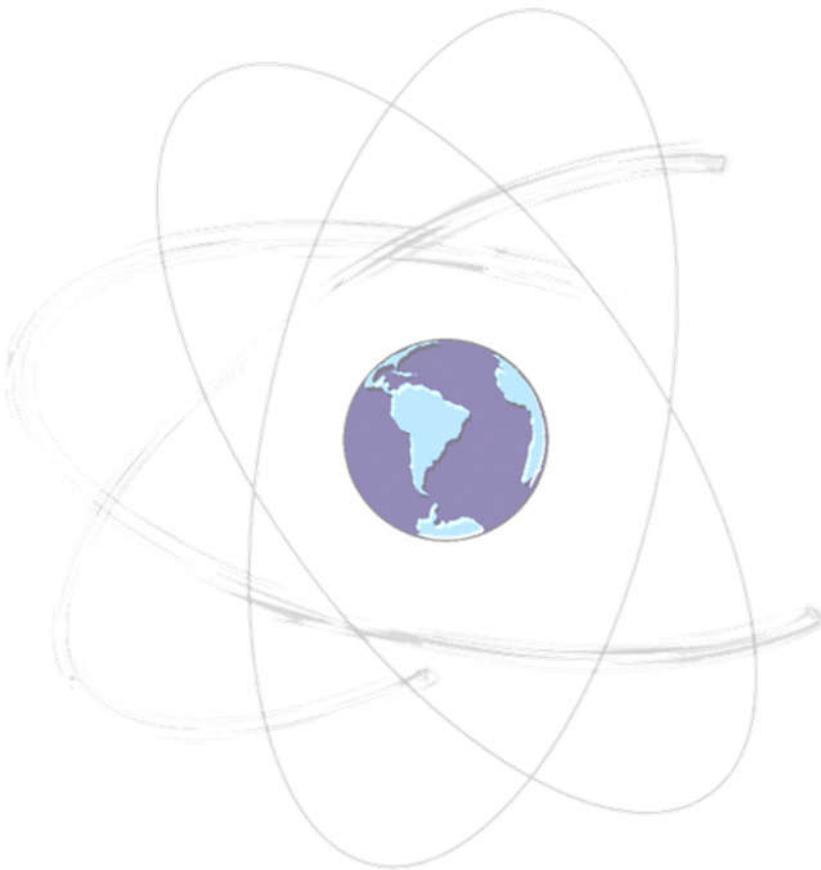


GABRIELA TAKAHASHI MIYOSHI

**EMERGENT TREE SPECIES IDENTIFICATION IN HIGHLY
DIVERSE BRAZILIAN ATLANTIC FOREST USING
HYPERSPETRAL IMAGES ACQUIRED WITH UAV**





UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Câmpus de Presidente Prudente

GABRIELA TAKAHASHI MIYOSHI

EMERGENT TREE SPECIES IDENTIFICATION IN HIGHLY
DIVERSE BRAZILIAN ATLANTIC FOREST USING
HYPERSPETRAL IMAGES ACQUIRED WITH UAV

Doctoral dissertation presented to the Graduate Program in Cartographic Sciences (PPGCC) at São Paulo State University (UNESP), School of Technology and Sciences, campus Presidente Prudente, SP, Brazil, for the partial fulfillment of the requirements for the grade of Doctor in Cartographic Sciences.

Advisors: Dr. Nilton Nobuhiro Imai
Dr. Antonio Maria Garcia Tommaselli
Dr. Eija Honkavaara

FICHA CATALOGRÁFICA

Miyoshi, Gabriela Takahashi.
M685e Emergent tree species identification in highly diverse Brazilian
Atlantic forest using hyperspectral images acquired with UAV /
Gabriela Takahashi Miyoshi. - Presidente Prudente, 2020
116 f.

Tese (doutorado) - Universidade Estadual Paulista (UNESP),
Faculdade de Ciências e Tecnologia, Presidente Prudente.

Orientador: Nilton Nobuhiro Imai

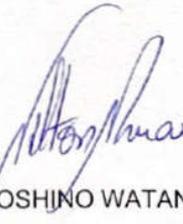
Coorientador: Antonio Maria Garcia Tommaselli

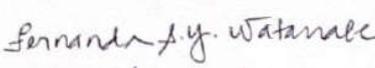
Coorientadora: Eija Honkavaara

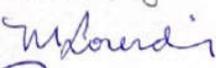
1. Tree species identification. 2. Atlantic forest; 3. multitemporal
spectral information. 4. Superpixels. 5. UAV. I. Título

ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE GABRIELA TAKAHASHI MIYOSHI, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS CARTOGRÁFICAS, DA FACULDADE DE CIÊNCIAS E TECNOLOGIA - CÂMPUS DE PRESIDENTE PRUDENTE.

Aos 11 dias do mês de março do ano de 2020, às 09:00 horas, no(a) Anfiteatro 7, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Prof. Dr. NILTON NOBUHIRO IMAI - Orientador(a) do(a) Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente, Profa. Dra. FERNANDA SAYURI YOSHINO WATANABE do(a) Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente, Profa. Dra. MARIA DE LOURDES BUENO TRINDADE GALO do(a) Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente, Prof. Dr. VERALDO LIESENBERG do(a) Centro de Ciências Agroveterinárias / Universidade do Estado de Santa Catarina, Dra. YUNSHENG WANG do(a) FGI / Finnish Geodetic Institute, sob a presidência do primeiro, a fim de proceder a arguição pública da TESE DE DOUTORADO de GABRIELA TAKAHASHI MIYOSHI, intitulada **Emergent tree species identification in highly diverse Brazilian Atlantic forest using hyperspectral images acquired with the use of UAV**. Após a exposição, a discente foi arguida oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: Aprovada _____. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.

Prof. Dr. NILTON NOBUHIRO IMAI 

Profa. Dra. FERNANDA SAYURI YOSHINO WATANABE 

Profa. Dra. MARIA DE LOURDES BUENO TRINDADE GALO 

Prof. Dr. VERALDO LIESENBERG 

Dra. YUNSHENG WANG

VIDEOCONFERÊNCIA

To my beloved ones. We did it.

ACKNOWLEDGMENT

One research can not be done by only one person, it is always a result of many people. I have changed ideas, experiences, have received a lot of helping and supporting. I am so grateful for you all, without whom I would not have made it through my doctoral degree.

I will always be grateful for my family for the invaluable assistance/love/patient that you all provided during my studies and in all my life. Mãe and Pai, you are my examples of resilience. Bia, Ju, Lão, and Jake thanks for always taking care of me, for sharing food, complaining and laughing, and for all the moments with you (I know it is not always and forever. Three days are enough lol). Pandora, Foguinho (*in memoriam*), Tiquinha (*in memoriam*), Megan, Pitico, Juninho, Luís Paulo, Brown, Faisca, Marvel and Mingau (without counting the other family members, Botas, Olívia, Miu and Ovelhinhas) you are my strength and peace. There are no words to describe how important you are to me. I love you all.

My deepest gratitude to Marcus Moraes and his family. Marcus, I have no words to describe how thankful I am for the last seven years of partnership, love, care, and evolution. Thank you for not letting me give up, for listening to me in the best and hardest moments, for accepting to share good and bad things during the masters and doctoral degree. Thank you for all the experiences and memories, for the unforgettable moments. You were essential during this journey.

I wish to acknowledge the support and great love of my family (Batians, Ditiam, tias, tios, primos e primas). You make this journey more pleasant.

To my colleagues and friends from the graduation program (PPGCC). It is wholeheartedly that I especially thank Fran Tolentino who helped me with the initial Weka steps, and Luiz Eduardo (Guil) who helped me with the Python challenges. Thank you also for your friendship. I cannot fail to mention Fer Watanabe. Thank you since the first year of the undergraduate and for being an example of person. I also thank the members of the Remote Sensing research group named SERTIE and of the Photogrammetry group for the help during the fieldworks.

Mari B. Thomaz (and Tião, for sure), Caio Fontoura, and Crislaine. I used to say that you were my gifts during this journey. Thank you for sharing moments, laughs, complaining, fun, a bit more complaining and problems, but mainly the friendship and endearments. Thank you for all the coffee and talkings at the balcony or your lab. Thank you, Cris, for the amazing trip, full of new experiences. Thank you, Mari and Caio, for the blessed moments.

Cesar (Czars), thank you not only for providing the ALS point cloud used in this dissertation, but also for your friendship and high spirits.

Thanks to my friends and colleagues from FGI who were essential during my internship in Finland. You made my 6 months in the freezing Finnish winter the warmest ones. Ziyi, Aimad and Karla, Paarvo, Yunsheng, Xinliang, Matti Lehtomäki, Wenxia, Yuwei, Deyu, and Olli Nevalainen, thank you. Raquel and Roope, I have no words to explain my

gratitude for you. There are no people in Finland with the biggest heart than you. I will always consider you my as Finnish family.

To my roomie in Helsinki. Kasia Kudła, I am so lucky to meet you and I hope we continue our friendship for always. Dziękuję.

I would like to thank the team from NUVEM UAV for the UX4 development and auxiliary, which made the images acquisition possible. Additionally, thanks to all UNESP staff who accompany me throughout these 11 years. Cinthia Onishi, thank you for all the patient, you are a fixer and a great person.

Waltinho and Baltazar, thank you for the great help during the fieldworks.

I wish to express my gratitude for the committee members who evaluate my dissertation. Dra. Yunsheng Wang, Dr. Veraldo Liesenberg, Dra. Maria de Lourdes B. T. Galo, and Dra. Fernanda Sayuri Y. Watanabe. Your comments and suggestions improved (a lot) the quality of this research.

Dr. Nilton N. Imai, Dr. Antonio M. G. Tommaselli, and especially, Dr. Eija Honkavaara. Thank you for the insights, knowledge and support shared with me during all these years. Dr. Eija, I am also grateful for you accepting me for 6 months at the Finnish Geospatial Research Institute (FGI).

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, in part by the Programa de Doutorado Sanduíche no Exterior - PDSE (process number 88881.187406/2018-01), in part by the Brazilian National Council for Scientific and Technological Development - CNPq (process number 153854/2016-2), in part by The São Paulo Research Foundation - FAPESP (process number 2013/50426-4), and in part by the Academy of Finland (grant number 327861).

Last, but not least, I would like to thank God for all the blesses and faith. Eleven years have passed so fast and now it's just the beginning of a new chapter in my history.

“Those who have a 'why' to live, can bear with almost any 'how'”
Viktor E. Frankl

ABSTRACT

The objective of this doctoral dissertation is to propose a new methodology to identify eight emergent tree species (i.e., that stood out from the canopy) belonging to highly diverse Brazilian Atlantic forest and with different ages and development stages. To achieve the objective, hyperspectral images were acquired in July/2017, June/2018, and July/2019 in a transect area located in the western part of São Paulo State. The area is in Ponte Branca ecological station, where the forest is classified as submontane semideciduous seasonal with different stages of succession. Images with a spatial resolution of 10 cm were acquired with a hyperspectral camera (500–900 nm) onboard unmanned aerial vehicle (UAV) and geometrically and radiometrically post-processed. In sequence, the individual tree crowns (ITCs) were manually delineated in each dataset to be used as reference in the experiments. From the performed experiments, it is highlighted the use of mean normalized spectra to reduce the within-species spectral variability, the use of region-based classification with the Random Forest algorithm, and the use of superpixels to automatically delineate the ITCs in each dataset. Additionally, the multitemporal superpixels with different multitemporal features (normalized spectra, texture and vegetation indexes) and structural features derived from the canopy height model, combined or not, were assessed to the tree species classification. The best result was achieved merging normalized spectra and vegetation indexes, where the value of area under the receiver operating characteristics curve (AUCROC) achieved values up to 0.964. From the obtained results it is pointed out the challenge when working with this type of forest due to the lack of emergent trees, which restrict the number of samples recognized in the field, and the existence of different ages and stages of development to the same tree species. Besides, the use of structural and textural features did not improve the tree species identification. Besides, the high spatial resolution of the images showed the slight differences in the spatial position of the tree crowns between the datasets. Finally, despite the challenges the results are promising and showed the feasibility to identify the tree species using multitemporal information.

Keywords: Tree species identification; Atlantic forest; multitemporal spectral information; superpixels, UAV.

RESUMO

O objetivo desse doutorado é propor uma nova metodologia para identificar oito espécies arbóreas emergentes (i.e., que se sobressaem do dossel florestal), em diferentes idades e estágios de desenvolvimento e pertencentes à Mata Atlântica brasileira. Para tal, imagens hiperespectrais foram adquiridas em Julho/2017, em Junho/2018, e em Julho/2019 em um transecto localizado no fragmento florestal Ponte Branca, localizado a Oeste do Estado de São Paulo, onde a floresta é considerada estacional semidecidual e submontana. As imagens com resolução espacial de 10 cm foram adquiridas com câmara hiperespectral (500–900 nm) acoplada em veículo aéreo não tripulado (VANT ou UAV, do inglês *Unmanned aerial vehicle*) e, posteriormente corrigidas geometricamente e radiometricamente. Em seguida, as copas arbóreas individuais (ITCs, do inglês *Individual tree crowns*) foram delineadas manualmente em cada conjunto de dados para serem utilizadas como referência para os experimentos. Dentre os experimentos realizados, destaca-se o uso do espectro normalizado para redução da variabilidade espectral intra-espécies, o uso da classificação baseada em regiões utilizando o algoritmo *Random Forest* e o uso de *superpixels* para delineamento automático das ITCs em cada conjunto de imagens. Além disso, avaliou-se o uso dos *superpixels* multitemporais com diferentes atributos multitemporais (espectro normalizado, textura e índices de vegetação) e estruturais (derivados do modelo de altura das copas), sozinhos ou combinados, para identificação das espécies arbóreas. O melhor resultado foi obtido a partir do uso combinado do espectro normalizado com os índices de vegetação, onde o valor da área sobre a curva característica de operação do receptor (AUCROC, do inglês *Area under the receiver operating characteristics curve*) atingiu valores de até 0.964. A partir dos resultados obtidos destaca-se o desafio ao trabalhar com esse tipo de floresta, devido à falta de árvores emergentes (que se sobressaem do dossel florestal), e a existência de árvores com diferentes idades e estágios de desenvolvimento, resultando em alta variabilidade espectral e estrutural para uma mesma espécie. Adicionalmente, foi verificado que o uso dos atributos estruturais e texturais não auxiliaram a tarefa de identificação de espécies e, que a alta resolução espacial das imagens mostrou as sutis diferenças de posição espacial das copas nas imagens dos diferentes anos. Por fim, apesar dos desafios, tem-se que os resultados são promissores e mostraram ser possível identificar espécies de árvores utilizando a informação multitemporal.

Palavras-chave: Identificação de espécies arbóreas; Mata Atlântica, informação espectral multitemporal, superpixels, VANT.

