and balance the trees against unidirectional stresses such as asymmetric canopy, and prevailing wind (Chapman et al., 1998; Zhiyuan et al., 2013). Another main hypothesis related to presence of buttresses on a tree is that they promote nutrient acquisition. Newbery et al. (2009) found that the degree of shallowness and spatial extension of buttresses is a negative proportion to nutrient availability. Moreover, Pandey et al. (2011) studied an Indian tropical rainforest and suggested that buttresses simultaneously build a pool of mineral nitrogen while increasing the supply of plant-available nitrogen. Furthermore, Tang et al. (2011) compared soil moisture on the uphill, downhill and lateral sides of buttresses over a year on a Chinese tropical rainforest. The study found that buttresses can serve as barriers to matter flow, and increase the contact area between the tree and the ground, as litter accumulation, soil moisture and nutrition, while species abundance were much higher on the uphill side.

Therefore, the overall objectives of this thesis are therefore to:

1) develop a novel machine learning model, based on Random Forest and eXtreme Gradient Boosting, to separate lianas from host trees, using geometric features derived from TLS point clouds

2) investigate the potential of different deep learning time series approaches to separateleaf and woody components from TLS data

3) build a non-destructive and reliable method for buttresses volume estimation basedon 3D point clouds

This thesis explores those objectives using three separate papers, all of which look at the potential of point clouds to solve scientific problems in forestry fields.

1.2 Outline

This thesis work is composed of five chapters, including this introductory chapter. Chapter 2 to 4 address the mentioned three objectives in the above section.

Extraction of liana Stems from Terrestrial LiDAR data using Geometric Features (Remote Sensing, https://doi.org/10.3390/rs14164039). Chapter 2 uses a new machine learning model, based on Random Forest (RF) and eXtreme Gradient Boosting (XGBoosting), to separate a liana from its host tree, using TLS point clouds offive tropical dry forest trees, which have different liana infestation levels. In this chapter, first, we use a multiple radius nearest neighbors approach to define the optimal radius of nine features, where a large difference is shown between the geometric features of liana andtree. Second, we compare the performance of RF and XGBoosting on the classification of liana and tree. Finally, we evaluate our model on independent data and compare our model with a liana/tree separation study. The key results of this chapter suggest that the performance of XGBoosting is similar to RF for liana and tree classification, and the proposed method is more efficient than the previous study. Given the irregular growth strategy of lianas, and TLS can provide detailed 3D information of forests, TLS could provide a reasonable way to study the impact of lianas on tropical forests.

A deep learning time-series approach for leaf and wood classification from terrestrial LiDAR point clouds (Remote Sensing, https://doi.org/10.3390/ rs14133157). Chapter 3 demonstrates the utilization of deep learning time-series approach for leaf and wood separation, based on TLS point clouds. Specifically, we use a multiple- radius nearest neighbors approach to obtain the time series of geometric features from point clouds, and compare the performance of Fully Convolutional Neural Network (FCN), Long Short-Term Memory Fully Convolutional Neural Network (LSTM-FCN), and Residual Network (ResNet) on leaf and wood classification. We also compare the effect of univariable (UTS) and multivariable (MTS) time series on classification accuracy. Additionally, we evaluate the performance of presented approach against an independent dataset to indicate its broader generality. Furthermore, we explore the utilization of the Class Activation Map (CAM) to explain the black-box effect of deep learning. In this chapter, we demonstrate deep learning coupled with the time series of geometric features can accurately separate leaf and woody components from point clouds.

A non-destructive approach to estimate buttress volume using 3D point cloud data. Chapter 4 presents the utilization of alpha shape algorithm (ASA) for buttress volume estimation. Specifically, we compare the performance of ASA and slice triangulation (ST) on estimating volume of 30 buttressed trees. Meanwhile, we use an independent dataset to indicate broader application of the ASA. In addition, we developed allometric models with the diameter above the buttress (DAB), the diameter calculated from the non-convex area ($D_{area130}$), and convex hull perimeter ($D_{convex130}$) at breast

height (1.3 m). Then, the developed allometric models are validated with an independent dataset collected using the terrestrial laser scanning (TLS) and destructive measurements. The main results in this chapter indicate the ASA is a more applicable and transferable approach than ST and allometric models for buttress modeling.

Chapter 5 illustrates the main conclusions and significant contributions of this thesis, along with challenges and future directions.

1.3 References

Abu Hanifa Mehedi, M., Kundu, C., Qumruzzaman Chowdhury, M., 2012. Patterns of tree buttressing at Lawachara National Park, Bangladesh. J. For. Res. 23, 461–466. https://doi.org/10.1007/s11676-012-0285-x

Bauwens, Sebastien, Adeline Fayolle, Sylvie Gourlet-Fleury, Leopold Mianda Ndjele, Coralie Mengal, and Philippe Lejeune (2017). "Terrestrial photogrammetry: a non-destructive method for modelling irregularly shaped tropical tree trunks". In: *Methods in Ecology and Evolution* 8.4, pp. 460–471.

Béland, Martin, Jean-Luc Widlowski, Richard A Fournier, Jean-François Côté, and Michel M Verstraete (2011). "Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements". In: *Agricultural and Forest Meteorology* 151.9, pp. 1252–1266.

Béland, M., Baldocchi, D. D., Widlowski, J. L., Fournier, R. A., & Verstraete, M. M. (2014). On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. Agricultural and Forest Meteorology, 184, 82–97. https://doi.org/10.1016/j.agrformet.2013.09.005

Bergesen, Helge Ole, Georg Parmann, and Oystein B. Thommessen (2019). Food and Agri- culture Organization (FAO). FAO, pp. 201–201. isbn: 9789251305614. doi: 10.4324/ 9781315066547-55.

Calders, Kim, Jennifer Adams, John Armston, Harm Bartholomeus, Sebastien Bauwens, Lisa Patrick Bentley, Jerome Chave, F Mark Danson, Miro Demol, Mathias Disney, et al. (2020). "Terrestrial laser scanning in forest ecology: Expanding the horizon". In: Remote Sensing of Environment 251, p. 112102.

Calders, Kim, Glenn Newnham, Andrew Burt, Simon Murphy, Pasi Raumonen, Martin Herold, Darius Culvenor, Valerio Avitabile, Mathias Disney, John Armston, et al. (2015). “Nondestructive estimates of above-ground biomass using terrestrial laser scan- ning”. In: Methods in Ecology and Evolution 6.2, pp. 198–208.

Calders, Kim, Glenn Newnham, Martin Herold, Simon Murphy, Darius Culvenor, Pasi Rau- monen, Andrew Burt, John Armston, Valerio Avitabile, and Mathias Disney (2013). “Estimating above ground biomass from terrestrial laser scanning in Australian Eu- calypt Open Forest”. In: Proceedings SilviLaser 2013, 9-11 October, Beijing, China, pp. 90–97.

Chapman, C. A., Kaufman, L., & Chapman, L. J. (1998). Buttress formation and directional stress experienced during critical phases of tree development. Journal of Tropical Ecology, 14(3), 341–349. https://doi.org/10.1017/S0266467498000261

Cushman, KC, Sarayudh Bunyavejchewin, Dairon C´ardenas, Richard Condit, Stuart J Davies, A´lvaro Duque, Stephen P Hubbell, Somboon Kiratiprayoon, Shawn KY Lum, and Helene C Muller-Landau (2021). “Variation in trunk taper of buttressed trees within and among five lowland tropical forests”. In: Biotropica 53.5, pp. 1442–1453.

Disney, Mathias I, Matheus Boni Vicari, Andrew Burt, Kim Calders, Simon L Lewis, Pasi Raumonen, and Phil Wilkes (2018). “Weighing trees with lasers: advances, challenges and opportunities”. In: Interface Focus 8.2, p. 20170048.

Gao, Yan, Alexander Quevedo, Zoltan Szantoi, and Margaret Skutsch (2021). “Monitoring forest disturbance using time-series MODIS NDVI in Michoac´an, Mexico”. In: Geocarto International 36.15, pp. 1768–1784.

Hackenberg, Jan, Marc Wassenberg, Heinrich Spiecker, and Dongjing Sun (2015). “Non destructive method for biomass prediction combining TLS derived tree volume and wood density”. In: Forests 6.4, pp. 1274–1300.

Harding, David J and Claudia C Carabajal (2005). “ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure”. In: Geophysical research letters 32.21.

Helmer, EH and MA Lefsky (2006). “Forest canopy heights in Amazon River basin forests as estimated with the Geoscience Laser Altimeter System (GLAS)”. In: In: Aguirre- Bravo, C.; Pellicane, Patrick J.; Burns, Denver P.; and Draggan, Sidney, Eds. 2006. Monitoring Science and Technology Symposium: Unifying Knowledge for Sustainability in the Western Hemisphere Proceedings RMRS-P-42CD. Fort Collins, CO: US De- partment of Agriculture, Forest Service, Rocky Mountain Research Station. p. 802-808. Vol. 42.

Holmgren, Johan and ˚Asa Persson (2004). “Identifying species of individual trees using airborne laser scanner”. In: Remote Sensing of Environment 90.4, pp. 415–423.

Holopainen, Markus, Mikko Vastaranta, Ville Kankare, Minna R¨aty, Matti Vaaja, Xinlian Liang, Xiaowei Yu, Juha Hyypp¨a, Hannu Hyyppa¨, R Viitala, et al. (2011). “Biomass estimation of individual trees using stem and crown diameter TLS measurements”. In: ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 3812, pp. 91–95.

Hosoi, Fumiki and Kenji Omasa (2006). “Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar”. In: IEEE transactions on geoscience and remote sensing 44.12, pp. 3610–3618.

Hussin, Y and Witske Bijker (2000). “Inventory of remote sensing applications in forestry for sustainable management”. In: International Archives of Photogrammetry and Re- mote Sensing 33.B7/2; PART 7, pp. 575–57.

Kankare, Ville, Marianna Joensuu, Jari Vauhkonen, Markus Holopainen, Topi Tanhuanp¨a¨a, Mikko Vastaranta, Juha Hyypp¨a, Hannu Hyyppa¨, Petteri Alho, Juha Rikala, et al. (2014). “Estimation of the timber quality of Scots pine with terrestrial laser scanning”. In: Forests 5.8, pp. 1879–1895.

Liang, Xinlian, Juha Hyyppa¨, Harri Kaartinen, Markus Holopainen, and Timo Melkas (2012). “Detecting changes in forest structure over time with bi-temporal terrestrial laser scanning data”. In: ISPRS International Journal of Geo-Information 1.3, pp. 242– 255.

Liang, Xinlian, Juha Hyyppa¨, Harri Kaartinen, Matti Lehtom¨aki, Jiri Py¨or¨al¨a, Norbert Pfeifer, Markus Holopainen, G´abor Brolly, Pirotti Francesco, Jan Hackenberg, et al. (2018). “International benchmarking of terrestrial laser scanning approaches for forest inventories”. In: ISPRS journal of photogrammetry and remote sensing 144, pp. 137– 179.

Liang, Xinlian, Ville Kankare, Juha Hyyppa¨, Yunsheng Wang, Antero Kukko, Henrik Haggr´en, Xiaowei Yu, Harri Kaartinen, Anttoni Jaakkola, Fengying Guan, et al. (2016). “Terrestrial laser scanning in forest inventories”. In: ISPRS Journal of Photogrammetry and Remote Sensing 115, pp. 63–77.

Main-Knorn, Magdalena, Warren B Cohen, Robert E Kennedy, Wojciech Grodzki, Dirk Pflugmacher, Patrick Griffiths, and Patrick Hostert (2013). “Monitoring coniferous for- est biomass change using a Landsat trajectory-based approach”. In: Remote Sensing of Environment 139, pp. 277–290.

Moorthy, Sruthi MK, Kim Calders, Manfredo Di Porcia e Brugnera, Stefan A Schnitzer, and Hans Verbeeck (2018). “Terrestrial laser scanning to detect liana impact on forest structure”. In: Remote Sensing 10.6, p. 810.

Newbery, D. M., Schwan, S., Chuyong, G. B., & Van Der Burgt, X. M. (2009). Buttress form of the central African rain forest tree Microberlinia bisulcata, and its possible role in nutrient acquisition. Trees - Structure and Function, 23(2), 219–234. https://doi.org/10.1007/s00468-008-0270-3

Newnham, Glenn J, John D Armston, Kim Calders, Mathias I Disney, Jenny L Lovell, Crystal B Schaaf, Alan H Strahler, and F Mark Danson (2015). “Terrestrial laser scan- ning for plot-scale forest measurement”. In: Current Forestry Reports 1.4, pp. 239–251.

Pandey, C. B., Singh, L., & Singh, S. K. (2011). Buttresses induced habitat heterogeneity increases nitrogen availability in tropical rainforests. Forest Ecology and Management, 262(9), 1679–1685. https://doi.org/10.1016/j.foreco.2011.07.019

Shimizu, Katsuto, Tetsuji Ota, and Nobuya Mizoue (2019). “Detecting forest changes using dense Landsat 8 and Sentinel-1 time series data in tropical seasonal forests”. In: Remote Sensing 11.16, p. 1899.

Srinivasan, Shruthi, Sorin C Popescu, Marian Eriksson, Ryan D Sheridan, and Nian-Wei Ku (2014). “Multi-temporal terrestrial laser scanning for modeling tree biomass change”. In: Forest Ecology and Management 318, pp. 304–317.

Suratno, Agus, Carl Seielstad, and Lloyd Queen (2009). “Tree species identification in mixed coniferous forest using airborne laser scanning”. In: ISPRS Journal of Pho- togrammetry and Remote Sensing 64.6, pp. 683–693.

Tang, Y., Yang, X., Cao, M., Baskin, C. C., & Baskin, J. M. (2011). Buttress trees elevate soil heterogeneity and regulate seedling diversity in a tropical rainforest. Plant and Soil, 338(1), 301–309. https://doi.org/10.1007/s11104-010-0546-4

Tao, S., Guo, Q., Xu, S., Su, Y., Li, Y., & Wu, F. (2015). A Geometric Method for Wood-Leaf Separation Using Terrestrial and Simulated Lidar Data. Photogrammetric Engineering & Remote Sensing, 81(10), 767–776. https://doi.org/10.14358/PERS.81.10.767

Wang, Yunsheng, Matti Lehtom¨aki, Xinlian Liang, Jiri Py¨or¨al¨a, Antero Kukko, Anttoni Jaakkola, Jingbin Liu, Ziyi Feng, Ruizhi Chen, and Juha Hyyppa¨ (2019). "Is field-measured tree height as reliable as believed–A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest". In: ISPRS Journal of Photogrammetry and Remote Sensing 147, pp. 132–145.

Waring, RH and SW Running (2007). "Forest Ecosystems. Analysis at Multiple Time and Space Scales". In: Forest Ecosystems (Third Edition). Academic Press, pp. 1–16.

Yang, Xiaoyuan, Alan H Strahler, Crystal B Schaaf, David LB Jupp, Tian Yao, Feng Zhao, Zhuosen Wang, Darius S Culvenor, Glenn J Newnham, Jenny L Lovell, et al. (2013). "Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®)". In: Remote sensing of environment 135, pp. 36–51.

Yu, Xiaowei, Xinlian Liang, Juha Hyypp¨a, Ville Kankare, Mikko Vastaranta, and Markus Holopainen (2013). "Stem biomass estimation based on stem reconstruction from ter- restrial laser scanning point clouds". In: Remote Sensing Letters 4.4, pp. 344–353.

Zhiyuan, H., Yong, T., Xiaobao, D., & Min, C. (2013). Buttress trees in a 20-hectare tropical dipterocarp rainforest in Xishuangbanna, SW China. Journal of Plant Ecology, 6(2), 187–192. https://doi.org/10.1093/jpe/rts031

Zhu, Xi, Andrew K Skidmore, Roshanak Darvishzadeh, K Olaf Niemann, Jing Liu, Yi- fang Shi, and Tiejun Wang (2018). "Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest". In: International journal of applied earth observation and geoinformation 64, pp. 43–50.

Chapter2

Extraction of liana Stems from Terrestrial LiDAR data using Geometric Features

Abstract

Lianas are self-supporting systems that are increasing their dominance in tropical forests due to climate change. As lianas increase tree mortality and reduce tree growth, one key challenge in ecological remote sensing is the separation of a liana and its host tree using remote sensing techniques. This separation can provide essential insights into how tropical forests respond, from the point of view of ecosystem structure to climate and environmental change. Here, we propose a new machine learning method, derived from Random Forest (RF) and eXtreme Gradient Boosting (XGBoosting) algorithms, to separate lianas and trees using Terrestrial Laser Scanning (TLS) point clouds. We test our method on five tropical dry forest trees with different levels of liana infestation. First, we use a multiple radius search method to define the optimal radius of six geometric features. Second, we compare the performance of RF and XGBoosting algorithms on the classification of lianas and trees. Finally, we evaluate our model against independent data collected by other projects. Our results show that the XGBoosting algorithm achieves an overall accuracy of 0.88 (recall of 0.66), and the RF algorithm has an accuracy of 0.85 (recall of 0.56). Our results also show that the optimal radius method is as accurate as the multiple radius method, with F1 scores of 0.49 and 0.48, respectively. The RF algorithm shows the highest recall of 0.88 on the independent data. Our method

provides a new flexible approach to extracting lianas from 3D point clouds, facilitating TLS to support new studies aimed to evaluate the impact of lianas on tree and forest structures using point clouds.

Keywords: Life forms, Classification, Machine Learning, Geometric Feature, Point Clouds

2.1 Introduction

Lianas are climber plants with woody stems, and they can climb to the forest canopy with the support of trees (Ingwell et al., 2010; Martinez-Izquierdo et al., 2016). Although lianas occupy a small proportion (less than 10%) of the above-ground biomass in tropical forests (Durán et al., 2013), they represent a large percent (up to 40%) of leaf productivity (Ingwell et al., 2010; Rodríguez-Ronderos et al., 2016; Schnitzer, 2005). Apart from competing with trees for available resources (Schnitzer, 2005), lianas also occupy gaps and illuminated areas on the upper part of the canopy more efficiently (Letcher et al., 2009; Sánchez-Azofeifa et al., 2006). During the last two decades, liana abundance and biomass have experienced increases in tropical forests (Schnitzer et al., 2020; Wright, 2005). These increases tend to be associated with changes in forest structure (Sánchez-Azofeifa et al., 2017). An increase in liana abundance can suppress tree generation, promote tree mortality, and decrease tree growth, thereby causing a cascade effect on carbon storage, biodiversity, and primary productivity (Gonzalez de Tanago et al., 2018; Schnitzer et al.,2011; Schnitzer et al., 2002).

Unlike other structure parasites such as epiphytes and hemi-epiphytes, lianas grow their roots into the ground (Gentry, 1991; Schnitzer, 2018). There are many growth strategies that lianas can link to their infested trees and grow to the forest canopy. Those strategies contain stem twining, clasping tendrils arising from the stem, leaf, and branch modification, down-ward-pointing adhesive hairs, adhesive adventitious roots, and thorns and spines that link lianas and trees (Schnitzer et al., 2002). The final climbing mechanismthat lianas use from the above strategies is determined by the forest successional stages or disturbance level (Dewalt et al., 2000; Schnitzer et al., 2002). Since lianas have irregular growth forms, they contribute considerably to the architectural complexity of a given forest (Moorthy et al., 2019; Schnitzer, 2005).

Many studies report the prevalence of lianas and their importance in tropical forests. However, few of them utilize quantitative methods to study their structure and their respective impact on an individual tree or forest plot (Londre et al., 2006; Moorthy et al., 2018). Recent advances in

remote sensing techniques, in particular, Terrestrial Laser Scanning (TLS), provide a great opportunity to study the relationships between tree and liana structures in an unprecedented manner (Calders et al., 2015; Calders et al., 2013; Sanchez-Azofeifa et al., 2017). For instance, Moorthy et al. (2018) investigated the potential utilization of TLS in monitoring changes in forest structure after liana removal, demonstrated that TLS could detect local structural changes after liana removal. Furthermore, Bao et al. (2018) used TLS data and Random Forest (RF) to classify lianas and trees with an overall accuracy of 94%. Although those studies present promising results, there are still uncertainties regarding the use of TLS to separate lianas from trees and extract them for their hosts. For example, as lianas grow towards a tree canopy, the stem diameter of most lianas is generally smaller than 10 cm, making it challenging to separate lianas and tree branches from a given point cloud. Also, lianas have a more irregular shape than trees, and they could grow around them in all directions randomizing the volumetric occupancy of the forest 3D space (Sanchez-Azofeifa et al., 2017). This attribute only increases a TLS point cloud's complexity, making lianas harder to separate and then study.

There are three approaches often used in previous studies to separate leaf and wood from TLS point clouds: (1) those using radiometric features (Beland et al., 2014; Beland et al., 2011), (2) those using geometric features (Ma et al., 2015; Tao et al., 2015a; Vicari et al., 2019), and (3) combined use of both features (Zhu et al., 2018). Methods based on radiometric features depend on the type of TLS system used to generate a point cloud since the optical properties of a tree (e.g., wood or leaves) may respond differently to the LiDAR wavelength (Vicari et al., 2019). The x-, y-, and z- coordinates of each point are basic information of all LiDAR systems. Additionally, the geometric feature approach proves to be better than the radiometric method when comparing classification accuracies (Ma et al., 2015). In the third method, a mixture of the two features mentioned above is used. Zhu et al. (2018) applied the mixed approach to separate leaves from trees, achieving an average overall accuracy of 84%. However, this mixture method still has the disadvantage of relying on radiometric features (Vicari et al., 2019).

Geometric feature methods are commonly derived from eigenvalues and eigenvectors calculated for each (x,y,z) point. A local neighborhood point set is required to obtain those features (Demantke et al., 2011). There are two approaches in terms of the definition of the local neighborhood of a point: k -nearest neighbors and radially bounded nearest neighbors. The former specifies a limited number of nearest neighbors for each point. In contrast, the later approach defines a spherical space, and local point sets are within a given radius (Zhu et al., 2018). When comparing the two approaches, the radially bounded nearest neighbors approach seems more advanced, since the k -nearest neighbors approach is affected by the density of point clouds (Ma et al., 2015; Vicari et al., 2019). For example, the canopy has

a lower point density than the understory because of occlusion effects and the physical distance between the object and scanner (Beland et al., 2014). As a result, using the same k -nearest neighbors approach, the geometry of leaves in the understory will be different from that in the canopy (Moorthy et al., 2019). Moreover, Thomas et al. (2018) indicated that the radially bounded nearest neighbors approach could allow the computed features (e.g., eigentropy, linearity, planarity, etc..) with a consistent geometric meaning, which is not possible for the k -nearest neighbors approach.

Concerning the radially bounded nearest neighbor approach, the size of the radius determines the local dimensional features of neighborhood points. For example, at a few centimeter scales, stem points are more likely to have surface features, while they tend to have linearly distributed ones at a large centimeter scale (Ma et al., 2015). Using a single radius randomly selected to search for neighborhood points, may not be reliable because the feature distribution of different objects varies in different situations (Zhu et al., 2018). This scale factor could be an important element to consider given how close lianas grow attached to a tree's branch and trunk. A combination of the geometric features from all scales is also used in Belton et al. (2013). This strategy shows a significant advantage compared to the single scale on leaf and wood classification. However, combing the features from all radiuses could produce many more of them than a single scale approach. Then, this multiple radius method would result in redundancy and low computational efficiency of the final model (Koenig et al., 2015). Selecting an optimal radius could be the right choice for classification purposes. Zhu et al. (2018) applied an adaptive radius near-neighbor search method to select an optimal radius for each point, and then extracted geometric features of those points to classify foliar and woody materials of a mixed forest. As their results suggested, the classification accuracy for adaptive radius near-neighbor search method was higher than that for the fixed radius near-neighbor search method.

There are many existing methods using machine learning to classify point clouds, while the purpose of most of those methods is to classify leaf and wood points (Ma et al., 2015; Tao et al., 2015a; Vicari et al., 2019; Wang et al., 2018). Only a few studies use TLS data to discriminate lianas and their host trees in tropical forests. Moorthy et al. (2019) presented a semi-automatic approach based on the Random Forest algorithm to separate liana from its host tree in Panama and French Guiana. In Moorthy et al. (2019), model recall of extracted liana stems increased after a manual intervention (54% to 90% and 65% to 70%, respectively). However, many lianas in their study are away from the main trunk, and with a vertical distribution. When lianas are close to a tree's main stem, and have more irregular shapes, their method cannot identify lianas successfully.

This work contributes to building a novel machine learning model based on Random Forests and

Gradient Boosting algorithms, that can be used to separate lianas from trees. This method uses geometric features derived from point clouds. Our method extends and builds on the previous work of Belton et al. (2013) and Moorthy et al. (2019) and Tao et al. (2015a). Furthermore, in this manuscript, we aim to answer the following questions: (1) Does the optimal radius method have significant advantages compared to the multiple radius method for liana/tree separation in Moorthy et al. (2019)? (2) Do different machine learning methods affect the classification accuracy of lianas and trees? And (3) Is our method more efficient than other existing methods for 3D point cloud classification? We answer those questions using a set of liana and trees in a Tropical Dry Forest (TDF) with different infestation levels.

2.2 Materials and Methods

2.2.1 Study area and data

The original data come from the Santa Rosa National Park - Environmental Monitoring Supersite (SRNP-EMSS) (Figure S2.1), Guanacaste, Costa Rica. The SRNP-EMSS has an average yearly temperature of 25 ℃, and average yearly precipitation of 1720 mm. The SRNP-EMSS also has five months of the dry season (December to April), and a high biodiversity of plant and animal species, including 96 species of trees of different life histories, and 18 species of lianas among them (Guzmán Q et al., 2018).

We randomly selected five trees with different levels of liana infestation, and then used a Leica C10 TLS scanner (Leica Geosystem AG, Sankt Gallen, Switzerland) in May 2015 (end of the dry season). The Leica C10 TLS scanner uses green laser light with a wavelength of 532 nm, a maximum vertical field of view (FOV) of 270°, and a horizontal FOV of 360° (Feliciano et al., 2014). The scanner collects 50,000 points per second at a range of up to 300 m (Gonzalez de Tanago et al., 2018).

Table 2. 1: Description of the trees with liana infestation collected in Santa Rosa National- Park Environmental Monitoring Supersite, Costa Rica; *, three DBH value for Tree 3 due to the three separated stems at the breast height (1.3 m).

Tree ID	Tree stem	DBH (cm)	Height (m)	Liana infestation levels	Liana points proportion	Number of total points
Tree 1	1	54.10	16.49	Low	1%	405262
Tree 2	1	30.70	14.51	Low	4%	121954
Tree 3	1	28.90/31.20/38.40*	16.57	High	9%	382599
Tree 4	1	55	17.10	High	15%	622967
Tree 5	1	20.50	13.96	Intermediate	37%	87346

We selected four or five scan positions at a radius of 10 m to collect the TLS data. The location of scan positions was determined by showing the highest visibility of the host tree and lianas. At least four reflective targets were used as control points to merge the point cloud from the four positions into a single high-resolution 3D data, with fewer shadow effects (Cote et al., 2012). This registration step was performed under the projected coordinate system (Cartesian) in Leica's Cyclone software using the Iterative Closet Point (ICP) method (mean error of 0.02 m) (Kankare et al., 2013). The data collection process was done on sunny days with low wind conditions. A detailed description of all five trees is given in Table 2.1. Specifically, we used the Point Picking and Height Histogram function in Cloud Compare software to obtain DBH and Height from the point cloud, respectively. Figure 2.1 presents the five trees with different infestation levels. The infestation levels were evaluated by professional dendrologists using standard forestry techniques.

We used an independent tree from Nouragues, French Guiana to test the performance of our method on unseen data (Moorthy et al., 2019). Specifically, Nouragues is a lowland moist tropical forest, which has a higher liana infestation (Schnitzer et al., 2006). Moorthy et al. (2019) utilized a Riegl VZ-1000 scanner (RIEGL, Horn, Austria) to scan the target tree (August – October 2017). The Riegl VZ-1000 scanner uses a narrow infrared laser light with a wavelength of 1550 nm, a maximum vertical field of view (FOV) of 100°, and a horizontal FOV of 360° (Schneider et al., 2019). Specifically, five scan positions at a radius of 15 m are selected to get better visibility of the tree with liana infestation. Around 20 reflective targets were used to co-register the point cloud from those five positions to obtain single high-resolution 3D data (Moorthy et al., 2019).

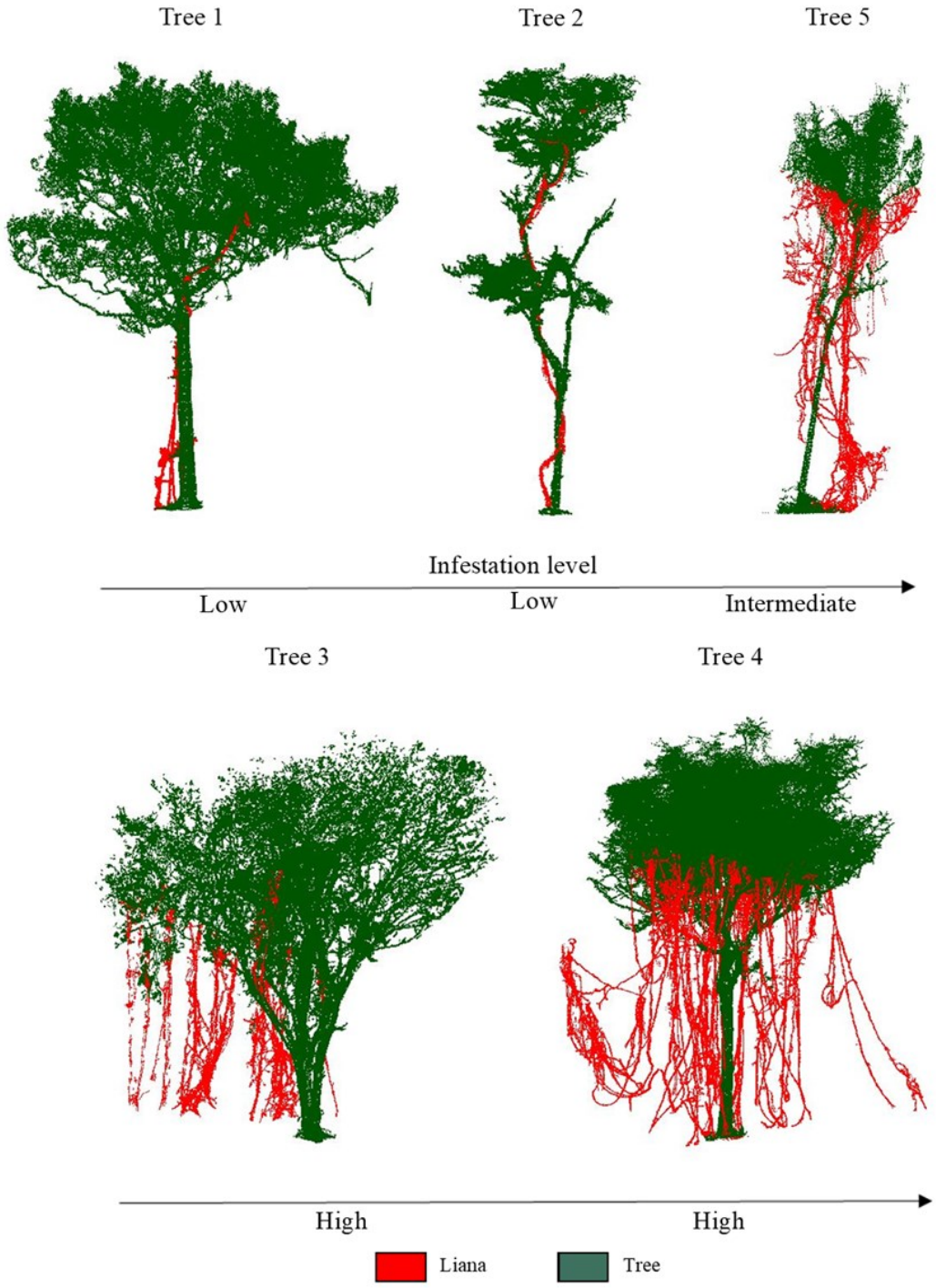


Figure 2. 1: Tree with different liana infestation collected at the SRNP–EMSS, Guanacaste, Costa Rica. The red color indicates the lianas stem, while the green color is the tree.

2.2.2 Data processing

Predictor variables

Geometric features estimate local geometry by characterizing the distribution of neigh- boring points (Moorthy et al., 2019; Zhu et al., 2018). The proper estimation of geometric features requires using a covariance matrix, calculated from a local point set within a certain radius. Then, the three positive and ordered eigenvalues computed from the covariance matrix (λ1 ≥ λ2 ≥ λ3) can express how the local point set is distributed in the 3D space (Ma et al., 2015).

Here, we used Cloud Compare (version 2.11, Cloud Compare, GPL software) to calculate the six geometric features (Table 2.2) to classify lianas and trees. These features are often used to separate leaf and woods from point clouds (Weinmann et al., 2015; Wang et al., 2017).

Table 2. 2: Six geometric features extracted from the point cloud. z represents the height of the point cloud, nv is the normal vector, and λi, the eigenvalues.

No.	Feature	Description
1	Omnivariance	$3\sqrt{(\lambda_1 * \lambda_2 * \lambda_3)}$
2	Anisotropy	$(\lambda_1 - \lambda_3)/\lambda_1$
3	Planarity	$(\lambda_2 - \lambda_3)/\lambda_1$
4	Linearity	$(\lambda_1 - \lambda_2)/\lambda_1$
5	Sphericity	λ_3/λ_1
6	Verticality	$1 - nv_z$

Data labeling

The most crucial part of building the liana/tree classification model regarding supervised machine learning methods is to label the training data. Here, we used Cloud Compare to manually mark the five tree point clouds into two classes, where class 1 and class 2 represented liana points, and tree points, respectively. To label the liana points, we visually follow lianas from the ground to the tree crown. This labelling process stops when we see lianas branch out to leaves. The data labelling process costs around 100 hours. The proportion of liana points (an average of 8% of all the points) was much less than the proportion of tree points. This imbalance between liana on tree points has the potential to cause an unbalance weight in the model development (Thomas et al., 2018). Therefore, we randomly sampled 25% of all tree points to improve our classifier's performance. As a result, tree points were two times as many as liana points in the final dataset for training and validation of the classifier.

Optimum radius for near-neighbor search

Some studies use a fixed radius to obtain the geometric features (Ma et al., 2016; Ma et al., 2015). For example, the diameter at breast height (DBH) of a tree was used as reference to select a suitable searching radius to separate leaf and wood from a point cloud (Ma et al., 2016). However, it is unlikely that the DBH is consistent across trees and plots, given that the DBH at the study site changes as a function of species and successional stage (Calvo-Rodriguez et al., 2020). As such, we used a multiple radius search solution to define the optimal radius of the nine features. This approach considers a heterogeneous DBH, while it could show a complete difference between the geometric features of liana and tree.

In this study, the geometric features in Table 2.2 were computed from 0.05 to 1 m, at a 0.01 m interval. A total of 96 values where created for each feature. We set up a 0.05 m as minimum radius because we used a voxel grid filter with a size of 0.04 m to downsample our point cloud. There are two reasons why we select the size of 0.04 m to filter the voxel grid in this study. First, downsampling ensures that the distribution of points is uniform, which means high computational efficiency, although this may cause some information loss (Burt et al., 2019). Second, the downsampling size in Moorthy et al. (2019) is also equal to 0.04 m, allowing us to compare our study with the aforementioned paper. The maximum points were six at 0.05 m in this study. Radius beyond 1 m was also not considered, as this would need high computational memory. We used R studio (R Core Team, 2021) to randomly select five times different numbers of liana and tree points. Then, we used Tree 2 to express how the geometric features

of liana and tree changed from 0.05 to 1 m. As a result, five different curves for each class were obtained. We selected 100, 200, 300, 400, 500 points in the study. The final number of points was determined by observing whether there was stability among those five curves for each class. Here, 500 points were used to plot the changing trend of liana and tree for all features. The optimum radius was defined where the most considerable difference was showed between liana and tree curves. Specifically, we calculated the mean difference between liana and tree for each feature, and then find out the best radius that has the largest difference (see supplementary materials). Finally, we used all features at the optimum radius for the following classification.

Classification

Two ensemble learning classifiers, Random Forest (RF) and eXtreme Gradient Boosting (XGBoosting), are used to develop the liana extraction model. Ensemble learning classifiers combine many weak learners, such as decision trees, to form a strong one (Rhys, 2020), and they show more advantages than other individual classifiers based on one strong learner (Moorthy et al., 2019). The RF has the ability to limit overfitting. Also, considerable performance is obtained by RF when applied to leaf and wood classifications (D. Wang et al., 2017; Zhu et al., 2018). XGBoosting is a well-known classifier among a number of Gradient boosting algorithms, while this classifier can show high performance on various tasks (Chen et al., 2016). The implementation of the RF and XGBoosting algorithms, to separate lianas and trees from the different point clouds, was done using the mlr R package (Rhys, 2020).

In terms of the RF algorithm, there are four important hyperparameters: (1) number of decision trees in the forest, (2) number of random features to be sampled at each node, (3) minimum number of cases allowed in a leaf, and (4) maximum number of leaves. The grid search method is often used as a tuning tool since it can always find the best-performing hyperparameters (Lakicevic et al., 2020). However, since we have four hyperparameters here, using the grid search over this four-dimensional space requires much time and computational budget. Therefore, we used the random search method to find the best-performing hyperparameters. It should be noted that random search cannot always guarantee getting the best set of hyperparameters. In other words, it could find a good combination of hyperparameter values that performs well if we provide enough iterations. As such, 500 combinations of hyperparameters were run using the random search. Concerning the XGBoosting algorithm, there are eight hyperparameters to be tuned: (1) *learning rate*, (2) *gamma*, (3) *max depth*, (4) *min child weight*, (5) *subsample*, (6) *colsample bytree*, (7) *nrounds*, (8) *eval metric*. Therefore, we run 1000 combinations of hyperparameters. We used 500 and 1000 combinations for RF and XGBoosting, respectively, based

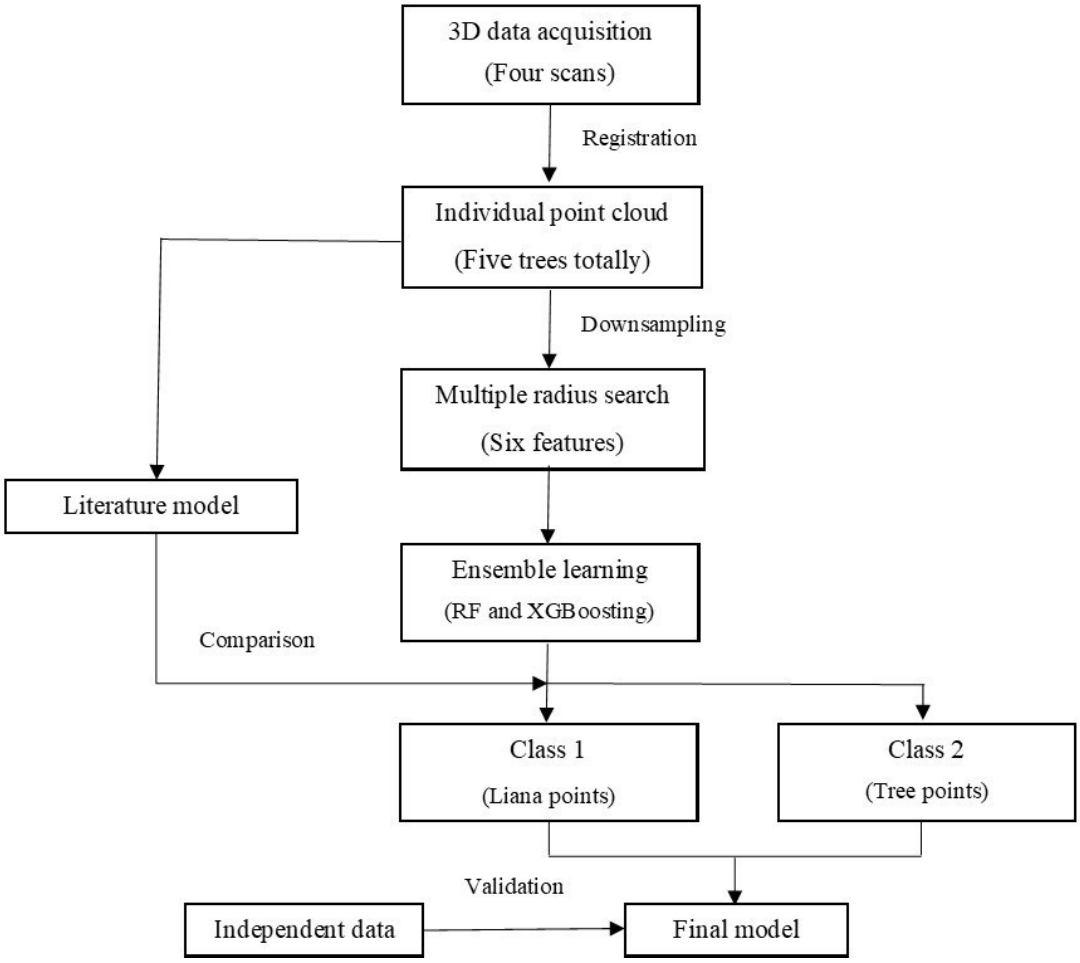


Figure 2. 2: The main steps used in presented geometric method based on RF and XGBoosting algorithms for liana/tree separation.

on the work of Rhys (2020). More details about hyperparameters and their optimization using RF and XGBoosting are given in supplementary materials (Section 2.6).

The performance of the model, including hyperparameter tuning, is evaluated using the k - fold spatial cross-validation method (Rhys, 2020). The model that has the optimal performance is therefore selected. The complete pipeline of our methods is outlined in Figure 2.2.

Performance assessment

We assess the performance of the presented liana/tree separation model using a k-fold spatial cross-validation method (Moorthy et al., 2019). The k-fold spatial cross-validation strategy randomly splits the data into k-folds, while one-fold at one time is used for validation, and the other unused k-1 folds are used for training. In this study, we performed a five-fold cross-validation strategy to assess the performance of model, as shown in Figure S2.2.

The following metrics: Precision, Recall, F1 score, and Precision-Recall (PR) curve, were chosen to estimate our model's performance (Tao et al., 2015b). The Precision is defined as the percentage of correctly classified liana points (TP) out of all classified liana points (TP and FP).

$$Precison = \frac{TP}{TP + FP} \quad (1)$$

At the same time, Recall is the percentage of correctly predicted liana points (TP) divided by all real liana points (TP and FN).

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

F1 score is the balanced value between precision and recall (0 to 1), while 0 means the worst performance, and 1 is the best performance.

$$F1\ score = 2 * \frac{Precison * Recall}{Precison + Recall} \quad (3)$$

Accuracy is the proportion of all corrected classified lianas and tree points (TP and TN) against all points.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Precision-Recall (PR) curve evaluates the trade-off between Precision and Recall for different liana point classification thresholds. The thresholds vary between 0 and 100% and are determined by the probability that the model estimates the positive class (liana point). The area under the PR curve shows the model's performance, while the high value means both high precision and recall of the classifier.

2.2.3 Intercomparison with the existing method

We compared our method with Moorthy et al. (2019), which used a multiple radius search method to obtain eigenvalues of each point at 0.1, 0.25, 0.5, 0.75, and 1.0 m. Then, the Random Forest method was selected to classify liana and tree based on those eigenvalues derived from point clouds. The details about running Moorthy et al. (2019) are on GitHub (https://github.com/sruthimoorthy/automated-liana-extraction.git).

Four postprocessing steps were conducted in the above study to correct the misclassified liana points predicted by the RF model. First, a Statistical Outlier Removal (SOR) filter was applied in Cloud Compare to filter noisy points from the predicted liana class. Second, the components that did not belong to liana were manually removed by visual inspection using the polygonal selection tool. Third, a density-based clustering algorithm named DBSCAN was used to correct the liana points misclassified as wood in the model prediction. Given the number of points in space, DBSCAN could group the points together if the neighboring points are close enough. An additional description of this algorithm can be found in Ferrara et al. (2018). Isolated clusters classified as wood were achieved by applying the DBSCAN. Finally, DBSCAN was also applied to the corrected liana points in the second step, and liana points clusters were obtained. Those liana clusters were tested for connectivity with the wood clusters in the third step. If the wood cluster were very close to any liana clusters, they would be merged into single cluster using DBSCAN. The processing pipeline mentioned above was applied to the point clouds used in this study to understand the generalization of Moorthy et al. (2019).

Concerning the postprocessing steps for our method, we only used the polygon selection tools

in Cloud Compare software to correct the misclassified liana points. We evaluated the performance of our model on the independent data from Nouragues, French Guiana (Moorthy et al., 2019). There are two reasons why we apply our method to this data. First, we want to indicate the broader utilization of our method, and second, we want to know whether our method shows differences with the former study.

2.3 Results

In this section, we first report the optimum radius for each geometric feature using the multiple radius search method. Then, we analyze the performance of the RF and XGBoosting models on liana and tree classification. Finally, we compare the generalization of our model with Moorthy et al. (2019), the only existing method available in the literature applied to liana/tree separation.