LIST OF FIGURES

| | | |
|-----------|--|----|
| Figure 1 | - Study area and tree species recognized in the field in the 2017 dataset. The red rectangle represents the imaged area, and the yellow rectangle is a zoom of the study area. Images acquired with RGB camera onboard aircraft (left image) and onboard UAV (right images). | 23 |
| Figure 2 | - Successional stages of Ponte Branca in the year of 2016. Letters followed by numbers represent the data used to validate this classification. | 24 |
| Figure 3 | - Pictures acquired over (a), inside (b) and outside (c) the study area. | 24 |
| Figure 4 | - Climograph of Paranapoema station from the National Institute of Meteorology. Bars represent the accumulated rainfall per month (mm), and dashed lines represent the mean monthly temperature (°C). | 25 |
| Figure 5 | - The spatial difference in the leaves of one sample of <i>Syagrus romanzoffiana</i> tree in each dataset (R: 690.28 nm; G: 565.10 nm; B: 519.94 nm; automatic contrast from QGIS software, version 3.0.0): (a) 2017, (b) 2018, and (c) 2019. ... | 26 |
| Figure 6 | - Canopy examples of each tree species identified in the field and delineated in the images acquired in 2017 (R: 780.49 nm; G: 565.10 nm; B: 506.22 nm; automatic contrast from QGIS software, version 3.0.0). (a) <i>Astronium graveolens</i> (AG), (b) <i>Apuleia leiocarpa</i> (AL), (c) <i>Aspidosperma polyneuron</i> (AP), (d) <i>Aspidosperma subincanum</i> (AS), (e) <i>Copaifera langsdorffii</i> (CL), (f) <i>Endlicheria paniculata</i> (EP), (g) <i>Helietta apiculata</i> (HA), (h) <i>Hymenaea courbaril</i> (HC), (i) <i>Inga vera</i> (IV), (j) <i>Pterodon pubescens</i> (PP), (k) <i>Syagrus romanzoffiana</i> (SR). | 27 |
| Figure 7 | - Profile of ALS data classified as ground and non-ground where the pink-colored points represent the ground and the one colored in yellow represents the surface objects (trees). Data acquired in November 2017 with a Riegl LMS-Q680i laser scanner..... | 32 |
| Figure 8 | - Canopy height model of the study area obtained from the ALS data. Colors values represent the tree heights. | 32 |
| Figure 9 | - Mean tree height versus the tree crown area for all samples identified in the field. Data are from the 2017 dataset. | 33 |
| Figure 10 | - (a) Rikola hyperspectral camera, model DT-0011, with its irradiance and GPS sensors; (b) UX4 UAV attached with Rikola camera. | 34 |
| Figure 11 | - Simulated normalized sensitivities of each spectral band set in the Rikola camera. Responses calculated from the central wavelength and its respective bandwidth. | 35 |
| Figure 12 | - 2018 mosaic of hyperspectral images before (a) and after the radiometric block adjustment (b) of the spectral band centered at 750.16 nm. Automatic contrast from QGIS software (version 3.0.0) in both images. | 38 |
| Figure 13 | - Flowchart of the experiments ¹ | 39 |

| | |
|---|----|
| Figure 14 - Differences in the proportions of sunlit and shadowed regions recorded by sensors depending on the geometry of view. The sensors' positions are represented by the number in parenthesis; the arrows represent the directions of the incident light (colored in orange) and of the reflected light (colored in green)..... | 40 |
| Figure 15 - Incident light in a forest resulting in different shadows density in each tree crown. | 41 |
| Figure 16 - (a) Random Forest principle, in which multiple decision trees are used to train the classifier. (b) After the classification model is built, the new data are classified in all decision trees and its class belongs to the most popular class among all decision trees..... | 44 |
| Figure 17 - False-colour CHM raster (a) without scale and (b) with scale in elevation of 1.5 times. The blue color represents the lower heights whereas the red color represents the highest values. | 49 |
| Figure 18 - Superpixels generated using initial parameter of (c) 2,000, (d) 10,000, and (e) 20,000. In the greyscale images, the darkest color represents the lower height values whereas the brightest color represents the higher values. | 49 |
| Figure 19 - (a) Example of manually delimited ITC and the superpixel representing the same ITC. (b) True positive, false negative and false positives areas..... | 51 |
| Figure 20 - Mean spectra of each tree species considering the: (a) Mean, and the (b) MeanNorm. | 55 |
| Figure 21 - Values for each tree species considering the mean reflectance factor spectra (Mean) and the mean normalized spectra (MeanNorm). The blue line represents the Mean, the red line represents the MeanNorm, and the shaded area represents the minimum and maximum values. | 56 |
| Figure 22 - Spectral differences between the tree species when using the Mean feature. The x-axis represents the tree species pairwise combination ¹ | 57 |
| Figure 23 - Spectral differences between the tree species when using the MeanNorm feature. The x-axis represents the tree species pairwise combination ¹ | 60 |
| Figure 24 - Spectra difference within-species considering the mean spectra and the mean normalized spectra for each sample ¹ | 61 |
| Figure 25 - ROC curves of the identification of each tree species from the application of RF to all imagery datasets (Dall_MeanNorm). (a) <i>Apuleia leiocarpa</i> (AL), (b) <i>Copaifera langsdorffii</i> (CL), (c) <i>Endlicheria paniculata</i> (EP), (d) <i>Helietta apiculata</i> (HA), (e) <i>Hymenaea courbaril</i> (HC), (f) <i>Inga vera</i> (IV), (g) <i>Pterodon pubescens</i> (PP), (h) <i>Syagrus romanzoffiana</i> (SR). | 65 |
| Figure 26 - Feature importance when using the RF for 8 tree species and all datasets (Dall_MeanNorm). The x-axis represents the spectral bands in each year. The black arrows point to the highest and lowest values of features importance. Wavelengths in nm. | 68 |

| | |
|--|----|
| Figure 27 - Flowchart of the segmentation experiments. | 69 |
| Figure 28 - Superpixels merged. The red polygons represent the superpixels after they were merged, and the blue lines represent the original superpixels from SLIC.... | 70 |
| Figure 29 - Segments generated by watershed after merging. The red polygons represent the segments after they were merged, and the blue lines represent the original ones..... | 71 |
| Figure 30 - Flowchart of the tree species identification experiments using multiple sets of features..... | 74 |

LIST OF TABLES

| | | |
|----------|---|----|
| Table 1 | - Tree species identified in the field and their characteristics..... | 28 |
| Table 2 | - The number of samples recognized in the field, their average number of pixels per crown, and its sum for each tree species. Pixels with 10 cm of spatial resolution. | 29 |
| Table 3 | - Rikola camera, model DT-0011, specifications. | 34 |
| Table 4 | - Spectral setting of Rikola camera, model DT-0011, with respective bandwidth in full width at half maximum (FWHM). | 34 |
| Table 5 | - Details of the image acquisition in each flight campaign..... | 36 |
| Table 6 | - Texture and vegetation indexes extracted from the mosaic of hyperspectral images. | 47 |
| Table 7 | - Summary of the extracted features. | 47 |
| Table 8 | - AUCROC values for each tree species identified in each classification experiment as well as its overall accuracy. Shaded cells indicate the highest AUCROC value of each tree species..... | 61 |
| Table 9 | - The number of spectral features used in each classification investigation..... | 62 |
| Table 10 | - AUCROC values for each tree species identified in each dataset. AUCROC values are from imagery data of (i) only 2017 (D17_MeanNorm); (ii) only 2018 (D18_MeanNorm); (iii) only 2019 (D19_MeanNorm); (iv) all years and the mean spectral values (Dall_Mean); and (v) all years and the mean normalized values (Dall_MeanNorm). | 63 |
| Table 11 | - Confusion matrix of the classification of eight tree species and all datasets (Dall_MeanNorm) and its user accuracy and producer accuracy..... | 65 |
| Table 12 | - Criteria to merge the segments ¹ | 70 |
| Table 13 | - Overall accuracies for the individual crown delineation using superpixels and watershed methods. Shaded cells highlight the highest F-Score values for each dataset. | 71 |
| Table 14 | - Classification investigations performed using different features combination and the number of features used in each test. | 75 |
| Table 15 | - AUCROC values for each tree species identification in each experiment. Shaded cells highlight the highest value of each row..... | 76 |

LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|----------|---|
| AG | - <i>Astronium graveolens</i> |
| AL | - <i>Apuleia leiocarpa</i> |
| ALS | - Aerial laser scanning |
| AP | - <i>Aspidosperma polyneuron</i> |
| AS | - <i>Aspidosperma subincanum</i> |
| ASM | - Angular second moment |
| ATM | - Airborne Thematic Mapper |
| AUC | - Area under curve |
| AUCROC | - Area under the receiver operating characteristics curve |
| BBA | - Bundle block adjustment |
| BRDF | - Bidirectional reflectance distribution function |
| CHM | - Canopy height model |
| CHMm | - Mean value of the canopy height model |
| CHMstd | - Standard deviation of the canopy height model |
| CL | - <i>Copaifera langsdorffii</i> |
| CON | - Contrast |
| COR | - Correlation |
| D | - Spectral distance |
| DBH | - Diameter at breast height |
| DIS | - Dissimilarity |
| DN | - Digital number |
| DSM | - Digital surface model |
| DTM | - Digital terrain model |
| EOP | - Exterior orientation parameter |
| EP | - <i>Endlicheria paniculata</i> |
| FGI | - Finnish Geospatial Research Institute |
| FPI | - Fabry–Pérot Interferometer |
| FPR | - False positive rate |
| FWHM | - Full width at half maximum |
| GCP | - Ground control point |
| GNSS | - Global navigation satellite system |
| GPS | - Global positioning system |
| GSD | - Ground sample distances |
| HA | - <i>Helietta apiculata</i> |
| HC | - <i>Hymenaea courbaril</i> |
| HOM | - Homogeneity |
| ID | - Identification abbreviation |
| IOP | - Interior orientation parameter |
| ITC | - Individual tree crown |
| IV | - <i>Inga vera</i> |
| KUR | - Kurtosis of the canopy height model |
| LOOCV | - Leave-one-out cross validation |
| Mean | - Mean reflectance factor spectra |
| MeanNorm | - Mean normalized reflectance factor spectra |
| NDVI | - Normalized difference vegetation index |
| NIR | - near-infrared part of the electromagnetic spectrum |
| OA | - Overall accuracy |
| p25 | - 25th percentile of the height |

| | |
|------|---|
| p50 | - 50th percentile of the height |
| p75 | - 75th percentile of the height |
| p90 | - 90th percentile of the height |
| PP | - <i>Pterodon pubescens</i> |
| PRI | - Photochemical reflectance index |
| PSRI | - Plant senescence reflectance index |
| REP | - Red-edge position |
| RF | - Random forest |
| ROC | - Receiver operating characteristics |
| SIPI | - Structure insensitive pigment reflectance index |
| SKE | - Skewness of the canopy height model |
| SLIC | - Simple linear iterative clustering |
| SR | - <i>Syagrus romanzoffiana</i> |
| SVM | - Support vector machine |
| SWIR | - Short-wave infrared |
| TPR | - True positive rate |
| UAV | - Unmanned aerial vehicle |
| VIS | - Visible part of the electromagnetic spectrum |

CONTENT

| | |
|--|------------|
| 1 INTRODUCTION | 17 |
| 1.1 HYPOTHESIS AND OBJECTIVE | 20 |
| 1.2 INTERNATIONAL COOPERATION | 20 |
| 1.3 CONTENT | 21 |
| 2 STUDY SITE..... | 23 |
| 2.1 REFERENCE DATA | 25 |
| 3 REMOTE SENSING DATA..... | 31 |
| 3.1 ALS POINT CLOUD..... | 31 |
| 3.2 MOSAICS OF HYPERSPECTRAL IMAGES..... | 33 |
| 4 METHODOLOGY | 39 |
| 5 RESULTS AND DISCUSSION | 53 |
| 5.1 SPECTRAL DIFFERENCES – ILLUMINATION DIFFERENCES REDUCTION | 53 |
| 5.2 COMPARISON OF PIXEL-BASED AND REGION-BASED APPROACHES | 60 |
| 5.3 CONTRIBUTION OF TEMPORAL DATA TO IDENTIFY THE TREE SPECIES..... | 62 |
| 5.3.1 Results of the feature importance..... | 68 |
| 5.4 ITC DELINEATION..... | 69 |
| 5.5 TREE SPECIES CLASSIFICATION USING SUPERPIXELS AND DIFFERENT SETS OF MULTITEMPORAL FEATURES | 73 |
| 6 CONCLUSION | 81 |
| 6.1 CONTRIBUTIONS AND FUTURE WORKS RECOMMENDATION | 82 |
| REFERENCES | 85 |
| APPENDIX A – WILCOXON-MANN-WHITNEY RESULTS | 97 |
| APPENDIX B – FEATURES IMPORTANCE AND CONFUSION MATRIXES WHEN USING DIFFERENT FEATURES COMBINATION | 107 |

1 INTRODUCTION

Forests play an important role in biodiversity, carbon stocks, the water cycle, and feedstock, but they are rapidly being deforested. In Brazil, they are targets of illegal loggers or even converted to crops, pasture, and urbanization. Knowledge about the tree species of a forest is a fundamental information. Tree species recognition can be performed through fieldwork campaigns, but generally, this practice has limitations, since it is expensive and laborious because of the forest density and forest access, which can be far from roads and thus, it is a time-consuming task. Remote sensing, together with automatic analysis techniques, has become a prominent tool for tree species mapping. Since the '80s, research papers related to “forest” and “Remote Sensing” exponentially increased (WEB-OF-SCIENCE, [s.d.]) showing that forest researches are a trending topic.

Most of the previous studies related to tree species identification using Remote Sensing have been performed in forests in the North hemisphere (FASSNACHT et al., 2016). There is a lack of studies in forests such as the Brazilian Atlantic forest, which encompasses different ecosystems, such as mixed ombrophilous, dense ombrophilous, open ombrophilous, semideciduous seasonal, and deciduous seasonal forests (BRASIL, 2006). Sothe et al. (2019) studied a mixed ombrophilous forest whose floristic compositions and forest structure characteristics differ from those of other types of Brazilian Atlantic forest, especially the semideciduous and deciduous seasonal forests (BRASIL, 2006), which highlights the importance of studying them separately.

In addition, most studies have investigated well-developed forests or forests in which trees with different heights are spatially distinguished such as coniferous forests. Plots containing tree species in different successional stages and ages can present similar heights, and thereby, cause spectral mixing due to leaf mixture and the effect of neighborhood spectra because the number of emergent trees, i.e., trees that stood out from the canopy, is lower than the number of smaller trees. Notwithstanding the importance of monitoring mature forests, monitoring fragments that are in the initial or intermediary regeneration process is considered a key element in the connection of forest patches, and contributes to the maintenance of biodiversity (LIRA et al., 2012; RIBEIRO et al., 2009). Emergent trees are equally important when it comes to tropical forests. From its importance it is highlighted their use for the movement of primates, who also use the emergent trees to sleep (ALEXANDER et al., 2018) and because of its transpiration rate when considering the water cycle (KUNERT et al., 2017).

Bearing the Remote Sensing concept, Jensen (2007) defines Remote Sensing as the art and science of acquiring information without the physical contact with the objects. The information is extracted by the acquisition and interpretation of the reflected energy from the objects (JENSEN, 2007). Considering the vegetation as a target, the amount of reflected light depends on the leaves' content, such as pigments and structure (PONZONI; KUPLICH; SHIMABUKURO, 2012). The reflected light can be registered by different sensors, which can be classified according to its platform as orbital, aerial or terrestrial sensors. Satellite sensors and airborne passive and/or active sensors, combined with the use of field spectroscopy, provide valuable information for the identification of tree species (COLGAN et al., 2012; HEINZEL; KOCH, 2012; IMMITZER; ATZBERGER; KOUKAL, 2012; WAGNER et al., 2018; ZHANG et al., 2012). Besides, the use of unmanned aerial vehicles (UAVs) has become a powerful tool to acquire forest information (NEVALAINEN et al., 2017; OTERO et al., 2018; SOTHE et al., 2020).