2.3.1 Geometric features analysis

The optimal radius for determining the differences between lianas and trees for the six features is detailed in Table 2.3. We can see that liana and tree points in Tree 2 tend to have a larger separability at the larger radius for all six features, while the verticality has the smallest radius of 0.68. The omnivariance is the only feature that shows the largest difference at 1 m across all trees collected due to the same increasing trend of liana and tree points. In addition, the optimal radii for Tree 2 are highlighted using red vertical lines in Figure 2.3.

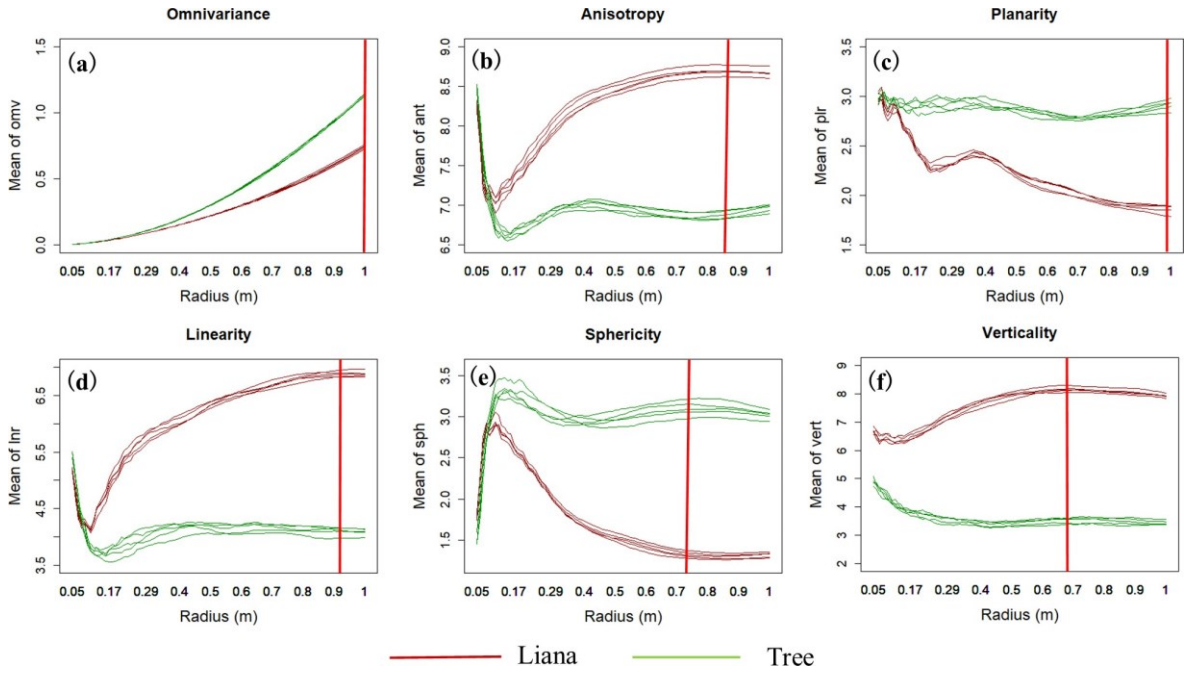


Figure 2. 3: The changing trends of Omnivariance, Planarity, Linearity, Sphericity, Verticality, and Anisotropy for Tree 2 across the various radius (y-axis indicates the mean value of each feature, and the x-axis shows the value of radius from 0.05 to 1 m; Red line: Tree; Green line: Liana); Red vertical line indicates the optimal radius used in this study (see Table 2.3). The multiple lines of same color show the changing trend of each class for different sample points.

Figure 2.3 presents the changing trends on the separation of lianas and trees for our selected six features for Tree 2 (Figure 2.1). The verticality is the only feature for which the mean value of liana is consistently higher than that of tree. Additionally, we can see that at the larger radius, the difference between liana and tree is similar (e.g., Anisotropy, Linearity, Sphericity and Verticality). A large number of points in the spherical space within a large radius could result in low computation efficiency when calculating geometric features. As such, we always select the smaller radius as the optimal one for classification (also see Figure S2.3). The changing trend of remaining trees can be found in Supplementary Materials.

Table 2. 3: Optimum radius for different features on the evaluated point clouds for all trees.

Feature	ID					
	Tree 1	Tree 2	Tree 3	Tree 4	Tree 5	Sruthi
Omnivariance	1	1	1	1	1	1
Anisotropy	1	0.86	0.98	0.5	0.45	0.37
Planarity	0.13	0.99	0.22	0.16	0.87	0.29
Linearity	0.18	0.91	1	0.4	0.21	0.33
Sphericity	1	0.73	1	0.51	0.21	0.37
Verticality	1	0.68	0.24	1	0.7	0.45

2.3.2 Model performance

To assess the effects of using an optimum radius for each feature, a five-fold cross- validation strategy that including hyperparameter tuning was applied. The details about the optimal hyperparameters of RF and XGBoosting are provided in the supplementary materials (Section 2.6).

Table 2.4 presents all the performance metrics for the five-fold cross-validation strategy used for the classification of each infested tree. As shown in Table 2.4, both ensemble classifiers did not perform well on average, while the XGBoosting model performed better than RF. First, the overall accuracy of XGBoosting was 0.88±0.07, which is higher than that of RF (0.85±0.07) (p-value = 0.55, t = -0.62). Second, the average recall of Random Forest was 0.56±0.17, 0.1 lower than the average recall of XGBoosting (0.66±0.19) (p- value = 0.61, t = -0.53). The XGBoosting model based on the optimum radius of the features listed in Table 2.4 showed an average F1 score

of 0.49±0.24, with an individual score ranging from 0.19 to 0.75. The F1 score of the Trees 1, 2 and 3 in Table 2.4 was lower than 0.42, while Tree 2 had the highest recall, whether based on RF or XGBoosting model. The recall for Tree 4 was 0.80, with an F1 score of 0.75 for XGBoosting, the highest value of all trees. The final classification results of the XGBoosting model of five trees are presented in Figure 2.4. Compared with Figure 2.1, most lianas in Tree 4 were successfully predicted. Figure 2.4 also revealed that some parts of the stem were misclassified as liana, especially in Tree 2.

Table 2. 4: Comparison of all performance metrics using RF and XGBoosting model; Aa, average area under the Precision-Recall curve; Ave, average; Sd: standard derivation.

ID	Random Forest					XGBoosting				
	Precision	Recall	F1 score	Accuracy	Aa	Precision	Recall	F1 score	Accuracy	Aa
Tree 1	0.05	0.44	0.09	0.90	0.09	0.12	0.47	0.19	0.95	0.51
Tree 2	0.20	0.72	0.31	0.88	0.35	0.29	0.77	0.42	0.90	0.86
Tree 3	0.32	0.37	0.34	0.82	0.21	0.30	0.44	0.36	0.85	0.5
Tree 4	0.68	0.74	0.71	0.91	0.73	0.7	0.80	0.75	0.91	0.86
Tree 5	0.67	0.52	0.59	0.73	0.69	0.65	0.82	0.73	0.77	0.89
Ave	0.38	0.56	0.41	0.85	0.41	0.41	0.66	0.49	0.88	0.72
Sd	0.28	0.17	0.24	0.07	0.29	0.25	0.19	0.24	0.07	0.20

The Precision-Recall (PR) curves of the above two machine learning algorithms is shown in Figure 2.5, while the area under the PR curve for each tree is indicated in Table 2.4. Here, we compute the standard derivation of the areas under the PR curve of two methods for each tree to determine a stable pattern (see Table S2.1). Figure 2.5 indicates PR curve of RF and XGBoosting were more stable in Trees 3, 4 and 5 than that on Trees 1 and 2. Visual inspection from Figure 2.1 indicates lianas in Trees 3, 4 and 5 are vertically distributed, while lianas in the remaining two trees have close contact with the main stem. In addition, Table 2.4 reveals that the average area under the PR curve of XGBoosting for all trees is 0.72±0.20, higher than that of the RF model (0.41±0.29) (p-value = 0.08, t = -1.99).

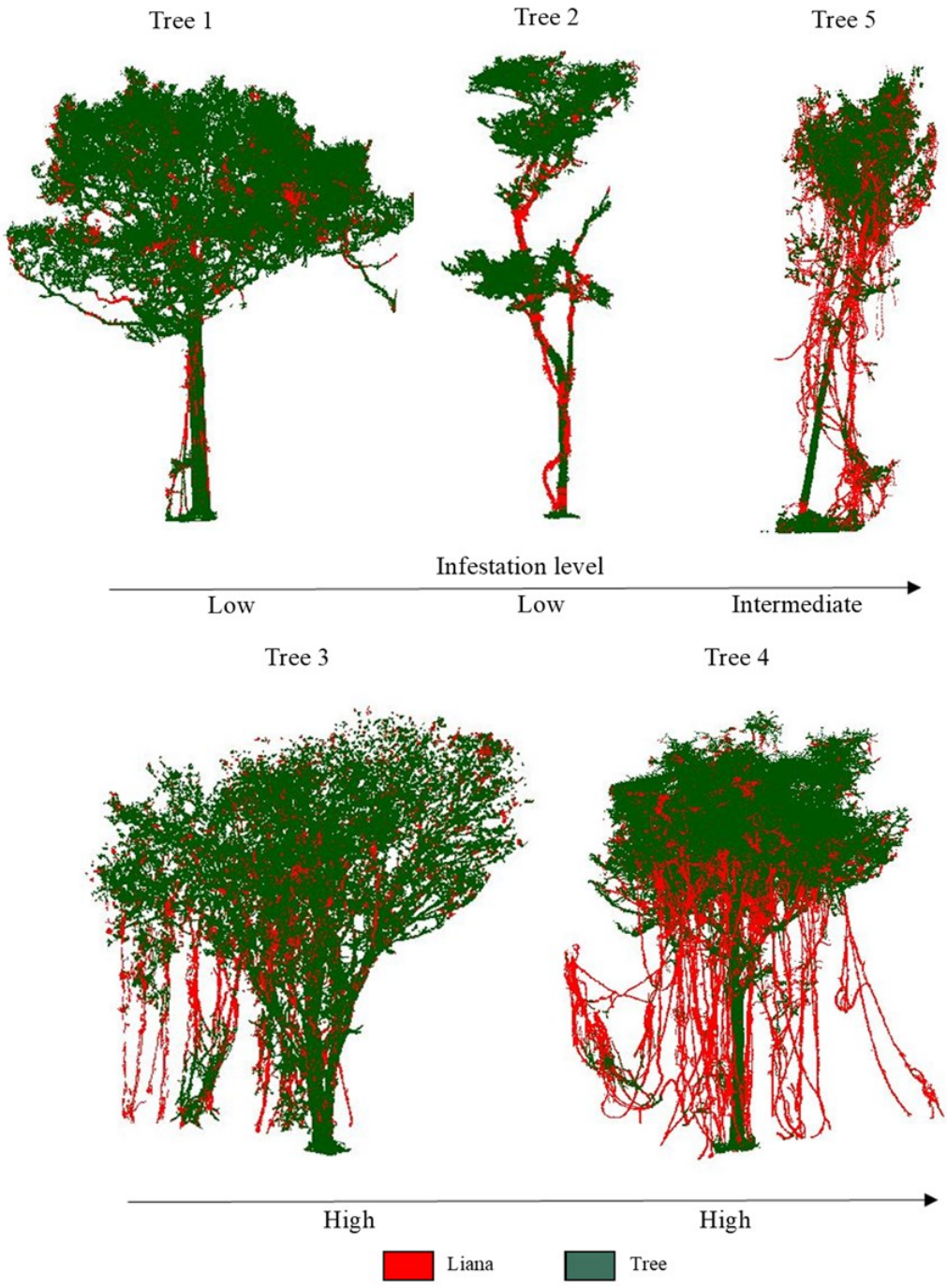


Figure 2. 4: XGBoosting model prediction for five trees. The prediction of Tree 4 tends to have a good consistency with the true data in Figure 1, while some parts of the stem in Tree 2 are misclassified as liana.

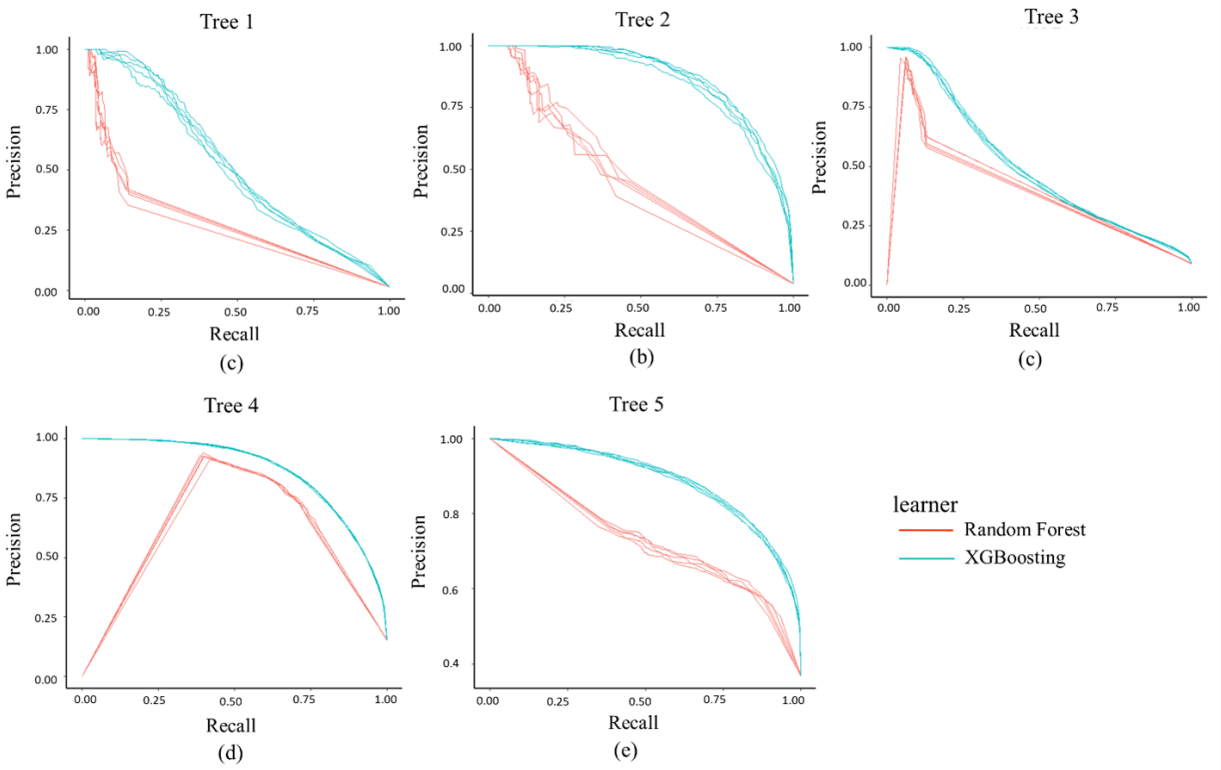


Figure 2. 5: Precision-Recall areas of each evaluated tree based on Random Forest and XGBoosting algorithms. The red and blue color indicate Random Forest, and XGBoosting model, respectively. The area under the Precision-Recall curve means the performance of the classifier.

2.3.3 Intercomparison

Our model's performance was evaluated and compared with the only method for liana and tree classification (Moorthy et al., 2019). Table 2.5 indicates the performance of their classifier on our datasets. Moorthy et al. (2019) showed lower performance for the five selected trees, the overall accuracy was 0.83±0.12. After postprocessing suggested by Moorthy et al. (2019), the method's recall improved, with an average recall changing from 0.19 to 0.56, while the overall accuracy increased to 0.88±0.07. Furthermore, the recall of Tree 4 and Tree 5 increased after the manual intervention, while Tree 4 had the highest value of 0.76. The method did not work for Tree 1 and Tree 2 since they had the lowest F1 score. Regarding postprocessing steps, DBSCAN generated a number of clusters ranging from 1, 600 to more than 17, 000 for our trees (Table

S2.2). Specifically, more than 17, 000 clusters were obtained for Tree 4, while the number of points in Tree 4 was the largest of all trees (Table 2.1). It should be also note that, the postprocessing step using the Moorthy et al. (2019) spends around 1-2 hours for each tree, while our manual intervention step only needs 5 minutes for each tree.

Table 2. 5: The performance of liana and tree classification on our data using Moorthy et al. (2019).

Tree ID	Without postprocessing				Postprocessing			
	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score	Accuracy
Tree 1	0.05	0.24	0.08	0.92	0.27	0.40	0.32	0.98
Tree 2	0.10	0.06	0.08	0.94	0.12	0.46	0.19	0.85
Tree 3	0.11	0.22	0.14	0.76	0.38	0.66	0.48	0.87
Tree 4	0.64	0.37	0.47	0.87	0.76	0.76	0.76	0.92
Tree 5	0.73	0.08	0.16	0.65	0.85	0.52	0.64	0.79
Ave	0.33	0.19	0.19	0.83	0.48	0.56	0.48	0.88
Sd	0.33	0.13	0.16	0.12	0.32	0.15	0.23	0.07

Compared with our method, the XGBoosting model performs similarly to Moorthy et al. (2019), which combined the features from multiple neighborhood sizes (F1 score: 0.49 vs. 0.48).

The independent dataset (Nouragues, French Guiana) was used to evaluate our model on new unseen data. Table 2.6 compared the final recall for the dataset from the literature based on different methods. It was clear that our models, whether based on RF or XGBoosting algorithm, showed higher recall than the existing model without postprocessing (0.67). Though the recall of Moorthy et al. (2019) improved to 0.87 after postprocessing, this showed similar performance to the presented RF model's recall (0.88).

Table 2. 6: Final recall for the dataset by Moorthy et al. (2019) after applying our model.

Methods	Model recall
Random Forest	0.88
Moorthy et al. (2019) with postprocessing	0.87
XGBoosting	0.82
Moorthy et al. (2019) without postprocessing	0.67

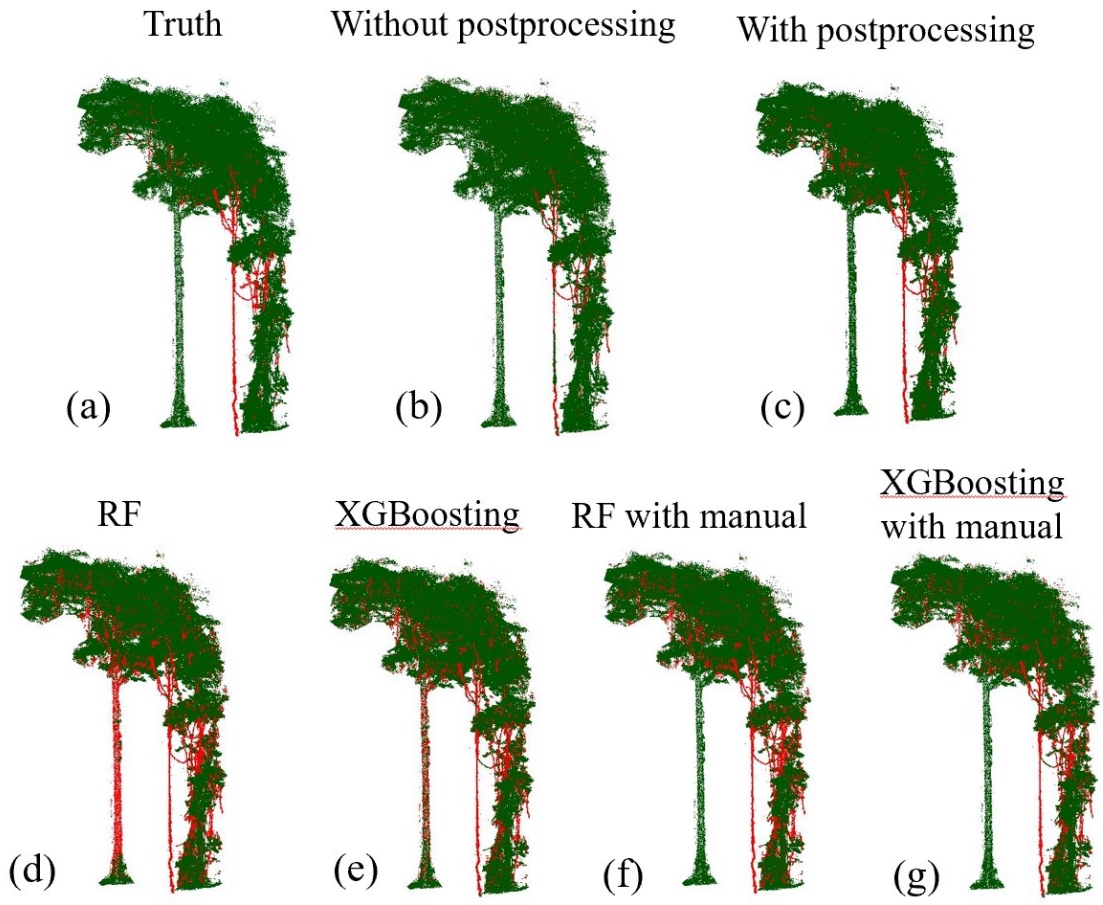


Figure 2. 6: The predications of different methods in Table 2.6, (a) Truth: manual labeled; Preprocessing: Moorthy et al. (2019) without manual intervention; (c) Postprocessing: Moorthy et al. (2019) with manual intervention; (d) Proposed method based on RF algorithm; (e) Proposed method based on XGBoosting algorithm; (f) RF predictions after manual intervention; (g) XGBoosting predictions after manual intervention.

Figure 2.6 indicates the final predictions of different methods in Table 2.6 against independent data. It can be observed that RF (Figure 2.6(d)) and XGBoosting (Figure 2.6(e)) methods can also extract most lianas accurately, compared to the postprocessing method (Figure 2.6(c)). Although RF and XGBoosting methods misclassified some parts of the main trunk as lianas, this can be easily corrected by manual intervention, which can be seen from Figures 2.6(f) and (g).

2.4 Discussion

Here, we proposed a liana and tree classification method coupled with the geometric features of a tree/liana point cloud at the optimal radius. Furthermore, we compared our approach with the only literature method for liana and tree separation (Moorthy et al., 2019). As shown in Table 2.4, the model based on XGBoosting showed a higher performance for all five trees. Table 2.6 further confirmed the applicability of our method to independent data. For example, RF performed with a recall of 0.88 when tested on the independent data from the tropical forest in Nouragues, French Guiana (Figure 2.6) without postprocessing, and this value was the same as the state-of-the-art method after postprocessing (0.87).

2.4.1 Geometric feature analysis

The optimal radius near-neighbor search method was used to obtain geometric features. This method showed similar performance to the multiple radius near-neighbor search on the discrimination of lianas and trees. According to Table 2.4, the average F1 score using the optimal radius based on the RF model was 0.41, which is higher than the multiple radius method (F1 score of 0.19) before the manual intervention. Although the average F1 score using the multiple radius method increased to 0.48 after using the postprocessing steps, this value was the same as our XGBoosting model without any postprocessing steps (F1 score of 0.49).

Our method suggests that the multiple radius method only works for Tree 4 from Table 2.5. The reason is that the lianas structure of Tree 4 is vertically distributed, thus having a similar shape to the independent data. Several reasons could explain why our method provides more advances when compared with Moorthy et al. (2019) and Zhu et al. (2018). First, although multiple

radius could produce much more features than using only one scale (Moorthy et al., 2019), it may lead to redundancy. The former means that combining features from all radius could cause high-dimensional data. The optimal radius search method chooses the scale with the largest difference between liana and tree points at various radius, thus avoiding high-dimensional data. Second, our method described more detailed geometric features. For instance, Zhu et al. (2018) computed the dimensional features for increasing radius values of 0.2, 0.3, and 0.4 m, and then determined the optimal radius based on Shannon entropy (Demantke et al., 2011). Searching only three radiuses may not be enough to define the optimum radius, which was a possible reason why their model obtained an overall accuracy of 70.4%. In this study, we determined the optimal radius from 0.04 to 1 m, at a 0.01 m interval. Then, we used a total of 96 radiuses to obtain the optimal radius, which had more quantities than the number of radius used in Zhu et al. (2018). As a result, an overall accuracy of 85% and 88% was achieved for Random Forest and XGBoosting, respectively.

Using the multiple radius search method, Moorthy et al. (2019) presented an improved supervised learning model to classify leaf and wood, achieving an overall accuracy of 94.2%. In addition, Ma et al. (2015) used the single radius method, based on the improved salient features, to classify leaf and wood for conifer and broadleaf trees. Those geometric features gave the model an overall accuracy of 95.4%. Even though the performance of our method is lower than those implemented by the previous studies, this is related to the fact that these two studies look at simple structures while our liana/tree structure provides several levels of complexity (Figure 1). Lianas are climber plants with irregular growth patterns, which means they could extend in all directions around their host trees (Schnitzer, 2018; Schnitzer et al., 2011). This characteristic confuses the 3D deterministic nature of more simple forests (Sanchez-Azofeifa et al., 2017), an element that is not considered on the trees analysed by Ma et al. (2015) and Moorthy et al. (2019).

2.4.2 Liana-Tree classification

Table 2.4 gives information about the performance of RF and XGBoosting on liana and tree classification. The XGBoosting method showed higher performance than the RF method for liana and tree classification on all five trees (0.88 vs. 0.85 and 0.49 vs. 0.41, respectively). Table 2.4 also reported the F1 score of Trees 1, 2 and 3 using the XGBoosting method was lower than 0.42. There are three reasons why those trees show worse performance. First, as liana stems move up to the canopy, their stems are to be of similar thickness as the tree components such as branches, and

thus are misclassified; the former seems to be one of the current and future limitations on the use of TLS to extract lianas from their hosts, especially on the upper parts of the canopy. Second, lianas have close contact with the tree; therefore, it is challenging to distinguish liana and tree point even the optimum radius is applied. Third, lianas have a more complex structure than those in Tree 4. Lianas grow around the tree and ascend to the canopy for Tree 2 (Figure 2.1), thus showing much more curve structure than the liana structure in Tree 4. By contrast, the liana structure in Tree 4 is relatively simple since most lianas on it are vertically distributed. This is the reason why Tree 4 shows the highest F1 score based on RF (0.71) or XGBoosting (0.75), same for Moorthy et al. (2019).

2.4.3 Intercomparison with the previous method

We also compared our method with the only literature liana extraction method available in the literature (Moorthy et al., 2019). The mentioned study used eigenvalues to separate liana and tree points, which showed some weaknesses compared to eigenvalue-based ratios used in this study. The structure of lianas from their study is relatively simple since their eigenvalues' variance is dominant in one direction. However, this is not always the case. Lianas can grow in many different strategies, meaning they tend to have more complex structures than stems or branches (Sanchez-Azofeifa et al., 2017). For example, the lianas in Tree 2 (Figure 2.1) grow around their host tree before moving up to the canopy. Therefore, the standard for liana and tree classification in their study may not work for our data. Their method only obtained an average recall of 0.19 before postprocessing steps, while this value increased to 0.56 after manual intervention (see Table 2.5). Instead, the ratios computed by eigenvalues could demonstrate how the lianas change in space. This is the one reason why our XGBoosting model shows higher performance, with an average recall of 0.66. When it comes to manual intervention, the density-based cluster algorithm (DBSCAN) spent much time recognizing the label of clusters (Ferrara et al., 2018). For example, DBSCAN produced over 17, 000 clusters for Tree 4 (Table S2.2), while the tree had the largest number of points of all trees (Table 2.1).

According to Table 2.6, our model's final recall, whether based on Random Forest or XGBoosting, is higher than the state-of-the-art method without postprocessing. The manual intervention improves the recall of their model to 0.87, showing similar performance to our model without postprocessing (0.88 of RF). Although our method misclassified some parts of the main

stem as lianas (Figure 2.6(d) and 2.6(e)), this can be easily corrected using the manual intervention. Figure 2.6 shows that our method, whether based on RF and XGBoosting algorithms, can extract most lianas from independent data after manual intervention. Therefore, our model's processing time (5 minutes vs. 1-2 hour for each tree) shows that the presented method coupled with the curvature is more efficient for liana and tree classification than Moorthy et al. (2019).

Our method's major advantages are avoiding many postprocessing steps (improve efficiency), while the performance of our method is also as accurate as previous studies. For instance, many isolate clusters (Table S2.2) need to be visually identified when using DBSCAN in Moorthy et al. (2019) to get a reasonable performance, while our method can obtain the same performance without postprocessing. Moreover, the presented method avoids combining the geometric features at multiple radius, thus producing high-dimensional data. We calculate the geometric features of 96 radius to define the optimal radius, and this does not need much labor force, which can be done using Cloud Compare and R programming. Also, we utilize geometric features to classify liana and tree cloud points, which means that the presented method only needs xyz coordinates, showing the broad applicability to the 3D point cloud collected from other sensors.

The highest F1 score of the proposed method is 0.49; the main reason is that the liana structure is complex, while they are generally smaller than almost all the stems and show similar diameter with branches. Also, our method is limited by sample size (6 trees), and developed on the individual tree level, which means our method may fails at the stand level. Since the lianas are increasing their dominance in the tropical forest, our method can provide a flexible method to extract lianas from point clouds, enabling researchers quantify the impact of lianas on tree and forest structure. Future work could also pay attention to improving the performance of the algorithms for liana and tree classification. Recently, the development of more sophisticated machine learning technologies, especially deep learning, provides a solution to unstructured point clouds. Deep learning could extract features automatically to build a classifier, although it requires a large amount of data and high computational power (LeCun et al., 2015). Therefore, deep learning could be one of the solutions to improve the algorithm's performance for liana and tree classification.

2.5 Conclusions

We present an open-source semi-automated liana extraction procedure (https://github.com/than2/liana-extraction) from point clouds derived from TLS data. Our approach avoids high-dimensional data and much manual intervention, making the entire procedure user-friendly and more efficient. Future research can pay more attention to improving the performance of liana extraction. For example, since we have the changing trend of liana and tree points from 0.05 to 1 m, time series classification can be used to separate them directly, which means more radius information can be used during the classification. Although time series classification may produce many more features, this problem can be handled well by Convolutional Neural Networks and Recurrent Neural Networks. Our method can help understand the contribution of lianas to forest structure by an accurate segmentation/classification from point clouds.

Furthermore, it may provide a flexible approach to continuously monitoring liana dominance, enabling a good understanding of their role in the forest dynamics.

Acknowledgement

This research was funded by the National Science and Engineering Research Council of Canada (NSERC) Discovery Grant. The LiDAR instrument and data processing workstation were provided by the Center for Earth Observation Sciences (CEOS) of U of A. We thank Felipe Alencastro for his assistance with data collection. The authors also thank Sruthi M. Krishna Moorthy for her data collected in Nouragues, French Guiana. The authors declare no conflicts of interest.

2.6 Supplementary materials

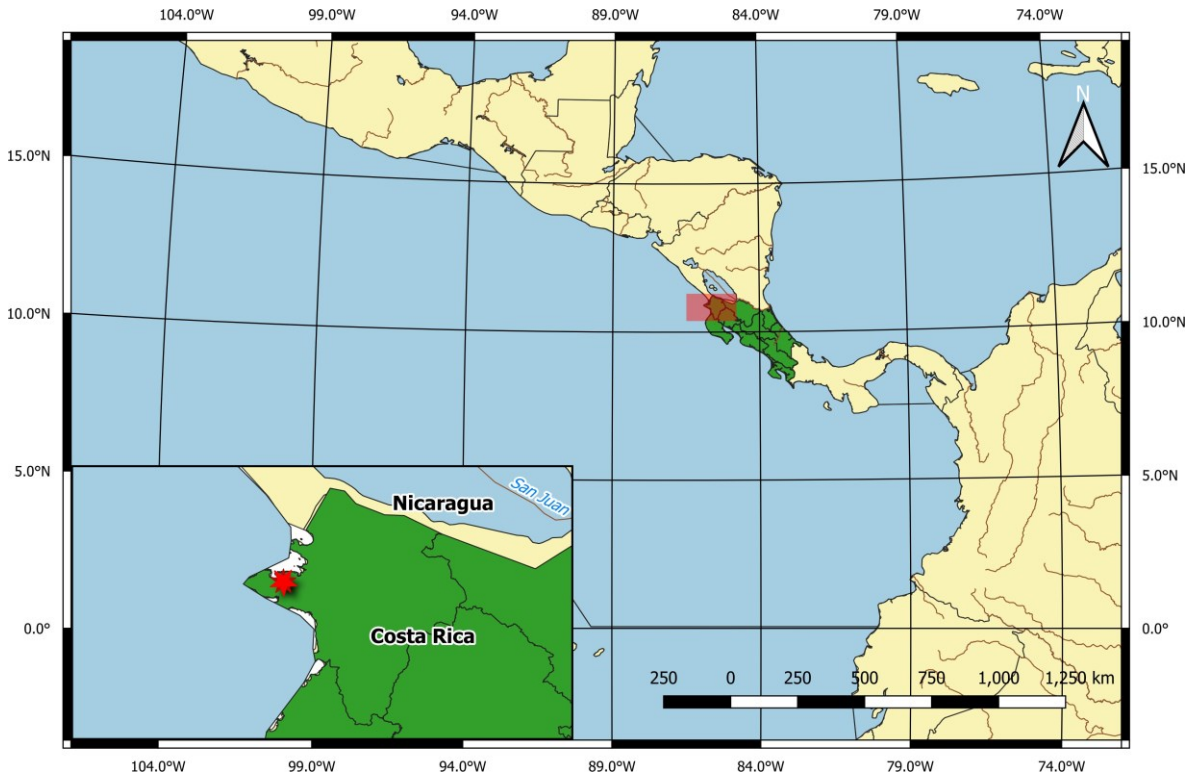


Figure S2. 1: Study area, the red star indicates the location of Santa Rosa National Park – Environmental Monitoring Supersite (SRNP-EMSS), Guanacaste, Costa Rica.

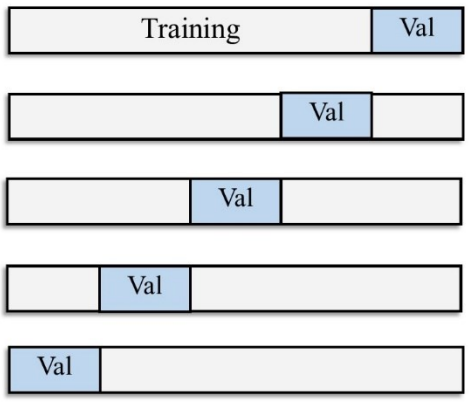


Figure S2. 2: Five-fold spatial cross-validation approach used in this chapter.

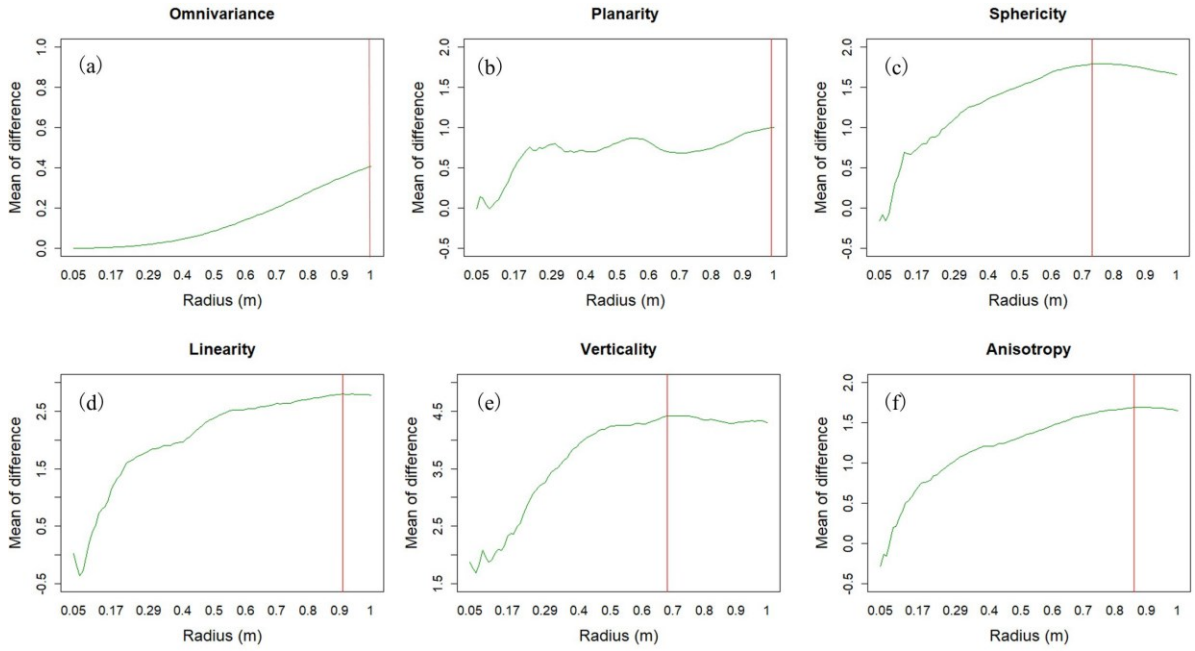


Figure S2. 3: The mean difference of Omnivariance, Planarity, Sphericity, Linearity, Verti- cality, and Anisotropy between liana and tree for Tree 2 across the various radius (y-axis indicate the mean difference between liana and tree, and x-axis shows the value of radius from 0.05 to 1 m; red vertical line means the largest difference of the feature).

Figure S2. 4: The actual photo where we scan our trees.

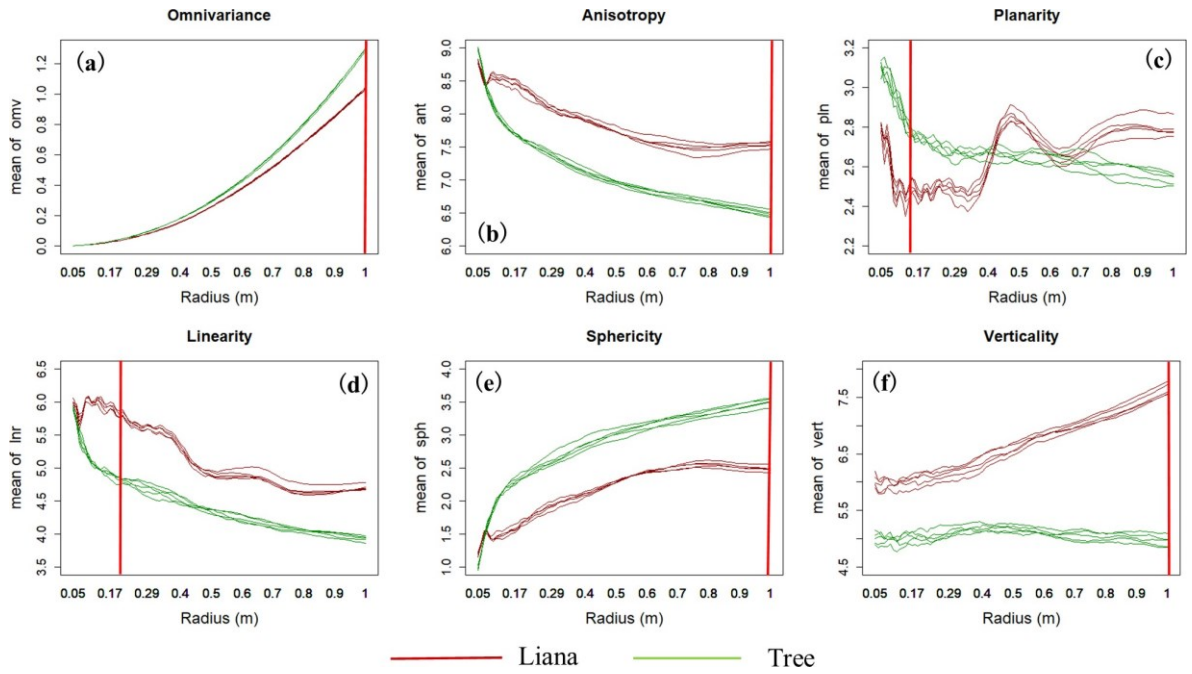


Figure S2. 5: The changing trend of Omnivariance, Planarity, Sphericity, Linearity, Verticality, and Anisotropy between liana and tree for Tree 1 across the various radius (y-axis indicate the metric values for liana and tree, and x-axis shows the value of radius from 0.05 to 1 m; red vertical line means the largest difference of the feature).

As indicated from Figure 2.1, lianas in Tree 1 have more linear structure in lower part of the stem than Tree 2, while we always see the lianas in Tree 2 has more irregular structure (curve structure here).

Also, if we see the lianas structure in Tree 3, most of them are more vertically distributed, while Tree 3 has higher liana density than Tree 1.

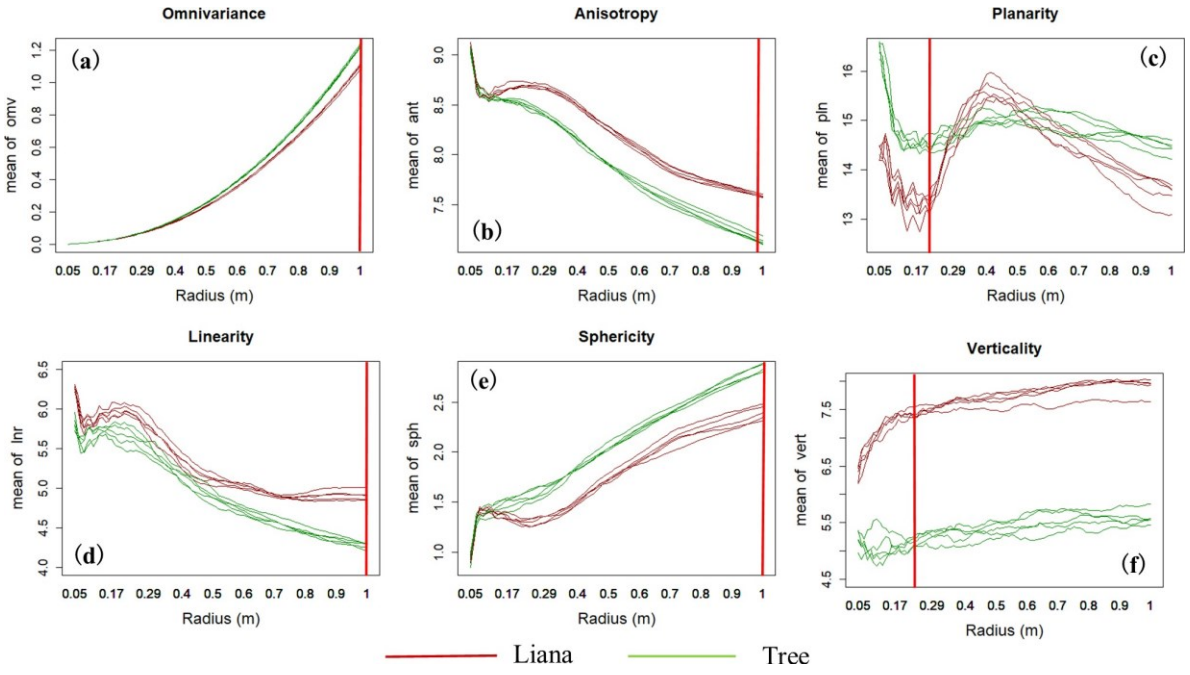


Figure S2. 6: The changing trend of Omnivariance, Planarity, Sphericity, Linearity, Verticality, and Anisotropy between liana and tree for Tree 3 across the various radius (y-axis indicate the metric values for liana and tree, and x-axis shows the value of radius from 0.05 to 1 m; red vertical line means the largest difference of the feature).

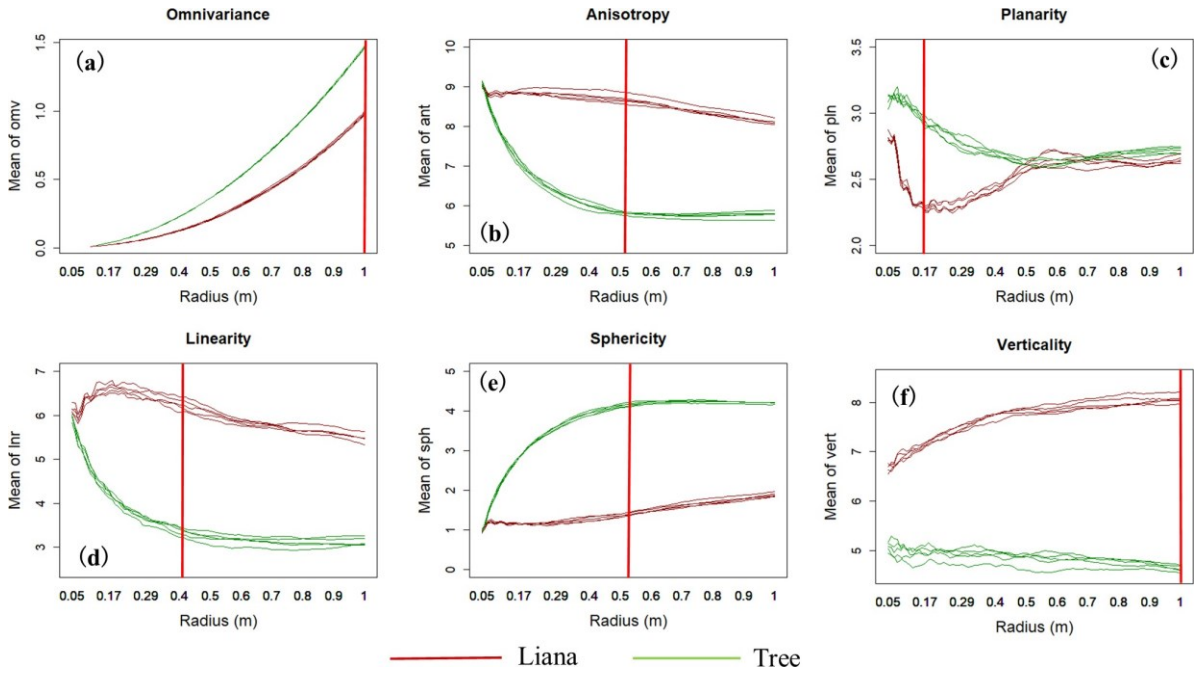


Figure S2. 7: The changing trend of Omnivariance, Planarity, Sphericity, Linearity, Verticality, and Anisotropy between liana and tree for Tree 4 across the various radius (y-axis indicate the metric values for liana and tree, and x-axis shows the value of radius from 0.05 to 1 m; red vertical line means the largest difference of the feature).

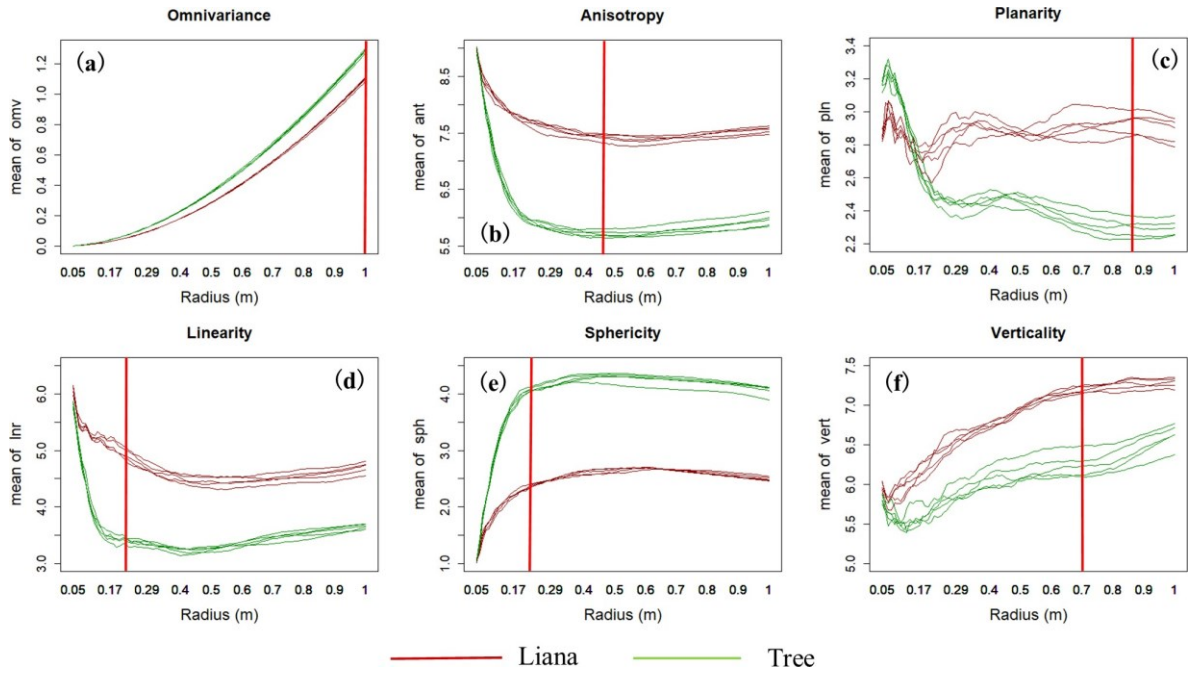


Figure S2. 8: The changing trend of Omnivariance, Planarity, Sphericity, Linearity, Verti- cality, and Anisotropy between liana and tree for Tree 5 across the various radius (y-axis indicate the metric values for liana and tree, and x-axis shows the value of radius from 0.05 to 1 m; red vertical line means the largest difference of the feature).

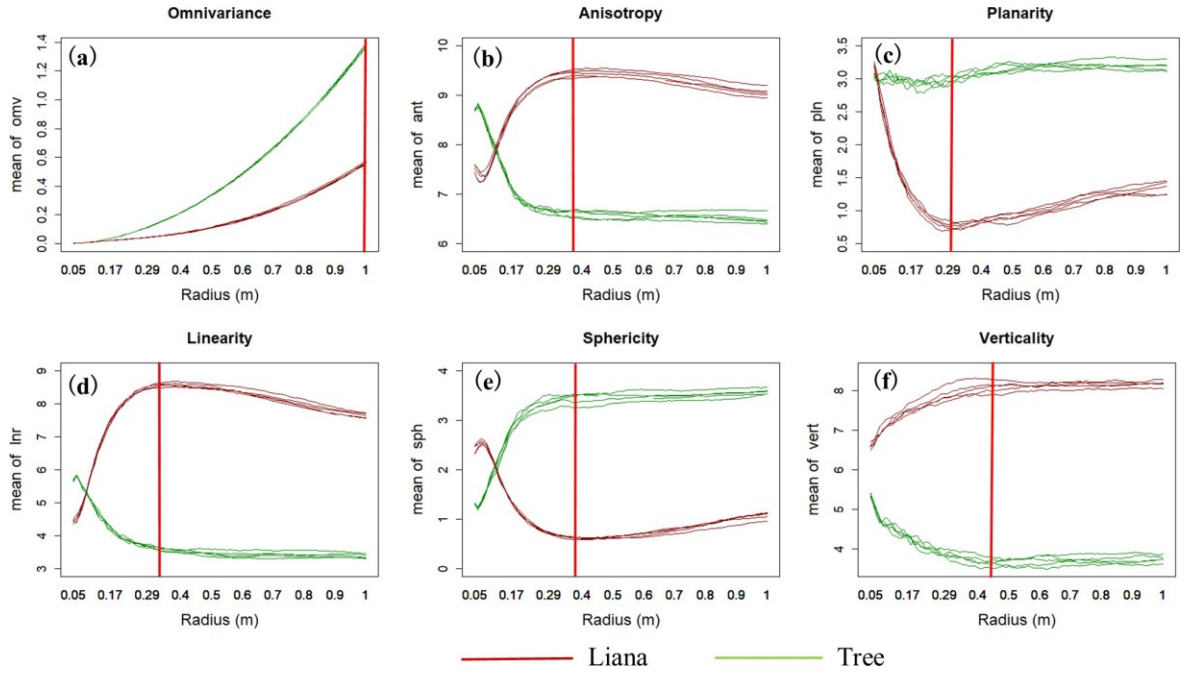


Figure S2. 9: The changing trend of Omnivariance, Planarity, Sphericity, Linearity, Verticality, and Anisotropy between liana and tree for independent dataset across the various radius (y-axis indicate the metric values for liana and tree, and x-axis shows the value of radius from 0.05 to 1 m; red vertical line means the largest difference of the feature).

Table S2. 1: The area under the PR curve of each tree using RF and XGB model.

ID	Tree 1		Tree 2		Tree 3		Tree 4		Tree 5	
Method	RF	XGB	RF	XGB	RF	XGB	RF	XGB	RF	XGB
	0.11	0.52	0.35	0.87	0.21	0.50	0.73	0.86	0.68	0.88
	0.09	0.52	0.33	0.84	0.20	0.50	0.74	0.86	0.7	0.88
Area	0.07	0.50	0.37	0.87	0.21	0.49	0.72	0.87	0.68	0.89
	0.07	0.53	0.32	0.86	0.21	0.49	0.74	0.87	0.69	0.89
	0.13	0.50	0.38	0.87	0.21	0.51	0.74	0.86	0.69	0.89
Ave	0.09	0.51	0.35	0.86	0.21	0.50	0.73	0.86	0.69	0.89
Max -min	0.04	0.03	0.06	0.03	0.01	0.02	0.02	0.01	0.02	0.01
Sd	0.03	0.01	0.03	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01

Table S2.1 gives information about the area under PR curve of RF and XGBoosting model for each tree. We can observe from the table that the standard derivation of Trees 3, 4 and 5 is lower than Trees 1 and 2.

Table S2. 2: The number of clusters generated by DBSCAN on five trees.

Tree ID	Liana	Tree	Total
Tree 1	966	11712	12678
Tree 2	75	1568	1643
Tree 3	4006	12293	16299
Tree 4	1768	16124	17892
Tree 5	169	2019	2188

The DBSCAN algorithm generate many clusters for liana and tree. Identifying the tree clusters that are close to lianas spend much time since there are a number of clusters for each tree, especially in Tree 4, which has the largest number of points of five trees.

Hyperparameter optimization of Random Forest and XGBoosting algorithms.

A: Random Forest

The following set of parameters were optimized for Random Forest. We show the range of values and the best value selected for each of the parameters as follows:

ntree, lower = 10, upper = 200, optimal = 84, the number of decision tree in the forest.

mtry, lower = 1, upper = 6, optimal = 5, the number of features in each node.

nodesize, lower = 1, upper = 5, optimal = 3, the minimum number of the data allowed in a leaf.

maxnodes, lower = 5, upper = 20, optimal = 20, the maximum number of leaves.

B: XGBoosting

The following set of parameters were optimized for XGBoosting. We show the range of values and the best value selected for each of the parameters as follows:

eta, lower = 0, upper = 1, optimal = 0.643, learning rate.

gamma, lower = 0, upper = 5, optimal = 3.98, the minimum number of splitting.

max depth, lower = 1, upper = 5, optimal = 5, the maximum depth of a decision tree.

min child weight, lower = 1, upper = 10, optimal = 4.31, the minimum total weight of all the data in a leaf node.

subsample, lower = 0.5, upper = 1, optimal = 0.864, the proportion of the data to be randomly sampled.

colsample bytree, lower = 0.5, upper = 1, optimal = 0.952, the fraction of predictors sampled by each tree.

nrounds, lower = 10, upper = 200, optimal = 188, the sequentially built trees to be used.

eval metric, values = c (”merror”, ”mlogloss”, ”auc”), optimal = ”auc”, the loss function used in the model.

2.7 References

Bao, Yunfei, Sruthi MK Moorthy, and Hans Verbeeck (2018). “Towards extraction of lianas from terrestrial lidar scans of tropical forests”. In: IGARSS 2018-2018 IEEE Interna- tional Geoscience and Remote Sensing Symposium. IEEE, pp. 7544–7547.

Béland, Martin, Dennis D Baldocchi, Jean-Luc Widlowski, Richard A Fournier, and Michel M Verstraete (2014). “On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR”. In: Agricultural and Forest Meteorology 184, pp. 82–97.

Béland, Martin, Jean-Luc Widlowski, Richard A Fournier, Jean-Francois Cote, and Michel M Verstraete (2011). “Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements”. In: Agricultural and Forest Meteorology 151.9, pp. 1252–1266.

Belton, David, Simon Moncrieff, and Jane Chapman (2013). “Processing tree point clouds using Gaussian Mixture Models”. In: ISPRS Annals of Photogrammetry, Remote Sens- ing and Spatial Information Sciences. Presented at the WG 3, pp. 43–48.

Burt, Andrew, Mathias Disney, and Kim Calders (2019). “Extracting individual trees from lidar point clouds using treeseg”. In: Methods in Ecology and Evolution 10.3, pp. 438– 445.

Calders, Kim, Glenn Newnham, Andrew Burt, Simon Murphy, Pasi Raumonen, Martin Herold, Darius Culvenor, Valerio Avitabile, Mathias Disney, John Armston, et al. (2015). “Nondestructive estimates of above-ground biomass using terrestrial laser scan- ning”. In: Methods in Ecology and Evolution 6.2, pp. 198–208.

Calders, Kim, Glenn Newnham, Martin Herold, Simon Murphy, Darius Culvenor, Pasi Raumonen, Andrew Burt, John Armston, Valerio Avitabile, and Mathias Disney (2013). "Estimating above ground biomass from terrestrial laser scanning in Australian Eu- calypt Open Forest". In: Proceedings SilviLaser 2013, 9-11 October, Beijing, China, pp. 90–97.

Calvo-Rodriguez, Sofia, Ralf Kiese, and G Arturo Sánchez-Azofeifa (2020). "Seasonality and budgets of soil greenhouse gas emissions from a tropical dry forest successional gradient in Costa Rica". In: *Journal of Geophysical Research: Biogeosciences* 125.9, e2020JG005647.

Chen, Tianqi and Carlos Guestrin (2016). "XGBoost: A scalable tree boosting system In Proceedings of the 22Nd ACM SIGKDD International Conference on Knowl- edge Discovery and Data Mining,(pp. 785–794)". In: *New York, NY, USA: ACM* 10.2939672.2939785.

Côté, Jean-François, Richard A Fournier, Gordon W Frazer, and K Olaf Niemann (2012). "A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure". In: *Agricultural and forest meteorology* 166, pp. 72–85.

Demantké, Jérôme, Clément Mallet, Nicolas David, and Bruno Vallet (2011). "Dimensionality based scale selection in 3D lidar point clouds". In: *Laser scanning*.

Dewalt, Saara J, Stefan A Schnitzer, and Julie S Denslow (2000). "Density and diversityof lianas along a chronosequence in a central Panamanian lowland forest". In: *Journal of Tropical Ecology* 16.1, pp. 1–19.

Durán, Sandra M and Ernesto Gianoli (2013). "Carbon stocks in tropical forests decrease with liana density". In: *Biology letters* 9.4, p. 20130301.

Feliciano, Emanuelle A, Shimon Wdowinski, and Matthew D Potts (2014). "Assessing mangrove above-ground biomass and structure using terrestrial laser scanning: A case study in the Everglades National Park". In: *Wetlands* 34.5, pp. 955–968.

Ferrara, Roberto, Salvatore GP Virdis, Andrea Ventura, Tiziano Ghisu, Pierpaolo Duce,and Grazia Pellizzaro (2018). “An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN”. In: *Agricultural and forest meteorology* 262, pp. 434–444.

Gentry, ALWYN H (1991). “The distribution and evolution of climbing plants”. In: *Thebiology of vines* 351.

Gonzalez de Tanago, Jose, Alvaro Lau, Harm Bartholomeus, Martin Herold, Valerio Avitabile, Pasi Raumonen, Christopher Martius, Rosa C Goodman, Mathias Disney, Solichin Manuri, et al. (2018). “Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR”. In: *Methods in Ecology and Evolution* 9.2, pp. 223–234.

Guzmán Q, J Antonio, Benoit Rivard, G Arturo Sánchez-Azofeifa, et al. (2018). “Discrimination of liana and tree leaves from a Neotropical Dry Forest using visible-near infraredand longwave infrared reflectance spectra”. In: *Remote Sensing of Environment* 219, pp. 135–144.

Ingwell, Laura L, S Joseph Wright, Kristen K Becklund, Stephen P Hubbell, and Stefan A Schnitzer (2010). “The impact of lianas on 10 years of tree growth and mortality on Barro Colorado Island, Panama”. In: *Journal of Ecology* 98.4, pp. 879–887.

Kankare, Ville, Mikko Vastaranta, Markus Holopainen, Minna Räty, Xiaowei Yu, Juha Hyyppä, Hannu Hyyppä, Petteri Alho, and Risto Viitala (2013). “Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR”. In: *Remote Sensing* 5.5, pp. 2257–2274.

Koenig, Kristina, Bernhard Höfle, Martin Hämmerle, Thomas Jarmer, Bastian Siegmann, and Holger Lilienthal (2015). “Comparative classification analysis of post-harvestgrowth detection from terrestrial LiDAR point clouds in precision agriculture”. In: *IS- PRS Journal of Photogrammetry and Remote Sensing* 104, pp. 112–125.

Lakicevic, Milena, Nicholas Povak, and Keith M Reynolds (2020). *Introduction to R for Terrestrial Ecology*. Springer.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning". In: nature 521.7553, pp. 436–444.

Letcher, Susan G and Robin L Chazdon (2009). "Lianas and self-supporting plants during tropical forest succession". In: Forest Ecology and Management 257.10, pp. 2150–2156.

Londre, Ronald A and Stefan A Schnitzer (2006). "The distribution of lianas and their change in abundance in temperate forests over the past 45 years". In: Ecology 87.12, pp. 2973–2978.

Ma, Lixia, Guang Zheng, Jan UH Eitel, Troy S Magney, and L Monika Moskal (2016). "Determining woody-to-total area ratio using terrestrial laser scanning (TLS)". In: Agricultural and forest meteorology 228, pp. 217–228.

Ma, Lixia, Guang Zheng, Jan UH Eitel, L Monika Moskal, Wei He, and Huabing Huang (2015). "Improved salient feature-based approach for automatically separating photo- synthetic and nonphotosynthetic components within terrestrial lidar point cloud data of forest canopies". In: IEEE Transactions on geoscience and remote sensing 54.2, pp. 679–696.

Martınez-Izquierdo, Laura, Marıa M Garcıa, Jennifer S Powers, and Stefan A Schnitzer (2016). "Lianas suppress seedling growth and survival of 14 tree species in a Panamanian tropical forest". In: Ecology 97.1, pp. 215–224.

Moorthy, Sruthi MK, Kim Calders, Manfredo Di Porcia e Brugnera, Stefan A Schnitzer, and Hans Verbeeck (2018). "Terrestrial laser scanning to detect liana impact on forest structure". In: Remote Sensing 10.6, p. 810.

Moorthy, Sruthi MK, Kim Calders, Matheus B Vicari, and Hans Verbeeck (2019). "Improved supervised learning-based approach for leaf and wood classification from LiDAR point clouds of forests". In: IEEE Transactions on Geoscience and Remote Sensing 58.5, pp. 3057–3070.

Moorthy, Sruthi MK, Yunfei Bao, Kim Calders, Stefan A Schnitzer, and Hans Verbeeck (2019). “Semi-automatic extraction of liana stems from terrestrial LiDAR point clouds of tropical rainforests”. In: ISPRS Journal of Photogrammetry and Remote Sensing 154, pp. 114–126.

R Core Team, other (2021). “R: A language and environment for statistical computing”. In: R project.

Rhys, Hefin (2020). Machine Learning with R, the tidyverse, and mlr. Simon and Schuster.

Rodrıguez-Ronderos, M Elizabeth, Gil Bohrer, Arturo Sanchez-Azofeifa, Jennifer S Powers, and Stefan A Schnitzer (2016). “Contribution of lianas to plant area index and canopy structure in a Panamanian forest”. In: Ecology 97.12, pp. 3271–3277.

Sanchez-Azofeifa, GA and K Castro-Esau (2006). “Canopy observations on the hyperspectral properties of a community of tropical dry forest lianas and their host trees”. In: International Journal of Remote Sensing 27.10, pp. 2101–2109.

Sanchez-Azofeifa, GA, J Antonio Guzman-Quesada, Mauricio Vega-Araya, Carlos Campos- Vargas, Sandra Milena Duran, Nikhil D’Souza, Thomas Gianoli, Carlos Portillo- Quintero, and Iain Sharp (2017). “Can terrestrial laser scanners (TLSs) and hemi- spherical photographs predict tropical dry forest succession with liana abundance?” In: Biogeosciences 14.4, pp. 977–988.

Schneider, Fabian D, Daniel Ku¨kenbrink, Michael E Schaepman, David S Schimel, and Felix Morsdorf (2019). “Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR”. In: Agricultural and Forest Meteorology 268, pp. 249–257.

Schnitzer, Stefan A (2005). “A mechanistic explanation for global patterns of liana abundance and distribution”. In: The American Naturalist 166.2, pp. 262–276.

Schnitzer, Stefan A (2018). “Testing ecological theory with lianas”. In: New Phytologist 220.2, pp. 366–380.

Schnitzer, Stefan A and Frans Bongers (2002). “The ecology of lianas and their role in forests”. In: Trends in Ecology & Evolution 17.5, pp. 223–230.

Schnitzer, Stefan A and Frans Bongers (2011). “Increasing liana abundance and biomass in tropical forests: emerging patterns and putative mechanisms”. In: Ecology letters 14.4, pp. 397–406.

Schnitzer, Stefan A, Saara J DeWalt, and Jerome Chave (2006). “Censusing and Measuring Lianas: A Quantitative Comparison of the Common Methods 1”. In: Biotropica 38.5, pp. 581–591.

Schnitzer, Stefan A, Sergio Estrada-Villegas, and S Joseph Wright (2020). “The response of lianas to 20 yr of nutrient addition in a Panamanian forest”. In: Ecology 101.12, e03190.

Tao, Shengli, Qinghua Guo, Shiwu Xu, Yanjun Su, Yumei Li, and Fangfang Wu (2015a). “A geometric method for wood-leaf separation using terrestrial and simulated lidar data”. In: Photogrammetric Engineering & Remote Sensing 81.10, pp. 767–776.

Tao, Shengli, Fangfang Wu, Qinghua Guo, Yongcai Wang, Wenkai Li, Baolin Xue, Xueyang Hu, Peng Li, Di Tian, Chao Li, et al. (2015b). “Segmenting tree crowns from terres- trial and mobile LiDAR data by exploring ecological theories”. In: ISPRS Journal of Photogrammetry and Remote Sensing 110, pp. 66–76.

Thomas, Hugues, François Goulette, Jean-Emmanuel Deschaud, Beatriz Marcotegui, and Yann LeGall (2018). “Semantic classification of 3D point clouds with multiscale spher- ical neighborhoods”. In: 2018 International conference on 3D vision (3DV). IEEE, pp. 390–398.

Vicari, Matheus B, Mathias Disney, Phil Wilkes, Andrew Burt, Kim Calders, and William Woodgate (2019). “Leaf and wood classification framework for terrestrial LiDAR point clouds”. In: Methods in Ecology and Evolution 10.5, pp. 680–694.

Wang, Di, Jasmin Brunner, Zhenyu Ma, Hao Lu, Markus Hollaus, Yong Pang, and Norbert Pfeifer (2018). “Separating tree photosynthetic and non-photosynthetic components from point cloud data using dynamic segment merging”. In: Forests 9.5, p. 252.

Wang, Di, Markus Hollaus, and Norbert Pfeifer (2017). “Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data”. In: ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4.

Weinmann, Martin, Steffen Urban, Stefan Hinz, Boris Jutzi, and Clément Mallet (2015). “Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas”. In: Computers & Graphics 49, pp. 47–57.

Wright, S Joseph (2005). “Tropical forests in a changing environment”. In: Trends in ecology & evolution 20.10, pp. 553–560.

Zhu, Xi, Andrew K Skidmore, Roshanak Darvishzadeh, K Olaf Niemann, Jing Liu, Yi- fang Shi, and Tiejun Wang (2018). “Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest”. In: International journal of applied earth observation and geoinformation 64, pp. 43–50.

Chapter3

A deep learning time-series approach for leaf and wood classification from terrestrial LiDAR point clouds

Abstract

The accurate separation between leaf and woody components from terrestrial laser scanning (TLS) data is vital for the estimation of leaf area index (LAI) and wood area index (WAI). Here, we present the application of deep learning time series separation of leaves and wood from TLS point clouds collected from broad-leaved trees. First, we use a multiple radius nearest neighbor approach to obtain a time series of the geometric features. Second, we compare the performance of Fully Convolutional Neural Network (FCN), Long Short-Term Memory Fully Convolutional Neural Network (LSTM-FCN), and Residual Network (ResNet) on leaf and wood classification. We also compare the effect of univariable (UTS) and multivariable (MTS) time series on classification accuracy. Finally, we explore the utilization of a class activation map (CAM) to reduce the black-box effect of deep learning. The average overall accuracy of the MTS method across the training data is 0.96, which is higher than the UTS methods (0.67 to 0.88). Meanwhile, ResNet spent much more time than FCN and LSTM-FCN in model development. When testing our method on an independent dataset, the MTS models based on FCN, LSTM-FCN, and ResNet all demonstrate similar performance. Our method indicates that the CAM can explain the black-box effect of deep learning and suggests that deep learning algorithms coupled

with geometric feature time series can accurately separate leaf and woody components from point clouds. This provides a good starting point for future research into estimation of forest structure parameters.

Keywords: Classification, Time Series, Point Clouds, Geometric Feature, Deep Learning

3.1 Introduction

Forest canopy structure affects the interactions between terrestrial ecosystems and the atmosphere (Hosoi et al., 2006). Green leaves in the canopy control vital ecological processes such as photosynthesis, gas exchange, and light interception, while the woody part contributes to tree biomass (Beland et al., 2011; Tao et al., 2015). Accurate separation of leaf and woody components from the canopy has the potential to improve our understanding and management of forests from a structural point of view (Zhu et al., 2018).

The development of remote sensing techniques during the last multiple decades, particularly terrestrial laser scanning (TLS), has been increasingly applied to forestry and ecology due to its ability to capture detailed 3D information about forest structure (Calders et al., 2020; Disney et al., 2018). TLS is an active remote sensing instrument that emits laser pulses, and then reads their return signal to generate 3D point clouds, allowing for the distance to surrounding objects to be calculated based on the pulse return time (Calders et al., 2015).

Many studies have investigated the potential of TLS point clouds to separate leaf and woody components from individual trees or plots using unsupervised and supervised machine learning algorithms. These separation methods can be subdivided into three groups: (1) by geometric features (Ma et al., 2015; Moorthy et al., 2019; Vicari et al., 2019), (2) by radiometric features (Beland et al., 2014), and (3) by a mixture of both (Zhu et al., 2018; Xi et al., 2020). The radiometric method depends on the type of LiDAR system (e.g., wavelength) which makes it sensor-specific (Ma et al., 2015). Conversely, the geometric method only needs the xyz coordinates of the points, which is the fundamental information provided by all sensors (Moorthy et al., 2019). Moreover, the geometric method consistently outperforms the radiometric techniques when comparing results across the literature (Tao et al., 2015). The mixed approach has been used to classify leaf and wood (Zhu et al., 2018) and has consistently

had the highest performance of all of the classification methods; however, the mixture method still relies on the type of LiDAR system due to the inclusion of radiometric features (Vicari et al., 2019).

Geometric features describe the local geometry of a point and are often estimated from a local neighborhood of points (Belton et al., 2013). Radially bounded nearest neighbor and k -nearest neighbor approaches typically define the local neighborhood of a point (Moorthy et al., 2019; Vicari et al., 2019; Wang et al., 2017). When comparing those two approaches, the radius-bounded nearest neighbor seems to be more advanced, while the density of the point clouds influences the latter. For example, the density of point cloud in the upper part of the tree is often lower than the points in the lower part, due to occlusion effects and physical distance between the target and scanner (Beland et al., 2011). Thus, the geometric features obtained by the same k -nearest neighbors approach in the canopy are different from those in the understory (Moorthy et al., 2019). By contrast, the features defined by the radius-bounded nearest neighbor approach can allow more consistent geometric meaning (Thomas et al., 2018).

Current separation geometric feature separation methodologies that use the radius- bounded nearest neighbor approach have relied on a fixed size or have defined an optimal size using the criterion such as Shannon entropy (Ma et al., 2015; Ma et al., 2016; Zhu et al.,2018). In contrast, using a multiple radius method eliminates the need to select a single sizeto obtain features (Moorthy et al., 2019). Belton et al. (2013) also presented a leaf and woodclassification method based on a multiple radius search approach, which shows advantages than the single radius search approach; however, there are some problems witha multiple radius search approach when classifying leaf and wood components (Moorthy et al., 2019; Xi et al., 2020; Belton et al., 2013). First, searching several sizes randomly may not completely capture the geometric information of each point at different scales. Second, the distribution of classes in the training and testing data affects the performance of machine learning models (Quiñonero-Candela et al., 2009), since testing data may not have the same proportion of training data compared to the real world (Wei et al., 2013).

Recent advances in deep learning are providing an excellent opportunity to quantify forest structure and classify leaf and wood components (Calders et al., 2020). Deep learning is believed to be the best solution for discovering complex architecture in high-dimensionaldata (LeCun et al., 2015). Compared with machine learning algorithms (random forest (RF), boosting), deep learning neural network (NN)-based algorithms are more complex and require high-performance hardware (Dargan et al., 2020; Qi et al., 2017). Additionally, deep learning algorithms are composed of black-box

networks, which makes them difficult to interpret (Fawaz et al., 2019).

A time series is defined as a series of data points ordered based on a uniform interval (Wang et al., 2016). Multiple radius nearest neighbors often search a set of radii from small (0.1 m) to large values (1 m) (Belton et al., 2013; Moorthy et al., 2019), and thus can be cast as time series data for classification. Generally, time series data can be grouped into univariable time series (UTC) and multivariable time series (MTC). MTC is the finite sequence of UTC, providing more patterns to understand the data (Wang et al., 2016;Zheng et al., 2016).

Time series classification (TSC) is regarded as one of the main challenges in data mining over the last several years (Esling et al., 2012). Also, time series data have foundmany applications in image classification (Liu et al., 2022), healthcare (Rajkomar et al., 2018), human recognition (Nweke et al., 2018), and the steel industry (Mehdiyev et al., 2017). This is because any classification problem using data that is arranged accordingto some notion of order can be taken as a TSC problem (Fawaz et al., 2019). With the increment of temporal data availability, many algorithms have been presented to tackle theTSC problem (Fawaz et al., 2019). At the early stage of TSC, distance-based methods use predefined similarity measures like Eulidean distance or dynamic time warping (DTW)to perform a classification on original time series directly (Keogh et al., 2005; Lines etal., 2015). DTW coupled with k-nearest neighbors have become popularized recently and are largely considered the gold standard in this field (Bagnall et al., 2017). Feature-based methods extract a number of features, which represent the local/global time series patterns, and then perform a classification (Baydogan et al., 2013); however, those methods require more individualized data preprocessing and feature engineering (Wang et al., 2017). Deep learning models, by contrast, apply an end-to-end approach that incorporates the feature engineering internally while the model makes decisions on its ownprocess (Nweke et al., 2018). As a result, deep learning models have the ability to extract information from the time series in a more efficient and complete manner (Dargan et al., 2020).

There are many methods presented primarily to handle satellite image time series classification problems (Dou et al., 2021; Wang et al., 2020; Xu et al., 2021). For example, Xu et al. (2021) proposed a recurrent neural network (RNN) coupled with random forest for scene classification using time series images, demonstrating higher performance than support vector machine (SVM). Meanwhile, Dou et al. (2021) used deep learning networks based on convolutional neural network (CNN) on mapping land use and land cover change, producing an overall accuracy of 0.83. Though those deep learning methods demonstrate promising

performance on time series image classification, their applications on TSC of point clouds are not well studied.

In the context of previous studies, this paper seeks to investigate the potential of deep learning algorithms to separate leaf and woody components using TLS data. Specifically, Section 3.2 introduces data source and deep learning methods, and Section 3.3 illustrates the performance of our methods on point clouds. Sections 3.4 and 3.5 present discussion and conclusions. Moreover, our study focuses on the following questions: (1) Can the multiple radius search method be used to create time series data? (2) Does MTC have significant advantages over UTC for leaf and wood classifications using TLS data? (3) How well do different deep learning methods perform the classification of leaf and woody components using point clouds? (4) Can the black-box effect of deep learning become interpretable? We explore the above questions using seven broad-leaved trees scanned by TLS during the leaf-on and leaf-off conditions.

3.2 Materials and Methods

3.2.1 Study area and data

To evaluate deep learning algorithms on leaf and wood classification, we collected TLS point clouds from seven broad-leaved trees (*Ulmus americana*) at the North Campus, University of Alberta (*N*53°31′42.4″, *W*113°31′26.5″). This site is located in Edmonton, Alberta, Canada, which experiences a humid continental climate (mean annual precipitation of 480 mm) with warmer summers (mean temperature up to 23.1 °C) and cold winter (mean temperature up to -14.8 °C) (Guzmán Q et al., 2020). The seven trees were scanned during the leaf-on and leaf-off conditions in October 2017 and January 2019, respectively.

We used a Rigel VZ-400i (Horn, Austria), to scan all trees. This TLS uses a 1550 nm shortwave infrared laser light, with a maximum vertical field of view (FOV) of 120° and a horizontal FOV of 360°. The beam divergence of this TLS is 0.35 mrad, with a range accuracy of 0.005 m at 100 m range (Calders et al., 2015). The scanning strategy took the form of a 30 m ×

50 m plot. Six-cylinder reflectors installed on 2 m poles outsides of the bound of the plot as control points were used for registration. To minimize the occlusion effect, 12 scan positions were used to cover each tree. The registration was performed using RiSCAN PRO software (RIEGL Laser Measurement Systems GmbH, Horn, Austria). All points that did not belong to target trees from the registered point clouds were removed. The description of the point cloud of all trees registered under the leaf-on and leaf-off conditions is illustrated in Figure 3.1.

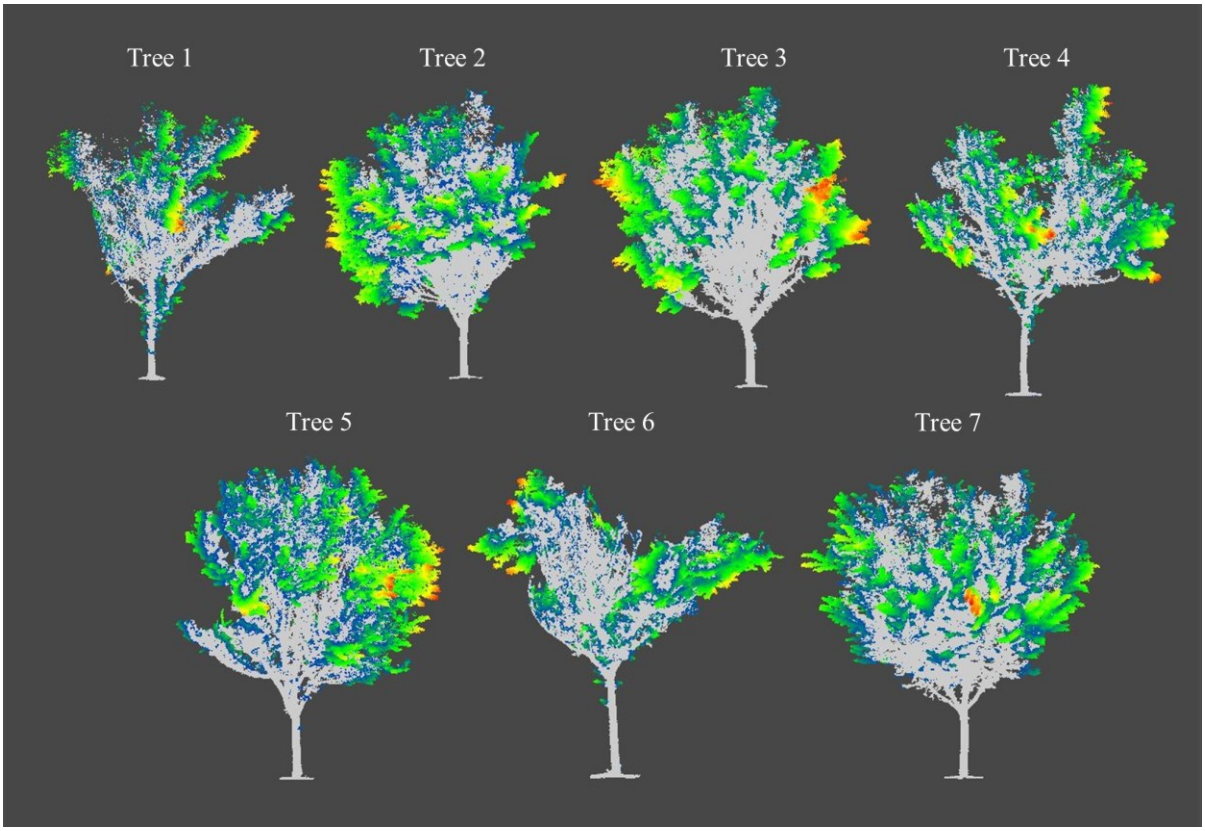


Figure 3. 1 Trees obtained after registering the point clouds under the two different phenological conditions. Color points: leaf. Gray points: wood.

3.2.2 Data processing

Predictor variables

Local dimensional features express how neighborhood points are distributed in space. They use geometric attributes derived from point clouds data to support classification problems (Ma et al., 2015; Zhen Wang et al., 2014). To obtain the local dimensional features, a covariance matrix is calculated based on several neighboring points within a specified radius. The spatial distribution of the point can then be represented by three positive eigenvalues (λ_1, λ_2, λ_3).

We used the geometric features in Table 3.1 as predictor variables to classify leaf and woody components, which is consistent with other literature (Vicari et al., 2019; Wang et al., 2017). Cloud Compare (version 2.11, Cloud Compare, GPL software) was used to obtain those features directly (Ma et al., 2015).

Table 3. 1: Three geometric features extracted from the point clouds, λ represents eigenvalues.

Feature	Name	Equation
1	Planarity	$(\lambda_2 - \lambda_3)/\lambda_1$
2	Linearity	$(\lambda_1 - \lambda_2)/\lambda_1$
3	Eigentropy	$-\sqrt[3]{\lambda_n log(\lambda_n)}$

Data labeling

To train a deep learning model, we used *registration by distance* function in Cloud Compare to identify clusters of points as training data. The distance obtained indicated the local changes due to loss of leaves and branches between the two phenological conditions.A large distance between the two phenological conditions often means leaves, while a small distance represents wood components (Figure 3.1)). Here, we labeled the points from the clouds of seven trees into two classes by hand, where class 1 and class 0 represented wood and leaf, respectively.

The classes are often unbalanced in leaf and wood classification, where the wood class is smaller than the leaf class (Moorthy et al., 2019). According to Wei et al. (2013), using balanced training data does lead to the highest overall accuracy, regardless of the percentage of the two classes in the testing data. Therefore, we randomly sampled about ten percent of the total points of each tree (about five percent per class) as our training data (see Table S3.1), indicating the broader applicability of our method. We used this sample size to balance the computational power and the number of points.

Time series data

Two types of time series, UTS and MTS, were used for the leaf and wood classification in this study. One of the challenges of applying deep learning to TSC is preparing the input time series (Van Kuppevelt et al., 2020). Here, the input data X should be a three-dimensional array (number of samples, number of timesteps, number of variables), while the output Y label is a two-dimensional array (number of samples, number of labels). Traditionally, a time series X_i is defined as a set of ordered real values of length T, witha corresponding class Y_i. The training data is composed of a number of n pairs (X_i, Y_i) (Fawaz et al., 2019). Therefore, the shape of UTS is (n, T), while MTS has the shape of (3n, T) in this case. As a result of this difference, we convert the two-dimensional shapeof UTS and MTS mentioned above into the three-dimensional time series (UTS: (n, T, 1); MTS: (n, T, 3)) using the python package *tsai* (Oguiza, 2020).

Before applying multiple radius nearest neighbors to define the number of timesteps, we used a voxel grid filter of size 0.01 m to down-sample the point clouds (Guzmán Q et al., 2020). Down-sampling ensures a uniform distribution of points in space, improving computation efficiency (Burt et al., 2019). The geometric features were computed froman increasing radius ranging from 0.05 m to 1 m, stepped by 0.05 m. As a result, the shapes of UTS and MTS are (n, 20, 1) and (n, 20, 3), respectively. Since we have a binaryclassification, the shape of the label is (n, 1). The only preprocess step for the time seriesis standardizing the input data. The values of time series in similar ranges can speed upthe computation of neural networks (Bagnall et al., 2017; Ruiz et al., 2021).

End-to-End deep learning model

There are two groups of models, in terms of the deep learning architecture, that can solve TSC: (1) deep learning algorithms with manually made features, and (2) end-to-end deep learning algorithms. This study only considers end-to-end deep learning models that incorporate feature engineering while optimizing the classifier. The main objective of deep learning models is to eliminate the bias of hand-engineered features, thus enabling the network to independently learn and make intelligent decisions (Ordóñez et al., 2016).

Convolutional neural networks (CNNs) have seen many successful applications for time series analysis (Gamboa, 2017; Karim et al., 2019), with some models obtaining accuracies of 99% (Karim et al., 2019). CNNs often demonstrate higher performance in time series classification, because applying multiple convolutional filters on an input time series results in multiple other time series, which mean multiple discriminate features can then be acquired for the classification task (Fawaz et al., 2019).

A. Fully Convolutional Neural Networks

Fully convolutional neural networks (FCNs) were originally proposed by Wang et al. (2017). This architecture outperformed other approaches (i.e., multilayer perceptrons, residual networks) for classifying UTS and MTS data according to the University of California Riverside/University of East Anglia (UCR/UEA) archive (Bagnall et al., 2018; Dau et al., 2019; Wang et al., 2017). There are three convolutional blocks in this network, where each block also has three operations:

$$\begin{aligned} y &= w \otimes x + b, \\ z &= BN\,(y), \\ h &= ReLU\,(z) \end{aligned} \tag{3.1}$$

Where x = input data; w = weight matrix; b = bias; y = output neuron; z = batch normalization on output neuron; h = activation of the output neuron, $\otimes$= convolutional operator, and BN = batch normalization BN.

BN can help speed up convergence and improve generalization. Those three blocks are performed as feature extractors and are fed into a global average pooling (GAP) layer instead of the traditional fully connected layer, thus reducing the number of parameters (Lin et al., 2013). Finally, a sigmoid function connects the GAP layer's output to produce the final label (Figure 3.2a).

B. Long Short-Term Memory-Fully Convolutional Neural Networks

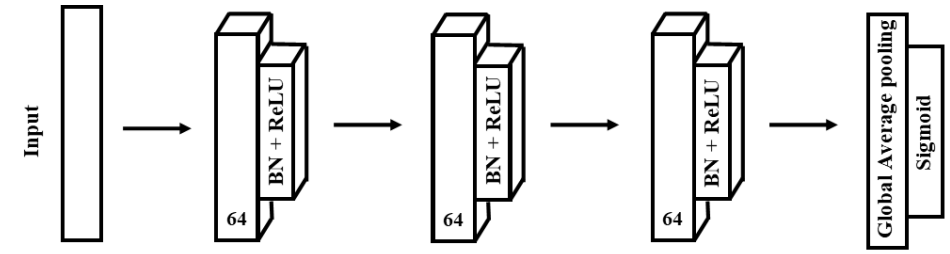
Long short-term memory (LSTM) is one of advanced RNN architecture, which handles vanishing gradient problems (Hochreiter et al., 1997). Karim et al. (2017) proposed combined utilization of LSTM and FCN on time series classification, and the obtained LSTM-FCN method further increased the performance of FCN on the UCR archive.

With the same network of FCN, LSTM-FCN incorporated a sub-module which contained two layers (Figure 3.2b). The original time series will be input into the first layer, also called the dimensional shuffle. This layer transforms a univariable time series with N timesteps into a multivariable time series (N variables with one timestep). The trans- formed time series is then passed into a LSTM layer followed by a dropout, with the dropout layer helping to prevent overfitting. The outputs of the LSTM and GAP layers are concatenated and then input into a final sigmoid layer.

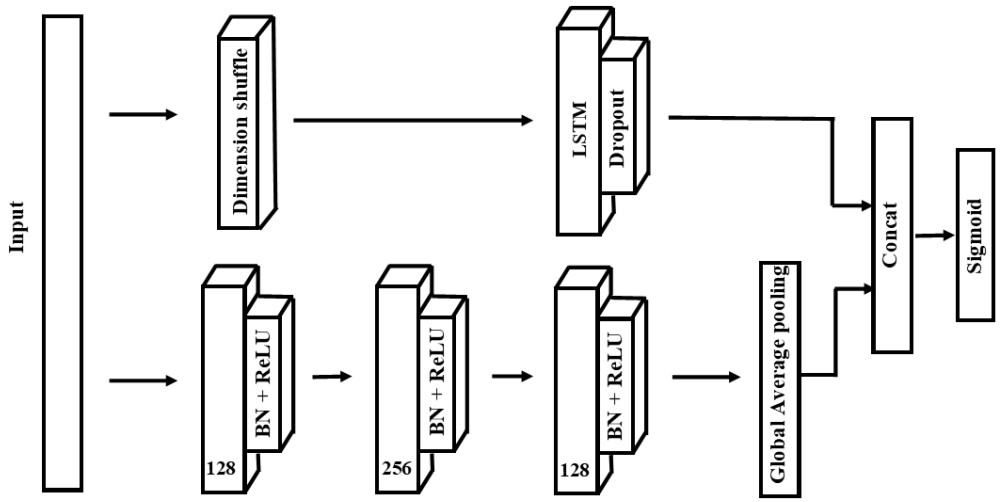
C. Residual Networks

ResNet is also a high-performance deep neural network in time series classification tasks (He et al., 2016). Xi et al. (2020) compared the performance of 15 machine learning and deep learning methods on leaf and wood classification and determined that ResNet is one of the most competitive classifiers among them.