UAVs enable fast information acquisition, and despite their constraints regarding the trade-off between resolution and coverage, they are low-cost alternatives for capturing information in areas that are endangered or need constant monitoring, such as mines or crops (COLOMINA; MOLINA, 2014; KANG et al., 2019; POPESCU et al., 2020; SHAKHATREH et al., 2019). UAVs can fly over many areas that are challenging for field data acquisition, such as water surfaces or dense forest areas. UAV missions can be quickly configured according to the user's needs. Furthermore, in the past few years, UAVs have been rapidly developed to fly for several hours; an example of such a platform is the fixed-wing Batmap II UAV, which can fly for more than 2 hours (NUVEM UAV, [s.d.]). UAVs can capture very high or ultrahigh spatial resolution data with ground sampling distances (GSD) ranging from centimeters to decimeters (AASEN et al., 2018; COLOMINA; MOLINA, 2014; PANEQUE-GÁLVEZ et al., 2014; SANCHEZ-AZOFEIFA et al., 2017) using small-format multispectral or hyperspectral cameras, such as MicaSense RedEdge-MX (MICASENSE, [s.d.]), Rikola hyperspectral imager (SENOP, [s.d.]), and Cubert FirefLEYE (CUBERT, [s.d.]). Beyond that, UAVs can acquire information of surface targets, such as trees, with high temporal frequency, which is a promising option in forest monitoring, since it can measure dynamic phenological behavior according to seasons and tree characteristics.

Besides the different platforms to acquire remotely sensed data, it is important to consider the need to interpret and label the registered information. This process is called as classification (RICHARDS; JIA, 2006). There are many methods to classify the data, where it

is highlighted the machine learning algorithms, which is a potential alternative to the traditional classification approaches (LARY et al., 2016). Santos et al. (2010) showed that genetic programming, a subset of the machine learning, presented better results to recognize coffee crops than using the maximum likelihood approach. Support vector machine (SVM) (MELGANI; BRUZZONE, 2004) and random forest (RF) (BREIMAN, 2001) are examples of machine learning algorithms that have been successfully applied to identify tree species in urban environments (LI et al., 2015), savannas (COLGAN et al., 2012), and different types of forests, including northern, boreal, temperate, and tropical forests (FERET; ASNER, 2013; FERREIRA et al., 2016; IMMITZER; ATZBERGER; KOUKAL, 2012; MASCHLER; ATZBERGER; IMMITZER, 2018; MATSUKI; YOKOYA; IWASAKI, 2015; WAGNER et al., 2018).

Moreover, efforts concerning the best features extracted to tree species classification is also highlighted. Spectral features comprised from the visible (VIS) to short-wave infrared (SWIR) region, texture, vegetation indexes, and structural features are among the most useful features to the tree species classification (BALDECK et al., 2015; DALPONTE et al., 2014; HEINZEL; KOCH, 2012; TUOMINEN et al., 2018). Textural and vegetation indexes can be extracted from the imagery information whereas structural features can be calculated from point clouds derived from aerial laser scanning (ALS), which can be used to obtain the canopy height model (CHM) of a forest (NEVALAINEN et al., 2017; SILVA et al., 2016). Besides, considering the vegetation context, relevant parameters can be extracted from multitemporal information. The differences in trees blossoming and defoliation depend on the season, weather conditions, and soil moisture. Consequently, the spectral response of crowns belonging to different tree species changes with the time. Although most of the previous studies conducted with seasonal/temporal information have not employed UAVs, they have shown spectral differences within tree species and reported whether the tree species classification was improved (CASTRO-ESAU et al., 2006; DEVENTER; CHO; MUTANGA, 2017; FERREIRA et al., 2019; HILL et al., 2010; IMMITZER et al., 2019; KARASIAK et al., 2019; KEY et al., 2001; LI et al., 2015; SOMERS; ASNER, 2014).

In this regard, considering that UAVs can fly over many areas acquiring fast information with high spatial resolution and temporally, the joint use of this information could be helpful to identify the tree species. However, at the same time, it would be challenging because all the variations within a tree crown would be recorded in the high spatial resolution of images. Differences in tree growth from one year/season to the next one can appear even

coregistering the images. Thus, methods to handle with such small variations would be needed, not to mention the bidirectional reflectance distribution function (BRDF) effects because of the sunlit variations and different crown geometries.

1.1 HYPOTHESIS AND OBJECTIVE

The hypothesis of this doctoral dissertation is based on the knowledge that tree species have different characteristics depending on the weather conditions, and the recent availability of UAVs, which can quickly acquire information and has been successfully applied in Northern forests to identify tree species. In this sense, the hypothesis is that tree species identification of a fragment from the Brazilian Atlantic forest can be improved by using temporal information acquired with sensors onboard UAV, integrated with structural data derived from ALS. Bearing the hypothesis, this doctoral dissertation aims to propose a new methodology to identify selected tree species belonging to the Brazilian Atlantic forest using temporal information acquired with sensor onboard UAV. Further objectives are to:

- Evaluate the spectral differences among the tree species;
- Evaluate the pixel-based and region-based classification approaches;
- Delineate the individual tree crowns (ITCs); and
- Identify the tree species.

1.2 INTERNATIONAL COOPERATION

This doctoral dissertation was developed under the framework of the international joint project called “*Unmanned Airborne Vehicle - Based 4D Remote Sensing for Mapping Rain Forest Biodiversity and its Change in Brazil (UAV_4D_Bio)*”. This Project is a partnership between researchers from São Paulo State University (UNESP), and Finnish Geospatial Research, part of the National Land Survey of Finland. UAV_4D_Bio project was supported in part by The São Paulo Research Foundation (FAPESP) (grant number 2013/50426-4) and in part by the Academy of Finland (AKA) (grant number 273806). The project aimed to develop technologies to map and detect biodiversity changes in Brazilian Atlantic forests.

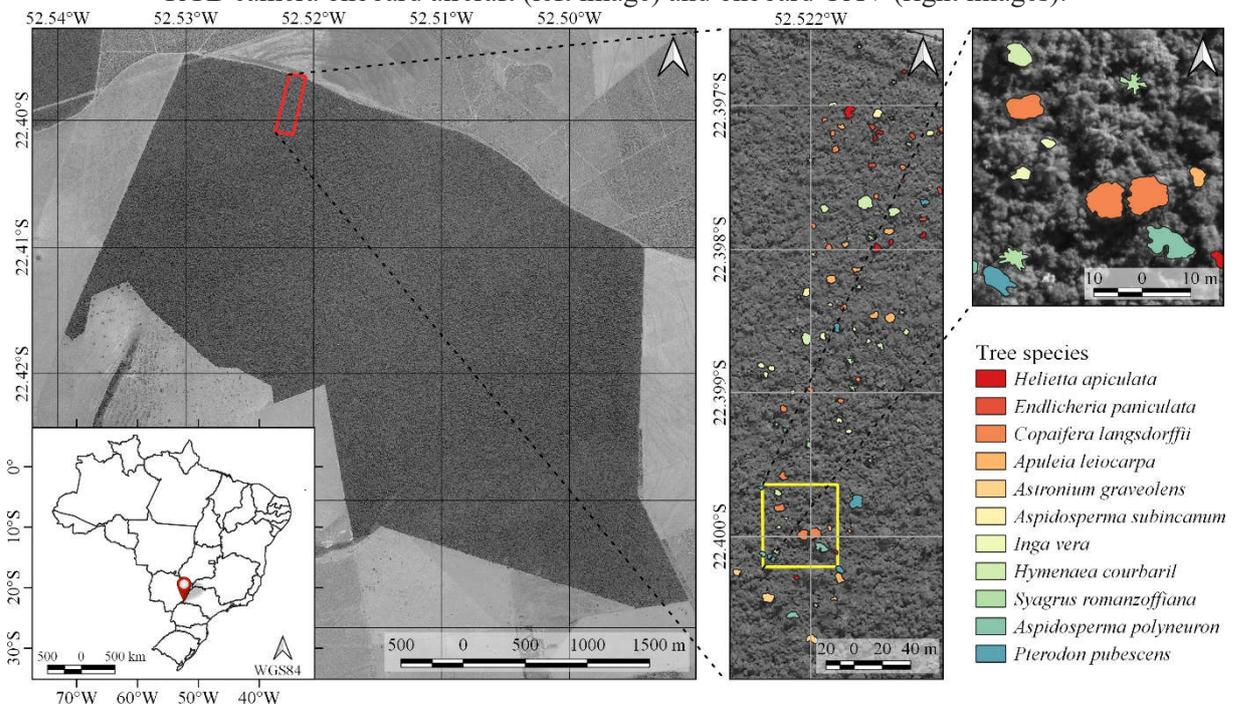
1.3 CONTENT

Section 1 introduced the objectives of the doctoral dissertation. Section 2 shows the study area, which is used in all experiments. Section 3 presents the Remote Sensing data used, i.e., the ALS and the hyperspectral imagery data, how they were acquired and processed. In Section 4 the developed methodology is described. Section 5 shows the results and discussion from the performed experiments. The first experiment (Section 5.1) is based on the papers of Miyoshi et al. ([s.d.], 2020) which show and evaluate the spectral differences between and within-species of trees belonging to the Brazilian Atlantic forest. The second experiment (Section 5.2) is an improvement of the work from Miyoshi et al. (2019), where the comparison of the pixel-based and region-based classification approaches when using the mean spectra and the mean normalized spectra as features are performed. Section 5.3 shows the third experiment, which is based on Miyoshi et al. (2020). This experiment evaluates the usefulness of multitemporal spectral information to identify tree species. Section 5.4 compares the superpixels and watershed methods to automatically delineate the ITCs in each imagery data. The last result is presented and discussed in Section 5.5. In this section, tree species identification using the findings from previous sections (5.1 to 5.4) and using additional set of features (vegetation indexes, texture, and structural features) is carried out. Finally, Section 6 shows the conclusion, contribution and recommendations of this doctoral dissertation.

2 STUDY SITE

The transect forest sample used in this study is centered at 22°23'55.21" S, 52°31'18.31" W, in the municipality of Euclides da Cunha Paulista, western part of São Paulo State, Brazil (Figure 1). This area has an approximate length and width of 500 m × 130 m and is established inside the Ponte Branca forest fragment, a protected area belonging to the Black Lion Tamarin Ecological Station. This area was chosen because it has the highest number of samples recognized in the field, as well as the highest number of images from different years available when compared with other datasets from the UAV_4D_Bio project.

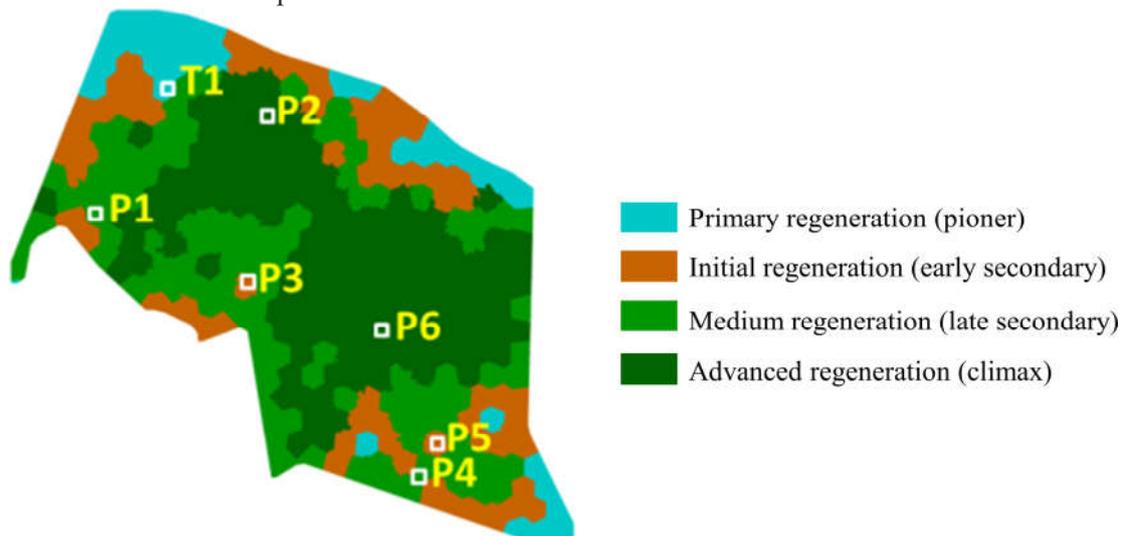
Figure 1 - Study area and tree species recognized in the field in the 2017 dataset. The red rectangle represents the imaged area, and the yellow rectangle is a zoom of the study area. Images acquired with RGB camera onboard aircraft (left image) and onboard UAV (right images).



Source: Miyoshi et al. (2020).

Until the end of the '70s Ponte Branca area suffered illegal logging, being later protected by governmental laws (BERVEGLIERI et al., 2018; BRASIL, 2006). As a result, it presents an initial to more advanced successional stages. According to Berveglieri et al. (2018), the Ponte Branca area has primary regeneration areas, where the dominant species present a uniform height, as well as more mature areas, where there are a larger number of tree species with a complex structure (Figure 2). Figure 3 shows pictures acquired over, inside and outside the study area, where it is possible to see its heterogeneity, with lianas covering trees from the overstory, taller and smaller trees.

Figure 2 - Successional stages of Ponte Branca in the year of 2016. Letters followed by numbers represent the data used to validate this classification.



Source: Adapted from Berveglieri et al. (2018).

Figure 3 - Pictures acquired over (a), inside (b) and outside (c) the study area.

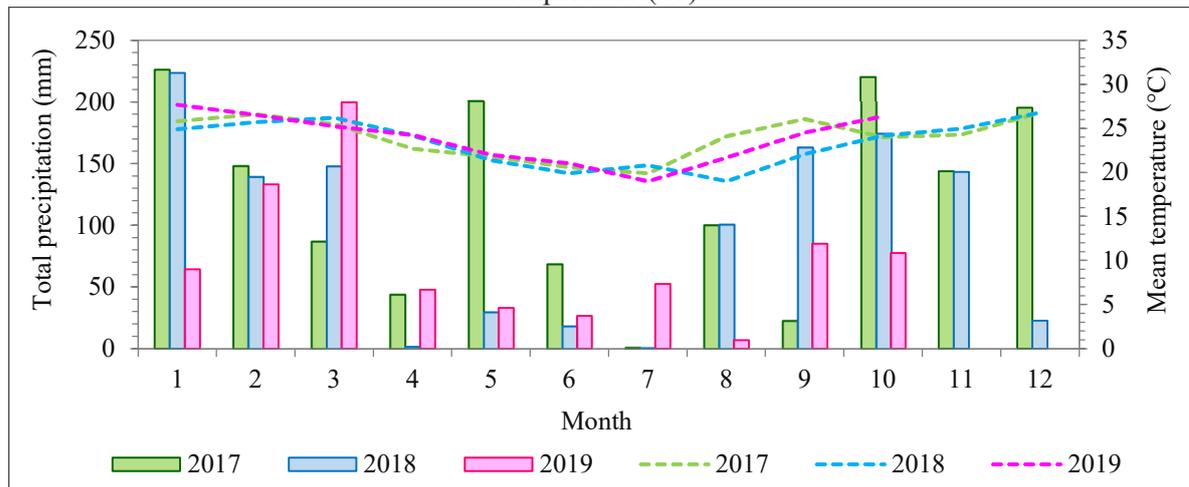


According to the Brazilian Institute for Geography and Statistics (IBGE), the vegetation of the study area is classified as a submontane semideciduous seasonal forest (IBGE, 2012). The regional climate is considered a tropical zone with dry winters (Aw) according to the Köppen classification (ALVARES et al., 2013). The mean temperature during the dry season is 21 °C, with less than 60 mm of total precipitation (INMET - INSTITUTO NACIONAL DE METEOROLOGIA, 2019).