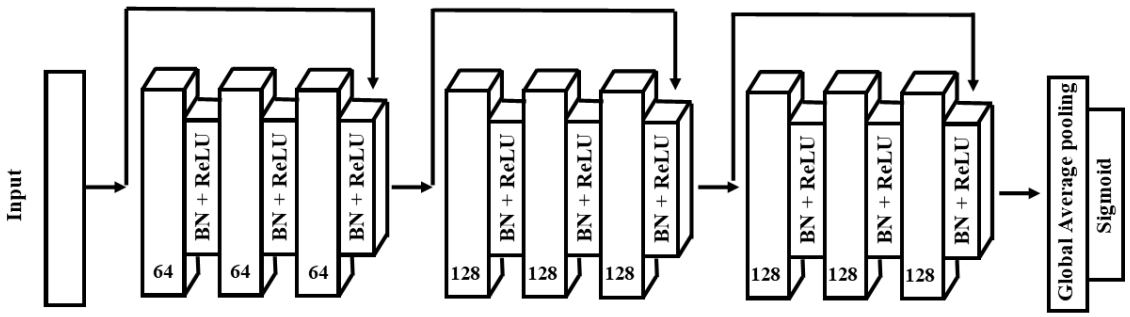
Figure 3.2c presents the network structure of ResNet used in this study. The ResNet network includes three residual blocks and connects with the GAP layer and a sigmoid function. We used the convolutional blocks in Equation 3.1 to develop each residual block. Here, the Blockn indicates the residual block with n filters, which have the following operations: x, input data; sn, output of each convolutional block; y, output neuron after adding shortcut connection; s, output of the residual block.



(a) Fully Convolutional Neural Network (FCN)



(b) Long Short Term Memory Fully Convolutional Neural Network (LSTM-FCN)



(c) Residual Network (ResNet)

Figure 3. 2: The network structure of FCN (a), LSTM-FCN (b), and ResNet (c); feature extractor: block with a convolutional layer plus a batch normalization (BN) layer and a ReLU activation function; output layer: global average pooling with a sigmoid activation function.

$$
\begin{aligned}
s_1 &= Block_{n1}(x), \\
s_2 &= Block_{n2}(s_1), \\
s_3 &= Block_{n3}(s_2), \\
y &= s_3 + x, \\
s &= ReLU\,(y)
\end{aligned}
\tag{3.2}
$$

Deep learning is often criticized for being a black-box network, meaning their network is harder to understand (Dargan et al., 2020). One of the benefits of the methodology employed here is that the GAP layer in the above network enables the utilization of class activation map (CAM) to reduce the black-box effect (Zhou et al., 2016). CAM highlights the contributing region in the raw data to the resulting classification. Therefore, CAM provides a possible explanation of how the convolution layers work for TSC (Wang et al., 2017).

When it comes to the hyperparameter of FCN, LSTM-FCN, and ResNet, the stride of all convolutions is equal to 1 with the same padding. This padding means that the length of the time series remains unchanged after the convolutions. The kernel size 3,5,8 and filter sizes 64, 128, 256 are determined by a random search method to find the best performing combinations. To make our method broadly applicable, we also add a regularization term to avoid overfitting in each convolution (Van Kuppevelt et al., 2020).

In this study, we used five trees for training and validation, while the remaining two trees served as testing data. The performance of our method, including parameter optimization, was evaluated via a k -fold cross-validation strategy (Rhys, 2020). Then, we applied Adam optimization to each method, and used binary cross-entropy as a loss function. Adam optimization is a commonly used gradient descent algorithm for training deep learning models (Kingma et al., 2014). We set up the learning rate as 0.001, the number of epochs as 100, and the batch size as 256, based on previous studies (Fawaz et al., 2019; Smith, 2018). Moreover, we used EarlyStopping to avoid overfitting (Raskutti et al., 2014). This approach keeps recording the loss on the validation data and stops training the classifier when there is no improvement in the performance of validation data. Finally, we applied the CAM method to identify the contribution region of an input time series for the predicted labels. More details about the

hyperparameter of FCN, LSTM-FCN, and ResNet are given in Section 3.6.

3.2.3 Evaluation

In addition to evaluating the performance of three deep learning networks on Trees 1 and 5, we tested our methods on six trees from Vicari et al. (2019) to indicate model generation on other tree species. Three of the trees came from Alice Holt, UK, named as alice 1, alice 2, and alice 3. The remaining three trees, called caxiuanaA 117, nouraguesH20 108, and pan 33 are from Caxiuana, Brazil, Nouragues, French Guiana and London, UK, respectively. More details about those trees can be obtained from Vicari et al. (2019). We randomly selected 10,000 points per class for each tree to evaluate our model generation.

To compare the performance of our method against other leaf and wood classification schemes in literature (Moorthy et al., 2019; Vicari et al., 2019; Zhu et al., 2018), we used Precision, Recall, F1 score, accuracy, and a receiver operating characteristic (ROC) curve to evaluate our model's performance. Precision measures the proportion of the true positive (TP) out of all predicted positive points (TP and FP).

$$Precison = \frac{TP}{TP + FP} \tag{3.3}$$

Recall is the proportion of the predicted true positive (TP) against all true positive points (TP and FN).

$$Recall = \frac{TP}{TP + FN} \tag{3.4}$$

F1 score is the trade-off between Precision and Recall, ranging from 0 to 1.0. A perfect classifier would give a 1.0 of F1 score.

$$F1\ score = 2 * \frac{Precison * Recall}{Precison + Recall} \tag{3.5}$$

Accuracy is the proportion of all corrected classified points (TP and TN) against all points.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.6}$$

We used Receiver Operating Characteristic (ROC) curve to understand the discrimination power

of three networks. The curve is generated by True Positive Rate (Recall) vs. False Positive Rate at different thresholds. The area under the ROC curve is between 0 and 1. A high area indicates the high prediction score of the classifier.

$$False\ Positive\ Rate\ = \frac{FP}{FP + TN} \tag{3.7}$$

3.3 Results

3.3.1 UTS vs. MTS

Figure 3.3 gives the time series of three features using the multiple radius search method, revealing the different changing patterns between a leaf and wood point. Concerning the leaf point, the eigentropy tended to have a lower standardized value (< -0.6) in the intermediate radius timesteps (6 to 16), while the radius value of the wood point (> 0) was larger.

To assess the effects of UTS and MTS on the model performance, a five-fold cross-validation strategy was applied to the training data. Table 3.2 demonstrates all the classification result metrics for both the UTS and MTS of three deep learning networks. As indicated in the table, our method based on MTS clearly outperformed the method based on UTS. When used with MTS, LSTM-FCN and ResNet demonstrated the same overall accuracy of 0.96, with the same F1 score of 0.96 for leaves and wood. FCN had the lowest accuracy of 0.92 among the three networks. Concerning UTS, the eigentropy had a better F1 score for leaf and wood and accuracy than planarity and linearity for all three networks.

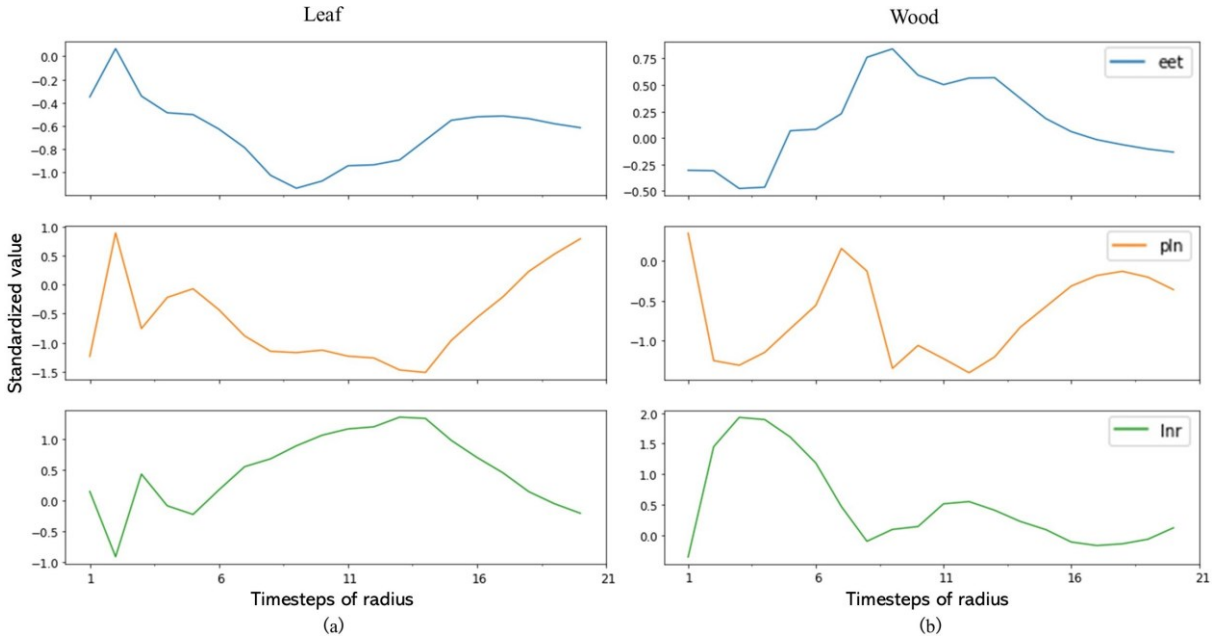


Figure 3. 3: Time series of a leaf and wood point from a cloud using multiple radius nearest neighbors; eet: eigentropy; pln: planarity; lnr: linearity; the Y-axis is the standardized value of each feature; the X-axis is the timesteps of radius (meter), here we have 20 radius or timesteps, then the value of X is ranging from 1 to 21.

For example, ResNet produced an overall accuracy of 0.88 on eigentropy, higher than planarity (0.80) and linearity (0.79). Moreover, the planarity and linearity demonstrated similar performances across all metrics for all three networks.

Table 3. 2: Comparison of univariable (UTS) and multivariable time series (MTS) model on FCN, LSTM-FCN, and ResNet using Precision, Recall, F1 score, and Accuracy for training data; Accuracy means the corrected classified leaf and wood points against the total points (UTS: Planarity; Linearity; Eigentropy, MTS: Overall).

Method	FCN			LSTM-FCN			ResNet		
	Accuracy	F1 score (leaf)	F1 score (wood)	Accuracy	F1 score (leaf)	F1 score (wood)	Accuracy	F1 score (leaf)	F1 score (wood)
Planarity	0.67	0.67	0.66	0.74	0.74	0.74	0.80	0.80	0.80
Linearity	0.69	0.65	0.72	0.76	0.76	0.76	0.79	0.79	0.79
Eigentropy	0.82	0.81	0.82	0.85	0.85	0.85	0.88	0.87	0.88
MTS	0.92	0.92	0.92	0.96	0.96	0.96	0.96	0.96	0.96

Table 3.3 provides information about the training process of the UTS and MTS methods of three networks. The average training time of ResNet (487 min) was around 7 and 14 times higher than FCN (68 min) and LSMT-FCN (35 min), respectively. Also, the average number of epochs and training time between UTS and MTS model were similar. For instance, the mean epochs and training time of UTS model based on ResNet were 225±34 and 487±74 min, while that of MTS model were 255 and 553 min.

Table 3. 3: The number of epochs and training time of univariable (UTS) and multivariable time series (MTS) model for FCN, LSMT-FCN, and ResNet; the number of epochs is the sum of the epochs of the fivefold cross-validation method; average is the mean of the number of epochs and training time of the three UTS methods (UTS: Planarity; Linearity; Eigentropy, MTS: Overall).

	FCN		LSTM-FCN		ResNet	
Method	Epochs	Time (mins)	Epochs	Time (mins)	Epochs	Time (mins)
Planarity	227	74	222	29	253	548
Linearity	226	75	335	45	187	405
Eigentropy	159	54	234	30	234	508
Average	204 ± 39	68 ± 12	264 ± 62	35 ± 9	225 ± 34	487 ± 74
MTS	185	62	223	54	255	553

3.3.2 Class activation map

Figure 3.4 demonstrates the result of applying CAM on the validation data using the UTS and MTS methods on FCN. We observed that the most-contributed regions of the time series for the predicted wood or leaf class were highlighted in a red color, while the blue regions are non-discriminatory. With respect to UTS, the small radius highly contributes the wood class for planarity, while the large radius contributes the wood class for linearity. This result is consistent with wood distribution in the 3D space. The wood point tends to have surface distribution in small scale, and the wood point tends to have linear distribution in large scale. Also, the CAM of eigentropy was more uniform distributed than that of planarity and linearity. In other words, the CAM of eigentropy contains darker red and blue regions, demonstrating that the eigentropy can filter out the most contributed region with higher confidence than planarity and linearity (See Table 3.2, eigentropy shows the highest performance among the three features). When comparing the eigentropy and MTS models, the red subsequence highlighted by the wood class of the MTS model was mainly located in a larger radius (Figure 3.4g). Conversely, the discriminatory regions of eigentropy can be observed in the middle and larger radius in the wood class of Figure 3.4e, demonstrating the overlap with the leaf class.

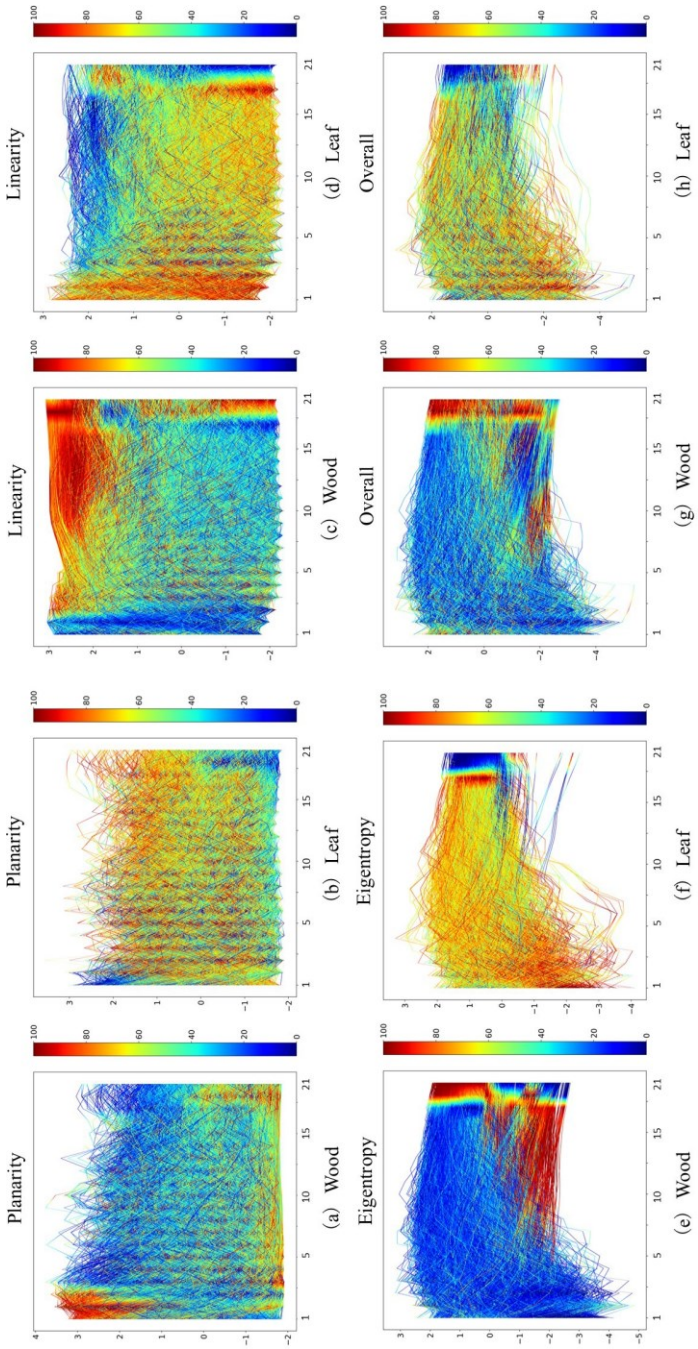


Figure 3. 4: Class Activation Map (CAM) of univariable (UTS) and multivariable time series (MTS) using fully convolutional networks (FCN) on the validation data; (a) CAM of Planarity on Wood, (b) CAM of Planarity on Leaf, (c) CAM of Linearity on Wood, (d) CAM of Linearity on Leaf, (e) CAM of Eigentropy on Wood, (f) CAM of Eigentropy on Leaf, (g) CAM of Overall on Wood, (h) CAM of Overall on Leaf. X- axis is the radius value, and Y-axis is the geometric feature values.

3.3.3 Model generalization

Table 3.4 compares the performance of the MTS model based on FCN, LSTM-FCN, and ResNet on the testing data. We did not find a significant difference between the accuracy of three deep learning networks (FCN vs. LSTM-FCN: t = 0.78, p-value = 0.45; FCN vs. ResNet: t = 1.01, p-value = 0.33; LSTM-FCN vs. ResNet: t = 0.24, p-value = 0.82). FCN generated the highest accuracy and F1 score of leaf and wood, whether for Tree 1 or Tree 5. For example, the overall accuracy of FCN on Tree 5 was 0.78, with F1 score of leaf (0.77) and wood (0.79), while the overall accuracy and F1 score of leaf and wood of LSTM-FCN (0.75, 0.74, and 0.75) and ResNet (0.73, 0.72, and 0.74) were lower. With respect to testing data from Vicari et al. (2019), all three networks demonstrated relatively poor performance on caxiuanA 117 and pan 33, while they generated the highest accuracy and F1 score on nouragueH20 108.

Table 3. 4: The performance of multivariable time series (MTS) of FCN, LSTM-FCN, and ResNet on independent data; accuracy means the corrected classified leaf and wood points against the total points (Ave, average; Sd, standard derivation).

	FCN			LSTM-FCN			ResNet		
Tree name	Accuracy	F1 score	F1 score	Accuracy	F1 score	F1 score	Accuracy	F1 score	F1 score
Tree 1	0.75	0.76	0.75	0.74	0.72	0.76	0.73	0.71	0.75
Tree 5	0.78	0.77	0.79	0.75	0.74	0.75	0.73	0.72	0.74
alice 1	0.76	0.75	0.83	0.73	0.72	0.69	0.65	0.66	0.64
alice 2	0.81	0.81	0.81	0.74	0.74	0.74	0.77	0.76	0.78
alice 3	0.79	0.78	0.81	0.75	0.74	0.78	0.78	0.77	0.82
caxiuanA 117	0.61	0.52	0.76	0.63	0.59	0.76	0.66	0.60	0.78
nouraguresH20 108	0.84	0.84	0.85	0.82	0.81	0.85	0.78	0.79	0.80
Pan 33	0.67	0.64	0.71	0.63	0.59	0.69	0.63	0.58	0.76
Ave	0.75	0.73	0.79	0.72	0.71	0.75	0.72	0.70	0.76
Sd	0.08	0.10	0.05	0.06	0.08	0.05	0.06	0.08	0.06

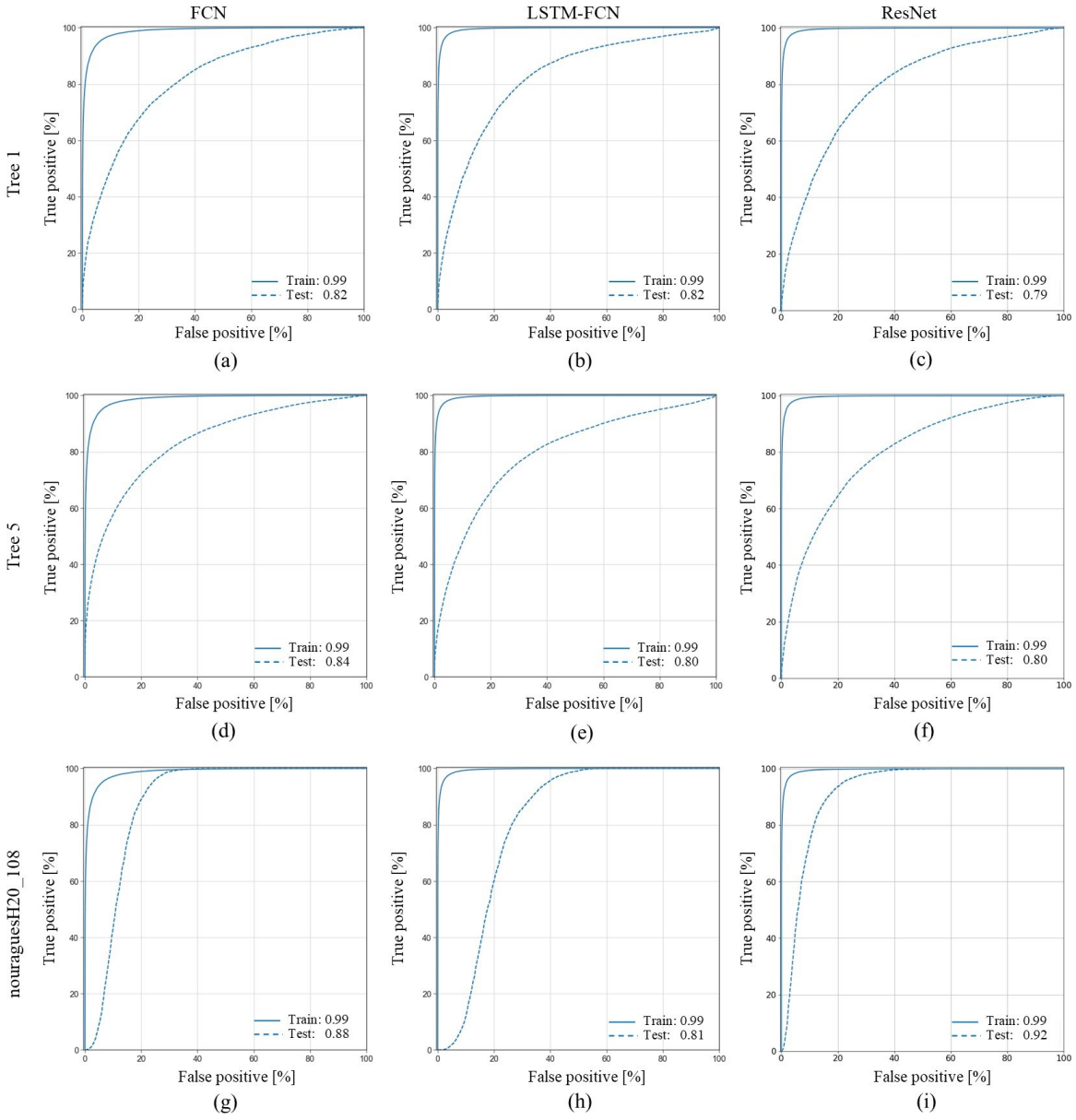


Figure 3. 5: Receiver operating characteristic (ROC) curve for training data and testing data (Tree 1, Tree 5, and nouraguesH20 108) using multivariable time series (MTS) method based on FCN, LSTM-FCN, and ResNet; (a) FCN on Tree 1, (b) LSTM-FCN on Tree 1, (c) ResNet on Tree 1, (d) FCN on Tree 5, (e) LSTM-FCN on Tree 5, (f) ResNet on Tree 5, (g) FCN on nouraguesH20 108, (h) LSTM-FCN on nouraguesH20 108, (i) ResNet on nouraguesH20 108.

Figure 3.5 displays the ROC curves of the MTS model of FCN, LSTM-FCN, and ReNet on Tree 1, Tree 5, and nouraguesH20 108. The area under the ROC curve of the training data was 0.99, indicating the solid discriminatory power of three networks. Although the area under ROC curve of three trees was lower than the training data, most of those areas were higher than 0.80, which still demonstrates the good discriminatory power of three deep learning networks. It should be note that the areas under ROC curve of Tree 1 (0.82, 0.82 and 0.79) and Tree 5 (0.84, 0.80 and 0.80) were similar for FCN, LSTM-FCN, and ResNet. When we evaluated nouraguesH20 108, ReNet (0.92) and FCN (0.88) had a higher area under the ROC curve than LSTM-FCN (0.81).

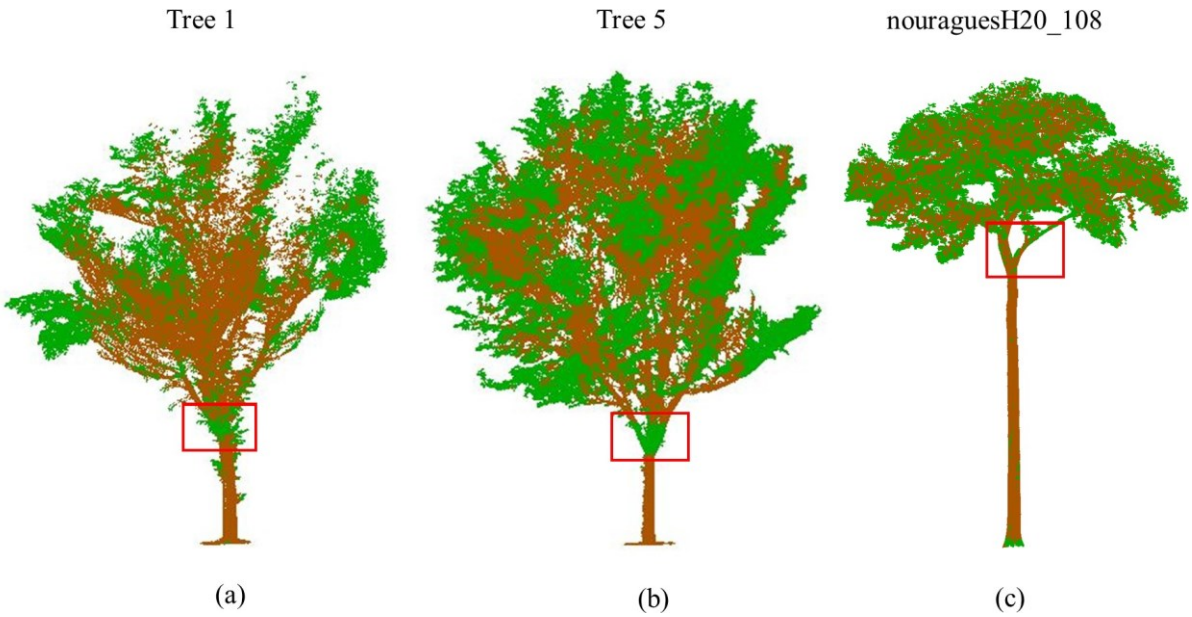


Figure 3. 6: The multivariable time series (MTS) model of fully convolutional networks (FCN) prediction for Tree 1 (a), Tree 5 (b), and nouraguesH20 108 (c); the red square highlights the misclassified areas. Green: leaf points. Brown: woody points.

The prediction of the MTS model based on FCN for Tree 1, Tree 5, and nouragueHs20 108 can be seen in Figure 3.6. Our visual inspection indicated that FCN could detect most of the big wood structures. Though some parts of the main trunk were misclassified as the leaf in Tree 1, Tree 5, and nouraguesH20 108 (indicated by red square in Figure 3.6), we can manually correct those misclassified leaf points in the main trunk as wood points for each tree (indicated by red square in Figure 3.7).

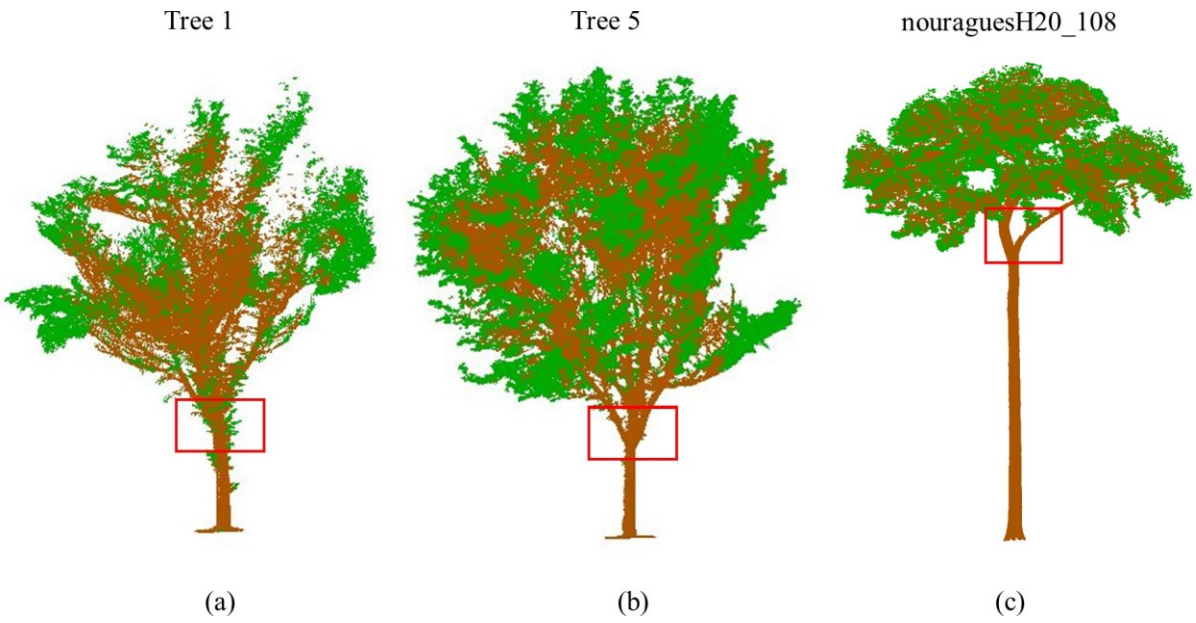


Figure 3. 7: The multivariable time series (MTS) model of fully convolutional networks (FCN) prediction for Tree 1 (a), Tree 5 (b), and nouragueH20 108 (c) after manual intervention; the red square highlights the area where we make manual intervention.

3.4 Discussion

In this paper, we proposed a time series classification method based on three deep learning networks to separate leaf and woody components of given trees. We also compared the effects of the UTS and MTS methods on the classification results. As indicated in Table 3.2, the MTS method outperformed the UTS method on the training data. Meanwhile, ResNet spent much more time than FCN and LSTM-FCN in the training model. Moreover, we found the three networks performed similarly on the testing data. Finally, we understood the utilization of the class activation map (CAM) can explain the black-box effect of deep learning networks.

3.4.1 UTS or MTS

We applied the UTS and MTS methods to separate leaf and woody components of point clouds. Compared with UTS, MTS always demonstrated higher performance on the training data. It is not a surprise since MTS is a finite sequence of the UTS, thus providing more information to help the classification process (Wang et al., 2016; Zheng et al., 2016). It also can be observed in Table 3.3 that the UTS and MTS methods demonstrate similar training time (e.g., ResNet: 487 ± 74 min vs. 553 min) and numbers of epochs (225 ± 34 vs. 255). Therefore, we recommend the MTS method for leaf and wood classification. With respect to the UTS method, the eigentropy (0.82 of FCN, 0.85 of LSTM-FCN, and 0.88 of ResNet) seems to be the most advanced feature for leaf and wood classification, while the overall accuracy of planarity and linearity range from 0.67 to 0.80. Therefore, we believe the eigentropy contributes most to the performance of MTS. The performance of the MTS method may be further improved if more power features like eigentropy are included.

CAM provides a method to understand what deep learning is doing for the time series (Fawaz et al., 2019; Wang et al., 2017). According to Figure 3.4, the main contributing regions of planarity for the wood points are located at a small radius, while the linearity for wood points is found in large values. The distribution of wood points across the different radii have been reported in Ma et al. (2015) and Zhu et al. (2018). For example, in a smaller radius, the wood points in the main trunk could have a surface distribution, while the point clouds of the trunk are more likely to be linearly distributed at a large radius (Ma et al., 2015). CAM also demonstrates that the discriminatory regions of planarity and linearity for both classes are less uniformly distributed than that of the eigentropy and MTS model for two classes. These leaf and wood points demonstrate similar discriminatory areas for planarity and linearity across the radius. This unequal distribution could be the reason for worse performance of the planarity and linearity models. Moorthy et al. (2019) applied a geometric approach to classify leaf and woody components from point clouds and found that the geometric properties of the leaves demonstrated similarity with the wood points in long branches in the upper part of a tree. Conversely, the contribution regions' boundary between leaf and wood in the MTS method is clear. The MTS method combines all features and extracts the most discriminatory regions from the time series, and thus has the highest F1 score and overall accuracy across training data.

3.4.2 Comparison with other leaf-wood classification methods

In this paper, we applied deep learning networks to separate leaf and woody components from the point clouds. There are many leaf and wood classification methods in the literature (Ma et al., 2015; Moorthy et al., 2019; Vicari et al., 2019; Zhu et al., 2018). Generally, those methods are mainly based on machine learning algorithms such as Gaussian mixture models (GMMs), random forest (RF) and boosting. Compared to deep learning algorithms, machine learning algorithms solve problems by decomposing a larger task into small tasks and combining the results. The features used in machine learning need to be accurately and precisely identified by users (Dargan et al., 2020). For instance, Zhu et al. (2018) applied an adaptive radius nearest neighbor search algorithm to drive the optimal radius for each point, and then used RF algorithms to classify leaf and wood, with an average overall accuracy of 70.4%. Ma et al. (2015) also conducted a sensitivity analysis to define the optimal radius for each point and then applied GMMs to separate leaf and wood, with an overall accuracy of 95.5%. Moreover, Moorthy et al. (2019) investigated the spatial distribution of leaf and wood points using eigenvalues and eigenvectors at five radii, with applied RF and boosting algorithms for classification. The eigenvalues and eigenvectors gave the model an average overall accuracy of 94.2%.

Furthermore, Xi et al. (2020) compared the performance of 15 machine learning and deep learning classifiers (i.e., RF, Ada Boost, PointNet, ResNet) on leaf and wood separation. This work computed geometric features of three randomly selected radii, and then combined radiometric features for classification. The results indicated that ResNet was one of the most competitive classifiers among the 15 approaches, while also providing a higher efficiency. In this study, we compared the performance of FCN, LSTM-FCN, and ResNet. As demonstrated in Table 3.3, ResNet spent much more time than the remaining two networks in developing the model, which means that FCN and LSTM-FCN were more efficient. When evaluating three networks on the testing data, we found FCN, LSTM-FCN, and ResNet performed similarly.

Here, we used the multiple radius search method to obtain the time series of leaf and wood points and input them into deep learning networks to make independent decisions. There is no need for prior knowledge, which means no heavy feature engineering is required to understand the feature distribution for deep learning algorithms. The only preprocessing step needed is standardizing the data. Deep learning algorithms can create new features through their own process and then make predictions (Fawaz et al., 2019).

3.4.3 Model generalization

Developing a appliable and transferable model for forests is one of the main challenges of point cloud classification (Moorthy et al., 2019). Though there are many methods proposed for leaf and wood classification in previous studies, the generalization of those methods is questionable. Moorthy et al. (2019) used imbalanced training data to classify leaf and wood, with the wood points accounting for about 15% of all the points. When testing the former method on the independent data, the average wood F1 score of field data was 0.69 (0.79 of FCN here). Wei et al. (2013) demonstrated that the training and testing of models on binary classification problems depends on the fraction of the two classes. In simpler terms, it is not easy to apply a model to independent data from the real world that is similar to the training data based on the fraction of the two classes. Using balanced training data can lead to the highest accuracy, regardless of the fraction of the two classes in the independent data. Although Zhu et al. (2018) applied a balanced method for classifying leaf and wood, only 1000 points per class are manually selected from the point clouds, much less than in the methodology applied here. We randomly selected 10% of the total points from each tree, thus creating 320,000 points for training (Table S3.1). The overall accuracy of Zhu et al. (2018) by geometric features on validation data was 70.4%, which is lower than our method (0.75 of FCN, see Table 3.4). Therefore, the presented method in this study has broader applicability than previous methods.

The main advantages of our method are as follows: First, we use a multiple radii search method to obtain a time series of the geometric features, which eliminates the work for a specific radius selection (Ma et al., 2015). Second, since we use deep learning models for classification, the model provides an end-to-end network to tackle the problems without heavy crafting on data preprocessing and feature engineering (Wang et al., 2017). Third, we take a multivariable time-series strategy coupled with deep learning algorithms to classify leaf and wood, which is more accurate than univariable time series strategy. Fourth, we add global average pooling (GAP) layer on our networks to classify leaf and wood points, enabling the utilization of CAM to reduce the black box effects of deep learning algorithms. Finally, since we use a balanced class for leaf and wood classification, the performance of our method on the independent data is higher than in previous studies (Moorthy et al., 2019), which indicates the broader applicability of our proposed model.

The balanced class also brings the limitation of our method; the leaf and wood points are still balanced in the predicted results, which is not the case for the trees in the natural environment (Ma et al., 2015; Moorthy et al., 2019). Therefore, we manually corrected the misclassified leaf points in the

main trunk of the predicted tree, which can be seen in Figures 3.6 and 3.7. Postprocessing steps may be needed to further improve the performance of our method; however, the presented leaf/wood separation model based on FCN demonstrated overall accuracies of 0.75, 0.78, and 0.84 on Tree 1, Tree 5, nouraguesH20 108, without any additional postprocessing, which means our model could provide a good starting point for the scientific community to using TLS to measure forest structure. It should be also noted that we do not include needle leaf trees in our study, as broad-leaved and needle leaf trees have different point arrangements. A more careful scanning protocol is needed since needles are often smaller than the footprint of the LiDAR instrument (Vicari et al., 2019; Zhu et al., 2018). We provide the entire leaf and wood classification procedure as an open-source python package on Github (https://github.com/than2/leafandwood), expecting more researchers to evaluate our methods on other forest types (e.g., needle-leaf trees, dry, and rain forests). Future work could focus on understanding the discriminatory regions of the time series for each class obtained by the CAM. In other words, the CAM may also provide a solid foundation for leaf/wood separation based on traditional machine learning algorithms. For example, machine learning algorithms could use the radius around the discriminatory regions of the time series for each class obtained by CAM to classify leaf and wood, which may help to improve their performance.

3.5 Conclusions

In this paper, we proposed an automatic deep learning time series method for leaf and wood classification from 3D point clouds. The presented method uses the multiple radius search method to obtain the time series of the geometric features, with the MTS method found to be more accurate than the UTS method. Our method eliminates the need for feature engineering, which is a requirement for most literature. Though the multiple radius search method results in high-dimensional data, this can be adequately handled by deep learning algorithms. The CAM of our deep learning time series networks can articulate the contributing region for leaf and wood points, and thus explain how the deep learning networks function. It is not a surprise that the generalization of our model is better than previous studies, since our model builds on and extends previous work. The presented method can help to understand forest structure and provide better estimates of forest structural parameters from TLS data such as LAI, WAI, or leaf area density.

Acknowledgement

This research was funded by the Chinese Scholarship Council (CSC): 201806850089, and the National Science and Engineering Research Council of Canada (NSERC) Discovery Grant. The LiDAR instrument and data processing workstation were provided by the Center for Earth Observation Sciences (CEOS) of U of A. Tao Han, and J. Antonio Guzman Q collected the data at the University of Alberta, Canada. The authors declare no conflicts of interest.

3.6 Supplementary materials

Table S3. 1: The sample size of seven trees for model training.

Tree ID	Sample points	Total points
Tree 1	32000	319037
Tree 2	70000	668950
Tree 3	96000	947704
Tree 4	56000	540369
Tree 5	80000	802740
Tree 6	32000	302494
Tree 7	66000	653643

Trees 2, 3, 4, 6, and 7 were selected as the training data, while the remaining trees (1 and 5) were regarded as the independent data to test the model generalization. The final number of points of the training data is 320, 000, and that of the testing data of Tree 1 and Tree 5 is 32 000 and 80 000, respectively.

There are many hyperparameters when using FCN, LSTM-FCN, and ResNet. The following values were used for those hyperparameters. The number of epochs is equal to 100, and it represents the number of times that the deep learning networks work through the entire training data. The batch size is equal to 256, and this hyperparameter deter-mines the number of samples to work through before updating the parameters of the network. A large batch size would lead to the significant degradation of the performance of the model, while a small batch size means more training time. The range of batch sizes used in this study is 128, 256, 512.

The callbacks in deep learning mean a set of functions that can control the training procedure. We used ReduceLROnPlateau and EarlyStopping to control the learning rate and the number of epochs during the training process. ReduceLROnPlateau means that when the loss of validation stops improving, the learning rate will decrease. We set up the factor, patience, and min lr as 0.5, 10, and 0.0001, respectively. This indicated that if the validation loss did not improve after ten epochs, the learning rate would decrease by 50% until it reached 0.0001. The patience in EarlyStopping was also equal to 10, which means that if the validation loss did not improve after ten epochs, the training procedure would stop. The verbose was equal to 1, demonstrating the output of our neural network when training the model.

3.7 References

Bagnall, Anthony, Hoang Anh Dau, Jason Lines, Michael Flynn, James Large, Aaron Bostrom, Paul Southam, and Eamonn Keogh (2018). "The UEA multivariate time series classification archive, 2018". In: *arXiv preprint arXiv:1811.00075*.

Bagnall, Anthony, Jason Lines, Aaron Bostrom, James Large, and Eamonn Keogh (2017). "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances". In: *Data mining and knowledge discovery* 31.3, pp. 606– 660.

Baydogan, Mustafa Gokce, George Runger, and Eugene Tuv (2013). “A bag-of-features framework to classify time series”. In: IEEE transactions on pattern analysis and ma- chine intelligence 35.11, pp. 2796–2802.

Beland, Martin, Dennis D Baldocchi, Jean-Luc Widlowski, Richard A Fournier, and Michel M Verstraete (2014). “On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR”. In: Agricultural and Forest Meteorology 184, pp. 82–97.

Beland, Martin, Jean-Luc Widlowski, Richard A Fournier, Jean-Fran¸cois Cˆot´e, and Michel M Verstraete (2011). “Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements”. In: Agricultural and Forest Meteorology 151.9, pp. 1252–1266.

Belton, David, Simon Moncrieff, and Jane Chapman (2013). “Processing tree point clouds using Gaussian Mixture Models”. In: ISPRS Annals of Photogrammetry, Remote Sens- ing and Spatial Information Sciences. Presented at the WG 3, pp. 43–48.

Burt, Andrew, Mathias Disney, and Kim Calders (2019). “Extracting individual trees from lidar point clouds using treeseg”. In: Methods in Ecology and Evolution 10.3, pp. 438– 445.

Calders, Kim, Jennifer Adams, John Armston, Harm Bartholomeus, Sebastien Bauwens, Lisa Patrick Bentley, Jerome Chave, F Mark Danson, Miro Demol, Mathias Disney, et al. (2020). “Terrestrial laser scanning in forest ecology: Expanding the horizon”. In: Remote Sensing of Environment 251, p. 112102.

Calders, Kim, Glenn Newnham, Andrew Burt, Simon Murphy, Pasi Raumonen, Martin Herold, Darius Culvenor, Valerio Avitabile, Mathias Disney, John Armston, et al. (2015). “Nondestructive estimates of above-ground biomass using terrestrial laser scan- ning”. In: Methods in Ecology and Evolution 6.2, pp. 198–208.

Dargan, Shaveta, Munish Kumar, Maruthi Rohit Ayyagari, and Gulshan Kumar (2020). “A survey of deep learning and its applications: a new paradigm to machine learning”. In: Archives of Computational Methods in Engineering 27.4, pp. 1071–1092.

Dau, Hoang Anh, Anthony Bagnall, Kaveh Kamgar, Chin-Chia Michael Yeh, Yan Zhu, Shaghayegh Gharghabi, Chotirat Ann Ratanamahatana, and Eamonn Keogh (2019). “The UCR time series archive”. In: IEEE/CAA Journal of Automatica Sinica 6.6, pp. 1293–1305.

Disney, Mathias I, Matheus Boni Vicari, Andrew Burt, Kim Calders, Simon L Lewis, Pasi Raumonen, and Phil Wilkes (2018). “Weighing trees with lasers: advances, challenges and opportunities”. In: Interface Focus 8.2, p. 20170048.

Dou, Peng, Huanfeng Shen, Zhiwei Li, and Xiaobin Guan (2021). “Time series remote sensing image classification framework using combination of deep learning and mul- tiple classifiers system”. In: International Journal of Applied Earth Observation and Geoinformation 103, p. 102477.

Esling, Philippe and Carlos Agon (2012). “Time-series data mining”. In: ACM Computing Surveys (CSUR) 45.1, pp. 1–34.

Fawaz, Hassan Ismail, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller (2019). “Deep learning for time series classification: a review”. In: Data mining and knowledge discovery 33.4, pp. 917–963.

Gamboa, John Cristian Borges (2017). “Deep learning for time-series analysis”. In: arXiv preprint arXiv:1701.01887.

Guzman Q, J Antonio, Iain Sharp, Felipe Alencastro, and G Arturo Sanchez-Azofeifa (2020). “On the relationship of fractal geometry and tree–stand metrics on point clouds derived from terrestrial laser scanning”. In: Methods in Ecology and Evolution 11.10, pp. 1309–1318.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2016). “Deep residual learning for image recognition”. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Hochreiter, Sepp and Ju¨rgen Schmidhuber (1997). “Long short-term memory”. In: Neural computation 9.8, pp. 1735–1780.

Hosoi, Fumiki and Kenji Omasa (2006). “Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar”. In: IEEE transactions on geoscience and remote sensing 44.12, pp. 3610–3618.

Karim, Fazle, Somshubra Majumdar, Houshang Darabi, and Shun Chen (2017). “LSTM fully convolutional networks for time series classification”. In: IEEE access 6, pp. 1662– 1669.

Karim, Fazle, Somshubra Majumdar, Houshang Darabi, and Samuel Harford (2019). “Multivariate LSTM-FCNs for time series classification”. In: Neural Networks 116, pp. 237– 245.

Keogh, Eamonn and Chotirat Ann Ratanamahatana (2005). “Exact indexing of dynamic time warping”. In: Knowledge and information systems 7.3, pp. 358–386.

Kingma, Diederik P and Jimmy Ba (2014). “Adam: A method for stochastic optimization”. In: arXiv preprint arXiv:1412.6980.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). “Deep learning”. In: nature 521.7553, pp. 436–444.

Lin, Min, Qiang Chen, and Shuicheng Yan (2013). “Network in network”. In: arXiv preprint arXiv:1312.4400.

Lines, Jason and Anthony Bagnall (2015). “Time series classification with ensembles of elastic distance measures”. In: Data Mining and Knowledge Discovery 29.3, pp. 565– 592.

Liu, Jia, Yuming Wu, Xing Gao, and Xuehua Zhang (2022). “A Simple Method of Mapping Landslides Runout Zones Considering Kinematic Uncertainties”. In: Remote Sensing 14.3, p. 668.

Ma, Lixia, Guang Zheng, Jan UH Eitel, Troy S Magney, and L Monika Moskal (2016). “Determining woody-to-total area ratio using terrestrial laser scanning (TLS)”. In: Agricultural and forest meteorology 228, pp. 217–228.

Ma, Lixia, Guang Zheng, Jan UH Eitel, L Monika Moskal, Wei He, and Huabing Huang (2015). “Improved salient feature-based approach for automatically separating photo- synthetic and nonphotosynthetic components within terrestrial lidar point cloud data of forest canopies”. In: IEEE Transactions on geoscience and remote sensing 54.2, pp. 679–696.

Mehdiyev, Nijat, Johannes Lahann, Andreas Emrich, David Enke, Peter Fettke, and Peter Loos (2017). “Time series classification using deep learning for process planning: A case from the process industry”. In: Procedia Computer Science 114, pp. 242–249.

Moorthy, Sruthi MK, Kim Calders, Matheus B Vicari, and Hans Verbeeck (2019). “Im- proved supervised learning-based approach for leaf and wood classification from LiDAR point clouds of forests”. In: IEEE Transactions on Geoscience and Remote Sensing 58.5, pp. 3057–3070.

Nweke, Henry Friday, Ying Wah Teh, Mohammed Ali Al-Garadi, and Uzoma Rita Alo (2018). “Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges”. In: Expert Systems with Applications 105, pp. 233–261.

Oguiza, Ignacio (2020). “tsai-A state-of-the-art deep learning library for time series and sequential data”. In: Retrieved April 20, p. 2021.

Ordonez, Francisco Javier and Daniel Roggen (2016). “Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition”. In: Sensors 16.1, p. 115.

Qi, Charles R, Li Yi, Hao Su, and Leonidas J Guibas (2017). “Pointnet++: Deep hierarchical feature learning on point sets in a metric space”. In: arXiv preprint arXiv:1706.02413.

Quin˜onero-Candela, Joaquin, Masashi Sugiyama, Neil D Lawrence, and Anton Schwaighofer (2009). Dataset shift in machine learning. Mit Press.

Rajkomar, Alvin, Eyal Oren, Kai Chen, Andrew M Dai, Nissan Hajaj, Michaela Hardt, Peter J Liu, Xiaobing Liu, Jake Marcus, Mimi Sun, et al. (2018). “Scalable and accurate deep learning with electronic health records”. In: NPJ Digital Medicine 1.1, pp. 1–10.

Raskutti, Garvesh, Martin J Wainwright, and Bin Yu (2014). “Early stopping and non- parametric regression: an optimal data-dependent stopping rule”. In: The Journal of Machine Learning Research 15.1, pp. 335–366.

Rhys, Hefin (2020). Machine Learning with R, the tidyverse, and mlr. Simon and Schuster.

Ruiz, Alejandro Pasos, Michael Flynn, James Large, Matthew Middlehurst, and Anthony Bagnall (2021). “The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances”. In: Data Mining and Knowledge Discovery 35.2, pp. 401–449.

Smith, Leslie N (2018). “A disciplined approach to neural network hyper-parameters: Part 1– learning rate, batch size, momentum, and weight decay”. In: arXiv preprint arXiv:1803.09820.

Tao, Shengli, Qinghua Guo, Shiwu Xu, Yanjun Su, Yumei Li, and Fangfang Wu (2015). “A geometric method for wood-leaf separation using terrestrial and simulated lidar data”. In: Photogrammetric Engineering & Remote Sensing 81.10, pp. 767–776.

Thomas, Hugues, Fran¸cois Goulette, Jean-Emmanuel Deschaud, Beatriz Marcotegui, and Yann LeGall (2018). “Semantic classification of 3D point clouds with multiscale spherical neighborhoods”. In: 2018 International conference on 3D vision (3DV). IEEE, pp. 390–398.

Van Kuppevelt, D, C Meijer, F Huber, A van der Ploeg, S Georgievska, and Vincent T van Hees (2020). “Mcfly: Automated deep learning on time series”. In: SoftwareX 12, p. 100548.

Vicari, Matheus B, Mathias Disney, Phil Wilkes, Andrew Burt, Kim Calders, and William Woodgate (2019). “Leaf and wood classification framework for terrestrial LiDAR point clouds”. In: Methods in Ecology and Evolution 10.5, pp. 680–694.

Wang, Di, Markus Hollaus, and Norbert Pfeifer (2017). “Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data”. In: ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4.

Wang, Lin, Zhigang Wang, and Shan Liu (2016). “An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm”. In: Expert Systems with Applications 43, pp. 237–249.

Wang, Peng, Liguo Wang, Henry Leung, and Gong Zhang (2020). “Super-resolution map- ping based on spatial–spectral correlation for spectral imagery”. In: IEEE Transactions on Geoscience and Remote Sensing 59.3, pp. 2256–2268.

Wang, Zhen, Liqiang Zhang, Tian Fang, P Takis Mathiopoulos, Xiaohua Tong, Huamin Qu, Zhiqiang Xiao, Fang Li, and Dong Chen (2014). “A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification”. In: IEEE Transactions on Geoscience and Remote Sensing 53.5, pp. 2409–2425.

Wang, Zhiguang, Weizhong Yan, and Tim Oates (2017). “Time series classification from scratch with deep neural networks: A strong baseline”. In: 2017 International joint conference on neural networks (IJCNN). IEEE, pp. 1578–1585.

Wei, Qiong and Roland L Dunbrack Jr (2013). “The role of balanced training and testing data sets for binary classifiers in bioinformatics”. In: PloS one 8.7, e67863.

Xi, Zhouxin, Chris Hopkinson, Stewart B Rood, and Derek R Peddle (2020). “See the forest and the trees: effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning”. In: ISPRS Journal of Photogrammetry and Remote Sensing 168, pp. 1–16.

Xu, Xiaowei, Yinrong Chen, Junfeng Zhang, Yu Chen, Prathik Anandhan, and Adhiyaman Manickam (2021). “A novel approach for scene classification from remote sensing im- ages using deep learning methods”. In: European Journal of Remote Sensing 54.sup2, pp. 383–395.

Zheng, Yi, Qi Liu, Enhong Chen, Yong Ge, and J Leon Zhao (2016). “Exploiting multi- channels deep convolutional neural networks for multivariate time series classification”. In: Frontiers of Computer Science 10.1, pp. 96–112.

Zhou, Bolei, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba (2016). “Learning deep features for discriminative localization”. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2921–2929.

Zhu, Xi, Andrew K Skidmore, Roshanak Darvishzadeh, K Olaf Niemann, Jing Liu, Yi- fang Shi, and Tiejun Wang (2018). “Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest”. In: International journal of applied earth observation and geoinformation 64, pp. 43–50.

Chapter4

A non-destructive approach to estimate buttress volume using 3Dpoint cloud data

Abstract

Buttressed trees provide mechanical support for themselves and offer essential ecological functions such as nutrient acquisition while having one of the largest sources of volumeor biomass estimation variation in tropical forests. To reduce this variation, we used the Alpha Shape Algorithm (ASA) and the Slice Triangulation (ST) method, based on 3D point clouds, to estimate the buttress volume of 30 trees. Also, we developed three allometric models: a diameter above the buttress-based (DAB) model, a diameter computed from non-convex area ($D_{area130}$) model, and the convex hull perimeter ($D_{convex130}$) at breast height model. The developed models were validated with independent data collected using a Terrestrial Laser Scanning (TLS) and destructive measurements to highlight the broader contextualization and application of these methods. Volume estimated by the ASA and ST showed a high agreement with the reference volume acquired using the Smalian formula (RRMSE of 0.07 and 0.11, respectively, regardless of species effect). The ASA was also robust when modeling trees with more and shallower horizontal buttresses. $D_{area130}$ was the most accurate predictor to estimate buttress volume, with a lower Akaike information criterion (AIC, -66.25) than DAB (-59.55) and

$D_{convex130}$ (30.56); however, the DAB (RRMSE of 0.23) and $D_{area130}$(RRMSE of 0.21) show similar performance when validated with in dependent datasets. Our results indicate that the ASA approach performs better than both the ST and the allometric models used in this paper. Furthermore, the ASA method can help to correct the bias in the present and past estimates of volume and biomass oflarge trees, which are foundational components to understanding biomass allocation and dynamics in tropical forests.

Keywords: Volume, Buttress, Alpha Shape, Diameter Above the Buttress, Point Clouds, Allometric Models

4.1 Introduction

Buttresses are large, wide roots on all sides of a shallowly rooted tree (Mehedi et al., 2012; Chapman et al., 1998) which prevent the trees from falling and balance the trees against unidirectional stresses such as asymmetrical canopies and prevailing winds (Chapman et al., 1998; Zhiyuan et al., 2013). Buttresses can act as barriers to matter flow, while also increasing the tree's contact area with the ground, resulting in higher litter accumulation, soil moisture, and nutrients (Pandey et al., 2011). Large trees are more likelyto have buttresses than small trees, especially in tropical forests (Zhiyuan et al., 2013). These large trees enhance the biodiversity of an area by providing a microenvironment where insects and other organisms can nest and seek shelter (Tang et al., 2011). Moreover, they fix a large amount of carbon due to their high wood volume, making them increasinglyof interest in climate adaptation research (Noelke et al., 2015).

The occurrence of buttresses has an enormous impact on the determination of the volume and biomass of large trees, leading to an error in the estimates of aboveground carbon in tropical forests (Nogueira et al., 2006). This error is due to the fact that measurements of diameter, basal area, and wood volume are consistently treating the trunk as a cylinder, thus assuming any that cross-section of the trunk is a circle (Cushman et al., 2014). Buttressed trees pose a particular challenge concerning the measurement of DBH. Many tree buttresses extend well above the standard breast height, which means trunks are not cylindrical at 1.3 m (Cushman et al., 2014). Therefore, assuming the trunk is a cylinder would lead to overestimation in diameter and volume measurement when it comes to buttressed trees (Clark, 2002; Nogueira et al., 2006).

Since buttressed trees have such complex structure, several methods have been developed to investigate their volume and biomass (Cushman et al., 2014; Cushman et al., 2014). The diameter above the buttress (DAB), where the stem reaches a relatively regular circle is one of the most widely used variables to estimate the wood volume of buttressed trees (Bauwens et al., 2017). There is no fixed definition of the height at which DAB should be measured. Generally, moving the point of measurement up 0.30 m, or 0.50 m above the buttresses is accepted in many studies (Cushman et al., 2014; Newbery et al., 2009; Noelke et al., 2015).

Inconsistencies and uncertainty introduced on tree volume and biomass estimations are in general due to the variable methods of measuring DAB. In general, field approaches often underestimate volume and biomass since they often use a lower value than the DBH (due to tree stem diameter decreasing with height) (Cushman et al., 2021). This underestimation has been accounted by using a scale factor and the non-convex area and convex hull perimeter of buttresses at breast height (Bauwens et al., 2017; Noelke et al., 2015). Furthermore, the diameter (Darea130) , derived from the non-convex area of buttresses at 1.3 m has been identified as a more accurate predictor of biomass when compared to the DAB approach, and the diameter calculated from a circular disc with the same perimeter as convex hull perimeter at 1.3 m (Dconvex130) (Bauwens et al., 2017).

Current methods for estimating volume and biomass of buttressed trees are labor intensive, expensive, slow, and under perform with a high degree of variability (Cushman et al., 2021; G. J. Newnham et al., 2015). These methods can be divided into destructive and non-destructive methods. Destructive methods include calculations of volume and biomass based on cross sections of logged trees, which are dried and weighted (Dean et al., 2006; Nogueira et al., 2006). These approaches, while accurate and used to calibrate allometric models, are not often conducted because of their cost, instrumentation issues, and harvesting restrictions (Calders et al., 2015; Chapman et al., 1998; Zhiyuan et al., 2013).

Non-destructive approaches can be divided into three categories: 1) wire methods (Ngomanda et al., 2012), 2) convex methods (Dean et al., 2006), and 3) terrestrial photogrammetry (TP) or laser scanning (TLS) methods (Calders et al., 2020; Cushman et al., 2021). Both the wire and convex methods produce inconsistent results and are slow due to the labor intensity. Concerning wire method, two wires are pressed against buttress to model the cross-section, while the convex–concave method need measure all convex (spurs) and concave parts (flutes) of the stem to model the stem. TP and TLS are more recent ways to deal with the buttress estimation problem as they can both retrieve a high accuracy 3D point cloud (Calders et al., 2015; Calders et al., 2020; Cushman et al., 2021).

With the increasing availability of point clouds, methods are developed to convert point clouds to volume. Quantitative Structure Model (QSM) is a recently used method to estimate tree volume (Raumonen et al., 2013). QSM assumes the tree is composed of cylinders with different diameter from ground to crown. This assumption meets problems when it comes to buttressed trees. For instance, Gonzalez de Tanago et al. (2018) used quantitative structure models (QSM) to estimate biomass of buttresses based on TLS point clouds, indicating QSM cylinders were unable to capture the detailed structure of buttresses, causing higher errors on the estimation of buttress volumes. Thus, a method which can handle the volume estimate of buttressed trees is needed to eliminate variations for aboveground carbon estimation.

Using a triangulation method might be a better option than cylinders to model buttresses (M. I. Disney et al., 2018). Meanwhile, Alpha Shape Algorithm is one of the most robust triangulation algorithms that can retain the most original surface features (Bonneau et al., 2019; Edelsbrunner et al., 1994). Though the alpha shape algorithm has been used to estimate crown attributes (Hadas et al., 2017) or canopy volume of the orange grove (Cola ̧co et al., 2017), its application on buttressed trees is not well studied (Bauwens et al., 2017; Cushman et al,. 2021, 2014; Nölke et al., 2015).

In this paper, we address some of the knowledge gaps in buttresses modeling indicated above to produce a consistent, accurate, and automated method to estimate tree buttress's volume and biomass. Here, in order to estimate buttress volumes, we test the Alpha Shape Algorithm (ASA), and a Slice Triangulation (ST) algorithm, two different Delaunay Triangulation (DT) methods for 3D point cloud surface reconstruction (Bonneau et al., 2019; M. I. Disney et al., 2018; Edelsbrunner et al., 1994; Su et al., 2020). Additionally, we develop allometric models for buttress volume estimation, and then identify the most accurate predictor to estimate buttresses volume from DAB, the diameters derived from the non-convex area (Darea130), and convex hull perimeter (Dconvex130) of buttressed trees at breast height. We used buttressed trees from three databases distributed around the world to reach the mentioned two goals.

4.2 Materials and Methods

4.2.1 Data sources

We used three databases of buttressed trees collected at different locations around the world to conduct this research: (1) the Yangambi Reserve (YR), Democratic Republic of Congo, (2) Bogor Botanical Garden (BBG), Indonesia. (3) Santiago de Puriscal (SP), Costa Rica (Table 4.1).

Table 4. 1: General description of two databases used in this research.

Database	Acronym	Location	Methods	n	Species
The Yangambi Reserve, Democratic Republic of Congo	YR	0° 46' 3" N 24° 26' 29" E	Photogrammetry	30	2
			Destructive	3	2
Bogor Botanical Garden, Indonesia	BBG	6°35' 51" S 106°47' 54" E	Terrestrial Laser Scanning	12	6
Santiago de Puriscal, Costa Rica	SP	9° 49' 55" N 84° 19' 60" W	Terrestrial Laser Scanning	6	6

The YR database consists of 30 point clouds of buttressed trees (Tree 1 to Tree 30 in Table S1), including two species: Celtis mildbraedii and Entandophragma cylindricum, generated using terrestrial photogrammetry (Bauwens et al., 2017). Also, YR database has extra three buttressed trees that were destructively measured (see Table S2). The BBG database has 12 buttressed trees (six species) generated using a multi-scan TLS method (Nölke et al., 2015). Here, 30 trees from YR database were used for the alpha shape and the slice triangulation reconstruction process. Of the 45-point clouds used from YR and BBG database, 36 trees (Table S1) containing four species were used for allometric model training, and nine (Table S2) including six species for testing. We applied six-fold cross-validation strategy on training data to assess model performance. The SP database consists of six point clouds of buttressed trees created by a multi-scan TLS method. We used those six trees to validate the alpha shape algorithm on buttress reconstruction.

4.2.2 Data preprocessing

Two preprocessing steps are used to clean the point clouds. First, a statistical outlier removal (SOR) filter was applied to remove the noisy points that are far away from the main trunk in Cloud Compare (version 2.11, Cloud Compare, GPL software). The following parameters were used to run the algorithm: the number of points used for mean distance estimation (nPoints), and the standard deviation multiplier threshold used to filter the points (nSigma). Here, we used the default setting (nPoints =6, nSigma = 1) in Cloud Compare. Second, the point cloud was downsampled to 0.01 meter to ensure uniform data distribution and improving computational efficiency (Guzman Q et al., 2020).

4.2.3 Slice Triangulation and Alpha Shape Algorithm

Surface reconstruction algorithms seek to construct a complete surface from a 3D point cloud (Su et al., 2020). The Delaunay Triangulation (DT) is one of the most widely used surface reconstruction algorithms, since it utilizes the simplest polygon meshes for a given feature, thereby improving computational efficiency in 3D modeling applications (Delaunay et al., 1934; Wang et al., 2019). Based on the DT, Slice Triangulation (ST) method fit the boundary curves of horizontal cross-sections of a given trunk, and then computes its volume (M. I. Disney et al., 2018; Raumonen et al., 2013). The ST algorithm separates the point cloud in continuous thin cross-sections, and uses curves (line segment) to fit those cross-sections. The vertices of the curves in continuous layers are then connected into a

uniform triangulated surface. The bottom and top planes are subsequently triangulated to close the surface. Finally, the buttress volume is calculated using the divergence theorem (Pfeffer, 1986). which uses the outward surface normal and area of triangles.

The ST method can meet errors in sloped areas since it uses a horizontal plane to model buttresses (Raumonen et al., 2013). Therefore, we applied a height transformation strategy to reduce the bias on the ST method. Specifically, the ground level model can be seen as a function G(x, y) = z, which gives the z-coordinate of the ground level for each (x, y). Also, the lowest z-coordinate of the vertices of the ground model is z0. Then a point Q = (x, y, z) is mapped to a point Q1 = (x, y, z – G (x, y) + z0), making the transformed point cloud at the ground level flat with a constant z coordinate equal to z0. Next, we apply the ST to produce the buttress triangulation model. Finally, we map the triangulation model back to the original coordinates and compute the buttress volume.

The Alpha Shape Algorithm (ASA) can use the DT to describe the shape of a limited number of points in a set with a high accuracy (Edelsbrunner et al., 1994). Moreover, the ASA can keep a balance between hole-filling and loss of detail, and therefore it is often used for surface reconstruction (Bonneau et al., 2019). Concerning the ASA, the output of the final surface shape is mainly affected by the parameter α (Hadas et al., 2017). The alpha value represents the level of refinement for a given set of points. With a small value of α, the shape reverts to the original point sets, while a large value of α often means the shape of a convex polygon (Vauhkonen et al., 2012).

Here, to obtain the accurate buttress volume, we selected the smallest α value when a watertight manifold mesh is produced. A mesh is watertight if there are no surface holes on the mesh (Bonneau et al., 2019), and this step can be visually inspected in R Studio. Also, we explore the relationship between α and DAB to determine whether DAB can be used as a reference to select α for buttresses volume estimation.

4.2.4 Structural complexity

Previous studies often use the fractal dimension to describe the structure complexity of a given forest's point (Guzman Q et al., 2020; Seidel, 2018); however, the fractal dimension mainly depends on how trees distribute their branches and leaves in 3D space (Guzman Q et al., 2020) making them insufficient for high structural complexity due to their buttresses. C. mildbraedii tends to generate more, higher, and shallower buttresses than E. cylindricum (Bauwens et al., 2017). As such, we

calculated the area-to-volume ratio for the above two species to define their respective structure complexity. If the tree presents more and shallower horizontal buttresses, it will mean that the tree tends to have a higher complexity. Here, we used a two-sample t-test to compare the mean value of α and area-to-volume ratio between C. mildbraedii and E. cylindricum (Muff et al., 2021). Also, we calculated the area-to-volume ratio of buttressed trees from SP database to define their structure complexity. Furthermore, we investigate whether there is a linear relationship between α and structure complexity.

4.2.5 Scale factor

The scale factor is commonly used metrics to estimate buttresses volume. Here, we used the same metric used in Bauwens et al. (2017) and Noelke et al. (2015) to define the proportion of buttress volume that is not considered when the volume is calculated as a cylinder with a diameter equal to DAB (Eqn. 4.1):

$$f = 1 - \frac{\pi DAB^2 H_{DAB}}{4V} \tag{4.1}$$

where V is the buttresses volume (m^3), and HDAB (m) is the measurement height of DAB. Three different volumes were calculated with Va as the volume estimated by the alpha shape algorithm, Vt as the volume estimated by slice triangulation and Vb as the buttresses volume calculated by Smalian formula. A two-sample t-test to find out whether there was any evidence that mean f value was different between alpha shape volume (f a), slice triangulation volume (f t) and the actual volume (f b).

The implementation of the ASA on buttress modeling was done using the R package rTLS (Guzman Q et al., 2021; Lafarge et al., 2014). In this package, the parameter of max.height in trunk volume() can extract all the points in the cloud lower than a given height, which provides an efficient way to obtain buttresses volume under the measurement height of DAB (HDAB). ST was applied in MATLAB (Raumonen et al., 2013).

4.2.6 Volume predictors

To define the most accurate predictor for volume estimation of the buttressed trees, we fitted functions to describe the relationships between the actual volume and the following predictors: DAB, Darea130, and Dconvex130. A log-transformation strategy was applied to meet the assumption of normality and homoscedasticity (Barbeito et al., 2017). The equation for calculating the buttress volume was (Eqn. 4.2):

$$ln(V_b) = a\ln(D^2 \times H_{DAB}) + b \quad (4.2)$$

Where D is one of the diameters mentioned above, and a and b are model parameters, and ln is the logarithm transformation.

A correction factor (CF) was used to correct the systematic bias generated by the log-transformation, when back-transforming the calculation into volume (Basuki et al., 2009; Sprugel, 1983). The equation was calculated as follows (Eqn. 3):

$$CF = exp^{(MSE^2/2)} \quad (4.3)$$

Where MSE is the mean squared error of the line fitted by the natural logarithm. We used six-fold cross-validation to get a robust idea of the error of our regression model (avoid overfitting). Then, the final model was trained on the whole training data. Also, we used the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to determine the best predictor fitted on the training data.

4.2.7 Validation

The absolute error of volume estimation using allometric equations is hard to calculate without destructive harvesting methods (Cushman et al., 2021); however, the YR database included measurements from three harvested trees, allowing for an assessment of absolute error in the model. To illustrate the broader application of our method, we validated the diameter-based models on these three harvested trees to obtain the absolute error of their volume estimates. Additionally, six TLS

measured trees were included for validation. The detailed information for those nine trees is provided in Table S4.2.

4.2.8 Modeling buttresses from SP database

The understory vegetation in dense tropical forests can obscure the target tree in point clouds even if a multi-scan strategy is applied (Calders et al., 2015). Meanwhile, a buttress structure can be diverse (e.g., more well-extended and shallow buttresses) because the buttress shape can be affected by local environment. Here, the YR database collected the point clouds of buttressed tree in such a way that all understory vegetation and small lianas up to 2 m high around the target tree were cleared making it our low-structural complexity control (Bauwens et al., 2017). By contrast, the SP database consists of six buttressed trees without pre-cleaning work. Also, those trees tend to have more shallow and well-extended buttresses. We use trees from SP database to test the effectiveness of the ASA method on complex buttresses in highly obscured forests.

4.3 Results

4.3.1 The relationship between the alpha and DAB, and the alpha and structure complexity

To determine the alpha value to estimate a buttress' volume, we explore the relationship between the alpha value and DAB, and the alpha value and structure complexity (Figure 4.1). There was significant difference in mean alpha value (p-value < 0.001) between E. cylindricum and C. mildbraedii. Also, we found a strong relationship (R2 = 0.89) between the alpha value and DAB for E. cylindricum and C. mildbraedii, while the relationship between the alpha value and structure complexity is lower (R2 = 0.81). The alpha value increased proportionally to DAB, while the DABs of E. cylindricum trees are higher than the C. mildbraedii (Figure 4.1a). By contrast, the alpha value was negative correlated with structure com- plexity according to Figure 4.1b. Moreover, the DABs of most C. mildbraedii trees were lower than 1.0 m (0 in the x-axis of Figure 4.1a) with an alpha value lower than 0.5.

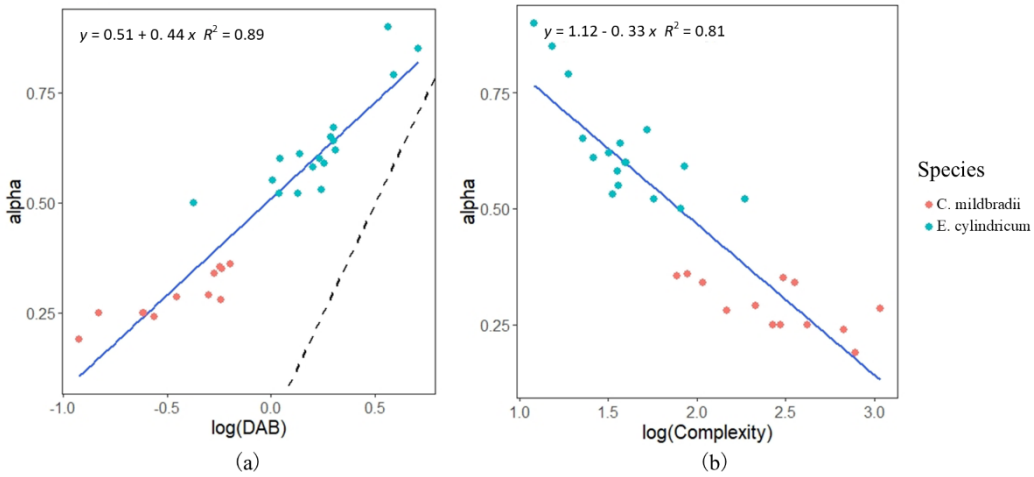


Figure 4. 1: The relationship between the diameter at breast height (DAB) and alpha (a), structure complexity (area-to-volume ratio) and alpha (b).

4.3.2 Slice Triangulation vs. Alpha Shape Algorithm

Overall, we find that the ASA and ST perform similarly. The ASA tends to have a lower relative RMSE than the ST method when the volume of both species is estimated (Table 4.2), 0.07 versus 0.11 respectively. The differences are not statistically significant (p-value = 0.60), In addition, both the ASA and the ST methods generate a larger RRMSE of C. mildbraedii than for E. cylindricum (0.15 vs. 0.05 and 0.15 vs 0.08, respectively). Moreover, E. cylindricum (10.06 m^3) generated five times larger volume than C. mildbraedii (2.12 m^3), while the ST (0.08) produced a larger RMSE on E. cylindricum than the ASA (0.05).

Table 4. 2: Results of regression analysis of reference volume (Vb) and V predicted by slice triangulation (Vt) and alpha shape algorithm (Va), respectively. RMSE: root mean squared error; RRMSE: relative RMSE, refers to the RMSE divided by mean values. (/, p-value > 0.05; *, 0.05 < p-value < 0.01; **, 0.01 < p-value < 0.001; ***, p-value < 0.001).

Method	Species	V	Intercept	RMSE (m^3)	Mean (Vb)	RRMSE
Alpha Shape Algorithm	E. cylindricum	***	.	0.5	10.06	0.05
	C.mildbraedii	***	.	0.32	2.12	0.15
	Both	***	*	0.49	6.62	0.07
Slice Triangulation	E. cylindricum	***	**	0.82	10.06	0.08
	C.mildbraedii	***	.	0.32	2.12	0.15
	Both	***	*	0.74	6.62	0.11

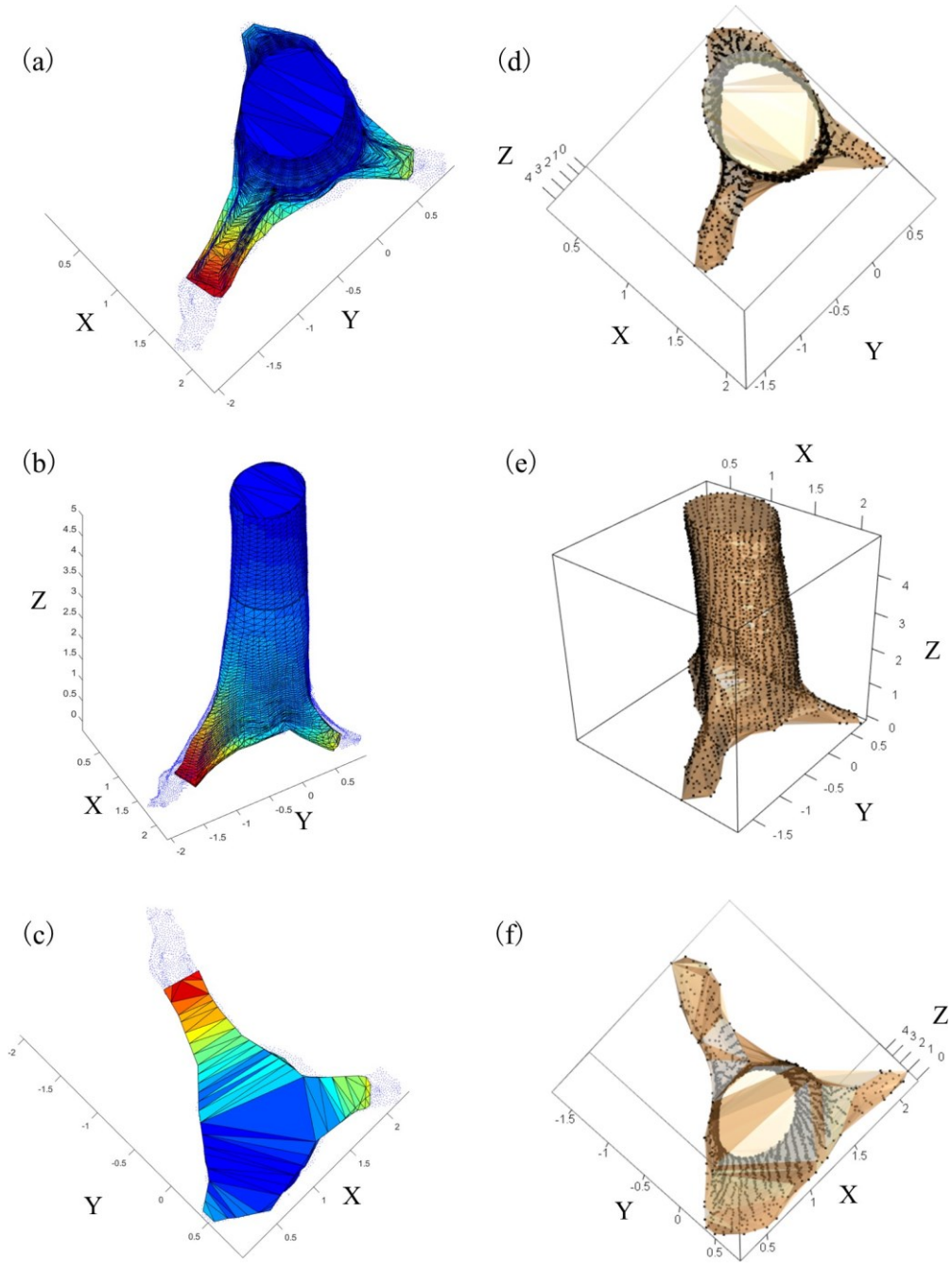


Figure 4. 2: The results of final buttress shape from the slice triangulation (ST) the alpha shape algorithm (ASA) for Tree 1, (a) Top view using ST; (b) Side view using ST; (c) Bottom view using ST; (d) Top view using ASA; (e) Side view using ASA; (f) Bottom view using ASA.

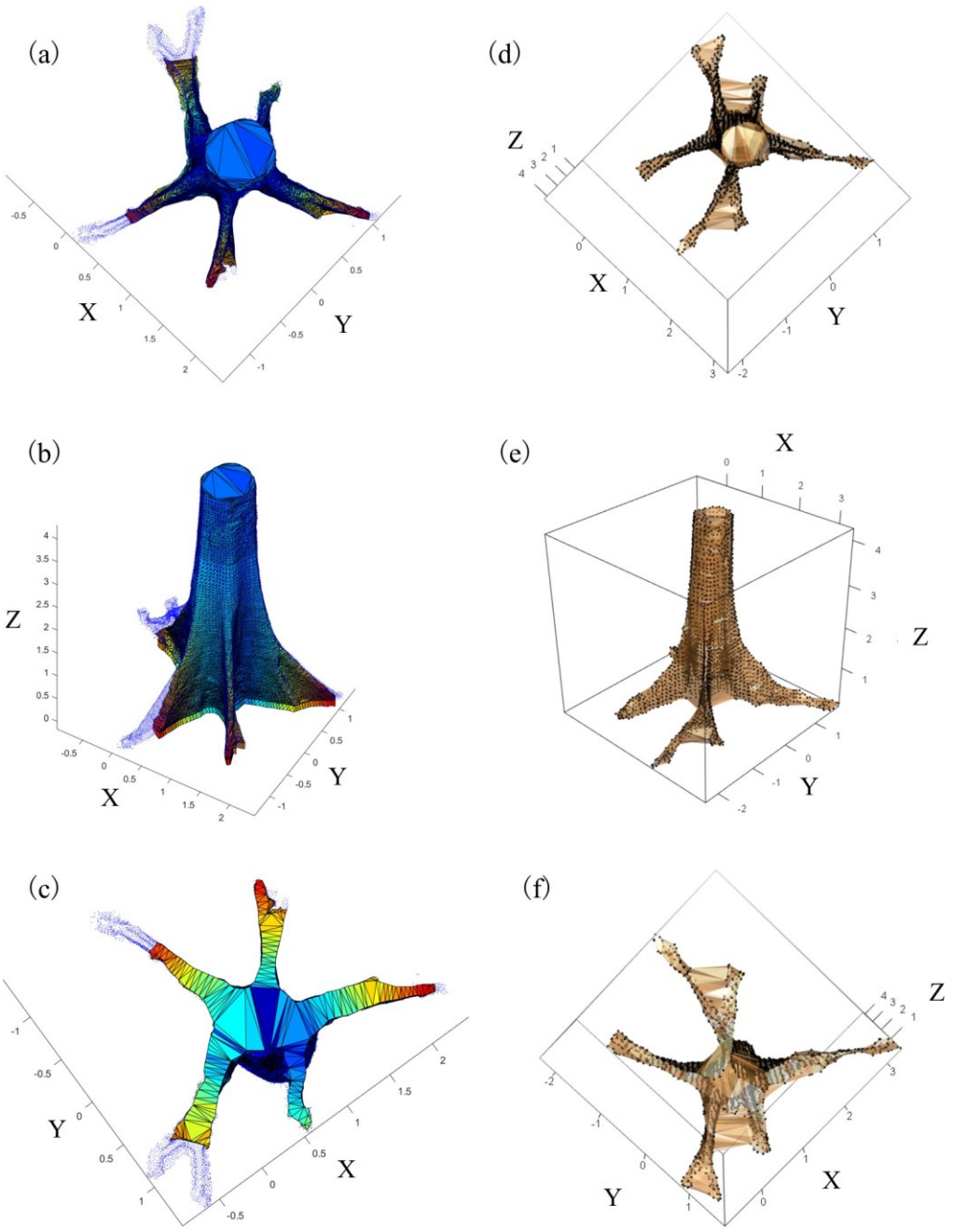


Figure 4. 3: The results of final buttress shape from the slice triangulation (ST) the alpha shape algorithm (ASA) for Tree 27, (a) Top view using ST; (b) Side view using ST; (c) Bottom view using ST; (d) Top view using ASA; (e) Side view using ASA; (f) Bottom view using ASA.

Figures 4.1 and 4.2 present the modeling results using the ASA and ST approaches on Tree 1 (E. cylindricum) and Tree 27 (C. mildbraedii) from Table S4.1. The ST method cannot capture the true shape of Tree 1 (Figure 4.2 a-c) and Tree 27 (Figure 4.3 a-c) when the length of buttress is longer. By contrast, the ASA creates better buttress model on Tree 1 (Figure 4.2 d-f); however, when the tree presents more and shallower horizontal buttresses, the ASA tends to generate more overlap areas (i.e., Tree 27; Figure 4.3 d-f) leading to an overestimation of the C. mildbraedii volume (also see Figure S4.1).

When the ASA (f a) and the ST (f t) volume are compared to the reference volume (f b), there is no evidence that their mean value is different in either species (see supplementary materials; C. mildbraedii: f a vs. f b, p=0.11, f t vs. fb, p=0.56; E. cylindricum: f a vs. f b, p=0.14, f t vs. f b, p=0.22). However, there is strong evidence that mean f a is different between the two species (p-value = 0.003). By contrast, we find no evidence that mean f b (p = 0.48), or f t (p = 0.56) is different between the two species (Figure S4.2).

4.3.3 Volume estimation using the standard predictor

Six buttressed trees collected using TLS from the BBG database were included on the development a more general allometric model for volume estimation. Therefore, we used 36 trees for model training and nine trees for validation. The six-fold cross-validation errors for Darea130, DAB, Dconvex130 models on the training data were given in Tables S4.4, S4.5 and S4.6, showing Darea130 and DAB models fitted well with the training data. Figure 4.3 indicates the final allometric models fitted on the whole training data. The allometric models developed with either Darea130 or DAB presented an identical R2 value of 0.99, while this value for Dconvex130 was 0.89.

Of the three models fit to the 36 buttressed trees, the model derived from Darea130 pro- vided the best fit with the lowest Akaike Information Criterion (AIC, -66.25) and Bayesian information criterion (BIC, -61.50) (Table 4.3). Meanwhile, the RRMSE of DAB model (0.11) was similar with the Darea130 model (0.08). Also, we found no systematic over or under-estimation of volume estimates given by DAB model (a slope coefficient of 1.00) from Figure 4.4(c). The Darea130 model was more likely to overestimate the volume with a slope of 0.91. Dconvex130 may not be a good predictor for volume estimation since this model generated the highest RMSE and RRMSE.

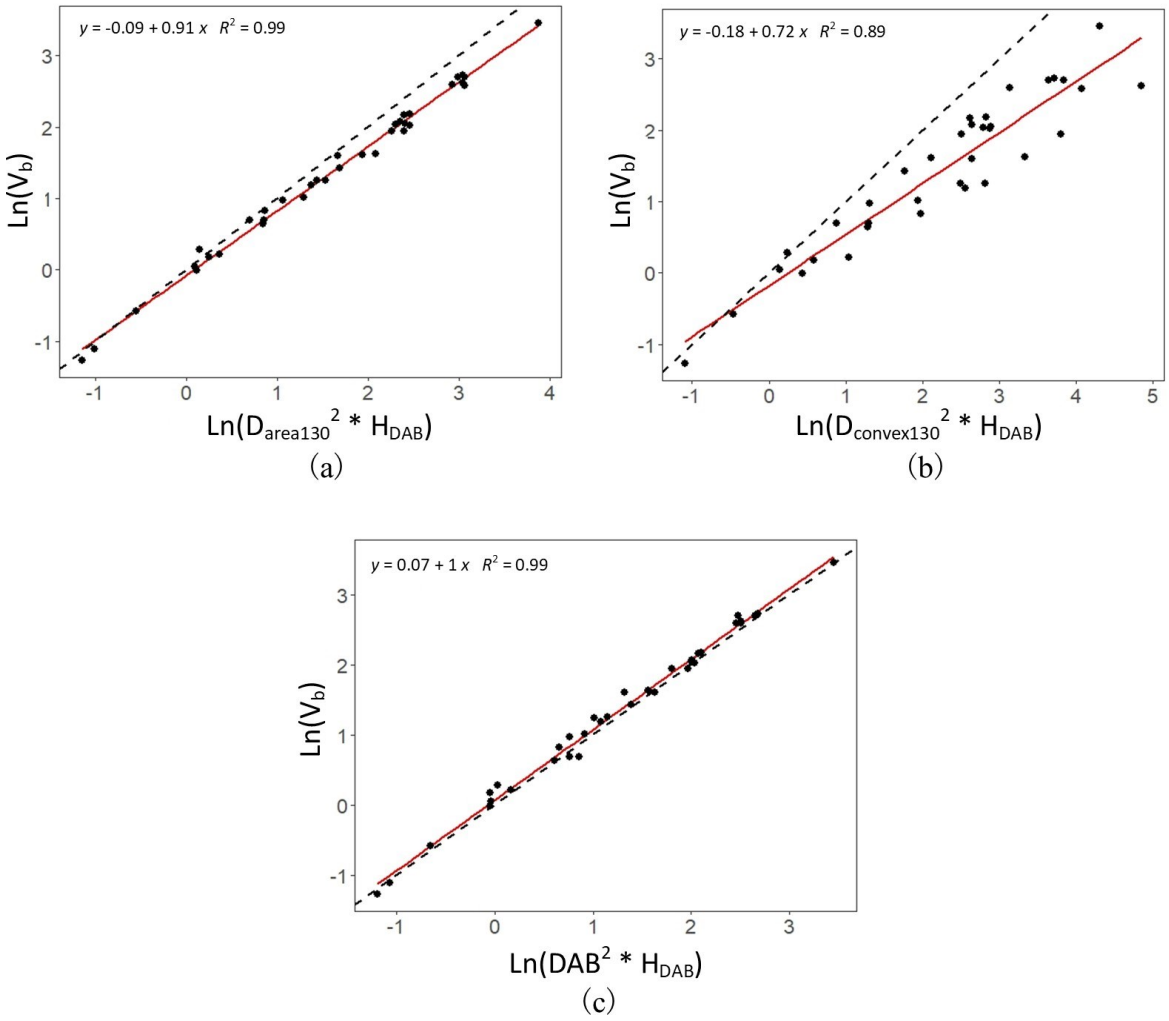


Figure 4. 4: (a) The relationship between Darea130 and reference volume; (b) relationship between Dconvex130 and reference volume; (c) relationship between DAB and reference volume.

Table 4. 3: Results of the regression analysis of the volume estimation equations (Eqn 2.2); D.f., degrees of freedom; AIC, Akaike information criterion; BIC: Bayesian information criterion; RMSE: root mean squared error after back transforming the results, refers to volume error (m3); RRMSE: relative RMSE.

Predictor	D_2	Intercept	D.f.	AIC	BIC	RMSE	RRMSE
$D_{area130}$	***	**	34	-66.25	-61.50	0.52	0.08
$D_{convex130}$	***	/	33	30.56	35.22	3.92	0.60
DAB	***	**	34	-59.55	-54.79	0.67	0.11

4.3.4 The allometric model validation

We evaluated the performance of the diameter-based models on three destructively harvested trees and six trees scanned with a TLS from the YR and BBG databases, respectively. Overall, DAB (0.21) and Darea130 (0.23) models generated similar RRMSEs on the validation data, which outperformed the Dconvex130 model to estimate volume (RRMSE = 0.80) (Table 4.4). Moreover, the above three models built with different diameters, showed similar performance on destructively sampled data and the TLS data. Concerning the TLS trees, T1-BBG generated the largest relative error among the six trees on Darea130 (0.24) and DAB (0.31) models. By contrast, Dconvex130 and DAB models provided the best fit on T6-BBG, with the lowest buttress volumes (1.72 m3) of all six trees. The overall RRMSE of TLS trees in Darea130 model (0.19) was slightly lower than DAB model (0.21). In terms of destructively harvested trees, the relative error of Darea130 model on all trees was larger than DAB model, except for T1-YR (0.14 vs. 0.20). Also, the DAB and Darea130 models showed a similar RRMSE (0.21 vs. 0.19). In addition, the f b value (Eqn. 2.1) of the 45 trees from YR and BBG databases was 0.28 (±0.09) (See supplementary materials), meaning the cylindrical volume calculated from DAB underestimated 28% of all 45 buttress volumes.

Table 4. 4: The validation results of different models on six TLS trees and three destructive harvest trees; the results of destructive measurement trees were italicized. Vp: predicted values of volume; RE: relative error; RRMSETLS: relative RMSE of TLS trees; RRMSEDes: relative RMSE of destructive harvest trees.