The study area comprises a variety of tree species in different development stages. More than 25 different tree species were recognized during fieldworks (BERVEGLIERI et al., 2016), and they can be divided into pioneer species such as *Eugenia piriformis* and *Dilodendon bipinnatum*, and in secondary species like *Hymenaea courbaril* and *Apuleia leiocarpa*, those last two can achieve height up to 20 m and 35 m, respectively (LORENZI, 1992a). Regarding the weather between the years 2017, 2018, and 2019 and between its flight campaigns (Figure 4 and Section 3.2), it is noticed different weather patterns. The season was wetter in 2017, with precipitation of 69 mm before the flight campaign, whereas the

precipitation was 18.6 mm before the flight campaign of 2018 and 51 mm before the flight campaign of 2019; however, the rain did not occur for at least eight days before image acquisition (INMET - INSTITUTO NACIONAL DE METEOROLOGIA, 2019).

Figure 4 - Climograph of Paranapoema station from the National Institute of Meteorology. Bars represent the accumulated rainfall per month (mm), and dashed lines represent the mean monthly temperature (°C).



Source: Miyoshi et al. (2020).

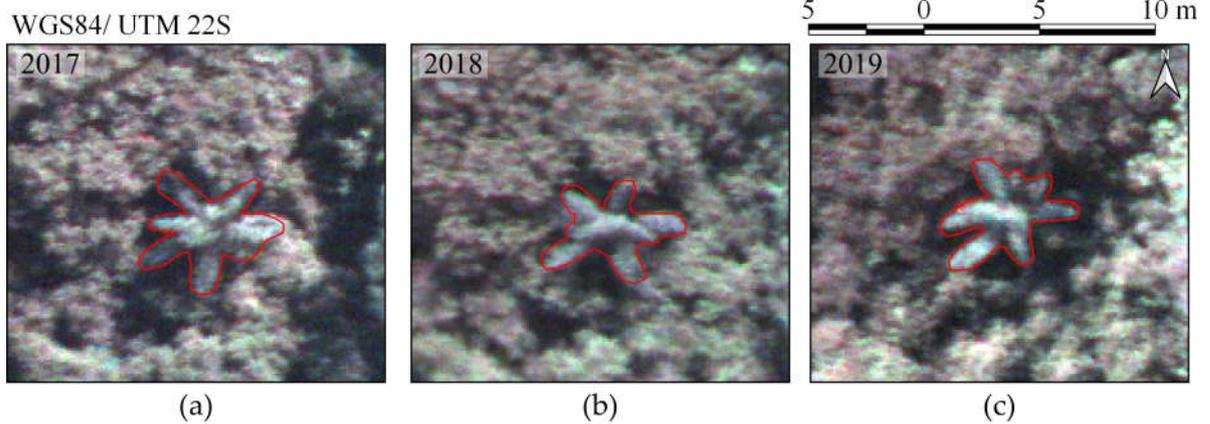
2.1 REFERENCE DATA

More than 25 tree species with a diameter at breast height (DBH) greater than 3.8 cm were identified during fieldwork (BERVEGLIERI et al., 2016). Tree species were in different development stages and ages, with the northernmost part of the study area containing trees in the initial stage of succession and the southernmost in a more advanced stage (BERVEGLIERI et al., 2018).

We located 101 trees of 11 species that emerged from the canopy (Table 1 and Table 2). The 101 tree samples were recognized during fieldworks carried out between 2015 and 2019 with the auxiliary of a guide who works in an arboretum near the study area. ITCs from the samples were manually delineated through visual interpretation of RGB image composites of each dataset (R: 628.73 nm; G: 550.39 nm; B: 506.22 nm, from Rikola camera data presented in Section 3.2), as were considered as the ground reference. A CHM of the area was also applied to improve the polygons delineation because they could provide a 3D view of the area. ITCs were delineated to each dataset described in Section 3.2, because of slightly different spatial position among the tree crowns. These differences are mainly caused by tree growth, changes in leaves with changing seasons, and weather conditions or projection

differences due to the characteristics of the surface used. Figure 5 shows the slight difference in the spatial distribution of the leaves of *Syagrus romanzoffiana* trees, a pioneer species which can be related with animals dispersion (SILVA et al., 2011).

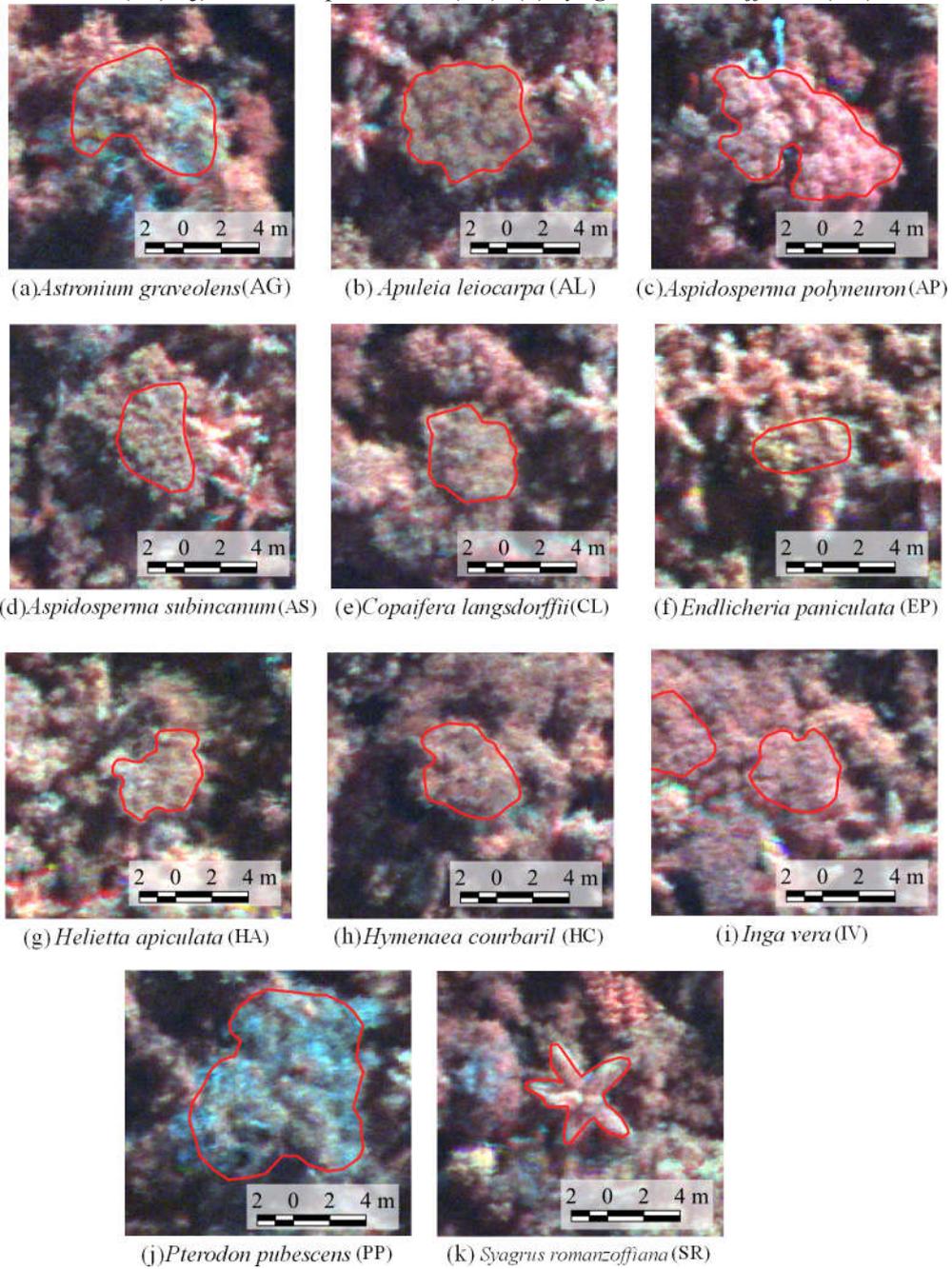
Figure 5 - The spatial difference in the leaves of one sample of *Syagrus romanzoffiana* tree in each dataset (R: 690.28 nm; G: 565.10 nm; B: 519.94 nm; automatic contrast from QGIS software, version 3.0.0): (a) 2017, (b) 2018, and (c) 2019.



Source: Miyoshi et al. (2020).

Figure 1 shows examples of delineated ITC polygons in the 2017 dataset, and Figure 6 shows canopy examples of each tree species in the mosaic of images acquired during the 2017 flight campaign. These tree species were chosen because they not only emerge from the canopy strata, but also because they are important for characterizing the successional stage of the forest, e.g., *Syagrus romanzoffiana*, which can be associated with the floristic composition (SILVA et al., 2011). It is important to note that smaller trees were excluded from analysis because lianas covering these trees and the overlap among individuals negatively affect the classification accuracy. From now on, tree species will be called by their abbreviations from Table 1 and Table 2.

Figure 6 - Canopy examples of each tree species identified in the field and delineated in the images acquired in 2017 (R: 780.49 nm; G: 565.10 nm; B: 506.22 nm; automatic contrast from QGIS software, version 3.0.0). (a) *Astronium graveolens* (AG), (b) *Apuleia leiocarpa* (AL), (c) *Aspidosperma polyneuron* (AP), (d) *Aspidosperma subincanum* (AS), (e) *Copaifera langsdorffii* (CL), (f) *Endlicheria paniculata* (EP), (g) *Helietta apiculata* (HA), (h) *Hymenaea courbaril* (HC), (i) *Inga vera* (IV), (j) *Pterodon pubescens* (PP), (k) *Syagrus romanzoffiana* (SR).



Source: Adapted from Miyoshi et al. (2020).

Table 1 - Tree species identified in the field and their characteristics.

| Abbreviation | Species | Family | Height (m)/Trunk Diameter (cm) | Characteristics ¹ | MeanH ² (m) |
|--------------|--------------------------------|-------------------------------|--------------------------------|--|------------------------|
| AG | <i>Astronium graveolens</i> | Anacardiaceae | 15–25/40–60 | Deciduous and heliophyte. It starts to bloom during the dry season, usually in July–October. | 14.32 |
| AL | <i>Apuleia leiocarpa</i> | Fabaceae: Caesalpinioideae | 25–35/60–90 | Deciduous and heliophyte. Blooms with the tree completely without leaves, usually in August–September | 12.45 |
| AP | <i>Aspidosperma polyneuron</i> | Apocynaceae | 20–30/60–90 | Evergreen and sciophyte. Blooms in September–November, fruits ripen in April–October. | 15.37 |
| AS | <i>Aspidosperma subincanum</i> | Apocynaceae | 5–20/40–50 | Deciduous, heliophyte, xerophyte selective. Blooms between September and November with fruits ripening in June to August. | 11.4 |
| CL | <i>Copaifera langsdorffii</i> | Fabaceae: Caesalpinioideae | 10–15/50–80 | Semideciduous, heliophyte, xerophyte selective. Blooms between December to March, fruits ripen in August–September with the tree almost without leaves | 10.45 |
| EP | <i>Endlicheria paniculata</i> | Lauraceae | 5–10/30–50 | Evergreen, cyophyte, and hygrophyte selective. Blooms during the summer, January–March, and fruits ripen in May–July depending on the season | 9.88 |
| HA | <i>Helietta apiculata</i> | Rutaceae | 10–18/30–50 | Evergreen, heliophyte, and hygrophytic selective. Blooms between November–December, and fruits ripen in March to May, outside the dry season | 11.69 |
| HC | <i>Hymenaea courbaril</i> | Fabaceae: Caesalpinioideae | 15–20/up to 100 | Semideciduous, heliophyte, xerophyte selective. Blooms in October–December, and fruits ripen from July | 12.46 |
| IV | <i>Inga vera</i> | Fabaceae: Mimosoideae | 5–10/20–30 | Semideciduous, heliophyte, pioneer, and hygrophyte selective. Blooms in August–November, and fruits ripen during the summer, December to February | 9.84 |
| PP | <i>Pterodon pubescens</i> | Fabaceae: Faboideae | 8–16/30–40 | Deciduous, heliophyte, xerophyte selective. Blooms between September to October and the fruits ripen with the tree almost without leaves | 12.31 |
| SR | <i>Syagrus romanzoffiana</i> | Arecaceae | 10–20/30–40 | Evergreen, heliophyte, and hygrophyte selective. Blooms almost during the entire year and fruits ripen mainly in February to August | 10.16 |

¹Information extracted from Lorenzi (1992a, 1992b, 1992c) . ²Mean height from the samples recognized in the field. Source: Adapted from Miyoshi et al. (2020).

Table 2 - The number of samples recognized in the field, their average number of pixels per crown, and its sum for each tree species. Pixels with 10 cm of spatial resolution.

| Abbreviation | Specie | ITCs | Average Pixels/Crown | Sum of Pixels |
|--------------|--------------------------------|------|----------------------|---------------|
| AG | <i>Astronium graveolens</i> | 4 | 2,965 | 11,861 |
| AL | <i>Apuleia leiocarpa</i> | 10 | 2,328 | 23,278 |
| AP | <i>Aspidosperma polyneuron</i> | 3 | 3,098 | 9,293 |
| AS | <i>Aspidosperma subincanum</i> | 4 | 2,533 | 10,132 |
| CL | <i>Copaifera langsdorffii</i> | 17 | 2,148 | 36,520 |
| EP | <i>Endlicheria paniculata</i> | 7 | 1,254 | 8,776 |
| HA | <i>Helietta apiculata</i> | 10 | 1,669 | 16,689 |
| HC | <i>Hymenaea courbaril</i> | 11 | 2,800 | 30,799 |
| IV | <i>Inga vera</i> | 8 | 1,288 | 10,302 |
| PP | <i>Pterodon pubescens</i> | 7 | 2,715 | 19,007 |
| SR | <i>Syagrus romanzoffiana</i> | 20 | 1,315 | 26,293 |

Source: Adapted from Miyoshi et al. (2020).

There was a low number of samples for some species because of challenges when acquiring reference data. First, our study area comprised different successional stages; thus, the species composition varied over the area. Second, we used tree samples that emerged from the canopy. Most of the area has trees in the initial to intermediate secondary stage of succession, which hindered the identification of trees that stood out from the canopy. Additionally, considering the trees of the secondary stage of succession, such as AP (which had 3 samples recognized in the field), it can be noticed the effect called “inverted J shape”, which shows that the number of trees per hectare decreases substantially as the DBH values of trees increase (D’OLIVEIRA et al., 2011; LIMA et al., 2017).

3 REMOTE SENSING DATA

3.1 ALS POINT CLOUD

One ALS point cloud was used to extract a CHM of the study area. The ALS data were acquired and provided by the company Fototerra (<http://www.fototerra.com.br/ingles/>). ALS data were acquired on 11 November 2017 using a Riegl LMS-Q680i laser scanner (RIEGL, Horn, Austria) onboard a manned aircraft at a flight height of 400 m, which resulted in an average density of 8.4 points·m⁻².

ALS point clouds contain XYZ coordinates related to the 3D-structure of a target (HYYPÄ et al., 2001). Based on this knowledge and using the LAStools software (Martin Isenburg, LAStools - efficient tools for LiDAR processing) (ISENBURG, 2014), the CHM was obtained by extracting the digital terrain model (DTM) from the digital surface model (DSM). The steps to obtain the CHM can be summarized in the following steps: (i) Clip data to the study area; (ii) Classify the points into ground and non-ground; (iii) Remove noisy points; (iv) Calculate the height above the ground (i.e., normalize the height values); (v) Create a uniform grid with points representing the CHM.