No.	V_b (m³)	$D_{area130}$		$D_{convex130}$		DAB	
		V_p (m³)	RE	V_p (m³)	RE	V_p (m³)	RE
T1-BBG	7.45	5.68	0.24	4.09	0.45	5.17	0.31
T2-BBG	7.97	9.67	0.21	15.78	0.98	5.91	0.26
T3-BBG	2.67	2.76	0.03	3.19	0.19	2.86	0.07
T4-BBG	20.5	19.28	0.06	19.66	0.04	21.64	0.06
T5-BBG	17.41	20.98	0.21	31.89	0.83	13.75	0.21
T6-BBG	1.72	1.70	0.01	1.67	0.03	1.70	0.01
T1-YR	*3.07*	*3.50*	*0.14*	*5.11*	*0.66*	*3.69*	*0.20*
T2-YR	*1.27*	*1.73*	*0.36*	*2.28*	*0.80*	*1.16*	*0.09*
T3-YR	*1.66*	*2.00*	*0.20*	*2.84*	*0.71*	*1.83*	*0.11*
RRMSE		0.21		0.80		0.23	
$RRMSE_{TLS}$		0.19		0.71		0.21	
$RRMSE_{Des}$		*0.21*		*0.74*		*0.19*	

4.3.5 Model buttressed tree from SP database

Figure 4.5 compares the mean area-to-volume ratio derived from the E. cylindricum, C. mildbraedii trees and the SP database. There is very strong evidence that the mean area-to- volume ratio is different between the E. cylindricum and the C. mildbraedii ($p < 0.001$). Moreover, we find no evidence that the mean area-to-volume ratio is different between the C. mildbraedii and the SP database, which indicate that buttressed trees from the SP database also have a higher structural complexity. Figure 4.6 gives two examples of the buttressed trees from the SP database. Compare with tree 27 in Figure 4.3, the trees in Figure 4.6 presented more and shallower horizontal buttresses. Visual inspection of Figure 4.6 indicates that the ASA could capture all buttresses, though the ASA tends to generate some overlap areas when the neighboring buttresses are close.

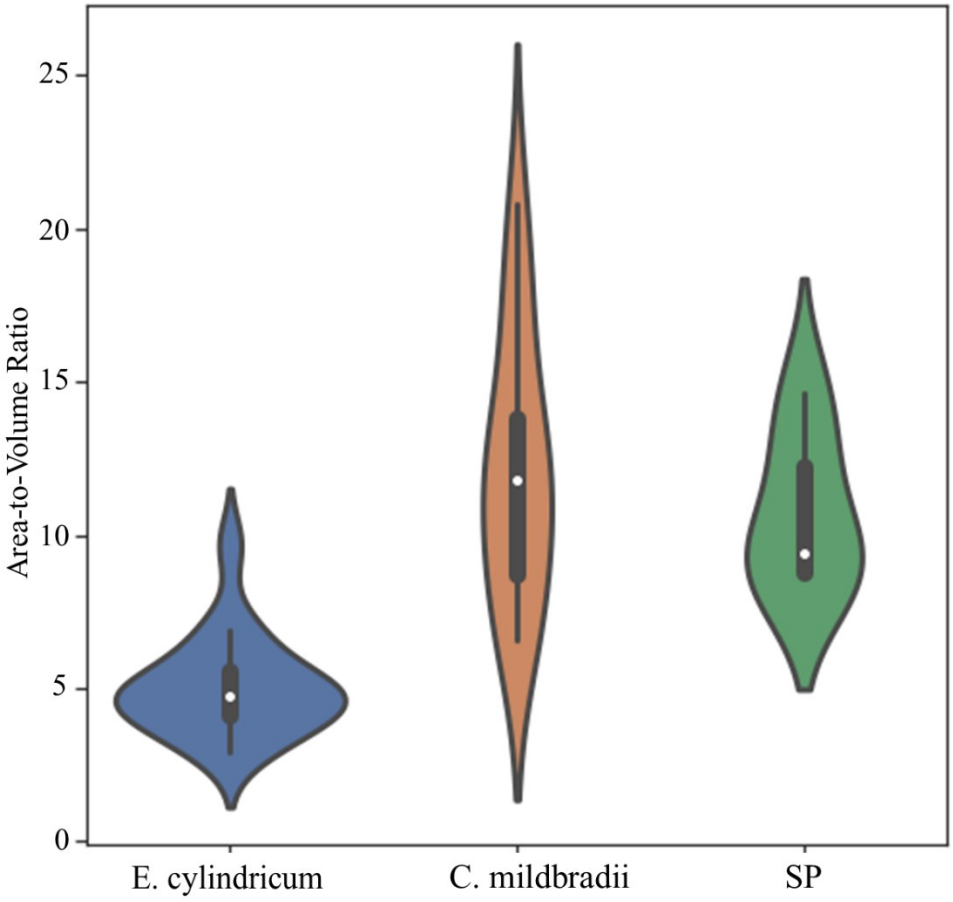


Figure 4. 5: Violin plot comparing mean area-to-volume ratio derived from E. cylindricum and C. mildbraedii (YR database) and SP database. YR database, the buttressed tree collected from Bauwens et al. (2017). SP database, the buttressed trees collected by this research.

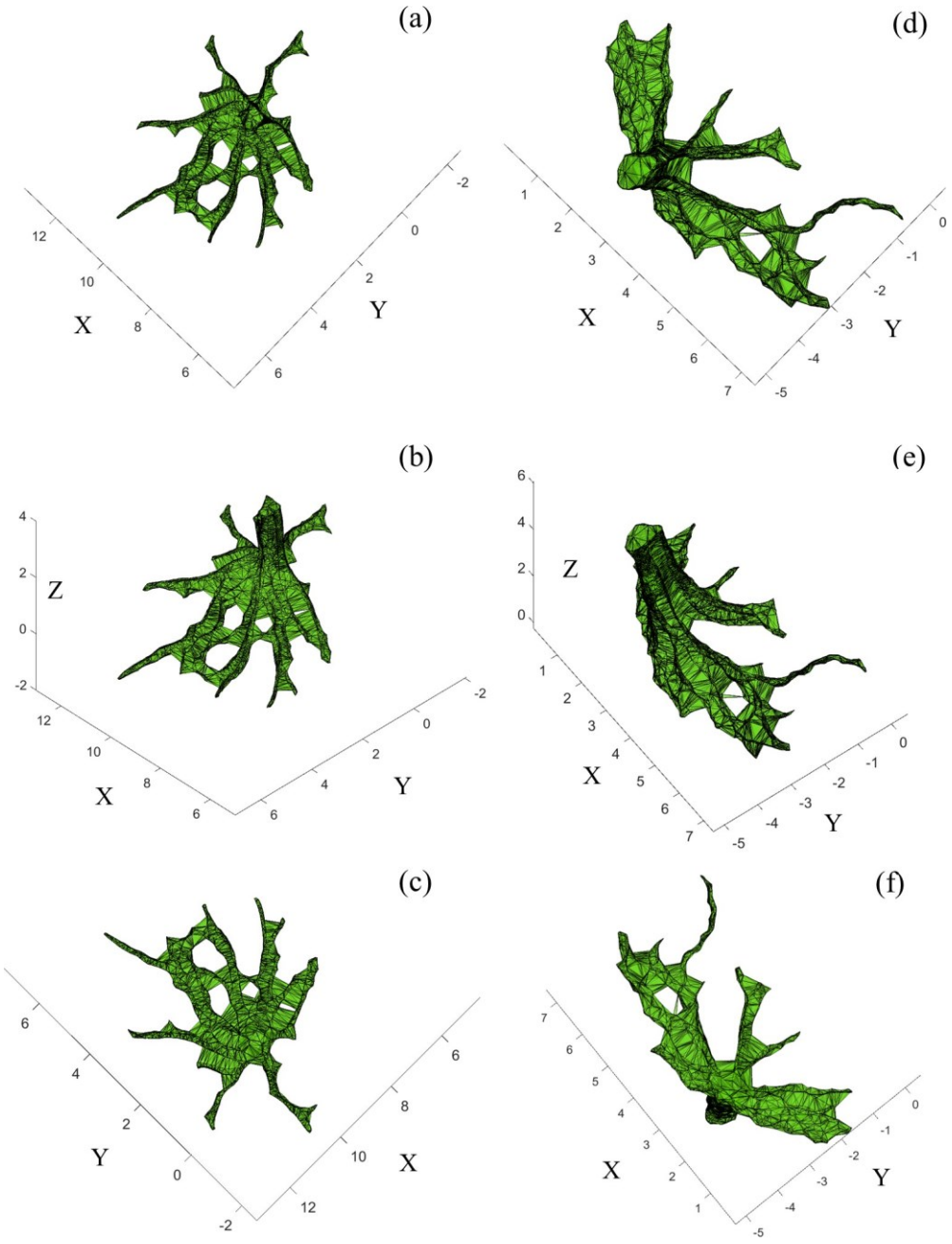


Figure 4. 6: The predictions of alpha shape algorithm (ASA) on trees from the Santiago de Puriscal,Costa Rica (SP database), (a) Top view of Tree 1 from the SP database; (b) Side view of Tree 1 from the SP database; (c) Bottom view of Tree 1 from the SP database; (d) Top view of Tree 2 from SP database; (e) Side view of Tree 2 from the SP database; (f) Bottom view of Tree 2 from the SP database.

4.4 Discussion

In this paper, we test the performance of the alpha shape algorithm (ASA) and slice triangulation (ST) on buttress volume estimation. Overall, the ASA and ST method tend to have better volume estimation than allometric models. Meanwhile, we demonstrate the ASA is a more applicable method than the ST for buttress volume estimation.

4.4.1 Alpha shape algorithm vs. Slice triangulation

Large trees exhibit higher variability in volume and biomass estimates, especially in tropical areas, where the trees show considerable stem irregularities, like buttressed trees (Noelke et al., 2015). The problems in estimating volume and biomass posed by buttressed trees have been widely reported but poorly understood (Cushman et al., 2014; Cushman et al., 2021). To eliminate this bias, Gonzalez de Tanago et al. (2018) used QSM cylinders to reconstruct tropical trees, including small trees (DBH < 70 cm) and large trees (DBH > 70 cm), where the latter always had buttresses. The QSM approach showed a high agreement with the reference data for the small trees, while it shown higher residuals in large trees. As such, we presented the utilization of the ASA and ST methods to estimate buttress volume from 30 point clouds from YR database. Our study has improved on previous methodologies by showing that the ASA and the ST perform similarly on volume estimation, and the ASA is a more applicable and transferable method than ST. Though the ASA has been shown to be good application in canopy architecture reconstruction (Cola¸co et al., 2017; Hadas et al., 2017), this is the first study to apply this model to buttressed trees.

It should be noted that the ASA method tends to overestimate the buttress volume, when the tree presents more and shallower horizontal buttresses (Figure 4.3d and Figure S4.1). However, the ASA can capture the true shape given the less and wider buttresses (Figure 4.2d). This is the reason why mean f a is statistically different between C. mildbraedii and E. cylindricum. The ST method cannot capture the true shape of longer horizontal buttresses (Figures 4.2a and 4.3a). Therefore, with larger volume in the E. cylindricum trees, the ST creates a larger error than the ASA. Furthermore, the buttresses' structure can be very complex due to the affection by local forest environments. The trees from the SP database do give 6 examples of how the buttresses can usually found in tropical forests. Even the ASA tends to generate many overlap areas between the neighboring buttresses (Figure 4.6),

the ASA does provide a solution to estimate the volume of those buttressed trees efficiently.

A key issue of applying the ASA to buttress modeling was the determination of the α, since the outputs of the alpha shape depend on the alpha value (Hadas et al., 2017). Previous studies often use the trial-and-error method to determine the value of α, and this is time-consuming when the sample size is larger (Cola¸co et al., 2017; Vauhkonen et al., 2012). Here, we found a stronger relationship between the DAB and the α, and this could help future researchers use the ASA to model buttress more efficiently.

4.4.2 Volume estimated by the diameter-based allometric models

In terms of the diameter-based allometric models for volume estimation, our study indicated that Darea130 model had the lowest AIC. This finding was the same as Bauwens et al. (2017), which demonstrated the biomass allometric model with Darea130 had the lowest AIC, compared to DAB and Dconvex130. However, no independent data was available in Bauwens et al. (2017) to assess the performance of the trained model. Concerning the testing process in this study, the DAB and Darea130 showed similar performance, more buttressed trees may be needed to determine the more accurate predictor for volume estimates. Though there are a limited number of buttressed trees that are scanned with TLS and destructively harvested trees for validation, the similar performance of allometric models on them does suggest that TLS is a reliable alternative method to destructive measurement without loss of accuracy, supported by Calders et al. (2015) and Calders et al. (2020) and M. I. Disney et al. (2018).

The allometric models are widely used for volume or biomass estimation in forestry, while large trees were often underrepresented in the calibration of such models (Clark, 2002; G. J. Newnham et al., 2015). The utilization of tree metrics like DAB or a scale factor (f index here) was recommended to estimate buttress volume (Ngomanda et al., 2012; Noelke et al., 2015). In this study, we not only built the allometric models with DAB and the diameter derived from the non-convex area and convex hull perimeter of buttressed trees at breast height, but we also calculated the scale factor, which is the proportion of the underestimation of buttresses volume, when DAB was used for cylindrical volume modeling. The RRMSE of the DAB or Darea130 based model in this study ranged from 0.21 to 0.23 on independent data. Since we only used nine buttressed trees for validation, we expect more studies in the future will assess the performance of our proposed allometric models. The mean f b (0.28±0.09) of all 45 trees indicated that around 28% of buttresses volume were not considered when the volume was

estimated with DAB. This value was similar with Bauwens et al. (2017) of 0.26, and lower than Noelke et al. (2015) of 0.35, while we used a total of 45 buttressed trees to obtain this underestimation factor, more than in the two mentioned studies.

4.4.3 The utilization of 3D point cloud data

Volume estimation using allometric models may not be a reliable indicator of volume for buttressed trees, compared with the ASA and the ST methods on 3D point cloud data. Here, the ASA and the ST methods generated lower RRMSEs (0.07 and 0.11, respectively) than the allometric models (0.21 for Darea130; 0.23 for DAB) on validation data. A similar issue was addressed by Calders et al. (2015),Kankare et al. (2013), and Lau et al. (2019). For example, Calders et al. (2015) identified that biomass estimates from the 3D point cloud with the QSM reconstruction approach showed a higher agreement (concordance correlation coefficient (CCC) of 0.98) with the reference data than the allometric models (CCC = 0.68-0.78). Utilizing TLS or terrestrial photogrammetry provides us with new insights into how to measure tree structure in a detailed 3D view (Calders et al., 2020). As a result, irregularly large trees, like buttressed trees, can also be modeled with higher accuracy. Since the ASA tends to overestimate a buttressed tree volume when the trees present more and shallower horizontal buttresses, more accurate, automatic tree construction methods are expected with the increasing availability of 3D point clouds. Future research could also focus on developing an automatic reconstruction method that optimizes the entire tree, not focusing on the buttressed sections alone. For instance, the algorithm can define the height of DAB, where the above and the below part can be modeled automatically, by QSMs and the alpha shape algorithm, respectively.

4.5 Conclusions

In this study, we used the Alpha Shape Algorithm and the Slice Triangulation method to estimate buttress volume based on 3D point clouds in order to reduce the above variation. The volume estimates of the alpha shape algorithm and the slice triangulation method shows a similar RRMSE, with both methods outperforming allometric models developed with DAB and Darea130. Meanwhile, the alpha shape algorithm tends to work better than the slice triangulation when the trees present more and shallower horizontal buttresses. With databases including trees from three continents, this model can be applied to tropical buttressed trees globally, increasing the applicability of this model to industry and field research. Additionally, large trees have been underrepresented in previous studies, providing an opportunity for the method presented here to better capture comprehensive values volume and biomass, and improve carbon storage estimations in tropical forests.

Acknowledgement

This research was funded by the Chinese Scholarship Council Grant: 201806850089, and the National Science and Engineering Research Council of Canada (NSERC) Discovery Grant. The data processing workstation was provided by the Center for Earth Observation Sciences (CEOS) of U of A. We thank Bauwens et al. (2017) and Noelke et al. (2015) for their data. We also thank J. Antonio Guzman Q for his code. The authors declare no conflicts of interest.

4.6 Supplementary materials

There are nine buttressed trees for validation. The first three trees in table S4.2 are measured by destructively harvest, while the remaining six are scanned by terrestrial laser scanning (TLS).

Table S4. 1: Buttressed trees used for allometric model development; Trees 1 to 30 and Trees 31 to 36 are from the YR and BBG database, respectively.

Tree	Species	DAB (m)	$D_{area130}$ (m)	$D_{convex130}$ (m)	HDAB (m)	BA (m^2)	*f*
1	E. cylindricum	1.01	1.18	1.29	5	1.09	0.2
2	C. mildbraedii	0.76	0.88	1.59	5.1	0.61	0.29
3	C. mildbraedii	0.79	0.95	1.81	5.1	0.71	0.29
4	E. cylindricum	1.36	1.67	2.45	7.7	2.2	0.25
5	C. mildbraedii	0.82	0.98	1.60	5.5	0.76	0.41
6	C. mildbraedii	0.57	0.63	0.73	2.9	0.31	0.25
7	E. cylindricum	1.35	1.76	2.42	6.5	2.42	0.37
8	E. cylindricum	1.75	2.16	2.40	4	3.68	0.28
9	E. cylindricum	1.35	1.62	1.94	4.5	2.05	0.27
10	E. cylindricum	0.69	0.73	0.75	2.05	0.42	0.27
11	E. cylindricum	1.80	2.16	3.02	4.5	3.67	0.25
12	C. mildbraedii	0.54	0.56	0.59	1.8	0.25	0.26
13	E. cylindricum	1.14	1.39	1.76	5.8	1.51	0.24
14	E. cylindricum	1.27	1.48	1.88	4.6	1.72	0.24
15	E. cylindricum	1.04	1.21	1.26	3.7	1.16	0.24
16	E. cylindricum	1.29	1.53	1.77	4.5	1.85	0.26
17	E. cylindricum	1.22	1.66	2.75	7.8	2.15	0.31
18	C. mildbraedii	0.54	0.63	0.73	3.3	0.31	0.36
19	C. mildbraedii	0.40	0.42	NA	2.15	0.14	0.18
20	E. cylindricum	2.02	2.51	3.10	7.7	4.94	0.22
21	C. mildbraedii	0.74	0.83	1.44	3.5	0.54	0.34
22	C. mildbraedii	0.78	0.97	1.64	4.5	0.74	0.38
23	E. cylindricum	1.26	1.46	1.65	4.5	1.67	0.2
24	E. cylindricum	1.04	1.09	1.10	1.99	0.92	0.15
25	E. cylindricum	1.33	1.57	1.74	4.5	1.93	0.28
26	C. mildbraedii	0.79	0.73	0.98	3.8	0.41	0.08
27	C. mildbraedii	0.76	0.92	1.27	4.3	0.67	0.29
28	C. mildbraedii	0.64	0.72	0.89	4.6	0.4	0.23
29	C. mildbraedii	0.44	0.44	0.46	1.6	0.15	0.15
30	E. cylindricum	1.15	1.42	1.75	5.8	1.59	0.21
31	Sterculia urceolata	0.81	1.06	1.97	7.23	0.88	0.27
32	Sterculia urceolata	0.95	1.27	2.57	6.78	1.27	0.31
33	Sterculia foetida	1.32	1.73	4.26	7.03	2.35	0.3
34	Ceiba pentandra	0.85	0.99	1.12	2.98	0.77	0.36
35	Bombax ceiba	0.67	0.74	1.04	2.63	0.43	0.25
36	Bombax valetonii	0.72	0.77	0.8	1.98	0.46	0.39

HDAB, the height of DAB, 50 cm higher than the buttresses; BA, basal area, cross-sectional area at breast height; f, volume efficient (Eqn. 2.1).

Table S4. 2: The description of nine buttressed trees for validation.

Tree	Species	DAB (m)	$D_{area130}$ (m)	$D_{convex130}$ (m)	H_{DAB} (m)	BA (m^2)	f_b
1	C. mildbraedii	0.86	0.98	1.58	4.6	0.75	0.13
2	C. mildbraedii	0.52	0.71	0.96	4	0.39	0.33
3	Cynometra hankei	0.71	0.83	1.22	3.4	0.54	0.19
4	Koompassia excelsa	1.4	1.75	1.86	2.43	2.41	0.5
5	Ficus robusta	0.88	1.38	2.8	7.07	1.49	0.46
6	Celtis rigescens	0.72	0.81	1.08	5.08	0.52	0.22
7	Ficus albipila	1.66	2	3.24	7.18	3.15	0.24
8	Shorea leprosula	1.18	1.87	4.04	9.08	2.73	0.43
9	Bombax valetonii	0.73	0.81	0.9	2.98	0.52	0.28

There are nine buttressed trees for validation. The first three trees are measured by destructively harvest from YR database, while the remaining six are scanned by terrestrial laser scanning (TLS) from SP database.

Table S4. 3: The description of six buttressed tree from Santiago de Puriscal, Costa Rica (SP database). * No DAB for Tree 1 from SP database since this tree is dead, 9.6 is the height of dead tree. Tree 6 has two DABs due to the two separate trunks.

Tree	DAB (m)	HDAB (m)	Alpha	Va (m3)	Area(m2)	Complexity
1	*	9.6*	0.4	17.8	73.07	4.11
2	0.68	3.2	0.24	2.39	34.96	14.6
3	0.75	3.3	0.25	4.17	51.01	12.23
4	0.91	14.8	0.25	16.12	141.05	8.75
5	0.86	3.8	0.24	2.1	18.33	8.73
6	0.55/0.61	3.5	0.24	5.51	51.61	9.38

Table S4. 4: Six-fold cross-validation for Darea130 model.

	Coefficient	Intercept	R^2	RMSE	Mean	RRMSE
Fold 1	0.89	-0.07	0.99	0.30	4.42	0.07
Fold 2	0.91	-0.09	0.99	0.78	5.81	0.13
Fold 3	0.91	-0.09	0.99	0.86	13.46	0.06
Fold 4	0.91	-0.09	0.99	0.16	2.78	0.06
Fold 5	0.9	-0.08	0.99	0.21	3.48	0.06
Fold 6	0.91	-0.10	0.99	0.48	8.30	0.06
Ave	0.91	-0.09	0.99	0.47	6.38	0.07
Final	0.91	-0.09	0.99	0.52	6.38	0.08

Table S4. 5: Six-fold cross-validation for DAB model.

	Coefficient	Intercept	R2	RMSE	Mean	RRMSE
Fold 1	1.00	0.07	0.99	0.27	4.42	0.06
Fold 2	1.00	0.07	0.99	0.95	5.81	0.16
Fold 3	1.01	0.07	0.99	1.73	13.46	0.13
Fold 4	1.01	0.05	0.99	0.21	2.78	0.08
Fold 5	0.99	0.10	0.99	0.38	3.48	0.11
Fold 6	1.01	0.06	0.99	0.45	8.30	0.05
Ave	1.00	0.07	0.99	0.67	6.38	0.10
Final	1.00	0.07	0.99	0.67	6.38	0.11

Table S4. 6: Six-fold cross-validation for Dconvex130 model.

	Coefficient	Intercept	R2	RMSE	Mean	RRMSE
Fold 1	0.70	-0.14	0.90	2.08	4.42	0.47
Fold 2	0.70	-0.23	0.91	3.86	5.81	0.66
Fold 3	0.68	-0.16	0.89	6.61	13.46	0.49
Fold 4	0.70	-0.10	0.89	1.31	2.78	0.47
Fold 5	0.71	-0.18	0.89	1.15	4.41	0.26
Fold 6	0.75	-0.24	0.89	7.38	8.30	0.89
Ave	0.71	-0.18	0.90	3.73	6.53	0.57
Final	0.72	-0.18	0.89	3.92	6.55	0.60

Table S4. 7: Two sample T-test of fa and fb for two species, the value that larger than 0.05 indicated no evidence of difference.

Species	E. cylindricum		C. mildbraedii	
Methods	f_a	f_b	f_a	f_b
Normality	0.55	0.64	0.35	0.89
Variance	0.71		0.64	
T.test	0.14		0.11	
Mean	0.22	0.25	0.34	0.27
Standard derivation	0.05	0.05	0.11	0.09

Table S4. 8: Two sample T-test of f t and f b for two species, the value that larger than 0.05 indicated no evidence of difference.

Species	E. cylindricum		C. mildbraedii	
Methods	f_t	f_b	f_t	f_b
Normality	0.92	0.64	0.08	0.89
Variance	0.06		0.27	
T.test	0.22		0.56	
Mean	0.22	0.25	0.24	0.27
Standard derivation	0.08	0.05	0.13	0.09

Table S4. 9: Two sample T-test of f t, f a and f b between E. cylindricum and C. mildbraedii.

Methods	f_t		f_a		f_b	
Species	E. cylindricum	C. mildbraedii	E. cylindricum	C. mildbraedii	E. cylindricum	C. mildbraedii
Normality	0.92	0.08	0.55	0.35	0.64	0.89
Variance	0.07		0.01**		0.02*	
T.test	0.56		0.003**		0.48	
Mean	0.22	0.24	0.22	0.34	0.25	0.27
Standard derivation	0.08	0.13	0.05	0.11	0.05	0.09

There is strong evidence that the f a (p = 0.003) is different between E. cylindricum and C. mildbraedii, which is different from the f t (p = 0.56) and f b (p = 0.48).

According to Figure S4.1, the ASA volume showed a strong relationship with the volume calculated by the Smalian formula for both species, while C. mildbraedii (R2 = 0.95) was weaker than E. cylindricum (R2 = 0.99). Without considering species effect, the ASA- Smalian volume relationship was also strong (R2 = 0.99, Figure S4.1(c)). The ST volume showed a same trend, with C. mildbraedii volume having a R2 of 0.95. The ASA method tends to overestimate buttress volume of C. mildbraedii, with a slope coefficient of 0.86; however, the ST volume (Vt) is relatively consistent with the reference volume (Vb) on C. mildbraedii (slope coefficient of 1.05; Figure S4.1(d)).

The assumptions of developed models from Figure S4.3. The top left figure indicates the linear relationship assumption, and the top right checks whether the residuals are normally distributed. The third figure at the bottom left checks the homoscedasticity of the residuals, while the last figure indicates influential values that affect the regression results.

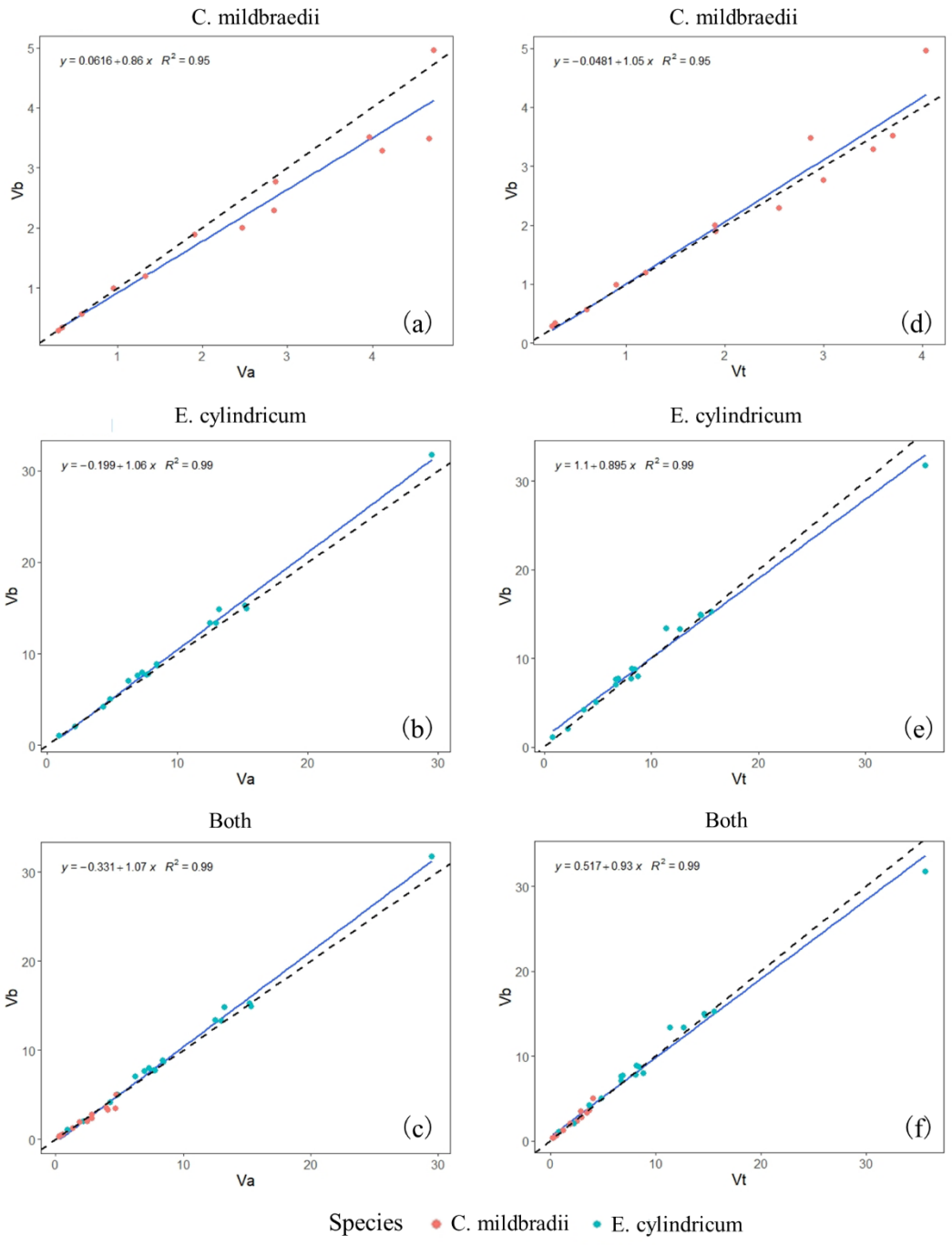


Figure S4. 1: The relationships of the slice triangulation volume (Vt) and alpha shape volume (Va) with the reference volume (Vb); (a) Vb vs. Va for C. mildbraedii; (b) Vb vs. Va for E. cylindricum; (c) Vb vs. Va for both species. (d) Vb vs. Vt for C. mildbraedii; (e) Vb vs. Vt for E. cylindricum; (f) Vb vs. Vt for both species.

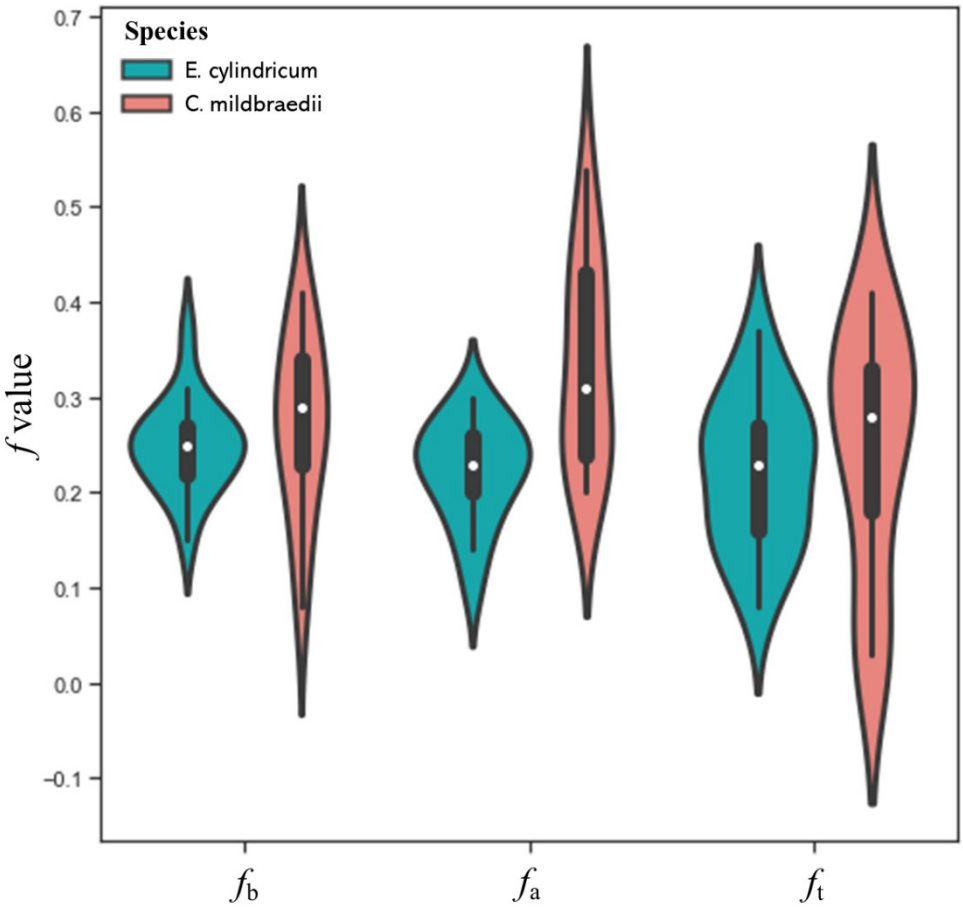


Figure S4. 2: Violin plot comparing the f b, f a and f t derived from reference data, alpha shape volume, and slice triangulation volume, respectively. f, the proportion of buttress volume that is not considered when the volume is calculated as a cylinder with a diameter equal to DAB.

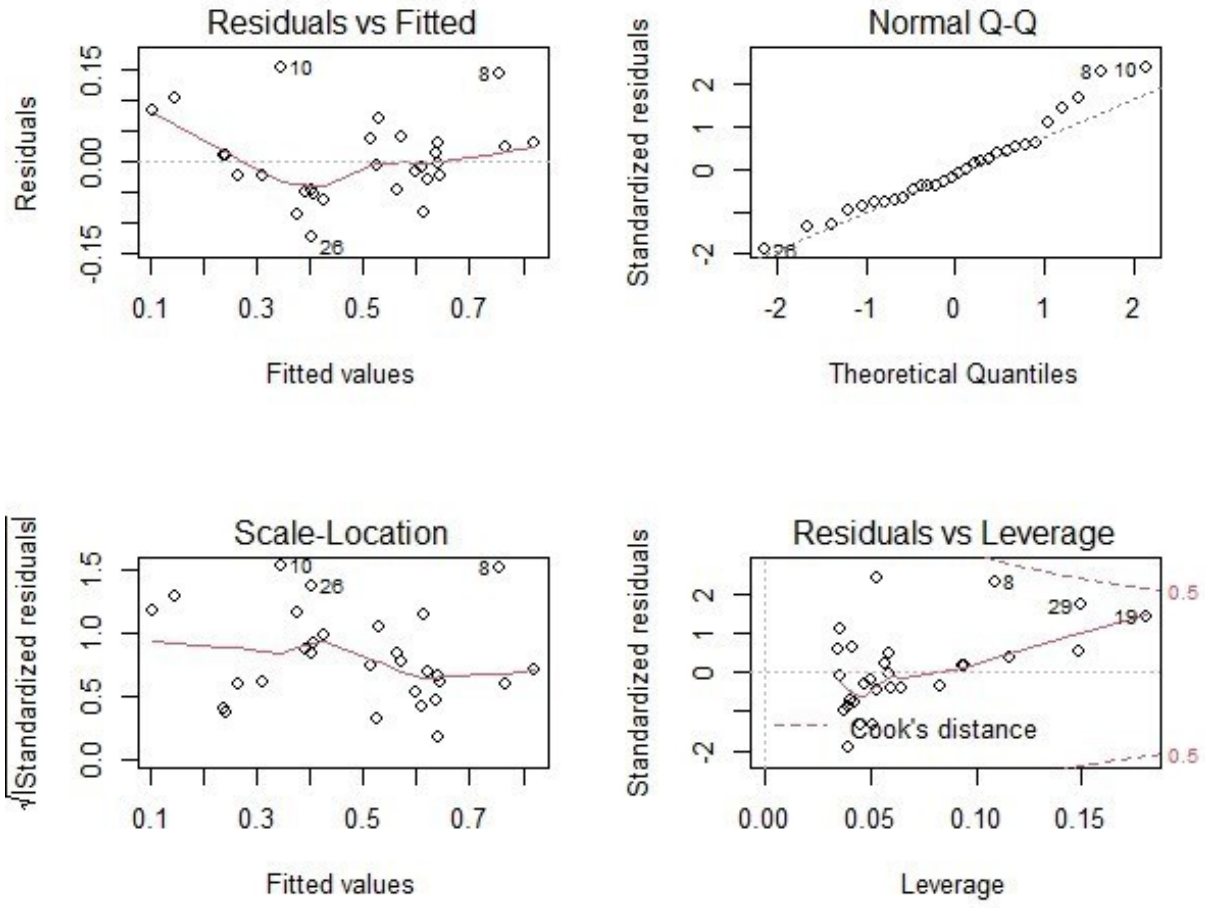


Figure S4. 3: The assumption of the alpha-DAB model. The Shapiro-Wilk normality test indicates p-value of 0.35 for this model.

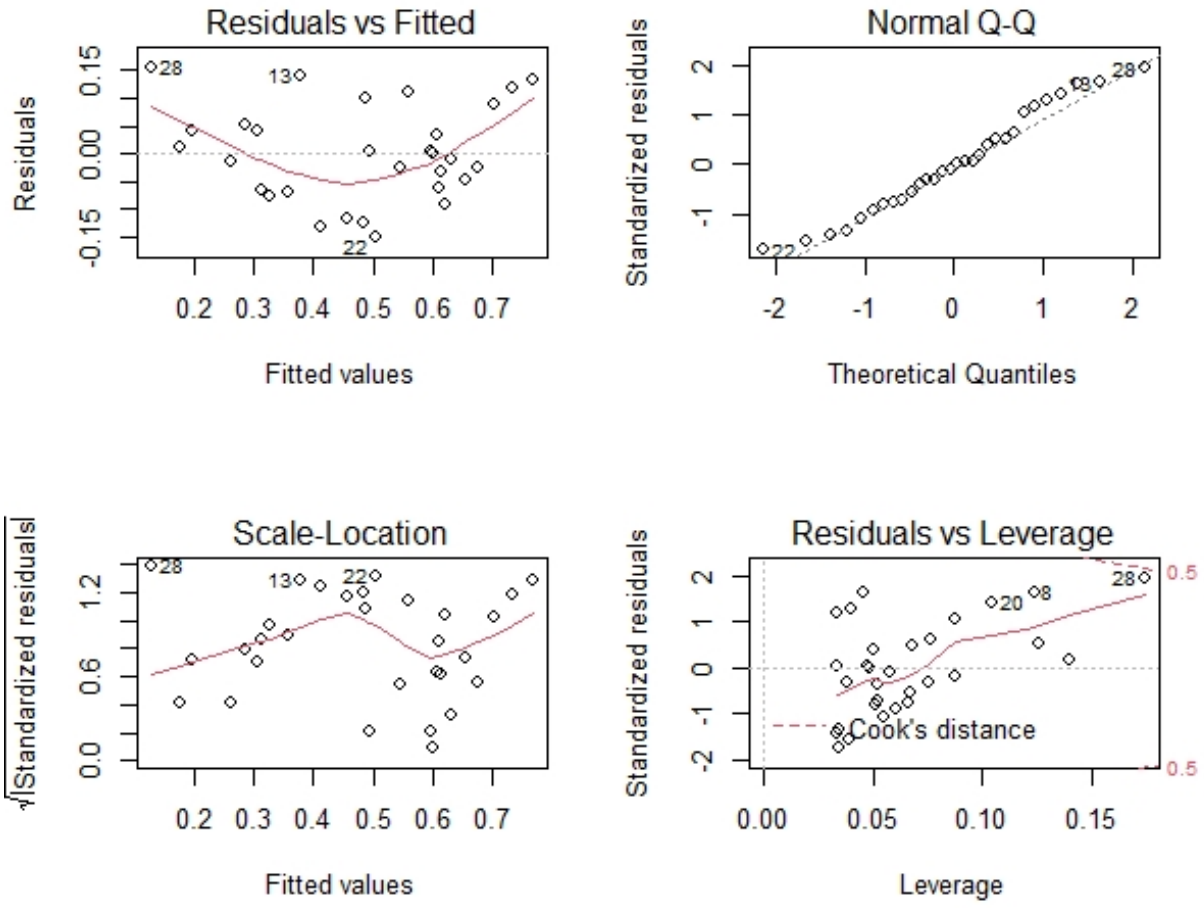


Figure S4. 4: The assumption of the alpha-complexity model. The Shapiro-Wilk normality test indicates p-value of 0.49 for this model.

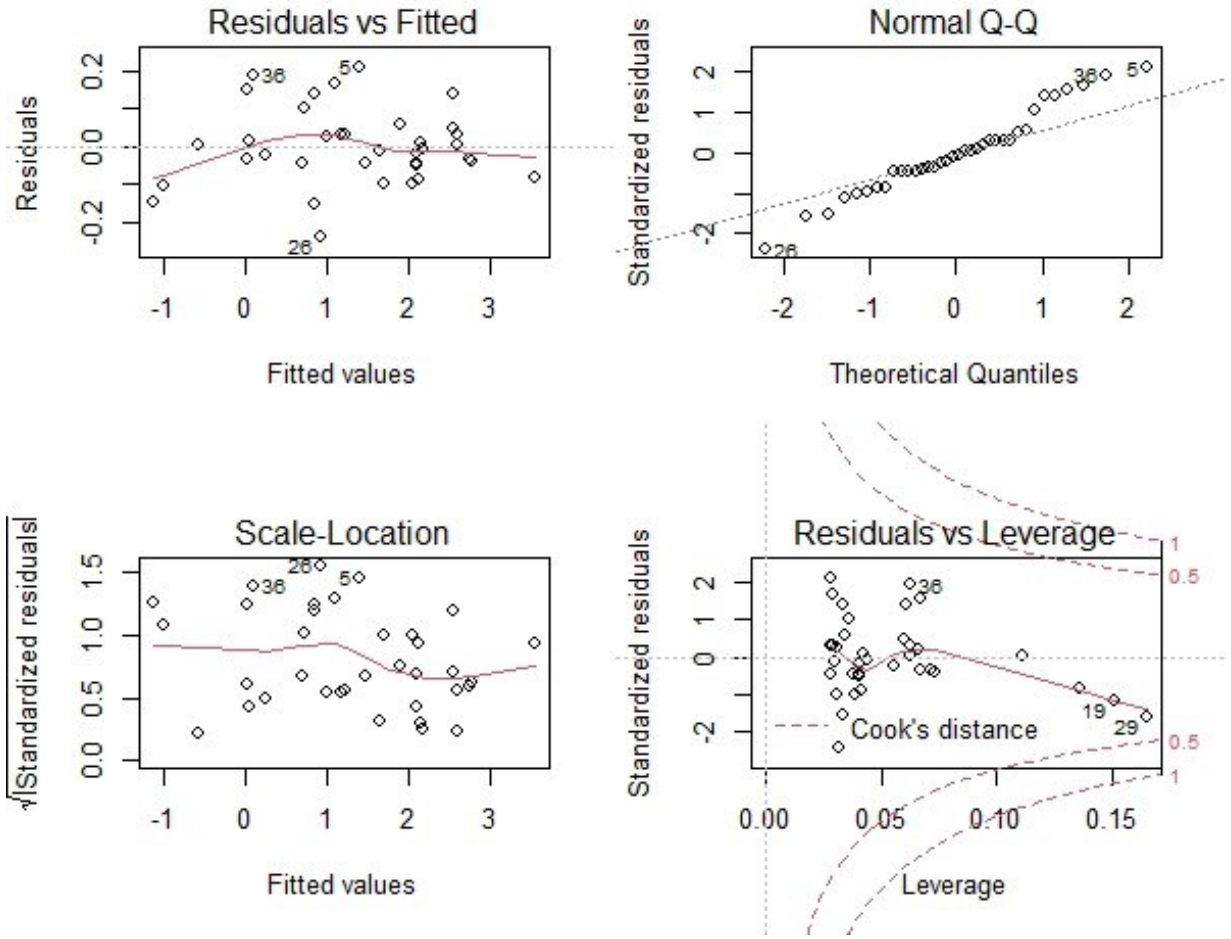


Figure S4. 5: The assumption of the DAB based model. The Shapiro-Wilk normality test indicates p-value of 0.43 for this model.