First, the ALS data were clipped to the interest study area covered by the hyperspectral camera. This was carried out using the *lasclip* tool using as input the boundaries of the imaged area. Next, the point cloud was classified as ground and non-ground (Figure 7) using the *lasground* tool, which considers only the last return of the laser pulse and it is based on the Axelsson (2000) algorithm. In sequence, 3 noisy points were automatically removed with the *lasnoise* tool. This tool was applied using its default parameters and aims to remove isolated points (ISENBURG, [s.d.]). Following, the height above the ground was calculated through the difference between the points classified as non-ground and the one previously classified as ground (step (ii) from the previous paragraph) using the *lasheight* tool. Finally, with the *lasthin* tool, the highest points from the ALS data were selected and used to create a uniform grid, of size 10 cm by 10 cm, to compose the final CHM. These points were assumed to represent the canopy heights. Figure 8 represents the CHM of the area, where it is possible to identify emergent trees from the canopy strata.

Figure 7 - Profile of ALS data classified as ground and non-ground where the pink-colored points represent the ground and the one colored in yellow represents the surface objects (trees). Data acquired in November 2017 with a Riegl LMS-Q680i laser scanner.

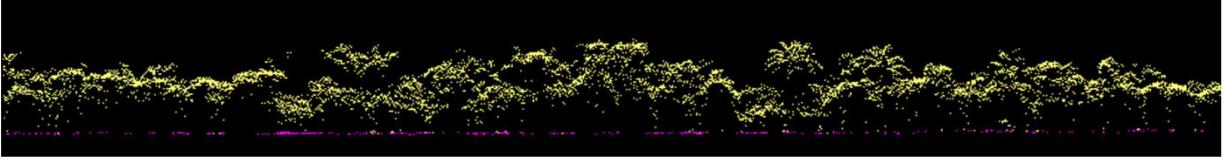


Figure 8 - Canopy height model of the study area obtained from the ALS data. Colors values represent the tree heights.

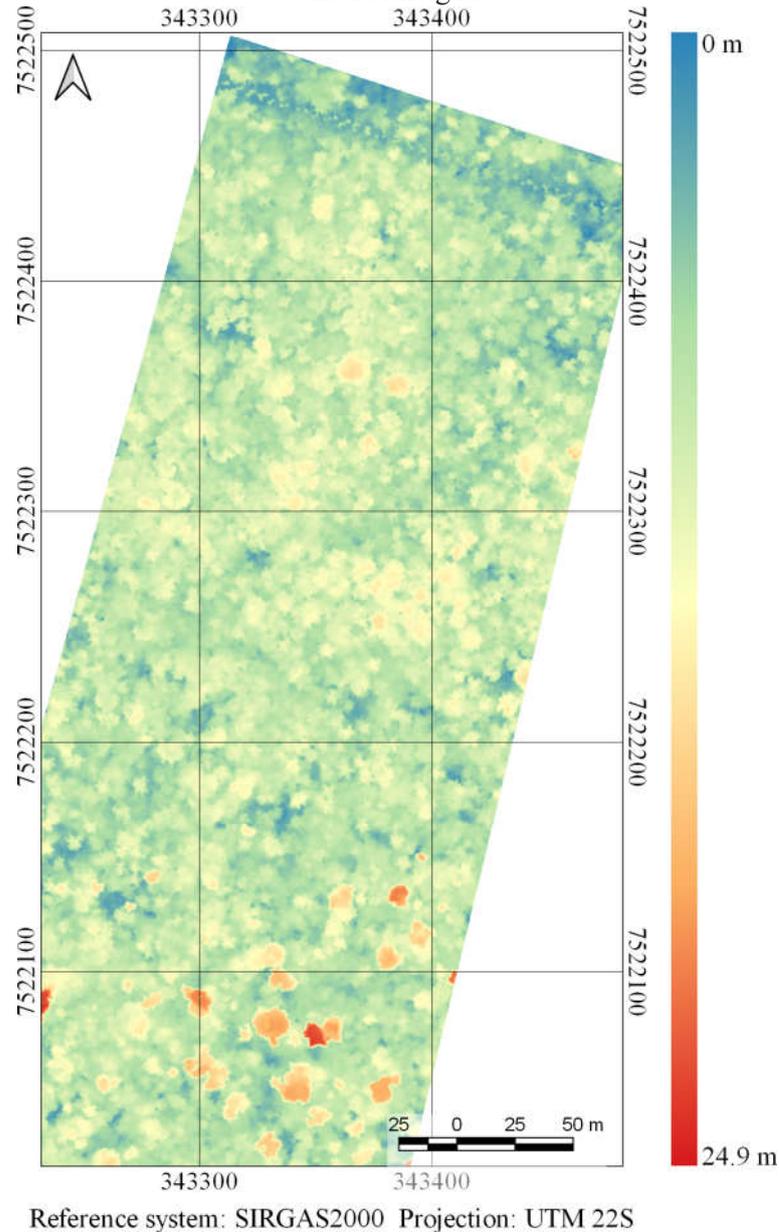
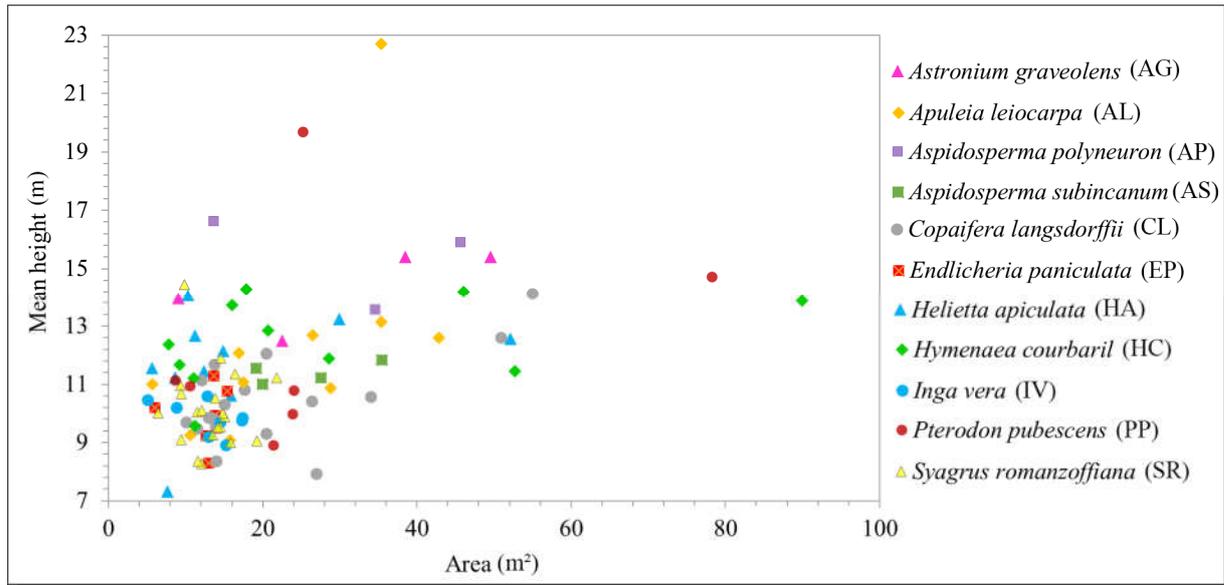


Figure 9 shows the mean height of each tree sample recognized in the field. Most of the observed samples fell within a similar height range. Additionally, taller trees were found in the more developed successional stage of the area. Trees of the same species varied in age and were found in regions of different successional stages. For example, PP trees had

crown areas of around 25 m² and mean heights of 10–20 m. Similarly, HC samples had a mean height of almost 14 m, with tree crown areas ranging from 16 to 90 m². In Figure 9 is also noted that the tree crown area was extracted from the 2017 imagery dataset because it is the imagery data closest to the ALS data.

Figure 9 - Mean tree height versus the tree crown area for all samples identified in the field. Data are from the 2017 dataset.



Source: Adapted from Miyoshi et al. (2020).

3.2 MOSAICS OF HYPERSPECTRAL IMAGES

Hyperspectral images were acquired with a 2D-format hyperspectral camera based on the tunable Fabry–Pérot Interferometer (FPI) from Senop Ltd, model DT-0011 and known as Rikola camera (MIYOSHI et al., 2019, 2018; OLIVEIRA; TOMMASELLI; HONKAVAARA, 2016) (Figure 10a). The camera has two sensors, both of which have 1017 pixels × 648 pixels with a pixel size of 5.5 μm both sides. The total weight of the camera is around 700 g with its accessories, which include an irradiance sensor and a Global Positioning System (GPS) receiver. Spectral bands can be selected from the VIS to near-infrared (NIR) region (500–900 nm). Table 3 summarizes the Rikola camera characteristics. It is important noting that the spectral range of the first and second sensors of the camera are 647–900 nm and 500–635 nm, respectively. Additionally, the spectral bands are acquired sequentially, i.e., the air gap of the FPI moves to acquire the different spectral bands of the same image, being necessary to perform their registration after the image acquisition (AASEN et al., 2018, p. 20; HONKAVAARA et al., 2017; OLIVEIRA; TOMMASELLI; HONKAVAARA, 2016).

Figure 10 - (a) Rikola hyperspectral camera, model DT-0011, with its irradiance and GPS sensors; (b) UX4 UAV attached with Rikola camera.

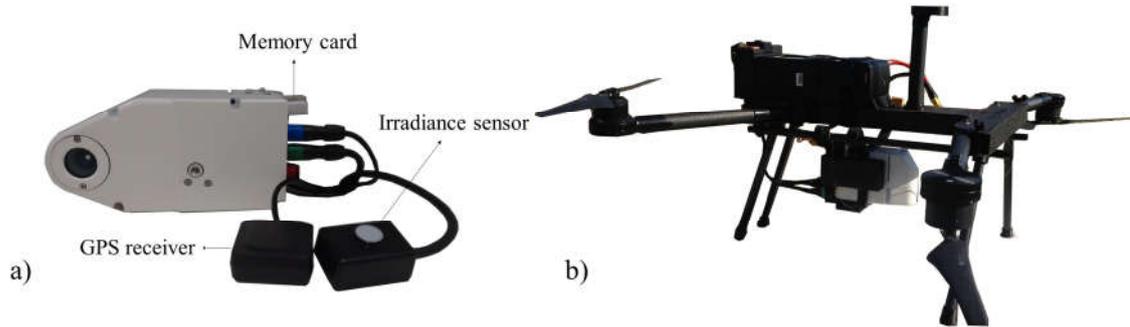


Table 3 - Rikola camera, model DT-0011, specifications.

| Parameter | DT-0011 |
|--------------------------------------|----------------------------|
| Horizontal/vertical FOV ¹ | 36.5°/36.5° |
| Spectral range | 500 nm to 900 nm |
| Minimum bandwidth ² | 10 nm |
| F-stop | ~2.8 |
| Sensors | CMV400 CMOS |
| Image size | 1017 pixels × 648 pixels |
| Pixel size | 5.5 μm × 5.5 μm |
| Approximated focal length | 9 mm |
| Weight with accessories | ~700 g |
| Camera size | 77 mm × 89.5 mm × 152 mm |
| Camera size with memory card | 77 mm × 89.5 mm × 172.7 mm |

¹FOV: field of view

²Bandwidth based on fullwidth at half maximum

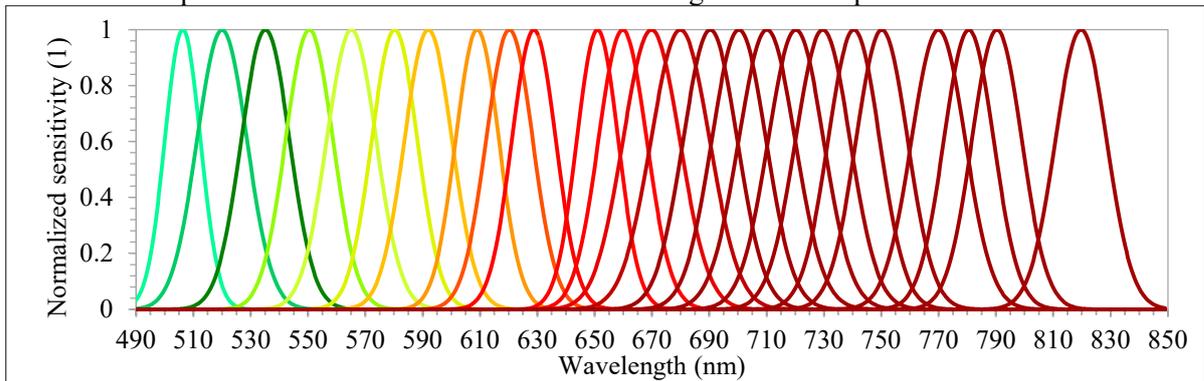
A total of 25 spectral bands were chosen, with the full-width at half maximum (FWHM) varying from 12.44 to 20.45 nm (Table 4 and Figure 11). For this spectral setting, each image cube needs 0.779 s to be acquired. Additionally, the exposure time was set to 5 ms, and the image blocks were divided into two flight strips, ensuring more than 70% and 50% forward and side overlaps, respectively.

Table 4 - Spectral setting of Rikola camera, model DT-0011, with respective bandwidth in full width at half maximum (FWHM).

| Band | λ (nm) | FWHM (nm) | Band | λ (nm) | FWHM (nm) | Band | λ (nm) | FWHM (nm) |
|------|--------|-----------|------|--------|-----------|------|--------|-----------|
| 1 | 506.22 | 12.44 | 10 | 628.73 | 15.30 | 18 | 720.17 | 19.31 |
| 2 | 519.94 | 17.38 | 11 | 650.96 | 14.44 | 19 | 729.57 | 19.01 |
| 3 | 535.09 | 16.84 | 12 | 659.72 | 16.83 | 20 | 740.42 | 17.98 |
| 4 | 550.39 | 16.53 | 13 | 669.75 | 19.80 | 21 | 750.16 | 17.97 |
| 5 | 565.10 | 17.26 | 14 | 679.84 | 20.45 | 22 | 769.89 | 18.72 |
| 6 | 580.16 | 15.95 | 15 | 690.28 | 18.87 | 23 | 780.49 | 17.36 |
| 7 | 591.90 | 16.61 | 16 | 700.28 | 18.94 | 24 | 790.30 | 17.39 |
| 8 | 609.00 | 15.08 | 17 | 710.06 | 19.70 | 25 | 819.66 | 17.84 |
| 9 | 620.22 | 16.26 | | | | | | |

Source: Miyoshi et al. (2020).

Figure 11 - Simulated normalized sensitivities of each spectral band set in the Rikola camera. Responses calculated from the central wavelength and its respective bandwidth.



Source: Miyoshi et al. (2020).

The Rikola camera was mounted onboard the UX4 UAV (Figure 10b), which is a rotary-wing quadcopter developed by the company Nuvem UAV (<https://www.nuvenuav.com/>). The UX4 UAV is almost 90 cm in diameter and 30 cm in height without counting the GPS antenna, which is approximately 15 cm. It is controlled by a PixHawk autopilot. The energy source for the UAV system and its sensors is one six-cell battery of 22 volts and one smaller three-cell battery of 11 volts, which allow the UAV to fly for up to 30 minutes, depending on payload, battery, and weather conditions. A flight speed of $4 \text{ m}\cdot\text{s}^{-1}$ was used to limit the maximum gap between the first and last band of the hyperspectral imager to 3.1 m in a single cube.

During the field campaigns, three radiometric reference targets were placed in the area to enable reflectance calibration. Flight campaigns were performed over the study area (Figure 1) on 1 July 2017, 16 June 2018, and 13 July 2019, with an above-ground flight height of approximately 160 m. Flights were performed during the winter because it had the most suitable weather conditions to fly, i.e., with a lower velocity of winds when compared with summer or spring. The flight height was selected so that a GSD of 10 cm was obtained. This ensured a good representation of tree crowns that were predominantly over 3 m in diameter. Table 5 provides more details about the flight time of each campaign and the mean zenith and azimuth angles of the Sun during the image acquisitions. Observing Table 5 it is possible to notice that the images were acquired in the same season, however, as shown in Figure 4, there were differences in the level of rain in each year as well as before each flight campaign.