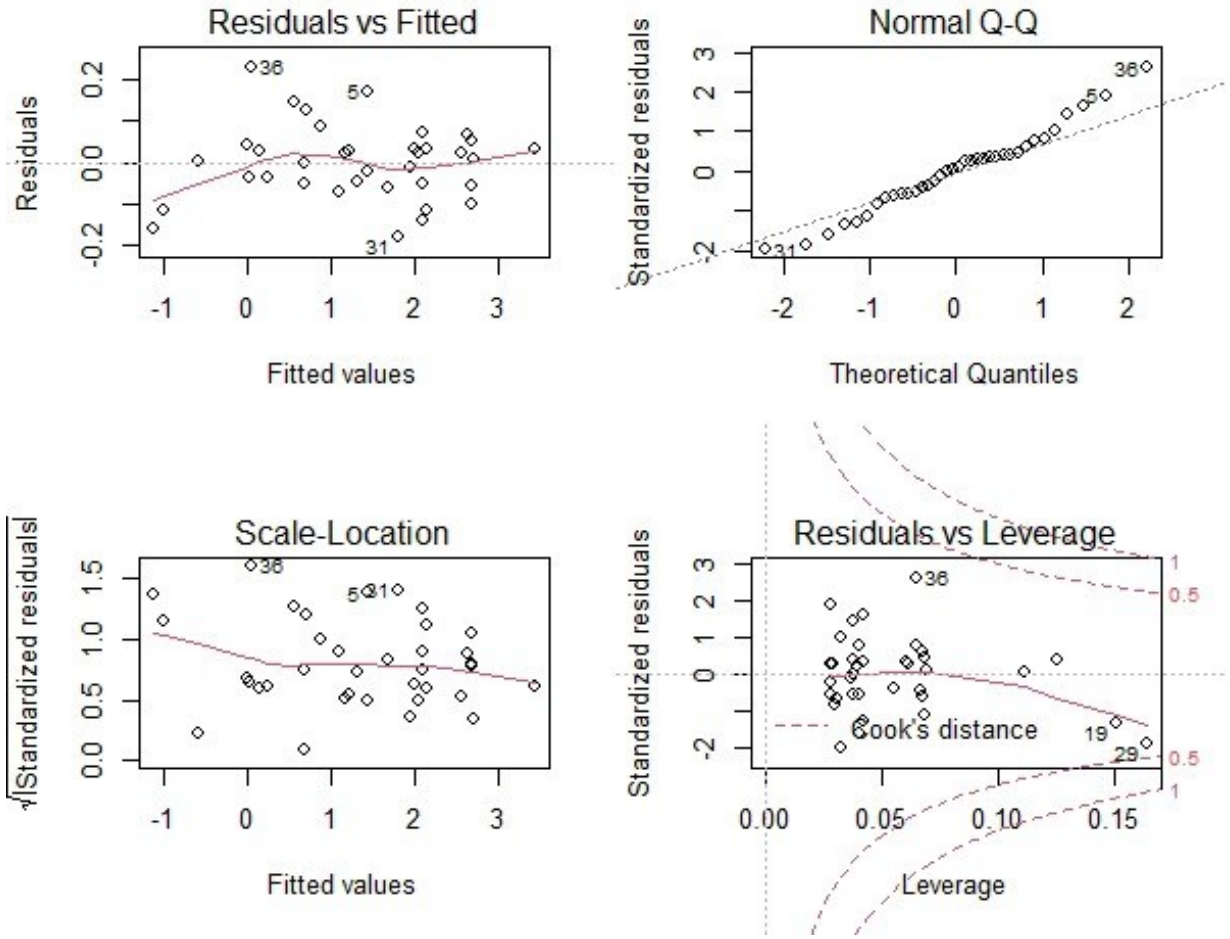


Figure S4. 6: The assumption of the Darea130 based model. The Shapiro-Wilk normality test indicates p-value of 0.70 for this model.

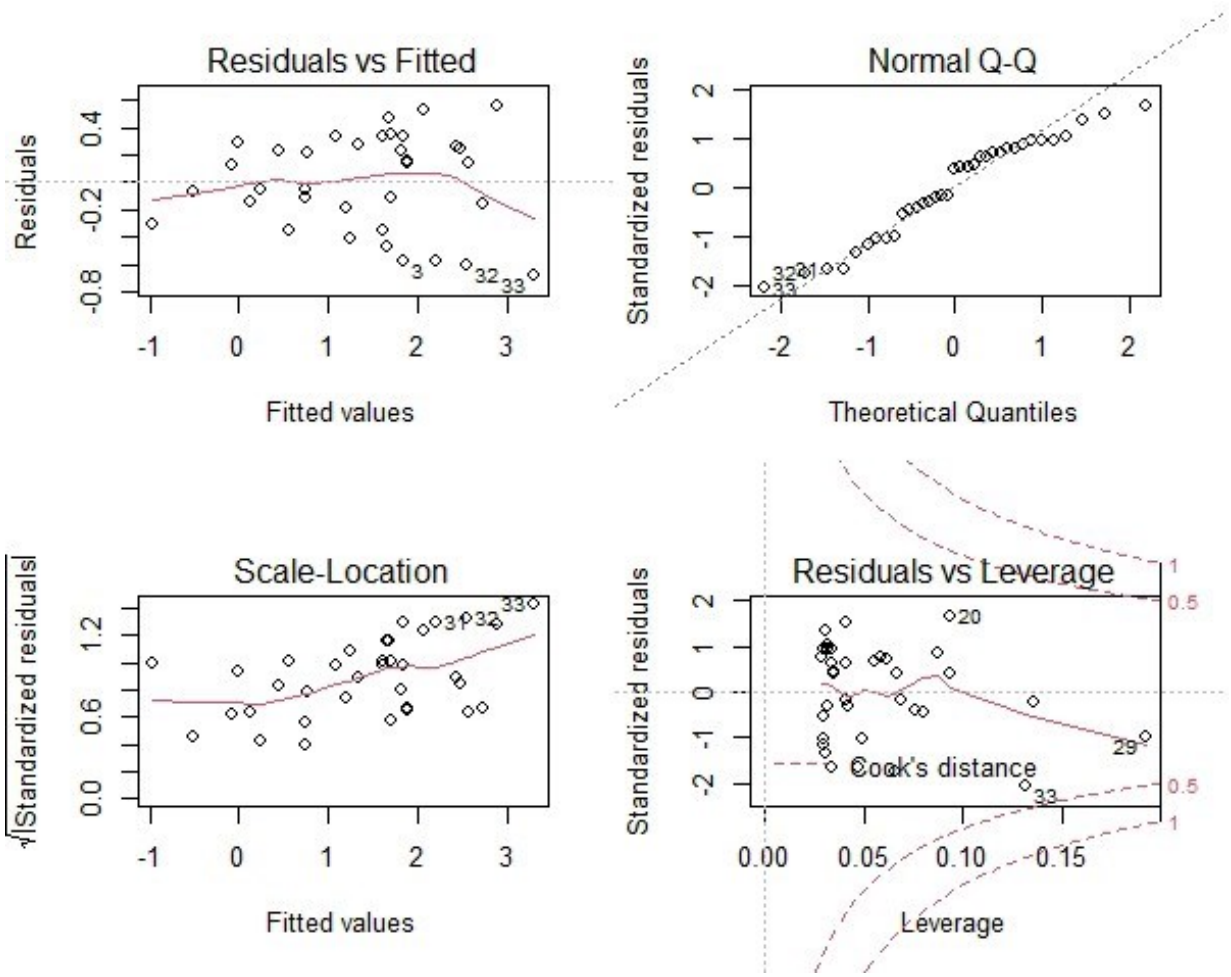


Figure S4. 7: The assumption of the Dconvex130 based model. The Shapiro-Wilk normality test indicates p-value of 0.09 for this model.

Figures S4.5, S4.6 and S4.7 revealed the residuals of DAB, $D_{area130}$ and $D_{convex130}$ models. DAB and $D_{area130}$ models were more evenly distributed across the different volumesthan that of $D_{convex130}$ model. It can be observed from Figure S4.7 that the residuals of $D_{convex130}$ model and volume were positively correlated. Specifically, the residuals of DAB and $D_{area130}$ models ranged from -0.2 to 0.2, while the residuals of $D_{convex130}$ model were from -0.8 to 0.6.

Also, the DAB and $D_{area130}$ based models satisfied the assumption of normality and homogenous variance of linear regression.

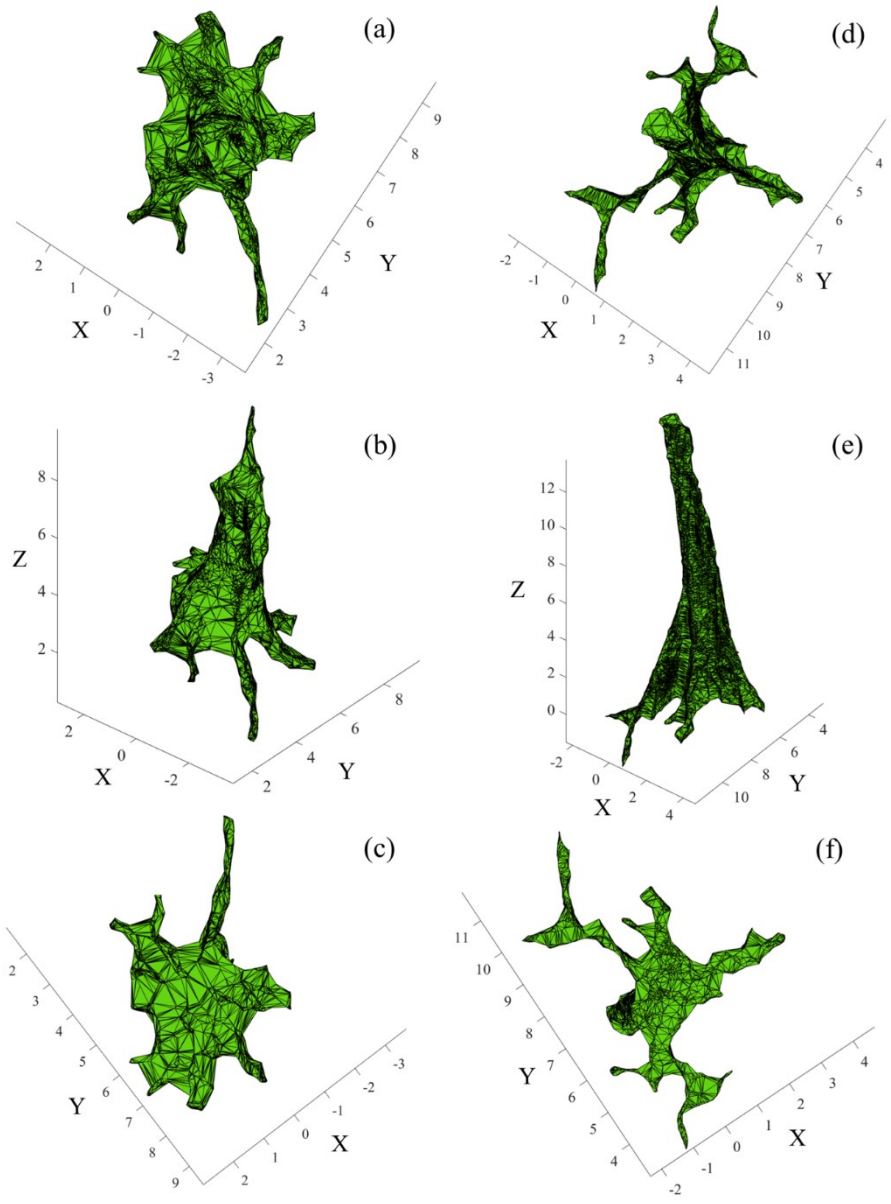


Figure S4. 8: The predictions of the alpha shape algorithm (ASA) on Tree 3 and Tree 4 from Santiago de Puriscal, Costa Rica (SP database), (a) Top view of Tree 3 from SP database; (b) Side view of Tree 3 from SP database; (c) Bottom view of Tree 3 from SP database; (d) Top view of Tree 4 from SP database; (e) Side view of Tree 4 from SP database; (f) Bottom view of Tree 4 from SP database. The dark area shows the buttresses, while the light area indicates the overlapped area.

It should be note that the alpha shape algorithm does not generate a good modeling on Tree 6 from SP database, since the two separate trunks in Tree 6 is closer, which create many overlapped areas (see Figure S4.9(e)). We better suggest separating the two trunks from the point cloud of Tree 6, and then run the alpha shape algorithm.

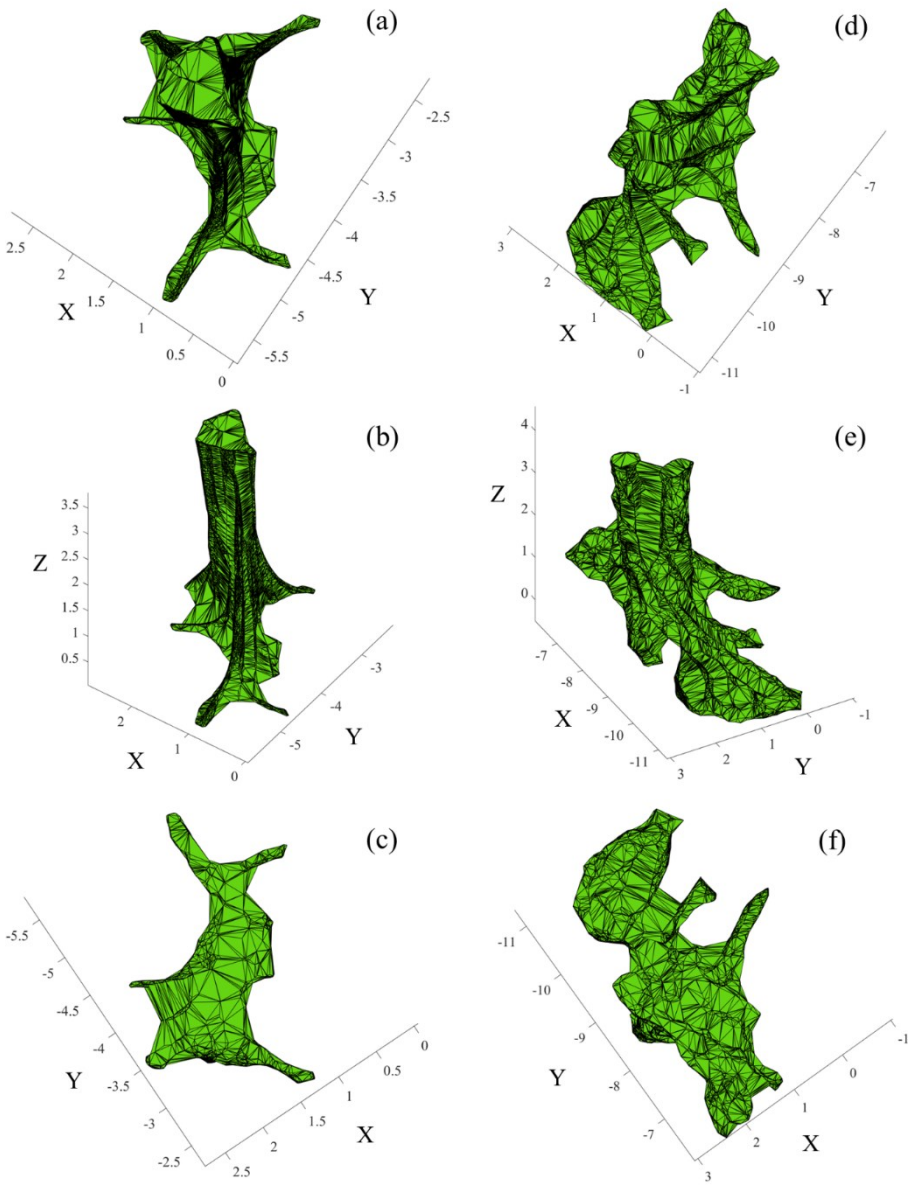


Figure S4. 9: The predictions of the alpha shape algorithm (ASA) on Tree 5 and Tree 6 from Santiago de Puriscal, Costa Rica (SP database), (a) Top view of Tree 5 from SP database; (b) Side view of Tree 5 from SP database; (c) Bottom view of Tree 5 from SP database; (d) Top view of Tree 6 from SP database; (e) Side view of Tree 6 from SP database; (f) Bottom view of Tree 6 from SP database. The area with no points indicates the overlapped area.

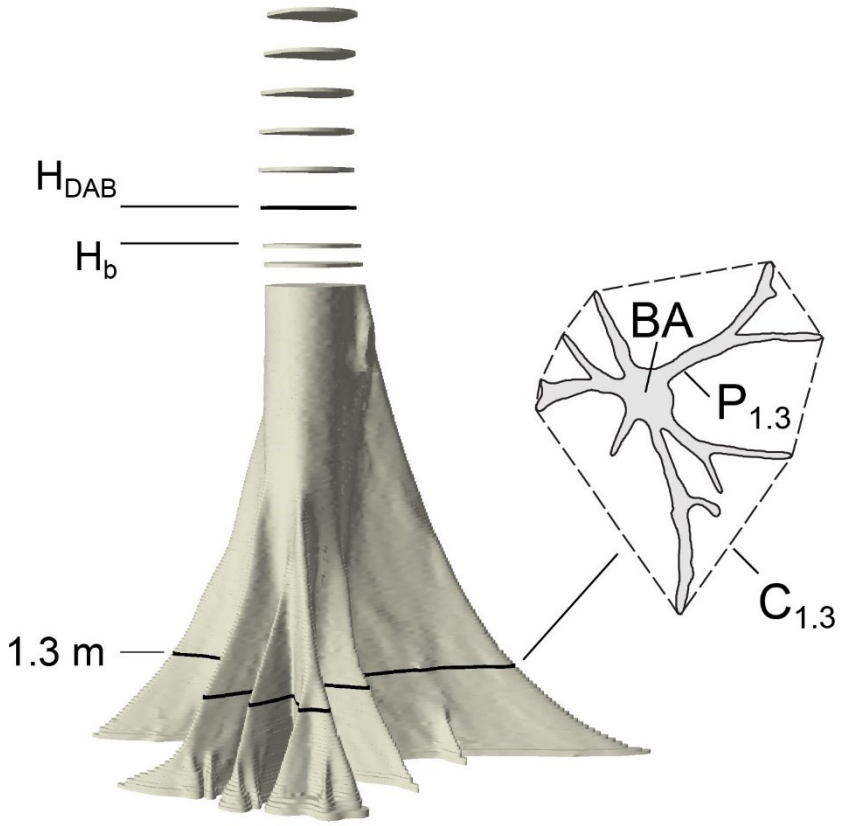


Table S4. 10: The buttress shape at the breast height (Nolke et al., 2015). BA is the basal area at the breast height; H_b is the buttress height; HDAB is the height of the DAB, P1.3 is the actual non-convex perimeter of the cross section at breast height; C 1.3 is the C1.3 is the perimeter of the convex hull (dashed line).

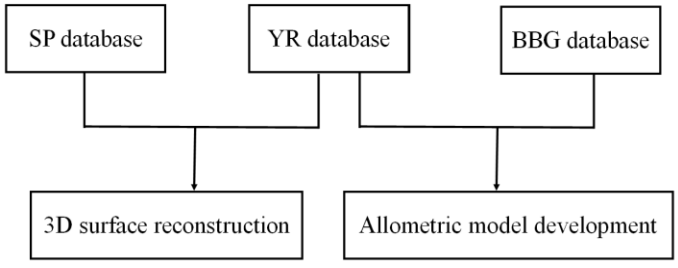


Table S4. 11: The data used strategy in this study.

4.7 References

Barbeito, Ignacio, Mathieu Dassot, Dominik Bayer, Catherine Collet, Lars Drössler, Mag- nus Löf, Miren del Rio, Ricardo Ruiz-Peinado, David I Forrester, Andrés Bravo-Oviedo, et al. (2017). "Terrestrial laser scanning reveals differences in crown structure of Fagus sylvatica in mixed vs. pure European forests". In: *Forest Ecology and Management* 405, pp. 381–390.

Basuki, TM, PE Van Laake, AK Skidmore, and YA Hussin (2009). "Allometric equations for estimating the above-ground biomass in tropical lowland Dipterocarp forests". In: Forest ecology and management 257.8, pp. 1684–1694.

Bauwens, Sebastien, Adeline Fayolle, Sylvie Gourlet-Fleury, Leopold Mianda Ndjele, Coralie Mengal, and Philippe Lejeune (2017). "Terrestrial photogrammetry: a non- destructive method for modelling irregularly shaped tropical tree trunks". In: Methods in Ecology and Evolution 8.4, pp. 460–471.

Bonneau, David, Paul-Mark DiFrancesco, and D Jean Hutchinson (2019). "Surface recon- struction for three-dimensional rockfall volumetric analysis". In: ISPRS International Journal of Geo-Information 8.12, p. 548.

Calders, Kim, Jennifer Adams, John Armston, Harm Bartholomeus, Sebastien Bauwens, Lisa Patrick Bentley, Jerome Chave, F Mark Danson, Miro Demol, Mathias Disney, et al. (2020). "Terrestrial laser scanning in forest ecology: Expanding the horizon". In: Remote Sensing of Environment 251, p. 112102.

Calders, Kim, Glenn Newnham, Andrew Burt, Simon Murphy, Pasi Raumonen, Martin Herold, Darius Culvenor, Valerio Avitabile, Mathias Disney, John Armston, et al. (2015). "Nondestructive estimates of above-ground biomass using terrestrial laser scan- ning". In: Methods in Ecology and Evolution 6.2, pp. 198–208.

Chapman, Colin A, Les Kaufman, and Lauren J Chapman (1998). “Buttress formation and directional stress experienced during critical phases of tree development”. In: Journal of Tropical Ecology 14.3, pp. 341–349.

Clark, Deborah A (2002). “Are tropical forests an important carbon sink? Reanalysis of the long-term plot data”. In: Ecological Applications 12.1, pp. 3–7.

Cola¸co, Andr´e F, Rodrigo G Trevisan, Jos´e P Molin, Joan R Rosell-Polo, et al. (2017). “A method to obtain orange crop geometry information using a mobile terrestrial laser scanner and 3D modeling”. In: Remote Sensing 9.8, p. 763.

Cushman, KC, Sarayudh Bunyavejchewin, Dairon C´ardenas, Richard Condit, Stuart J Davies, A´lvaro Duque, Stephen P Hubbell, Somboon Kiratiprayoon, Shawn KY Lum, and Helene C Muller-Landau (2021). “Variation in trunk taper of buttressed trees within and among five lowland tropical forests”. In: Biotropica 53.5, pp. 1442–1453.

Cushman, KC, Helene C Muller-Landau, Richard S Condit, and Stephen P Hubbell (2014). “Improving estimates of biomass change in buttressed trees using tree taper models”. In: Methods in Ecology and Evolution 5.6, pp. 573–582.

Dean, Christopher and SH Roxburgh (2006). “Improving visualisation of mature, high- carbon-sequestering forests”. In: Forest Biometry, Modelling and Information Sciences 1, pp. 48–69.

Delaunay, Boris, S Vide, A Lam´emoire, and V De Georges (1934). “Bulletin de l’Acad´emie des Sciences de l’URSS”. In: Classe des sciences math´ematiques et naturelles 6, pp. 793– 800.

Disney, Mathias I, Matheus Boni Vicari, Andrew Burt, Kim Calders, Simon L Lewis, Pasi Raumonen, and Phil Wilkes (2018). “Weighing trees with lasers: advances, challenges and opportunities”. In: Interface Focus 8.2, p. 20170048.

Edelsbrunner, Herbert and Ernst P Mu¨cke (1994). “Three-dimensional alpha shapes”. In: ACM Transactions on Graphics (TOG) 13.1, pp. 43–72.

Gonzalez de Tanago, Jose, Alvaro Lau, Harm Bartholomeus, Martin Herold, Valerio Avitabile, Pasi Raumonen, Christopher Martius, Rosa C Goodman, Mathias Disney, Solichin Manuri, et al. (2018). “Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR”. In: Methods in Ecology and Evolution 9.2, pp. 223–234.

Guzm´an Q, J Antonio, Hernandez Ronny, and Sanchez-Azofeifa Arturo (2021). rTLS: Tools to Process Point Clouds Derived from Terrestrial Laser Scanning.

Guzm´an Q, J Antonio, Iain Sharp, Felipe Alencastro, and G Arturo S´anchez-Azofeifa (2020). “On the relationship of fractal geometry and tree–stand metrics on point clouds derived from terrestrial laser scanning”. In: Methods in Ecology and Evolution 11.10, pp. 1309–1318.

Hadas, Edyta, Andrzej Borkowski, Javier Estornell, and Przemyslaw Tymkow (2017). “Au- tomatic estimation of olive tree dendrometric parameters based on airborne laser scan- ning data using alpha-shape and principal component analysis”. In: GIScience & Re- mote Sensing 54.6, pp. 898–917.

Kankare, Ville, Mikko Vastaranta, Markus Holopainen, Minna R¨aty, Xiaowei Yu, Juha Hyyppa¨, Hannu Hyyppa¨, Petteri Alho, and Risto Viitala (2013). “Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR”. In: Remote Sensing 5.5, pp. 2257–2274.

Lafarge, T and B Pateiro-Lopez (2014). alphashape3d: Implementation of the 3D Alpha- Shape for the Reconstruction of 3D Sets from a Point Cloud.

Lau, Alvaro, Kim Calders, Harm Bartholomeus, Christopher Martius, Pasi Raumonen, Martin Herold, Matheus Vicari, Hansrajie Sukhdeo, Jeremy Singh, and Rosa C Good- man (2019). “Tree biomass equations from terrestrial LiDAR: A case study in Guyana”. In: Forests 10.6, p. 527.

Mehedi, Md Abu Hanifa, Chandan Kundu, and Md Qumruzzaman Chowdhury (2012). “Patterns of tree buttressing at Lawachara National Park, Bangladesh”. In: Journal of Forestry Research 23.3, pp. 461–466.

Muff, Stefanie, Erlend B Nilsen, Robert B O’Hara, and Chlo´e R Nater (2021). “Rewriting results sections in the language of evidence”. In: Trends in ecology & evolution.

Newbery, David M, Sarah Schwan, George B Chuyong, and Xander M van der Burgt (2009). “Buttress form of the central African rain forest tree Microberlinia bisulcata, and its possible role in nutrient acquisition”. In: Trees 23.2, pp. 219–234.

Newnham, Glenn J, John D Armston, Kim Calders, Mathias I Disney, Jenny L Lovell, Crystal B Schaaf, Alan H Strahler, and F Mark Danson (2015). “Terrestrial laser scan- ning for plot-scale forest measurement”. In: Current Forestry Reports 1.4, pp. 239–251.

Ngomanda, Alfred, Quentin Moundounga Mavouroulou, Nestor Laurier Engone Obiang, Donald Midoko Iponga, Jacques-Fran¸cois Mavoungou, Nicaise L´epengu´e, Nicolas Pi- card, and Bertrand Mbatchi (2012). “Derivation of diameter measurements for but- tressed trees, an example from Gabon”. In: Journal of tropical ecology 28.3, pp. 299– 302.

Noelke, Nils, Lutz Fehrmann, Surati Jaya I Nengah, Tatang Tiryana, Dominik Seidel, and Christoph Kleinn (2015). “On the geometry and allometry of big-buttressed trees-a challenge for forest monitoring: new insights from 3D-modeling with terrestrial laser scanning”. In: iForest-Biogeosciences and Forestry 8.5, p. 574.

Nogueira, Euler Melo, Bruce Walker Nelson, and Philip M Fearnside (2006). “Volume and biomass of trees in central Amazonia: influence of irregularly shaped and hollow trunks”. In: Forest Ecology and Management 227.1-2, pp. 14–21.

Pandey, CB, Lalita Singh, and SK Singh (2011). “Buttresses induced habitat heterogeneity increases nitrogen availability in tropical rainforests”. In: Forest ecology and manage- ment 262.9, pp. 1679–1685.

Pfeffer, WF (1986). “The divergence theorem”. In: Transactions of the American Mathe- matical Society 295.2, pp. 665–685.

Raumonen, Pasi, Mikko Kaasalainen, Markku ˚Akerblom, Sanna Kaasalainen, Harri Kaarti- nen, Mikko Vastaranta, Markus Holopainen, Mathias Disney, and Philip Lewis (2013). “Fast automatic precision tree models from terrestrial laser scanner data”. In: Remote Sensing 5.2, pp. 491–520.

Seidel, Dominik (2018). “A holistic approach to determine tree structural complexity based on laser scanning data and fractal analysis”. In: Ecology and Evolution 8.1, pp. 128–134.

Sprugel, DG (1983). “Correcting for bias in log-transformed allometric equations”. In:
Ecology (Durham) 64.1, pp. 209–210.

Su, Tianyun, Wen Wang, Haixing Liu, Zhendong Liu, Xinfang Li, Zhen Jia, Lin Zhou, Zhuanling Song, Ming Ding, and Aiju Cui (2020). “An adaptive and rapid 3D Delaunay triangulation for randomly distributed point cloud data”. In: The Visual Computer, pp. 1–25.

Tang, Yong, Xiaofei Yang, Min Cao, Carol C Baskin, and Jerry M Baskin (2011). “Buttress trees elevate soil heterogeneity and regulate seedling diversity in a tropical rainforest”. In: Plant and soil 338.1, pp. 301–309.

Vauhkonen, Jari, Anne Sepp¨anen, Petteri Packal´en, and Timo Tokola (2012). “Improving species-specific plot volume estimates based on airborne laser scanning and image data using alpha shape metrics and balanced field data”. In: Remote Sensing of Environment 124, pp. 534–541.

Wang, Wen, Tianyun Su, Haixing Liu, Xinfang Li, Zhen Jia, Lin Zhou, Zhuanling Song, and Ming Ding (2019). “Surface reconstruction from unoriented point clouds by a new triangle selection strategy”. In: Computers & Graphics 84, pp. 144–159.

Zhiyuan, He, Tang Yong, Deng Xiaobao, and Cao Min (2013). “Buttress trees in a 20- hectare tropical dipterocarp rainforest in Xishuangbanna, SW China”. In: Journal of Plant Ecology 6.2, pp. 187–192.

Chapter5

Conclusions

5.1 Conclusions and significant contributions

The primary objectives of this thesis are to 1) build a new machine learning model (RF and XGBoosting) coupled with the geometric features derived from point clouds, to separate lianas from trees, and then compare the performance of the presented method with Moorthy et al. (2019), 2) assess the performance of different deep learning time-series algorithms for leaf and wood classification, and investigate the potential of Class ActivationMap (CAM) for explanation of black-box effects of deep learning models, 3) understand the performance of alpha shape algorithm (ASA) and slice triangulation (ST) method on buttress volume estimation, and develop allometric models between different tree metrics and buttress volume. This thesis fills important knowledge gaps in respective scientific fields. Also, with more availability of 3D point clouds of forests, this work can facilitate new studies on the utilization of TLS point clouds to manage and measure forests. We give main conclusions and contributions of chapter 2, chapter 3, and chapter 4 as follows.

The second chapter of this thesis looks at using TLS data coupled with machine learning algorithms to separate lianas from five tropical dry forest trees. Here, we demonstrate that XGBoosting algorithm (accuracy of 0.88, recall of 0.66) shows higher performance than RF algorithm (accuracy of 0.85, recall of 0.56) for liana and tree separation. We also find that the using the optimal radius search method (F1 score of 0.49) to obtain geometric features of liana and tree for classification, showing similar performance with the multiple radius search method (F1 score of 0.48) in Moorthy et al. (2019). The optimal radius search method chooses the scale with the largest difference between liana and tree

points across different radius, and this can avoid high dimensional data (D. Wang et al., 2018). We validate our method on the independent data from Moorthy et al. (2019). The RF model showed a high recall of 0.88, having a similar performance with Moorthy et al. (2019) after manual intervention (0.87). It is not surprise that our model is a more effective and reliable approach than the only one published liana/tree separation study, since our method build on and extend the previous work by Moorthy et al. (2019) and Thomas et al. (2018).

Chapter 2 contributes to providing a flexible approach to extract lianas from 3D point clouds. Since lianas continue to proliferate in tropical forests (Schnitzer et al., 2011), our method can facilitate new studies on evaluate the impact of lianas on tree and forest structure. The presented method can help to quantify liana abundance and biomass in forests, which are often ignored in allometric models (Miao et al., 2016). Also, our model could obtain a high recall of 0.88 on the independent data without any manual intervention, this indicates that researchers could have a good start point to have a liana-free point clouds. Then, those researchers could estimate other forest metrics such as tree volume or biomass in a more efficient way. Furthermore, the present method use geometric features to classify liana and tree, which means the presented method only needs xyz coordinates, indicating that the method can be easily applied to other sensors (Vicari et al., 2019; Moorthy et al., 2019; Tao et al., 2015). We provide the entire liana/tree separation model on (https://github.com/than2/liana-extraction) for scientific community to evaluate and make contributions for improvements of the model in the future.

In chapter 3, we present the utilization of a deep learning time-series approach coupled with TLS point clouds to separate leaf and woody component. The presented method uses a multiple radius search method to obtain the time-series of eigentropy, planarity and linearity for classification. Then, we use Fully Convolutional Neural Network (FCN), Long Short-Term Memory Fully Convolutional Neural Network (LSTM-FCN), and Residual Network (ResNet) to separate leaf and woody components, and compare the effect of univariable (UTS) and multivariable (MTS) time-series on classification accuracy. We find that the MTS based method (0.96) outperforms the UTS based method (0.67 to 0.88) concerning the classification accuracy. The MTS based model could combine all UTS features for classification, providing more patterns to understand data (L. Wang et al., 2016; Zheng et al., 2016). With regard to the UTS based model, we find the eigentropy (0.88) is the most advanced feature for leaf and wood classification, while the overall accuracy of planarity and linearity are 0.67 and 0.80, respectively. When testing our method on an independent dataset, the MTS models based on FCN, LSTM-FCN, and ResNet all demonstrate similar performance. Furthermore, we find that the utilization of CAM can help to explain the black-box effects of deep learning algorithms. We conclude that deep

learning algorithms coupled with time-series of the geometric features can accurately separate leaf and woody components from point clouds.

Chapter 3 contributes to providing a more applicable and transferable leaf and wood classification model than previous studies (Zhu et al., 2018; Moorthy et al., 2019). Accord- ing to Wei et al. (2013), the balanced training data can lead to the highest accuracy of the classification task, no matter what the proportion of the two classes. As such, we used balanced training data to classify leaf and wood, and then tested the presented model on the independent data. The MTS method of FCN shows F1 score (wood) of 0.79, without any additional postprocessing, higher than Moorthy et al. (2019) (average F1 score (wood) of 0.69), which used imbalanced training data for leaf/wood separation. Moreover, the total number of point per class for model training are 160,000, this value is much larger than Zhu et al. (2018) (1000 per class). The overall accuracy of Zhu et al. (2018) on validation data was 0.70, also lower than the presented model (FCN of 0.75). Those key results demonstrate that our model could provide a good start point for the scientific community to use TLS point clouds to measure forest metrics such as LAI and WAI. Furthermore, we provide the entire leaf and wood classification pipeline as an open-source python pack- age on Github (https://github.com/than2/leafandwood), enabling more studies on the evaluation of our model on other forest types (e.g., needle-leaf trees, dry, and rain forests).

In chapter 4, we propose a non-destructive approach to estimate buttresses volume using 3D point clouds from different databases distributed around the world. The alpha shape algorithm (ASA) and slice triangulation (ST) method shows a similar RRMSE (0.07 and 0.11 regardless of species effect, respectively) on buttress volume estimation. Meanwhile, the ASA is more accurate than ST when the tree present more and shallower horizontal buttresses. Concerning the allometric models, DAB (RRMSE of 0.23) and Darea130 (RRMSE of 0.21) show similar performance on an independent dataset. As such, we demonstrate that the ASA used in this chapter is a more applicable approach than ST and allometric models for buttresses volume estimation. Though the final shape of the ASA depend on the alpha value, DAB can be used as reference to select the optimal alpha value (R2 of 0.89). In addition, we understand that around 0.28 (±0.09) of buttress volume is underestimated if assuming a cylindrical volume with DAB. This underestimation value is lower than Noelke et al. (2015) of 0.35, and similar with Bauwens et al. (2017) of 0.26, while we use a total of 45 buttressed trees here to obtain this factor, more than the number of buttressed trees in the above two studies.

Chapter 4 contributes to develop a non-destructive and automatic method to estimate buttresses volume. As indicated in chapter 4, the ST method cannot capture the true shape when the trees

presented more and shallower horizontal buttresses. The scenario of buttresses in tropical forests can be diverse. The trees from SP database do give six examples that how the buttresses can be in tropical forests. Although the ASA tend to generate many overlap areas between the neighboring buttresses, the ASA did provide a solution to estimate the volume of those buttressed trees efficiently. Utilizing TLS or terrestrial photogrammetry in forests provides us with new insight to measure tree structure in a detailed 3D view (Calders et al., 2020). Therefore, irregularly large trees like buttressed trees can also be modeled using either ASA or ST with a higher accuracy than allometric models. Additionally, we use more numbers of buttressed trees (45) than Bauwens et al. (2017) and Noelke et al. (2015) to develop allometric models for volume estimation. More studies are expected to evaluate the performance of the proposed allometric models. Due to the under-representation of buttressed trees in past studies (Tang et al., 2011), our method could help to enhance present and previous estimates of volume and biomass of buttressed trees, which are keystone components to understanding biomass allocation and dynamics in tropical forests.

In this thesis, we explore the potential of TLS for forest monitoring using three different chapters. During the last decade, the development in close-range remote sensing technique such as TLS has fundamentally changed the way we measure the forests (Hackenberg et al., 2014; Lau et al., 2019; Liang et al., 2022). The most impactful change is that turning the traditional cost and inefficient manual forest data collections into automatic, detailed, efficient and comprehensive observations (Calders et al., 2015). The insufficient processing power, and geometric accuracy of data limited the advanced and practical applications of TLS (Calders et al., 2020). For example, occlusion effects and wind effects would directly affect the data quality (Béland et al., 2014). Also, the upper parts of forests are often missing in the point cloud especially in dense tropical forests (Disney et al., 2018). In such way, the fusion of TLS of Unmanned Aerial Vehicles Laser scanning can significantly reduce occlusions (Schneider et al., 2019). The chapters 2 and 3 would benefit from such fusional data. Also, it should mention that our methods in chapters 2 and 3 still involves manual intervention. Those semi-automatic extraction process is quite common in TLS field (Moorthy et al., 2019; Vicari et al., 2019). The robust automate processing power to extract structural features from 3D point clouds for is required for the broader application of TLS in forest monitoring. Also, automatic pipeline would lower the barrier for researchers from different fields to use the point clouds. We would also suggest that multiple spectral laser scanning (MS-LS) would also further the research in chapters 2 and 3. The radiometric information enable the quantification of leaf biochemical contents (water content, chlorophyll content., etc.) (Junttila et al., 2021). It should also mention that the method in chapter 4 is limited since we only evaluate 3D reconstruction method on buttressed part alone. Developing a method to rebuild the whole

trees with buttress may be of greater significance. In summary, the three chapters in this thesis focus on the individual tree level, how to develop methods to upscale our studies into large scale are important for future research, while the fusion of TLS with other laser scanning platforms (spaceborne, airborne, unmanned aerial vehicles) may provide solutions.

5.2 Challenges and future directions

Chapter 2 reveals the average F1 score of five tropical dry forest trees is 0.49. The main reason is that lianas have more irregular structure than trees, while they are generally smaller than almost all the stem and show similar diameter with branches. Future studies could investigate the utilization of more advanced algorithms to improve the performance for liana and tree classification, and therefore reduce the work on manual intervention. For example, recent development of Deep Learning algorithms provides a solution for classification on unstructured point clouds. Deep Learning algorithms apply an end-to- end approach to solve the problem, which means they could extract features automatically to build a classifier, and then making decisions on its own process (LeCun et al., 2015; Calders et al., 2020). Specifically, since we have the changing trend of liana and tree points from 0.05 to 1 m, time series classification can be used to separate them directly. Thus, there is no need to discard any radius, avoid losing information. Although time series classification may produce much more features, this problem can be well handled by deep learning algorithms such as Recurrent Neural Networks or Convolutional Neural Networks.

The balanced training data also brings the limitation of the presented method in Chapter 3. The leaf and wood points are still balanced in the prediction, which means the postprocessing steps are needed to further improve the classification results. Also, we use planarity, linearity and eigentropy in the MTS based model for classification, while the planarity and linearity features in the UTS method show worse performance than the eigentropy. As such, if more advanced features like the eigentropy are provided in the MTS method, the performance of our model for leaf/wood classification can be improved more. Future work could also focus on understanding the discriminatory regions of the time series for each class obtained by the CAM. In other word, the CAM may also provide a solid foundation for leaf/wood separation based on traditional machine learning algorithms. For example, machine learning algorithms could use the radius around the discriminatory regions of the time series for each class of CAM to classify leaf and wood, which may help to improve their performance.

Chapter 4 illustrates that the ASA tends to create many overlapped area when buttress structure is more complex. As such, with increasingly availability of 3D point clouds of forests, more accurate, automatic tree reconstruction method than the ASA are expected to further improve the volume estimation of buttresses. Also, DAB and Darea130 show similar performance on validation data, while we only use nine buttressed trees for validation. Therefore, we expect more studies in the future to assess the performance of the proposed allometric models. Future research could focus on developing an automatic reconstruction method that optimizes the entire tree, not focusing on the buttressed sections alone, which is also one of the limitations in chapter 4. In other word, the new algorithm can define the height of DAB, where the above and the below part can be modeled automatically, with the QSM cylinder fitting and ASA, respectively.

5.3 References

Bauwens, Sebastien, Adeline Fayolle, Sylvie Gourlet-Fleury, Leopold Mianda Ndjele, Coralie Mengal, and Philippe Lejeune (2017). “Terrestrial photogrammetry: a non-destructive method for modelling irregularly shaped tropical tree trunks”. In: *Methods in Ecology and Evolution* 8.4, pp. 460–471.

Béland, M., Baldocchi, D.D., Widlowski, J.L., Fournier, R.A., Verstraete, M.M., 2014. On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. Agric. For. Meteorol. 184, 82–97. https://doi.org/10.1016/j.agrformet.2013.09.005

Calders, Kim, Jennifer Adams, John Armston, Harm Bartholomeus, Sebastien Bauwens, Lisa Patrick Bentley, Jerome Chave, F Mark Danson, Miro Demol, Mathias Disney, et al. (2020). “Terrestrial laser scanning in forest ecology: Expanding the horizon”. In: Remote Sensing of Environment 251, p. 112102.

Calders, K., Schenkels, T., Bartholomeus, H., Armston, J., Verbesselt, J., Herold, M., 2015. Monitoring spring phenology with high temporal resolution terrestrial LiDAR measurements. Agric. For. Meteorol. 203, 158–168. https://doi.org/10.1016/j.agrformet.2015.01.009

Disney, M.I., Boni Vicari, M., Burt, A., Calders, K., Lewis, S.L., Raumonen, P., Wilkes, P., 2018. Weighing trees with lasers: Advances, challenges and opportunities. Interface Focus 8, 20170048. https://doi.org/10.1098/rsfs.2017.0048

Hackenberg, J., Morhart, C., Sheppard, J., Spiecker, H., Disney, M., 2014. Highly accurate tree models derived from terrestrial laser scan data: A method description. Forests 5, 1069–1105. https://doi.org/10.3390/f5051069

Junttila, S., Hölttä, T., Puttonen, E., Katoh, M., Vastaranta, M., Kaartinen, H., Holopainen, M., Hyyppä, H., 2021. Terrestrial laser scanning intensity captures diurnal variation in leaf water potential. Remote Sens. Environ. 255. https://doi.org/10.1016/j.rse.2020.112274

Lau, A., Calders, K., Bartholomeus, H., Martius, C., Raumonen, P., Herold, M., Vicari, M., Sukhdeo, H., Singh, J., Goodman, R.C., 2019. Tree biomass equations from terrestrial LiDAR: A case study in Guyana. Forests 10, 1–18. https://doi.org/10.3390/f10060527

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). “Deep learning”. In: nature 521.7553, pp. 436–444.

Liang, X., Kukko, A., Balenovic, I., Saarinen, N., Junttila, S., Kankare, V., Holopainen, M., Mokros, M., Surovy, P., Kaartinen, H., Jurjevic, L., Honkavaara, E., Nasi, R., Liu, J., Hollaus, M., Tian, J., Yu, X., Pan, J., Cai, S., Virtanen, J.-P., Wang, Y., Hyyppa, J., 2022. Close-Range Remote Sensing of Forests: The state of the art, challenges, and opportunities for systems and data acquisitions. IEEE Geosci. Remote Sens. Mag. 10, 32–71. https://doi.org/10.1109/MGRS.2022.3168135

Miao, Zewei, Sally E Koerner, Vincent P Medjibe, and John R Poulsen (2016). “Wanted: new allometric equations for large lianas and African lianas”. In: Biotropica 48.5, pp. 561–564.

Moorthy, Sruthi MK, Kim Calders, Matheus B Vicari, and Hans Verbeeck (2019). “Im- proved supervised learning-based approach for leaf and wood classification from LiDAR point clouds of forests”. In: IEEE Transactions on Geoscience and Remote Sensing 58.5, pp. 3057–3070.

Moorthy, Sruthi MK, Yunfei Bao, Kim Calders, Stefan A Schnitzer, and Hans Verbeeck (2019). “Semi-automatic extraction of liana stems from terrestrial LiDAR point clouds of tropical rainforests”. In: ISPRS Journal of Photogrammetry and Remote Sensing 154, pp. 114–126.

Noelke, Nils, Lutz Fehrmann, Surati Jaya I Nengah, Tatang Tiryana, Dominik Seidel, and Christoph Kleinn (2015). “On the geometry and allometry of big-buttressed trees-a challenge for forest monitoring: new insights from 3D-modeling with terrestrial laser scanning”. In: iForest-Biogeosciences and Forestry 8.5, p. 574.

Schnitzer, Stefan A and Frans Bongers (2011). “Increasing liana abundance and biomass in tropical forests: emerging patterns and putative mechanisms”. In: Ecology letters 14.4, pp. 397–406.

Schneider, F.D., Kükenbrink, D., Schaepman, M.E., Schimel, D.S., Morsdorf, F., 2019. Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR. Agric. For. Meteorol. 268, 249–257. https://doi.org/10.1016/j.agrformet.2019.01.033

Tang, Yong, Xiaofei Yang, Min Cao, Carol C Baskin, and Jerry M Baskin (2011). “Buttress trees elevate soil heterogeneity and regulate seedling diversity in a tropical rainforest”. In: Plant and soil 338.1, pp. 301–309.

Tao, Shengli, Qinghua Guo, Shiwu Xu, Yanjun Su, Yumei Li, and Fangfang Wu (2015). “A geometric method for wood-leaf separation using terrestrial and simulated lidar data”. In: Photogrammetric Engineering & Remote Sensing 81.10, pp. 767–776.

Thomas, Hugues, Fran¸cois Goulette, Jean-Emmanuel Deschaud, Beatriz Marcotegui, and Yann LeGall (2018). “Semantic classification of 3D point clouds with multiscale spher- ical neighborhoods”. In: 2018 International conference on 3D vision (3DV). IEEE, pp. 390–398.

Vicari, Matheus B, Mathias Disney, Phil Wilkes, Andrew Burt, Kim Calders, and William Woodgate (2019). “Leaf and wood classification framework for terrestrial LiDAR point clouds”. In: Methods in Ecology and Evolution 10.5, pp. 680–694.

Wang, Di, Jasmin Brunner, Zhenyu Ma, Hao Lu, Markus Hollaus, Yong Pang, and Norbert Pfeifer (2018). “Separating tree photosynthetic and non-photosynthetic components from point cloud data using dynamic segment merging”. In: Forests 9.5, p. 252.

Wang, Lin, Zhigang Wang, and Shan Liu (2016). “An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm”. In: Expert Systems with Applications 43, pp. 237–249.

Wei, Qiong and Roland L Dunbrack Jr (2013). “The role of balanced training and testing data sets for binary classifiers in bioinformatics”. In: PloS one 8.7, e67863.

Zheng, Yi, Qi Liu, Enhong Chen, Yong Ge, and J Leon Zhao (2016). “Exploiting multi- channels deep convolutional neural networks for multivariate time series classification”. In: Frontiers of Computer Science 10.1, pp. 96–112.

Zhu, Xi, Andrew K Skidmore, Roshanak Darvishzadeh, K Olaf Niemann, Jing Liu, Yi- fang Shi, and Tiejun Wang (2018). “Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest”. In: International journal of applied earth observation and geoinformation 64, pp. 43–50.

Bibliography

Alder, Denis and Timothy J Synnott (1992). Permanent sample plot techniques for mixed tropical forest. Oxford Forestry Institute, University of Oxford.

Bagnall, Anthony, Hoang Anh Dau, Jason Lines, Michael Flynn, James Large, Aaron Bostrom, Paul Southam, and Eamonn Keogh (2018). "The UEA multivariate time series classification archive, 2018". In: arXiv preprint arXiv:1811.00075.

Bagnall, Anthony, Jason Lines, Aaron Bostrom, James Large, and Eamonn Keogh (2017). "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances". In: Data mining and knowledge discovery 31.3, pp. 606– 660.

Bao, Yunfei, Sruthi MK Moorthy, and Hans Verbeeck (2018). "Towards extraction of lianas from terrestrial lidar scans of tropical forests". In: IGARSS 2018-2018 IEEE Interna- tional Geoscience and Remote Sensing Symposium. IEEE, pp. 7544–7547.

Barbeito, Ignacio, Mathieu Dassot, Dominik Bayer, Catherine Collet, Lars Dr¨ossler, Mag- nus L¨of, Miren del Rio, Ricardo Ruiz-Peinado, David I Forrester, Andr´es Bravo-Oviedo, et al. (2017). "Terrestrial laser scanning reveals differences in crown structure of Fagus sylvatica in mixed vs. pure European forests". In: Forest Ecology and Management 405, pp. 381–390.

Basuki, TM, PE Van Laake, AK Skidmore, and YA Hussin (2009). “Allometric equations for estimating the above-ground biomass in tropical lowland Dipterocarp forests”. In: Forest ecology and management 257.8, pp. 1684–1694.

Bauwens, Sebastien, Adeline Fayolle, Sylvie Gourlet-Fleury, Leopold Mianda Ndjele, Coralie Mengal, and Philippe Lejeune (2017). “Terrestrial photogrammetry: a non- destructive method for modelling irregularly shaped tropical tree trunks”. In: Methods in Ecology and Evolution 8.4, pp. 460–471.

Baydogan, Mustafa Gokce, George Runger, and Eugene Tuv (2013). “A bag-of-features framework to classify time series”. In: IEEE transactions on pattern analysis and ma- chine intelligence 35.11, pp. 2796–2802.

Beland, Martin, Dennis D Baldocchi, Jean-Luc Widlowski, Richard A Fournier, and Michel M Verstraete (2014). “On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR”. In: Agricultural and Forest Meteorology 184, pp. 82–97.

Beland, Martin, Jean-Luc Widlowski, Richard A Fournier, Jean-Fran¸cois Cˆot´e, and Michel M Verstraete (2011). “Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements”. In: Agricultural and Forest Meteorology 151.9, pp. 1252–1266.

Belton, David, Simon Moncrieff, and Jane Chapman (2013). “Processing tree point clouds using Gaussian Mixture Models”. In: ISPRS Annals of Photogrammetry, Remote Sens- ing and Spatial Information Sciences. Presented at the WG 3, pp. 43–48.

Berger, Ambros, Thomas Gschwantner, Ronald E McRoberts, and Klemens Schadauer (2014). “Effects of measurement errors on individual tree stem volume estimates for the Austrian National Forest Inventory”. In: Forest Science 60.1, pp. 14–24.

Bergesen, Helge Ole, Georg Parmann, and Oystein B. Thommessen (2019). Food and Agri- culture Organization (FAO). FAO, pp. 201–201. isbn: 9789251305614. doi: 10.4324/ 9781315066547-55.

Bonneau, David, Paul-Mark DiFrancesco, and D Jean Hutchinson (2019). “Surface reconstruction for three-dimensional rockfall volumetric analysis”. In: ISPRS International Journal of Geo-Information 8.12, p. 548.

Burt, Andrew, Mathias Disney, and Kim Calders (2019). “Extracting individual trees from lidar point clouds using treeseg”. In: Methods in Ecology and Evolution 10.3, pp. 438– 445.

Calders, Kim, Jennifer Adams, John Armston, Harm Bartholomeus, Sebastien Bauwens, Lisa Patrick Bentley, Jerome Chave, F Mark Danson, Miro Demol, Mathias Disney, et al. (2020). “Terrestrial laser scanning in forest ecology: Expanding the horizon”. In: Remote Sensing of Environment 251, p. 112102.

Calders, Kim, Glenn Newnham, Andrew Burt, Simon Murphy, Pasi Raumonen, Martin Herold, Darius Culvenor, Valerio Avitabile, Mathias Disney, John Armston, et al. (2015). “Nondestructive estimates of above-ground biomass using terrestrial laser scan- ning”. In: Methods in Ecology and Evolution 6.2, pp. 198–208.

Calders, Kim, Glenn Newnham, Martin Herold, Simon Murphy, Darius Culvenor, Pasi Rau- monen, Andrew Burt, John Armston, Valerio Avitabile, and Mathias Disney (2013). “Estimating above ground biomass from terrestrial laser scanning in Australian Eu- calypt Open Forest”. In: Proceedings SilviLaser 2013, 9-11 October, Beijing, China, pp. 90–97.

Calvo-Rodriguez, Sofia, Ralf Kiese, and G Arturo S´anchez-Azofeifa (2020). “Seasonality and budgets of soil greenhouse gas emissions from a tropical dry forest successional gradient in Costa Rica”. In: Journal of Geophysical Research: Biogeosciences 125.9, e2020JG005647.

Chapman, Colin A, Les Kaufman, and Lauren J Chapman (1998). “Buttress formation and directional stress experienced during critical phases of tree development”. In: Journal of Tropical Ecology 14.3, pp. 341–349.

Chave, Jerome, Christophe Andalo, Sandra Brown, Michael A Cairns, Jeffrey Q Chambers, Derek Eamus, Horst Fo¨lster, Fran¸cois Fromard, Niro Higuchi, Tatuo Kira, et al. (2005). “Tree allometry and improved estimation of carbon stocks and balance in tropical forests”. In: Oecologia 145.1, pp. 87–99.

Chave, Jerome, Maxime Rejou-Mechain, Alberto Burquez, Emmanuel Chidumayo, Matthew S Colgan, Welington BC Delitti, Alvaro Duque, Tron Eid, Philip M Fearnside, Rosa C Goodman, et al. (2014). “Improved allometric models to estimate the above- ground biomass of tropical trees”. In: Global change biology 20.10, pp. 3177–3190.

Chen, Tianqi and Carlos Guestrin (2016). “XGBoost: A scalable tree boosting system In Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,(pp. 785–794)”. In: New York, NY, USA: ACM 10.2939672.2939785.

Clark, David B and James R Kellner (2012). “Tropical forest biomass estimation and the fallacy of misplaced concreteness”. In: Journal of Vegetation Science 23.6, pp. 1191– 1196.

Clark, Deborah A (2002). “Are tropical forests an important carbon sink? Reanalysis of the long-term plot data”. In: Ecological Applications 12.1, pp. 3–7.

Cola¸co, Andre F, Rodrigo G Trevisan, Jose P Molin, Joan R Rosell-Polo, et al. (2017). “A method to obtain orange crop geometry information using a mobile terrestrial laser scanner and 3D modeling”. In: Remote Sensing 9.8, p. 763.

Condit, Richard (1998). Tropical forest census plots: methods and results from Barro Col- orado Island, Panama and a comparison with other plots. Springer Science & Business Media.

Cote Jean-Francois, Richard A Fournier, Gordon W Frazer, and K Olaf Niemann (2012). “A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure”. In: Agricultural and forest meteorology 166, pp. 72–85.

Cushman, KC, Sarayudh Bunyavejchewin, Dairon Cardenas, Richard Condit, Stuart J Davies, A´lvaro Duque, Stephen P Hubbell, Somboon Kiratiprayoon, Shawn KY Lum, and Helene C Muller-Landau (2021). “Variation in trunk taper of buttressed trees within and among five lowland tropical forests”. In: Biotropica 53.5, pp. 1442–1453.

Cushman, KC, Helene C Muller-Landau, Richard S Condit, and Stephen P Hubbell (2014). “Improving estimates of biomass change in buttressed trees using tree taper models”. In: Methods in Ecology and Evolution 5.6, pp. 573–582.

Dargan, Shaveta, Munish Kumar, Maruthi Rohit Ayyagari, and Gulshan Kumar (2020). “A survey of deep learning and its applications: a new paradigm to machine learning”. In: Archives of Computational Methods in Engineering 27.4, pp. 1071–1092.

Dau, Hoang Anh, Anthony Bagnall, Kaveh Kamgar, Chin-Chia Michael Yeh, Yan Zhu, Shaghayegh Gharghabi, Chotirat Ann Ratanamahatana, and Eamonn Keogh (2019). “The UCR time series archive”. In: IEEE/CAA Journal of Automatica Sinica 6.6, pp. 1293–1305.

Dean, Christopher and SH Roxburgh (2006). “Improving visualisation of mature, high- carbon-sequestering forests”. In: Forest Biometry, Modelling and Information Sciences 1, pp. 48–69.

Delaunay, Boris, S Vide, A Lam´emoire, and V De Georges (1934). “Bulletin de l’Acad´emie des Sciences de l’URSS”. In: Classe des sciences math´ematiques et naturelles 6, pp. 793– 800.

Demantk´e, J´erˆome, Cl´ement Mallet, Nicolas David, and Bruno Vallet (2011). “Dimension- ality based scale selection in 3D lidar point clouds”. In: Laserscanning.

Dewalt, Saara J, Stefan A Schnitzer, and Julie S Denslow (2000). “Density and diversity of lianas along a chronosequence in a central Panamanian lowland forest”. In: Journal of Tropical Ecology 16.1, pp. 1–19.

Disney, Mathias I, Matheus Boni Vicari, Andrew Burt, Kim Calders, Simon L Lewis, Pasi Raumonen, and Phil Wilkes (2018). “Weighing trees with lasers: advances, challenges and opportunities”. In: Interface Focus 8.2, p. 20170048.

Dou, Peng, Huanfeng Shen, Zhiwei Li, and Xiaobin Guan (2021). “Time series remote sensing image classification framework using combination of deep learning and mul- tiple classifiers system”. In: International Journal of Applied Earth Observation and Geoinformation 103, p. 102477.

Duran, Sandra M and Ernesto Gianoli (2013). “Carbon stocks in tropical forests decrease with liana density”. In: Biology letters 9.4, p. 20130301.

Edelsbrunner, Herbert and Ernst P Mu¨cke (1994). “Three-dimensional alpha shapes”. In:
ACM Transactions on Graphics (TOG) 13.1, pp. 43–72.

Esling, Philippe and Carlos Agon (2012). “Time-series data mining”. In: ACM Computing Surveys (CSUR) 45.1, pp. 1–34.

Fawaz, Hassan Ismail, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller (2019). “Deep learning for time series classification: a review”. In: Data mining and knowledge discovery 33.4, pp. 917–963.

Feliciano, Emanuelle A, Shimon Wdowinski, and Matthew D Potts (2014). “Assessing mangrove above-ground biomass and structure using terrestrial laser scanning: A case study in the Everglades National Park”. In: Wetlands 34.5, pp. 955–968.

Ferrara, Roberto, Salvatore GP Virdis, Andrea Ventura, Tiziano Ghisu, Pierpaolo Duce, and Grazia Pellizzaro (2018). “An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN”. In: Agricultural and forest meteorology 262, pp. 434–444.

Ferreira, Artur J and M´ario AT Figueiredo (2012). “Efficient feature selection filters for high-dimensional data”. In: Pattern Recognition Letters 33.13, pp. 1794–1804.

Gamboa, John Cristian Borges (2017). “Deep learning for time-series analysis”. In: arXiv preprint arXiv:1701.01887.

Gao, Yan, Alexander Quevedo, Zoltan Szantoi, and Margaret Skutsch (2021). “Monitoring forest disturbance using time-series MODIS NDVI in Michoac´an, Mexico”. In: Geocarto International 36.15, pp. 1768–1784.

Gentry, ALWYN H (1991). “The distribution and evolution of climbing plants”. In: The biology of vines 351.

Gonzalez de Tanago, Jose, Alvaro Lau, Harm Bartholomeus, Martin Herold, Valerio Avitabile, Pasi Raumonen, Christopher Martius, Rosa C Goodman, Mathias Disney, Solichin Manuri, et al. (2018). “Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR”. In: Methods in Ecology and Evolution 9.2, pp. 223–234.

Gourlet-Fleury, Sylvie, Fr´ed´eric Mortier, Adeline Fayolle, Fid`ele Baya, Dakis Ou´edraogo, Fabrice B´en´edet, and Nicolas Picard (2013). “Tropical forest recovery from logging: a 24 year silvicultural experiment from Central Africa”. In: Philosophical Transactions of the Royal Society B: Biological Sciences 368.1625, p. 20120302.

Guzm´an Q, J Antonio, Benoit Rivard, G Arturo S´anchez-Azofeifa, et al. (2018). “Discrimi- nation of liana and tree leaves from a Neotropical Dry Forest using visible-near infrared and longwave infrared reflectance spectra”. In: Remote Sensing of Environment 219, pp. 135–144.

Guzman Q, J Antonio, Hernandez Ronny, and Sanchez-Azofeifa Arturo (2021). rTLS: Tools to Process Point Clouds Derived from Terrestrial Laser Scanning.

Guzman Q, J Antonio, Iain Sharp, Felipe Alencastro, and G Arturo S´anchez-Azofeifa (2020). “On the relationship of fractal geometry and tree–stand metrics on point clouds derived from terrestrial laser scanning”. In: Methods in Ecology and Evolution 11.10, pp. 1309–1318.

Hackenberg, Jan, Marc Wassenberg, Heinrich Spiecker, and Dongjing Sun (2015). “Non destructive method for biomass prediction combining TLS derived tree volume and wood density”. In: Forests 6.4, pp. 1274–1300.

Hadas, Edyta, Andrzej Borkowski, Javier Estornell, and Przemyslaw Tymkow (2017). “Au- tomatic estimation of olive tree dendrometric parameters based on airborne laser scan- ning data using alpha-shape and principal component analysis”. In: GIScience & Re- mote Sensing 54.6, pp. 898–917.

Hansen, Endre Hofstad, Terje Gobakken, and Erik Næsset (2015). “Effects of pulse density on digital terrain models and canopy metrics using airborne laser scanning in a tropical rainforest”. In: Remote Sensing 7.7, pp. 8453–8468.

Harding, David J and Claudia C Carabajal (2005). “ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure”. In: Geophysical research letters 32.21.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2016). “Deep residual learning for image recognition”. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Helmer, EH and MA Lefsky (2006). “Forest canopy heights in Amazon River basin forests as estimated with the Geoscience Laser Altimeter System (GLAS)”. In: In: Aguirre- Bravo, C.; Pellicane, Patrick J.; Burns, Denver P.; and Draggan, Sidney, Eds. 2006. Monitoring Science and Technology Symposium: Unifying Knowledge for Sustainability in the Western Hemisphere Proceedings RMRS-P-42CD. Fort Collins, CO: US De- partment of Agriculture, Forest Service, Rocky Mountain Research Station. p. 802-808. Vol. 42.

Hochreiter, Sepp and Ju¨rgen Schmidhuber (1997). “Long short-term memory”. In: Neural computation 9.8, pp. 1735–1780.

Holmgren, Johan and ˚Asa Persson (2004). “Identifying species of individual trees using airborne laser scanner”. In: Remote Sensing of Environment 90.4, pp. 415–423.

Holopainen, Markus, Mikko Vastaranta, Ville Kankare, Minna R¨aty, Matti Vaaja, Xinlian Liang, Xiaowei Yu, Juha Hyypp¨a, Hannu Hyyppa¨, R Viitala, et al. (2011). “Biomass estimation of individual trees using stem and crown diameter TLS measurements”. In: ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 3812, pp. 91–95.

Hosoi, Fumiki and Kenji Omasa (2006). “Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar”. In: IEEE transactions on geoscience and remote sensing 44.12, pp. 3610–3618.

Hussin, Y and Witske Bijker (2000). “Inventory of remote sensing applications in forestry for sustainable management”. In: International Archives of Photogrammetry and Re- mote Sensing 33.B7/2; PART 7, pp. 575–579.

Ingwell, Laura L, S Joseph Wright, Kristen K Becklund, Stephen P Hubbell, and Stefan A Schnitzer (2010). “The impact of lianas on 10 years of tree growth and mortality on Barro Colorado Island, Panama”. In: Journal of Ecology 98.4, pp. 879–887.

Kankare, Ville, Marianna Joensuu, Jari Vauhkonen, Markus Holopainen, Topi Tanhuanp¨a¨a, Mikko Vastaranta, Juha Hyypp¨a, Hannu Hyyppa¨, Petteri Alho, Juha Rikala, et al. (2014). “Estimation of the timber quality of Scots pine with terrestrial laser scanning”. In: Forests 5.8, pp. 1879–1895.

Kankare, Ville, Mikko Vastaranta, Markus Holopainen, Minna R¨aty, Xiaowei Yu, Juha Hyyppa¨, Hannu Hyyppa¨, Petteri Alho, and Risto Viitala (2013). “Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR”. In: Remote Sensing 5.5, pp. 2257–2274.

Karim, Fazle, Somshubra Majumdar, Houshang Darabi, and Shun Chen (2017). “LSTM fully convolutional networks for time series classification”. In: IEEE access 6, pp. 1662– 1669.

Karim, Fazle, Somshubra Majumdar, Houshang Darabi, and Samuel Harford (2019). “Mul- tivariate LSTM-FCNs for time series classification”. In: Neural Networks 116, pp. 237– 245.

Keogh, Eamonn and Chotirat Ann Ratanamahatana (2005). “Exact indexing of dynamic time warping”. In: Knowledge and information systems 7.3, pp. 358–386.

Kingma, Diederik P and Jimmy Ba (2014). “Adam: A method for stochastic optimization”.
In: arXiv preprint arXiv:1412.6980.

Koenig, Kristina, Bernhard Ho¨fle, Martin Ha¨mmerle, Thomas Jarmer, Bastian Sieg- mann, and Holger Lilienthal (2015). “Comparative classification analysis of post-harvest growth detection from terrestrial LiDAR point clouds in precision agriculture”. In: IS- PRS Journal of Photogrammetry and Remote Sensing 104, pp. 112–125.

Lafarge, T and B Pateiro-Lopez (2014). alphashape3d: Implementation of the 3D Alpha- Shape for the Reconstruction of 3D Sets from a Point Cloud.

Lakicevic, Milena, Nicholas Povak, and Keith M Reynolds (2020). Introduction to R for Terrestrial Ecology. Springer.

Lau, Alvaro, Lisa Patrick Bentley, Christopher Martius, Alexander Shenkin, Harm Bartholomeus, Pasi Raumonen, Yadvinder Malhi, Tobias Jackson, and Martin Herold (2018). “Quantifying branch architecture of tropical trees using terrestrial LiDAR and 3D modelling”. In: Trees 32.5, pp. 1219–1231.

Lau, Alvaro, Kim Calders, Harm Bartholomeus, Christopher Martius, Pasi Raumonen, Martin Herold, Matheus Vicari, Hansrajie Sukhdeo, Jeremy Singh, and Rosa C Good- man (2019). “Tree biomass equations from terrestrial LiDAR: A case study in Guyana”. In: Forests 10.6, p. 527.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). “Deep learning”. In: nature 521.7553, pp. 436–444.

Letcher, Susan G and Robin L Chazdon (2009). “Lianas and self-supporting plants during tropical forest succession”. In: Forest Ecology and Management 257.10, pp. 2150–2156.

Lewis, Simon L, Gabriela Lopez-Gonzalez, Bonaventure Sonk´e, Kofi Affum-Baffoe, Tim- othy R Baker, Lucas O Ojo, Oliver L Phillips, Jan M Reitsma, Lee White, James A Comiskey, et al. (2009). “Increasing carbon storage in intact African tropical forests”. In: Nature 457.7232, pp. 1003–1006.

Li, Zhan, Michael Schaefer, Alan Strahler, Crystal Schaaf, and David Jupp (2018a). “On the utilization of novel spectral laser scanning for three-dimensional classification of vegetation elements”. In: Interface Focus 8.2, p. 20170039.

Li, Zhan, Alan Strahler, Crystal Schaaf, David Jupp, Michael Schaefer, and Pontus Olof- sson (2018b). “Seasonal change of leaf and woody area profiles in a midlatitude decid- uous forest canopy from classified dual-wavelength terrestrial lidar point clouds”. In: Agricultural and Forest Meteorology 262, pp. 279–297.

Liang, Xinlian, Juha Hyyppa¨, Harri Kaartinen, Markus Holopainen, and Timo Melkas (2012). “Detecting changes in forest structure over time with bi-temporal terrestrial laser scanning data”. In: ISPRS International Journal of Geo-Information 1.3, pp. 242– 255.

Liang, Xinlian, Juha Hyyppa¨, Harri Kaartinen, Matti Lehtom¨aki, Jiri Py¨or¨al¨a, Norbert Pfeifer, Markus Holopainen, G´abor Brolly, Pirotti Francesco, Jan Hackenberg, et al. (2018). “International benchmarking of terrestrial laser scanning approaches for forest inventories”. In: ISPRS journal of photogrammetry and remote sensing 144, pp. 137– 179.

Liang, Xinlian, Ville Kankare, Juha Hyyppa¨, Yunsheng Wang, Antero Kukko, Henrik Haggr´en, Xiaowei Yu, Harri Kaartinen, Anttoni Jaakkola, Fengying Guan, et al. (2016). “Terrestrial laser scanning in forest inventories”. In: ISPRS Journal of Photogrammetry and Remote Sensing 115, pp. 63–77.

Lim, Seng Poh and Habibollah Haron (2014). “Surface reconstruction techniques: a review”. In: Artificial Intelligence Review 42.1, pp. 59–78.

Lin, Min, Qiang Chen, and Shuicheng Yan (2013). “Network in network”. In: arXiv preprint arXiv:1312.4400.

Lines, Jason and Anthony Bagnall (2015). “Time series classification with ensembles of elastic distance measures”. In: Data Mining and Knowledge Discovery 29.3, pp. 565– 592.

Liu, Jia, Yuming Wu, Xing Gao, and Xuehua Zhang (2022). “A Simple Method of Mapping Landslides Runout Zones Considering Kinematic Uncertainties”. In: Remote Sensing 14.3, p. 668.

Londre, Ronald A and Stefan A Schnitzer (2006). “The distribution of lianas and their change in abundance in temperate forests over the past 45 years”. In: Ecology 87.12, pp. 2973–2978.

Ma, Lixia, Guang Zheng, Jan UH Eitel, Troy S Magney, and L Monika Moskal (2016). “Determining woody-to-total area ratio using terrestrial laser scanning (TLS)”. In: Agricultural and forest meteorology 228, pp. 217–228.

Ma, Lixia, Guang Zheng, Jan UH Eitel, L Monika Moskal, Wei He, and Huabing Huang (2015). “Improved salient feature-based approach for automatically separating photo- synthetic and nonphotosynthetic components within terrestrial lidar point cloud data of forest canopies”. In: IEEE Transactions on geoscience and remote sensing 54.2, pp. 679–696.

Main-Knorn, Magdalena, Warren B Cohen, Robert E Kennedy, Wojciech Grodzki, Dirk Pflugmacher, Patrick Griffiths, and Patrick Hostert (2013). “Monitoring coniferous for- est biomass change using a Landsat trajectory-based approach”. In: Remote Sensing of Environment 139, pp. 277–290.

Mart´ınez-Izquierdo, Laura, Mar´ıa M Garc´ıa, Jennifer S Powers, and Stefan A Schnitzer (2016). “Lianas suppress seedling growth and survival of 14 tree species in a Panamanian tropical forest”. In: Ecology 97.1, pp. 215–224.

Mehdiyev, Nijat, Johannes Lahann, Andreas Emrich, David Enke, Peter Fettke, and Peter Loos (2017). “Time series classification using deep learning for process planning: A case from the process industry”. In: Procedia Computer Science 114, pp. 242–249.

Mehedi, Md Abu Hanifa, Chandan Kundu, and Md Qumruzzaman Chowdhury (2012). “Patterns of tree buttressing at Lawachara National Park, Bangladesh”. In: Journal of Forestry Research 23.3, pp. 461–466.

Metcalf, C Jessica E, James S Clark, and Deborah A Clark (2009). “Tree growth inference and prediction when the point of measurement changes: modelling around buttresses in tropical forests”. In: Journal of Tropical Ecology 25.1, pp. 1–12.

Miao, Zewei, Sally E Koerner, Vincent P Medjibe, and John R Poulsen (2016). “Wanted: new allometric equations for large lianas and African lianas”. In: Biotropica 48.5, pp. 561–564.

Momo Takoudjou, St´ephane, Pierre Ploton, Bonaventure Sonk´e, Jan Hackenberg, S´ebastien Griffon, Fran¸cois De Coligny, Narcisse Guy Kamdem, Moses Libalah, Gislain II Mofack, Gilles Le Mogu´edec, et al. (2018). “Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach”. In: Methods in Ecology and Evolution 9.4, pp. 905–916.

Moorthy, Sruthi MK, Kim Calders, Manfredo Di Porcia e Brugnera, Stefan A Schnitzer, and Hans Verbeeck (2018). “Terrestrial laser scanning to detect liana impact on forest structure”. In: Remote Sensing 10.6, p. 810.

Moorthy, Sruthi MK, Kim Calders, Matheus B Vicari, and Hans Verbeeck (2019). “Im- proved supervised learning-based approach for leaf and wood classification from LiDAR point clouds of forests”. In: IEEE Transactions on Geoscience and Remote Sensing 58.5, pp. 3057–3070.

Moorthy, Sruthi MK, Yunfei Bao, Kim Calders, Stefan A Schnitzer, and Hans Verbeeck (2019). “Semi-automatic extraction of liana stems from terrestrial LiDAR point clouds of tropical rainforests”. In: ISPRS Journal of Photogrammetry and Remote Sensing 154, pp. 114–126.

Muff, Stefanie, Erlend B Nilsen, Robert B O’Hara, and Chlo´e R Nater (2021). “Rewriting results sections in the language of evidence”. In: Trends in ecology & evolution.

Newbery, David M, Sarah Schwan, George B Chuyong, and Xander M van der Burgt (2009). “Buttress form of the central African rain forest tree Microberlinia bisulcata, and its possible role in nutrient acquisition”. In: Trees 23.2, pp. 219–234.

Newnham, Glenn J, John D Armston, Kim Calders, Mathias I Disney, Jenny L Lovell, Crystal B Schaaf, Alan H Strahler, and F Mark Danson (2015). “Terrestrial laser scan- ning for plot-scale forest measurement”. In: Current Forestry Reports 1.4, pp. 239–251.

Ngomanda, Alfred, Quentin Moundounga Mavouroulou, Nestor Laurier Engone Obiang, Donald Midoko Iponga, Jacques-Fran¸cois Mavoungou, Nicaise L´epengu´e, Nicolas Pi- card, and Bertrand Mbatchi (2012). “Derivation of diameter measurements for but- tressed trees, an example from Gabon”. In: Journal of tropical ecology 28.3, pp. 299– 302.

Noelke, Nils, Lutz Fehrmann, Surati Jaya I Nengah, Tatang Tiryana, Dominik Seidel, and Christoph Kleinn (2015). “On the geometry and allometry of big-buttressed trees-a challenge for forest monitoring: new insights from 3D-modeling with terrestrial laser scanning”. In: iForest-Biogeosciences and Forestry 8.5, p. 574.

Nogueira, Euler Melo, Bruce Walker Nelson, and Philip M Fearnside (2006). “Volume and biomass of trees in central Amazonia: influence of irregularly shaped and hollow trunks”. In: Forest Ecology and Management 227.1-2, pp. 14–21.

Nweke, Henry Friday, Ying Wah Teh, Mohammed Ali Al-Garadi, and Uzoma Rita Alo (2018). “Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges”. In: Expert Systems with Applications 105, pp. 233–261.

Oguiza, Ignacio (2020). “tsai-A state-of-the-art deep learning library for time series and sequential data”. In: Retrieved April 20, p. 2021.

Ordonez, Francisco Javier and Daniel Roggen (2016). “Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition”. In: Sensors 16.1, p. 115.

Pandey, CB, Lalita Singh, and SK Singh (2011). “Buttresses induced habitat heterogeneity increases nitrogen availability in tropical rainforests”. In: Forest ecology and manage- ment 262.9, pp. 1679–1685.

Pfeffer, WF (1986). “The divergence theorem”. In: Transactions of the American Mathe- matical Society 295.2, pp. 665–685.

Pueschel, Pyare, Glenn Newnham, Gilles Rock, Thomas Udelhoven, Willy Werner, and Joachim Hill (2013). “The influence of scan mode and circle fitting on tree stem de- tection, stem diameter and volume extraction from terrestrial laser scans”. In: ISPRS journal of photogrammetry and remote sensing 77, pp. 44–56.

Qi, Charles R, Li Yi, Hao Su, and Leonidas J Guibas (2017). “Pointnet++: Deep hierarchical feature learning on point sets in a metric space”. In: arXiv preprint arXiv:1706.02413.

Quinonero-Candela, Joaquin, Masashi Sugiyama, Neil D Lawrence, and Anton Schwaighofer (2009). Dataset shift in machine learning. Mit Press.

R Core Team, other (2021). “R: A language and environment for statistical computing”. In: R project.

Rajkomar, Alvin, Eyal Oren, Kai Chen, Andrew M Dai, Nissan Hajaj, Michaela Hardt, Peter J Liu, Xiaobing Liu, Jake Marcus, Mimi Sun, et al. (2018). “Scalable and accurate deep learning with electronic health records”. In: NPJ Digital Medicine 1.1, pp. 1–10.

Raskutti, Garvesh, Martin J Wainwright, and Bin Yu (2014). “Early stopping and non- parametric regression: an optimal data-dependent stopping rule”. In: The Journal of Machine Learning Research 15.1, pp. 335–366.

Raumonen, Pasi, Mikko Kaasalainen, Markku ˚Akerblom, Sanna Kaasalainen, Harri Kaarti- nen, Mikko Vastaranta, Markus Holopainen, Mathias Disney, and Philip Lewis (2013). “Fast automatic precision tree models from terrestrial laser scanner data”. In: Remote Sensing 5.2, pp. 491–520.

Rhys, Hefin (2020). Machine Learning with R, the tidyverse, and mlr. Simon and Schuster.

Rodrıguez-Ronderos, M Elizabeth, Gil Bohrer, Arturo Sanchez-Azofeifa, Jennifer S Powers, and Stefan A Schnitzer (2016). “Contribution of lianas to plant area index and canopy structure in a Panamanian forest”. In: Ecology 97.12, pp. 3271–3277.

Ruiz, Alejandro Pasos, Michael Flynn, James Large, Matthew Middlehurst, and Anthony Bagnall (2021). “The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances”. In: Data Mining and Knowledge Discovery 35.2, pp. 401–449.

Sanchez-Azofeifa, G Arturo, Margaret Kalacska, Mario Marcos do Esp´ırito-Santo, G Wil- son Fernandes, and Stefan Schnitzer (2009). “Tropical dry forest succession and the contribution of lianas to wood area index (WAI)”. In: Forest ecology and management 258.6, pp. 941–948.

Sanchez-Azofeifa, GA and K Castro-Esau (2006). “Canopy observations on the hyperspec- tral properties of a community of tropical dry forest lianas and their host trees”. In: International Journal of Remote Sensing 27.10, pp. 2101–2109.

Sanchez-Azofeifa, GA, J Antonio Guzm´an-Quesada, Mauricio Vega-Araya, Carlos Campos- Vargas, Sandra Milena Dur´an, Nikhil D’Souza, Thomas Gianoli, Carlos Portillo- Quintero, and Iain Sharp (2017). “Can terrestrial laser scanners (TLSs) and hemi- spherical photographs predict tropical dry forest succession with liana abundance?” In: Biogeosciences 14.4, pp. 977–988.

Schneider, Fabian D, Daniel Ku¨kenbrink, Michael E Schaepman, David S Schimel, and Felix Morsdorf (2019). “Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR”. In: Agricultural and Forest Meteorology 268, pp. 249–257.

Schnitzer, Stefan A (2005). “A mechanistic explanation for global patterns of liana abun- dance and distribution”. In: The American Naturalist 166.2, pp. 262–276.

Schnitzer, Stefan A (2018). “Testing ecological theory with lianas”. In: New Phytologist 220.2, pp. 366–380.

Schnitzer, Stefan A and Frans Bongers (2002). “The ecology of lianas and their role in forests”. In: Trends in Ecology & Evolution 17.5, pp. 223–230.

Schnitzer, Stefan A (2011). “Increasing liana abundance and biomass in tropical forests: emerging patterns and putative mechanisms”. In: Ecology letters 14.4, pp. 397–406.

Schnitzer, Stefan A, Saara J DeWalt, and J´erˆome Chave (2006). “Censusing and Measuring Lianas: A Quantitative Comparison of the Common Methods 1”. In: Biotropica 38.5, pp. 581–591.

Schnitzer, Stefan A, Sergio Estrada-Villegas, and S Joseph Wright (2020). “The response of lianas to 20 yr of nutrient addition in a Panamanian forest”. In: Ecology 101.12, e03190.

Seidel, Dominik (2018). “A holistic approach to determine tree structural complexity based on laser scanning data and fractal analysis”. In: Ecology and Evolution 8.1, pp. 128–134.

Shimizu, Katsuto, Tetsuji Ota, and Nobuya Mizoue (2019). “Detecting forest changes using dense Landsat 8 and Sentinel-1 time series data in tropical seasonal forests”. In: Remote Sensing 11.16, p. 1899.

Slik, JW Ferry, Gary Paoli, Krista McGuire, Ieda Amaral, Jorcely Barroso, Meredith Bas- tian, Lilian Blanc, Frans Bongers, Patrick Boundja, Connie Clark, et al. (2013). “Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics”. In: Global ecology and biogeography 22.12, pp. 1261–1271.

Smith, Leslie N (2018). “A disciplined approach to neural network hyper-parameters: Part 1–learning rate, batch size, momentum, and weight decay”. In: arXiv preprint arXiv:1803.09820.

Sprugel, DG (1983). “Correcting for bias in log-transformed allometric equations”. In:
Ecology (Durham) 64.1, pp. 209–210.

Srinivasan, Shruthi, Sorin C Popescu, Marian Eriksson, Ryan D Sheridan, and Nian-Wei Ku (2014). “Multi-temporal terrestrial laser scanning for modeling tree biomass change”. In: Forest Ecology and Management 318, pp. 304–317.

Stephenson, Nathan L, AJ Das, R Condit, SE Russo, PJ Baker, Noelle G Beckman, DA Coomes, ER Lines, WK Morris, Nadja Ru¨ger, et al. (2014). “Rate of tree carbon accu- mulation increases continuously with tree size”. In: Nature 507.7490, pp. 90–93.

Su, Tianyun, Wen Wang, Haixing Liu, Zhendong Liu, Xinfang Li, Zhen Jia, Lin Zhou, Zhuanling Song, Ming Ding, and Aiju Cui (2020). “An adaptive and rapid 3D Delaunay triangulation for randomly distributed point cloud data”. In: The Visual Computer, pp. 1–25.

Suratno, Agus, Carl Seielstad, and Lloyd Queen (2009). “Tree species identification in mixed coniferous forest using airborne laser scanning”. In: ISPRS Journal of Pho- togrammetry and Remote Sensing 64.6, pp. 683–693.

Taheriazad, Leila, Hamid Moghadas, and Arturo Sanchez-Azofeifa (2019). “Calculation of leaf area index in a Canadian boreal forest using adaptive voxelization and terrestrial LiDAR”. In: International Journal of Applied Earth Observation and Geoinformation 83, p. 101923.

Tang, Yong, Xiaofei Yang, Min Cao, Carol C Baskin, and Jerry M Baskin (2011). “Buttress trees elevate soil heterogeneity and regulate seedling diversity in a tropical rainforest”. In: Plant and soil 338.1, pp. 301–309.

Tao, Shengli, Qinghua Guo, Shiwu Xu, Yanjun Su, Yumei Li, and Fangfang Wu (2015a). “A geometric method for wood-leaf separation using terrestrial and simulated lidar data”. In: Photogrammetric Engineering & Remote Sensing 81.10, pp. 767–776.

Tao, Shengli, Fangfang Wu, Qinghua Guo, Yongcai Wang, Wenkai Li, Baolin Xue, Xueyang Hu, Peng Li, Di Tian, Chao Li, et al. (2015b). “Segmenting tree crowns from terres- trial and mobile LiDAR data by exploring ecological theories”. In: ISPRS Journal of Photogrammetry and Remote Sensing 110, pp. 66–76.

Thomas, Hugues, Fran¸cois Goulette, Jean-Emmanuel Deschaud, Beatriz Marcotegui, and Yann LeGall (2018). “Semantic classification of 3D point clouds with multiscale spher- ical neighborhoods”. In: 2018 International conference on 3D vision (3DV). IEEE, pp. 390–398.

Van Kuppevelt, D, C Meijer, F Huber, A van der Ploeg, S Georgievska, and Vincent T van Hees (2020). “Mcfly: Automated deep learning on time series”. In: SoftwareX 12, p. 100548.

Vauhkonen, Jari, Anne Sepp¨anen, Petteri Packal´en, and Timo Tokola (2012). “Improving species-specific plot volume estimates based on airborne laser scanning and image data using alpha shape metrics and balanced field data”. In: Remote Sensing of Environment 124, pp. 534–541.

Vicari, Matheus B, Mathias Disney, Phil Wilkes, Andrew Burt, Kim Calders, and William Woodgate (2019). “Leaf and wood classification framework for terrestrial LiDAR point clouds”. In: Methods in Ecology and Evolution 10.5, pp. 680–694.

Wang, Di, Jasmin Brunner, Zhenyu Ma, Hao Lu, Markus Hollaus, Yong Pang, and Norbert Pfeifer (2018). “Separating tree photosynthetic and non-photosynthetic components from point cloud data using dynamic segment merging”. In: Forests 9.5, p. 252.

Wang, Di, Markus Hollaus, and Norbert Pfeifer (2017). “Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data”. In: ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4.

Wang, Lin, Zhigang Wang, and Shan Liu (2016). “An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm”. In: Expert Systems with Applications 43, pp. 237–249.

Wang, Peng, Liguo Wang, Henry Leung, and Gong Zhang (2020). “Super-resolution map- ping based on spatial–spectral correlation for spectral imagery”. In: IEEE Transactions on Geoscience and Remote Sensing 59.3, pp. 2256–2268.

Wang, Wen, Tianyun Su, Haixing Liu, Xinfang Li, Zhen Jia, Lin Zhou, Zhuanling Song, and Ming Ding (2019). “Surface reconstruction from unoriented point clouds by a new triangle selection strategy”. In: Computers & Graphics 84, pp. 144–159.

Wang, Yunsheng, Matti Lehtom¨aki, Xinlian Liang, Jiri Py¨or¨al¨a, Antero Kukko, Anttoni Jaakkola, Jingbin Liu, Ziyi Feng, Ruizhi Chen, and Juha Hyyppa¨ (2019). “Is field-measured tree height as reliable as believed–A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest”. In: ISPRS Journal of Photogrammetry and Remote Sensing 147, pp. 132–145.

Wang, Zhen, Liqiang Zhang, Tian Fang, P Takis Mathiopoulos, Xiaohua Tong, Huamin Qu, Zhiqiang Xiao, Fang Li, and Dong Chen (2014). “A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification”. In: IEEE Transactions on Geoscience and Remote Sensing 53.5, pp. 2409–2425.

Wang, Zhiguang, Weizhong Yan, and Tim Oates (2017). “Time series classification from scratch with deep neural networks: A strong baseline”. In: 2017 International joint conference on neural networks (IJCNN). IEEE, pp. 1578–1585.

Waring, RH and SW Running (2007). “Forest Ecosystems. Analysis at Multiple Time and Space Scales”. In: Forest Ecosystems (Third Edition). Academic Press, pp. 1–16.

Wei, Qiong and Roland L Dunbrack Jr (2013). “The role of balanced training and testing data sets for binary classifiers in bioinformatics”. In: PloS one 8.7, e67863.

Weinmann, Martin, Steffen Urban, Stefan Hinz, Boris Jutzi, and Cl´ement Mallet (2015). “Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas”. In: Computers & Graphics 49, pp. 47–57.

Wright, S Joseph (2005). “Tropical forests in a changing environment”. In: Trends in ecology & evolution 20.10, pp. 553–560.

Xi, Zhouxin, Chris Hopkinson, Stewart B Rood, and Derek R Peddle (2020). “See the forest and the trees: effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning”. In: ISPRS Journal of Photogrammetry and Remote Sensing 168, pp. 1–16.

Xu, Xiaowei, Yinrong Chen, Junfeng Zhang, Yu Chen, Prathik Anandhan, and Adhiyaman Manickam (2021). “A novel approach for scene classification from remote sensing im- ages using deep learning methods”. In: European Journal of Remote Sensing 54.sup2, pp. 383–395.

Yang, Xiaoyuan, Alan H Strahler, Crystal B Schaaf, David LB Jupp, Tian Yao, Feng Zhao, Zhuosen Wang, Darius S Culvenor, Glenn J Newnham, Jenny L Lovell, et al. (2013). “Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®)”. In: Remote sensing of environment 135, pp. 36–51.

Yao, Tian, Xiaoyuan Yang, Feng Zhao, Zhuosen Wang, Qingling Zhang, David Jupp, Jenny Lovell, Darius Culvenor, Glenn Newnham, Wenge Ni-Meister, et al. (2011). “Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar”. In: Remote sensing of Environment 115.11, pp. 2965–2974.

Yu, Xiaowei, Xinlian Liang, Juha Hyypp¨a, Ville Kankare, Mikko Vastaranta, and Markus Holopainen (2013). “Stem biomass estimation based on stem reconstruction from ter- restrial laser scanning point clouds”. In: Remote Sensing Letters 4.4, pp. 344–353.

Zheng, Yi, Qi Liu, Enhong Chen, Yong Ge, and J Leon Zhao (2016). “Exploiting multi- channels deep convolutional neural networks for multivariate time series classification”. In: Frontiers of Computer Science 10.1, pp. 96–112.

Zhiyuan, He, Tang Yong, Deng Xiaobao, and Cao Min (2013). “Buttress trees in a 20- hectare tropical dipterocarp rainforest in Xishuangbanna, SW China”. In: Journal of Plant Ecology 6.2, pp. 187–192.

Zhou, Bolei, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba (2016). “Learning deep features for discriminative localization”. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2921–2929.

Zhou, Qian-Yi, Jaesik Park, and Vladlen Koltun (2018). “Open3D: A modern library for 3D data processing”. In: arXiv preprint arXiv:1801.09847.

Zhu, Xi, Andrew K Skidmore, Roshanak Darvishzadeh, K Olaf Niemann, Jing Liu, Yi- fang Shi, and Tiejun Wang (2018). “Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest”. In: International journal of applied earth observation and geoinformation 64, pp. 43–50.