Table 5 - Details of the image acquisition in each flight campaign.

| Flight Campaign | Time (UTC-3) | Sun Zenith¹ | Sun Azimuth¹ |
|------------------------|---------------------|-------------------------------|--------------------------------|
| 1 July 2017 | 10:14–10:24 | 56.35° | 38.46° |
| 16 June 2018 | 11:47–11:54 | 46.75° | 12.55° |
| 13 July 2019 | 14:27–14:34 | 52.32° | 325.61° |

¹Data from <<https://www.esrl.noaa.gov/gmd/grad/solcalc/>>

Images were geometrically and radiometrically processed to obtain hyperspectral image orthomosaics. First, the images were radiometrically corrected from the dark current and nonuniformity of sensors using a dark image acquired before each flight and laboratory parameters (HONKAVAARA et al., 2013; MIYOSHI et al., 2018).

The geometric processing was performed using the Agisoft PhotoScan software (version 1.3) (Agisoft LLC, St. Petersburg, Russia). In the orientation process, for each year, the exterior orientation parameters (EOPs) of four reference bands (band 3: 550.39 nm; band 8: 609.00 nm; band 14: 679.84 nm; and band 22: 769.89 nm) were estimated in the same Agisoft PhotoScan project to reduce misregistration between the datasets. The EOPs of the other bands were calculated using the method developed in Honkavaara et al. (2013, 2017). Positions from the camera GPS were used as initial values and refined using a bundle block adjustment (BBA) and ground control points (GCPs). The number of GCPs varied between datasets, with 3, 3, and 4 used in 2017, 2018, and 2019, respectively. GCPs were placed outside the forest since it was not possible to see the ground from imagery acquired over the forested area. A base station was defined near the study area, and the global navigation satellite system (GNSS) observations from GCPs were collected and processed in differential mode.

A self-calibrating bundle adjustment was used to estimate the interior orientation parameters (IOPs) of each sensor and for each year of the dataset. After initial image alignment, parameters estimation was optimized with automatic outlier removal using a gradual selection of tie points based on reconstruction uncertainty and reprojection error, together with the manual removal of points. The final products of this step were the calibrated IOPs, EOPs, sparse and dense point clouds, and DSM of the area with a GSD of 10 cm. These were used in the following radiometric block adjustment and mosaic generation.

Radiometric adjustment processing aims to correct the digital number (DN) of pixels of images from the BRDF effects and differences caused by the different geometries of acquisition due to the UAV and Sun movements. Thus, nonuniformities among images were compensated by the radBA software, developed at the Finnish Geospatial Research Institute

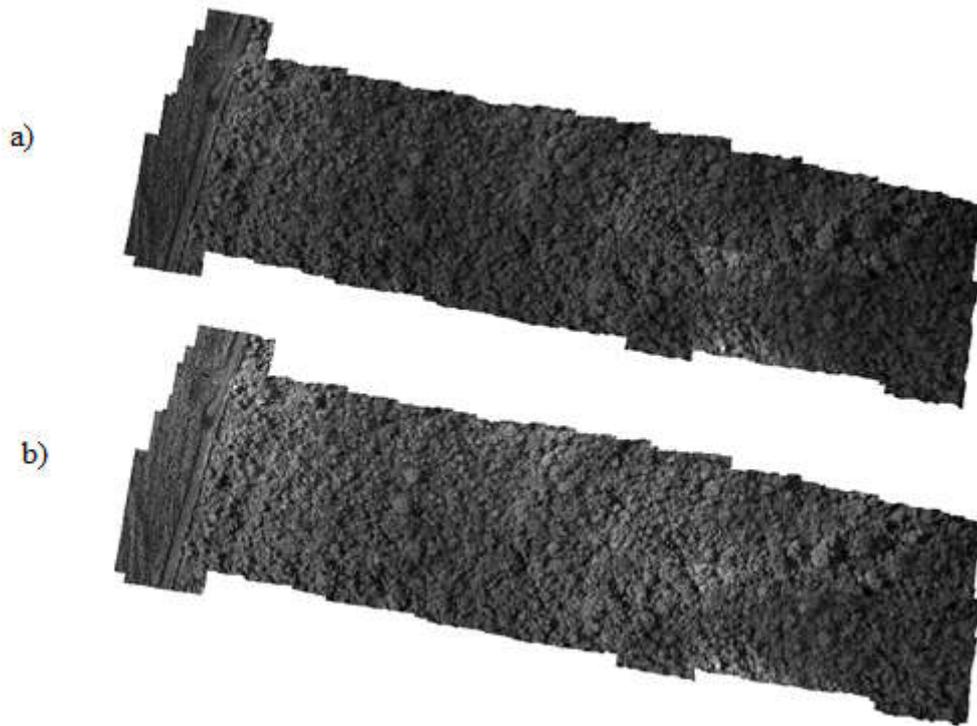
(FGI) (HONKAVAARA et al., 2013; HONKAVAARA; KHORAMSHAHI, 2018). Equation (1) shows the model used in the software to extract the reflectance value from the DN of each pixel.

$$DN_{jk} = a_{relj}(a_{abs} \cdot R_{jk}(\theta_i, \theta_r, \varphi) + b_{abs}), \quad (1)$$

where DN_{jk} is the digital number of pixel k in image j ; $R_{jk}(\theta_i, \theta_r, \varphi)$ is the corresponding reflectance factor with respect to the zenithal angle θ of the incident and reflected light, i and r , respectively, and with the relative azimuthal angle φ ($\varphi_r - \varphi_i$), where φ_r refers to the reflected azimuthal angle, and φ_i denotes the incident azimuthal angle; a_{relj} is the relative correction factor of illumination differences for the reference image; and a_{abs} and b_{abs} are the empirical line parameters for the linear transformation between reflectance and DNs.

According to a previous study of Miyoshi et al. (2018), for the study area, the best initial relative correction factor (a_{relj}) is the one (1) value, with a standard deviation of 0.05. Miyoshi et al. (2018) evaluated whether the radiometric block adjustment provides better results, comparable with the reflectance factor of reference targets. They showed that the lack of radiometric adjustment or the use of irradiance values from the camera sensor, or from spectroradiometer in the field, provide the poorest results, being recommended the use of one (1) as the initial value of a_{relj} . It is worth noting that an exception was necessary for the dataset from 2018 because of higher density differences in cloud covering. The 2017 and 2019 flights were carried out in almost blue-sky conditions, with slight differences compensated by the radiometric block adjustment. The radiometric block adjustment was performed in two steps for the 2018 dataset. First, an initial radiometric block adjustment was performed using initial values of a_{relj} of 1 (one). In sequence, the final values of a_{relj} were used as the initial values for the second radiometric block adjustment. Figure 12 shows the mosaic of band 21 ($\lambda = 750.16$ nm) before and after the radiometric processing, where it is possible to see that the illumination differences were minimized.

Figure 12 - 2018 mosaic of hyperspectral images before (a) and after the radiometric block adjustment (b) of the spectral band centered at 750.16 nm. Automatic contrast from QGIS software (version 3.0.0) in both images.

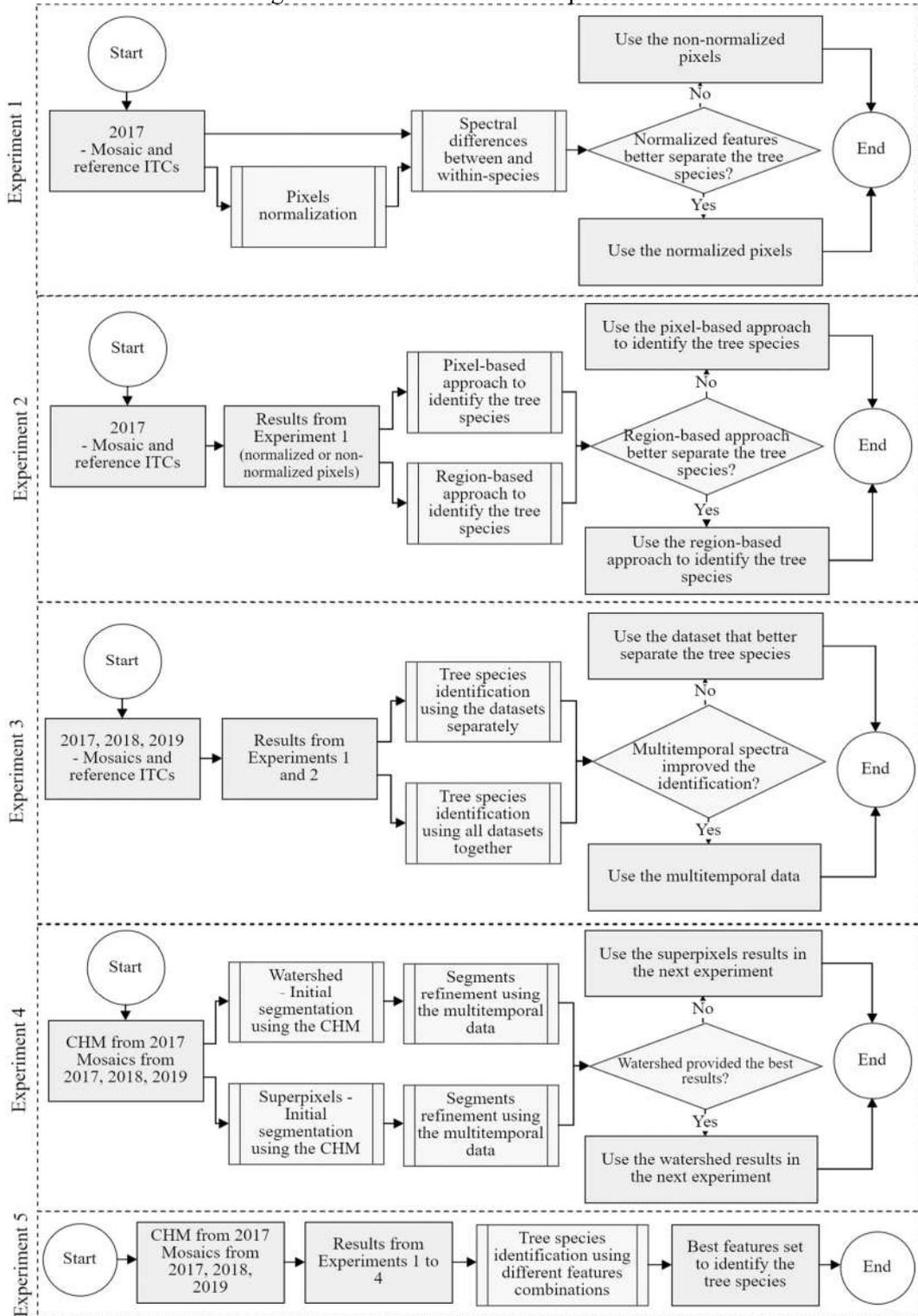


The reflectance factor values were estimated using the empirical line method (SMITH; MILTON, 1999). The empirical line parameters (a_{abs} and b_{abs}) were estimated from the linear relationship between the DN values of three radiometric reference targets and its reflectance factor. The radiometric reference targets had at least $90 \text{ cm} \times 90 \text{ cm}$ and were composed of synthetic material. Colors of the radiometric reference targets were light-grey, grey, and black with average reflectance factor of, respectively, 37%, 11% and 4%. Thus, the mosaics of hyperspectral images for each dataset representing the reflectance factor values were obtained.

4 METHODOLOGY

After the data acquisition and processing, five different experiments were performed to achieve the main goal of this research, they are summarized in Figure 13.

Figure 13 - Flowchart of the experiments¹.

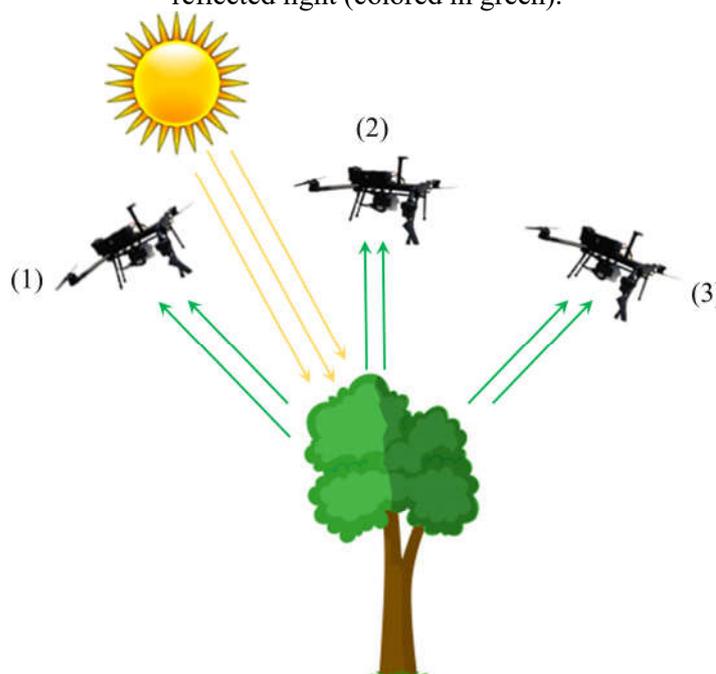


¹Results from Experiment 1 were published in Miyoshi et al. ([s.d.], 2020), results from Experiment 2 is an improvement of the work from Miyoshi et al. (2019), and results from Experiment 3 are in Miyoshi et al. (2020).

The experiments were divided into: i) Spectral differences of tree species; ii) Pixel-based \times region-based classification approaches; iii) Use of temporal spectral information; iv) Multitemporal ITC delineation; v) Tree species identification using temporal data and ITCs. They considered the importance of defining the best spectral feature to identify the tree species because of the BRDF and geometry of view effect when recording the images, the best approach to classify the tree species, the ITC delineation and the use of multitemporal information. and it is important to note that they were performed sequentially, being the results of the next experiment based on the results of the previous experiments.

Initially, the knowledge of the geometry effects of the Sun and sensor positions in vegetation targets were considered. Depending on the geometry of view, vegetation targets can register different amounts of shadowed and sunlit pixels as illustrated in Figure 14 (DALPONTE; FRIZZERA; GIANELLE, 2019; NEVALAINEN et al., 2017; OLIVEIRA; GALVÃO; PONZONI, 2019). In Figure 14, the first sensor position (1) would register higher proportions of sunlight whereas, in position (3), the sensor would record higher proportions of shadows, and the sensor in position (2) would register similar proportions of shadow and sunlight. Additionally, there is the influence of the vegetation structure because depending on the distribution of the crown and its foliage, the density of solar irradiance reaching the canopy varies (WANG; NI-MEISTER, 2019).

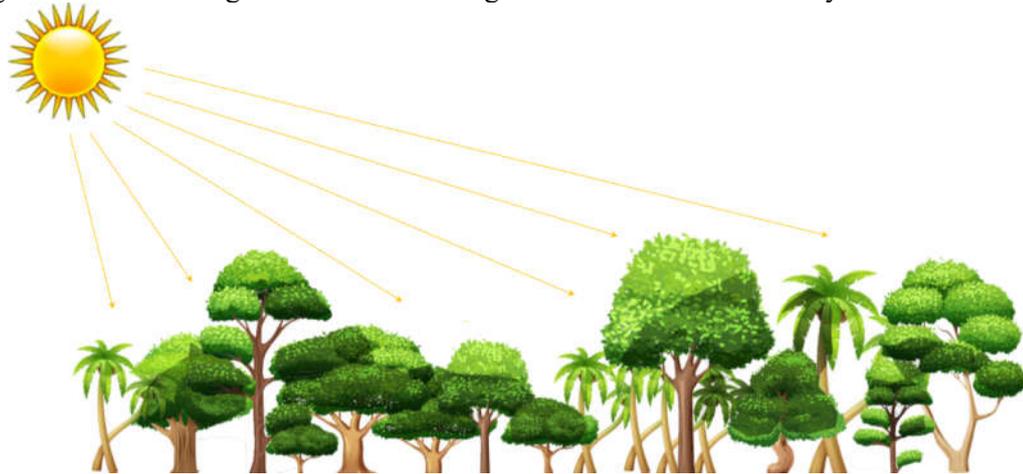
Figure 14 - Differences in the proportions of sunlit and shadowed regions recorded by sensors depending on the geometry of view. The sensors' positions are represented by the number in parenthesis; the arrows represent the directions of the incident light (colored in orange) and of the reflected light (colored in green).



Source: Adapted from Liesenberg (2005).

Besides, the different tree species with different structures and ages, not counting the forest density also affect the amount of incident and reflected light. In Figure 15 the arrows present the direction of the incident light over a forest and the darker green color represent how the shadows can be different in each ITC. The alternative to reduce these differences in the images acquired over the canopy is to apply the pixel normalization procedure. The normalization process reduces the differences between the sunlit and shadowed pixels, assuming a uniform distribution across the crown (DALPONTE; FRIZZERA; GIANELLE, 2019; NEVALAINEN et al., 2017). The normalized pixel value is achieved by dividing the pixel value of a band by the sum of values of this pixel in all bands (Equation (2)) (DALPONTE; FRIZZERA; GIANELLE, 2019). Hereof the use of the mean values of the non-normalized pixels will be referred to as Mean whereas the mean value of the normalized pixels will be referred to as MeanNorm.

Figure 15 - Incident light in a forest resulting in different shadows density in each tree crown.



$$Npix_{ij} = \frac{pix_{ij}}{\sum_{k=1}^n pix_{ik}}, \quad (2)$$

where, $Npix_{ij}$ is the i normalized pixel value in spectral band j ; pix_{ij} is the i pixel value in spectral band j ; n is the number of spectral bands.

From the normalized and non-normalized pixels, it was possible to calculate the spectral distances between-species (inter-classes) and within-species (intra-classes). The method adopted to calculate the spectral differences were based on the distance from Price (1994) (Equation (3)). This difference is given by the root mean square difference between two spectra and averaged through the number of spectral bands since they are uniformly distributed for all tested spectra (PRICE, 1994). Price (1994), who evaluated this metric for distinguished targets, found out that D can have values up to 7.56% within-corn samples,

which could be a potential problem to distinguish this target to other vegetation types, such as sunflower and alfalfa. Castro-Esau et al. (2006) calculated D using different pairwise combinations including different sites and seasons. They observed that leaves from the same tree species may have different spectral characteristics due to its content.

$$D = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [S_1(\lambda_i) - S_2(\lambda_i)]^2}, \quad (3)$$

where, D is the spectral distance between spectra S_1 and S_2 ; N is the number of spectral bands and λ_i represents each spectral band where the difference will be calculated.

The between-species spectral distance was calculated for each pairwise combination of tree species, e.g. AG with AP, AG with PP, AP with PP and so on. Similarly, the within-species spectral distance was calculated for each pairwise combination of samples belonging to the same tree species.

Following, the two-tailed Wilcoxon-Mann-Whitney test (MANN; WHITNEY, 1947; WILCOXON, 1945) was applied to verify whether the between-species spectra were statistically different. Wilcoxon-Mann-Whitney test is applied when there are not enough samples to affirm that the data follow a normal distribution or when they do not follow the normal distribution (MANN; WHITNEY, 1947; NACHAR, 2008). Wilcoxon-Mann-Whitney test is a non-parametric test which verifies if two groups belong to the same population or not. The null hypothesis states that the two groups belong to the same population whereas the alternative hypothesis states that the groups belong to distinct populations. The hypothesis can be calculated by the medians of each group (NACHAR, 2008):

$$H_0: \theta_{sp1} = \theta_{sp2},$$

$$H_1: \theta_{sp1} \neq \theta_{sp2},$$

where, θ is the median of each tree species to be compared sp1 and sp2 to each spectral band.

The test was applied with 95% of confidence level ($\alpha = 5\%$), indicating that when the p-value is lower than 0.05 the null hypothesis is rejected, i.e., there is no spectral difference between the tested pairwise combination.

In sequence, to assess if the findings about the use of normalized or non-normalized pixel values improve the tree species classification, two classification approaches were evaluated because the classification method is inherently important when working with tree species identification. Pixel-based and region-based classification approaches are widely applied (FERREIRA et al., 2016; HEINZEL; KOCH, 2012; WAGNER et al., 2018; ZHANG

et al., 2006). Results from the pixel-based method can be evaluated using the pixel values or using a majority-vote rule, in which the class of a sample is defined by the most popular class of a set of pixels. When working with the region-based method, the ITC delineation method is of high importance. Both image and ALS data-based methods have been proposed for automatic ITC delineation (DAI et al., 2018; EYSN et al., 2015; HYYPPÄ et al., 2001; KAARTINEN et al., 2012; LI et al., 2012)

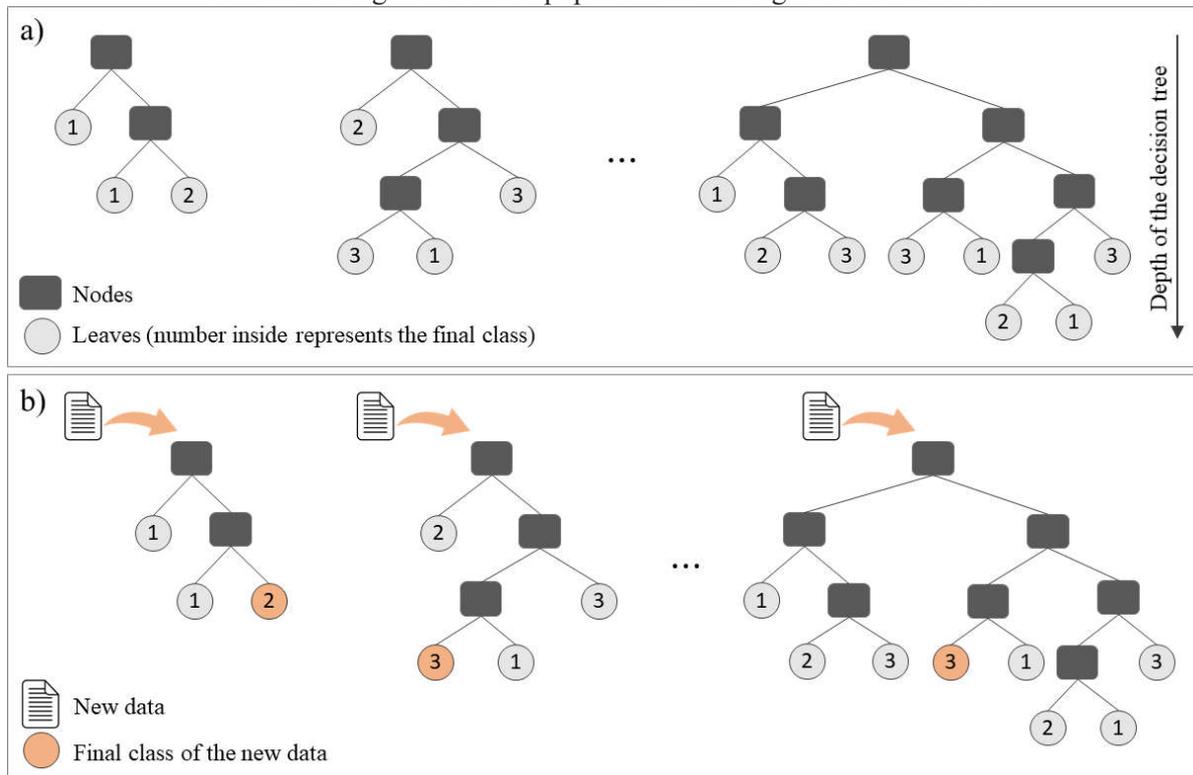
The first evaluated approach considered the use of a pixel-based approach and the second one considered the use of a region-based approach. Previous studies showed that the use of pixel and region-based approaches can provide acceptable results with accuracies higher than 70%. Ferreira et al. (2016) and Wagner et al. (2018) used the majority vote rule of pixels belonging to segments to identify tree-species of well-developed Brazilian Atlantic Forest. Tuominen et al. (2018) and Nevalainen et al. (2017) performed the tree species classification using the region-based approach in Finnish forests. Immitzer, Atzberger and Koukal (2012) compared pixel-based and object-based approaches of four tree species classification in a temperate forest in Austria.

Despite the accuracies achieved by different researches, it is known that when using the pixel-based approach there is no need to image segmentation. However, the pixel-based approach can result in noisy classifications (ZHANG et al., 2006). The existence of shadowed and sunlit pixels may confuse the classifier because of the differences in the density of shadowed and sunlit pixels in each ITC. Overexposed and shadowed pixels tend to have a similar spectral response among tree species hindering the performance of the classifier model. Previous studies from Miyoshi et al. (2019) applied the pixel-based and region-based approaches to verify which one is more suitable for the study area used in this research. The authors used the mean reflectance factor and applied the identification of only four tree species belonging to the Brazilian Atlantic forest.

The classifier method chosen was the RF from Breiman (2001) which is based on multiple decision trees, where the class is determined by the most popular vote. Decision trees are composed of different features, which are drawn with replacement, i.e., one feature can belong to more than one tree (BELGIU; DRĂGUȚ, 2016). Figure 16 exemplifies the principle of RF where multiple decision trees are created using different sets of features samples. In the example, there are three classes divided in s samples with f features. The first step of the classifier is the construction of multiple decision trees (Figure 16a). They are built using different features and can have different depths, nodes and leaves. After the decision

trees are created, it is the classification phase (Figure 16b). New data is used as input and classified in each decision tree. Then, the most popular class of the decision trees, i.e., the most frequent class, is assigned as the class for the input data. Figure 16b shows this principle, where the most popular class of the new data is classified as class “3”, which was the most popular class among the decision trees.

Figure 16 - (a) Random Forest principle, in which multiple decision trees are used to train the classifier. (b) After the classification model is built, the new data are classified in all decision trees and its class belongs to the most popular class among all decision trees.



According to Belgiu and Drăguț (2016) RF does not overfit when increasing the number of decision trees. RF performs an indirect feature selection through a criteria, usually the entropy or the Gini impurity, and thus, providing the features' importance and increasing the RF performance (MENZE et al., 2009). However, RF can be sensitive to the unbalanced distribution of samples per class (CHEN et al., 2004). To reduce this issue, it is possible to under-sample the data or to attribute different weights to the classes based on the number of samples (FARQUAD; BOSE, 2012). Additionally, RF was chosen because it showed good performance to classify targets using hyperspectral data (IMMITZER; ATZBERGER; KOUKAL, 2012; NEVALAINEN et al., 2017; TUOMINEN et al., 2018).

In our work, the RF was applied in the Weka software version 3.8.3 (The University of Waikato, Hamilton, New Zealand) (HALL et al., 2009). Similar to Nevalainen

et al. (2017), the default parameters were used for the classification step. To circumvent the low and unbalanced number of samples, the tree species with the lowest number of samples (i.e., AG, AP and AS) were removed from the classification experiments, being used the 90 ITCs of the remaining tree species (see Table 2). Moreover, the leave-one-out cross validation (LOOCV) method was chosen for the region-based classification. LOOCV is a particular case of k-folds cross validation, where k is equal to the total number of samples of the dataset (BROVELLI et al., 2008; NEVALAINEN et al., 2017). The classification model is trained “k” times. In each iteration, the model is trained using k – 1 samples and tested with the remaining sample. The final accuracy values are obtained by averaging the accuracy values of each iteration (BROVELLI et al., 2008). LOOCV has been successfully applied in tree species classification studies with a small sample size (e.g., less than 10 samples per class (SOTHE et al., 2019)) or with an unbalanced number of samples per class (NEVALAINEN et al., 2017). When applying the pixel-based classification, the 10-folds cross validation was used because of the higher number of samples (more than 200,000 pixels representing all the samples).

The results were evaluated through the area under the receiver operating characteristic curve, known as AUC (area under the curve) ROC (receiver operating characteristics) or AUCROC (BRADLEY et al., 2006; EVANGELISTA et al., 2009; FAN; UPADHYE; WORSTER, 2006). ROC is the relationship between the false positive rate (FPR), or “1-specificity”, and the true positive rate (TPR), or sensitivity, and it is useful when working with unbalanced classes because it is independent of the class distribution (EVANGELISTA et al., 2009; FAWCETT, 2006). When using classifiers such as RF that provide probabilities or scores, thresholds can be applied to acquire different points in the ROC space to form a ROC curve (FAWCETT, 2006). AUCROC is the area under the ROC curve and represents the probability of the classification model correctly classifying a random sample in a specific class. AUCROC varies from 0 to 1 for each class, where a value of 0.5 indicates that the specific classification model is no better than a random assignment, and a value of 1 represents perfect discrimination of a class from the remaining ones (EVANGELISTA et al., 2009). In addition, the overall accuracy (OA) (i.e., the percentage of correctly classified instances of the total number of samples) was calculated.

Besides the use of spectral features, different researches showed the use of structural features derived from the CHM as well as vegetation indexes and textural features

(FERREIRA et al., 2016; MASCHLER; ATZBERGER; IMMITZER, 2018; MICHEZ et al., 2016; NEVALAINEN et al., 2017; SOTHE et al., 2019, 2020; TUOMINEN et al., 2018).

The structural features were extracted from the CHM described in Section 3.1. using the polygons from the ITCs from the 2017 dataset and the *lascanopy* tool from LAStools (ISENBURG, 2014). The extracted features were: (i) mean height (CHMm); (ii) standard deviation of height (CHMstd); (iii) skewness (SKE); (iv) kurtosis (KUR); (v) 90th percentile of height (p90); (vi) 75th percentile of height (p75); (vi) 50th percentile of height (p50); and (vii) 25th percentile of height (p25). These structural features were chosen because they were previously applied by different researches to identify tree species such as in Sothe et al. (2019) and Tuominen et al. (2018).

Haralick, Shanmugam and Dinstein (1973) proposed 14 texture metrics extracted from the grey level co-occurrence matrix (GLCM) widely used for vegetation mapping purposes (DIAN; LI; PANG, 2014; KIM; MADDEN; WARNER, 2009; SOTHE et al., 2019, 2020). Five texture features were available in the *scikit-image 0.13.1* (VAN DER WALT et al., 2014) library for Python and were extracted from the datasets. Texture features were: (i) angular second moment (ASM); (ii) contrast (CON); (iii) correlation (COR); (iv) dissimilarity (DIS); and (v) homogeneity (HOM). The GLCM was calculated with a window size empirically chosen as 5 pixels \times 5 pixels. The texture of each segment was extracted for each spectral band and dataset, summarizing a total of 375 features (5 texture metrics \times 25 spectral bands \times 3 datasets).

Five vegetation indexes were chosen based on the knowledge that the tree species present different spectral responses and based on the previous experiments of Miyoshi (2016), who evaluated leaves and ITC spectra of trees belonging to Ponte Branca area. The vegetation indexes were: (i) normalized difference vegetation index (NDVI) (ROUSE et al., 1974); (ii) red-edge position (REP) (GUYOT; BARET, 1988); (iii) photochemical reflectance index (PRI) (GAMON; PEÑUELAS; FIELD, 1992; SIMS; GAMON, 2002); (iv) structure insensitive pigment reflectance index (SIPI) (PENUELAS; BARET; FILELLA, 1995); and (v) plant senescence reflectance index (PSRI) (MERZLYAK et al., 1999). Similar to the texture features, the vegetation indexes were extracted for each dataset (i.e., from the mosaics of 2017, 2018 and 2019). Table 6 shows the equations from the texture and vegetation indexes and Table 7 summarizes the total of features in each category.

Table 6 - Texture and vegetation indexes extracted from the mosaic of hyperspectral images.

| Type/Name | Equation ¹ | Reference |
|--|--|---|
| Angular second moment (ASM) | $\sum_{i,j=0}^{levels-1} P_{ij}^2$ | |
| Contrast (CON) | $\sum_{i,j=0}^{levels-1} P_{ij} \cdot (i - j)^2$ | |
| Texture Correlation (COR) | $\sum_{i,j=0}^{levels-1} P_{ij} \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}$ | Haralick; Shanmugam; Dinstein (1973) |
| Dissimilarity (DIS) | $\sum_{i,j=0}^{levels-1} P_{ij} i - j $ | |
| Homogeneity (HOM) | $\sum_{i,j=0}^{levels-1} \frac{P_{ij}}{1+(i-j)^2}$ | |
| Normalized difference vegetation index (NDVI) | $NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$ | Rouse et al. (1974) |
| Red-edge position (REP) | $REP = 700 + 40 \frac{\rho_{rededge} - \rho_{700}}{\frac{\rho_{740} - \rho_{700}}{\rho_{670} + \rho_{780}} + \frac{\rho_{780} - \rho_{700}}{\rho_{670} + \rho_{780}}}$ | Guyot; Baret (1988) |
| Vegetation index Photochemical reflectance index (PRI) | $PRI = \frac{\rho_{531} - \rho_{570}}{\rho_{531} + \rho_{570}}$ | Gamon; Peñuelas; Field, (1992); Sims; Gamon, (2002) |
| Structure insensitive pigment reflectance index (SIPI) | $SIPI = \frac{\rho_{800} - \rho_{500}}{\rho_{800} - \rho_{680}}$ | Penuelas; Baret; Filella, (1995) |
| Plant senescence reflectance index (PSRI) | $PSRI = \frac{\rho_{680} - \rho_{500}}{\rho_{750}}$ | Merzlyak et al. (1999) |

¹ P is the grey-level co-occurrence histogram for which texture metric; i is the row number and j is the column number of the grey-level of P ; and $levels$ is the maximum digital number of the tested image (HALL-BEYER, 2017; HARALICK; SHANMUGAM; DINSTEIN, 1973).

Table 7 - Summary of the extracted features.

| Type of feature/ Description | Number of features in each dataset | Number of datasets (years) | Total of features |
|---|------------------------------------|----------------------------|-------------------|
| Spectral: mean normalized reflectance factor for each spectral band | 25 | 3 | 75 |
| Structural: mean, standard deviation, 25 th , 50 th , 75 th , and 90 th percentiles, kurtosis and skewness of the CHM | 8 | 1 | 8 |
| Textural: angular second moment, contrast, correlation, dissimilarity and homogeneity for each spectral band | 125 | 3 | 375 |
| Vegetation indexes: normalized difference vegetation index, red-edge position, photochemical reflectance index, structure insensitive pigment reflectance index, plant senescence reflectance index | 5 | 3 | 15 |
| Total: | | | 473 |

Last, but not least, it is noticed the importance to automatically delineate the ITCs to improve the efficiency to identify the tree species. ITCs detection and delineation are highly important to the production of tree species maps (CLARK; ROBERTS; CLARK, 2005; DALPONTE et al., 2014). Additionally, the use of individual tree polygons showed to increase the accuracy when classifying the tree species (CLARK; ROBERTS; CLARK, 2005; DALPONTE et al., 2014; FERET; ASNER, 2013). However, the automatic ITC delineation is particularly challenging, especially in tropical forests, where the crown boundaries cannot be well defined due to the mixture of leaves and branches of neighbor trees, without considering the lianas and understory influence (COLGAN et al., 2012). Besides, when using multitemporal data with very high spatial resolution, the spatial differences of the same tree crown in the different years are visible as seen in Figure 5 from Section 3.2 and should be considered in the segmentation task.

Two segmentation approaches were assessed: the simple linear iterative clustering (SLIC) algorithm to generate superpixels, and the watershed segmentation. Superpixels were successfully applied to segment and classify hyperspectral images and to evaluate the successional stage of forests (BERVEGLIERI et al., 2018; ZHANG et al., 2017). Watershed showed to provide acceptable results to ITCs delineation in boreal forests (KAARTINEN et al., 2012; NÄSI et al., 2015; TANHUANPÄÄ et al., 2014). Both segmentation approaches were divided into two steps: (i) initial segments generation using the CHM from the ALS data acquired in 2017; (ii) segments refinement considering the multitemporal spectral information, which varies in the different years as seen in the results from Section 5.3.

The SLIC algorithm was applied using the *scikit-image 0.13.1* Python library (VAN DER WALT et al., 2014). This algorithm uses the CIELAB space color and the pixel positions to calculate a 5D distance (ACHANTA et al., 2012). Three of the distances correspond to the values of the pixel color in the CIELAB space and the other two correspond to the xy pixel position. The algorithm works interactively, updating clusters based on the k-means clustering, which uses a $2S \times 2S$ window search size. S is the grid size of the initial superpixels defined by an approximated number of superpixels, which is the only input parameter needed in the algorithm.

Since SLIC uses an RGB image, the CHM presented in Section 3.1 was exported to a raster format using the *las2dem* tool from LAsTools (ISENBURG, 2014). As parameters, we used a false-color image, created based on the elevation values, in meters. In this image, the DN of each pixel was multiplied by 1.5 to increase the elevation differences (Figure 17).

Following, different numbers of superpixels were tested to minimize the error on the ITC delineation. 2,000, 10,000 and 20,000 were some of the tested values (Figure 18). It is worth noting, that those values can be chosen according to the expected number of trees. In our case the number of superpixels was chosen in order to generate smaller segments and thus, avoid under-segmentation (i.e., one segment containing more than one tree crown). Moreover, the algorithm considers a rectangular area delimited by the upper left and bottom right corners of the study area and not the four corners of the area.

Figure 17 - False-colour CHM raster (a) without scale and (b) with scale in elevation of 1.5 times. The blue color represents the lower heights whereas the red color represents the highest values.

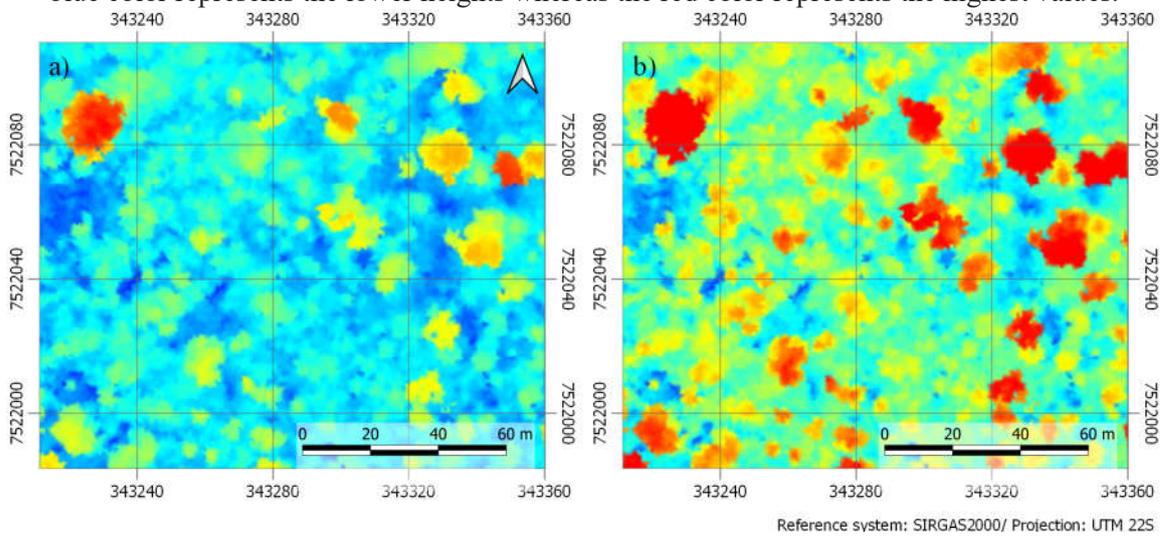
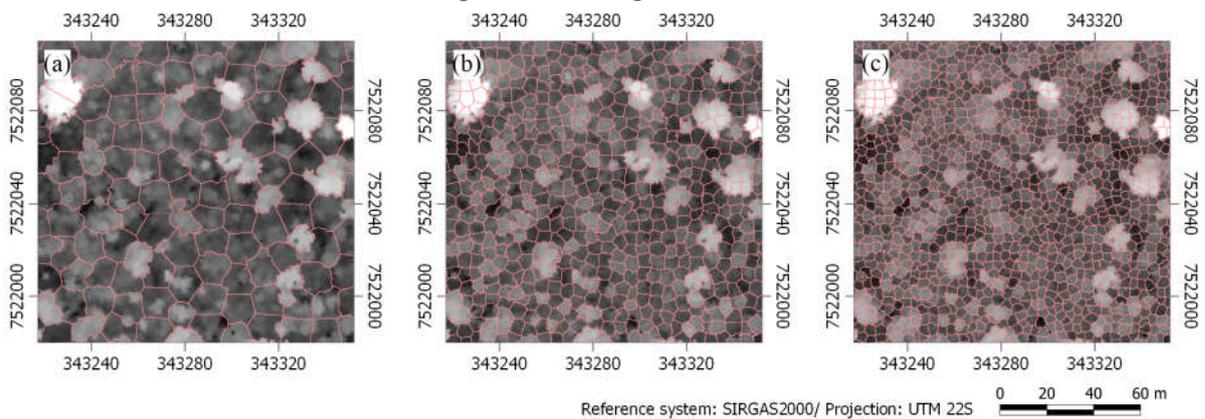


Figure 18 - Superpixels generated using initial parameter of (c) 2,000, (d) 10,000, and (e) 20,000. In the greyscale images, the darkest color represents the lower height values whereas the brightest color represents the higher values.



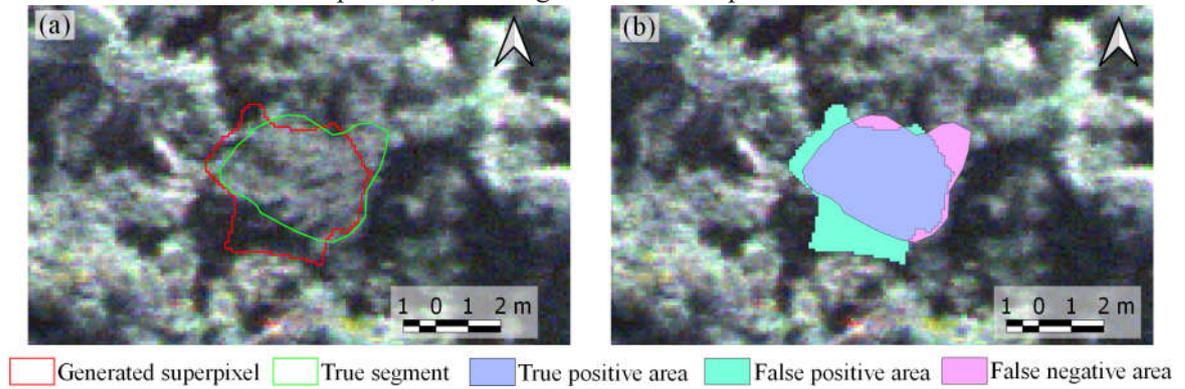
Using the initial value of 20,000 superpixels and excluding those belonging to the non-imaged area, around 10,000 superpixels was generated. In sequence, the refinement process was carried out to merge adjacent superpixels which could belong to the same tree crown. This step was required because the wider tree crowns were in more than one segment, resulting in over-segmentation. The criteria to merge the superpixels considered the minimum

adjacency length between superpixels, their absolute maximum heights difference, the maximum and mean height values of each superpixel, and the spectral distance D presented in Equation (3) from Section 5.1.

Watershed was the second method applied and it is based on mathematical morphology (TARABALKA et al., 2008). The method considers the topographic relief to create catchment basins (TARABALKA et al., 2008). Based on points with minimum height values, the influence zones are created (LI; XIAO, 2007). In our study, we used the watershed from *ForestTools 0.2.0* package of R, proposed and implemented by Plowright (2018). The watershed implemented in the package is based on Popescu and Wyne (2004) and in Beucher and Meyer (1993). The method considers an inverted CHM, being the treetops the seed points used to create the influence zones. A mean smooth filter sized 3×3 pixels was applied three times to remove small gaps in the CHM. The small gaps can be caused by the detailed information recorded in the ALS data, which can show the different leaves layouts of a tree crown. Moreover, the smooth filter is widely applied in different researches that detect and delineate the tree crowns such as in Näsi et al. (2015) and Silva et al. (2016). Following, a linear function to define the radius search of the treetops was set as 0.05 from the CHM values. This value was empirically chosen and considered that taller trees usually have wider crowns (POPESCU; WYNNE, 2004). Besides, to avoid tag lower trees as a “treetop”, the minimum height of 8 m was set. In sequence, with the detected treetops, the watershed was applied using the criterion that pixels belonging to the influence zone (i.e., to the ITC polygons) should be greater than 7.5 m.

To evaluate the segmentations, the 90 samples of the 8 tree species with the highest number of samples were used in the classification experiments (Sections 2.1, 5.2, and 5.3) were used. The minimum distance between centroids from the ground reference and the generated segments was the criterion applied to select the corresponding segments. Centroids and matrix distance were calculated with QGIS software version 3.0.0 (QGIS DEVELOPMENT TEAM, 2009). In sequence, the number of pixels classified as true positive, true negatives and false negatives was used to calculate the user accuracy, producer accuracy, F-Score, omission error, and commission error (CONGALTON, 1991; LI et al., 2012). Figure 19 shows an example of areas with pixels representing the true positives, the false negatives and the false positives.

Figure 19 - (a) Example of manually delimited ITC and the superpixel representing the same ITC. (b) True positive, false negative and false positives areas.



5 RESULTS AND DISCUSSION

5.1 SPECTRAL DIFFERENCES – ILLUMINATION DIFFERENCES REDUCTIONⁱ

This section presents the spectra assessment between-species (inter-classes) and within-species (intra-class) since the study area comprises different trees with different ages and stages of development. Besides, the variability of samples belonging to the same tree species may affect the classification, causing confusion among the classes. In this sense, this experiment aims to compare the mean reflectance factor spectra and the mean normalized reflectance factor spectra to verify which one presents higher spectral separability between the tree species. All tests were performed using the mosaic of hyperspectral images acquired in July 2017 and the manually delineated ITCs.

Spectral differences within-species were calculated for all pairwise combinations, summarizing a total number of combinations as following (tree species and number of combinations in parenthesis): AG (6); AL (45); AP (3); AS (6); CL (136); EP (21); HA (90); HC (55); IV (28); PP (21); SR (190). Additionally, a total of 55 pairwise combinations were calculated for the between-species spectral difference. The Wilcoxon-Mann-Whitney test was applied to each pairwise combination between-species and for each spectral band with a confidence level of 95% ($\alpha = 95\%$). Considering 11 tree species, there are a total of 55 pairwise combinations for each spectral band, summing up a total of 1375 tests (25 spectral bands times 55 pairwise combinations). This number of hypothesis tests was applied to the mean spectra (i.e., Mean) as well as to the mean normalized spectra (i.e., MeanNorm) using the *SciPy 1.1.0* package for Python (VIRTANEN et al., 2019).

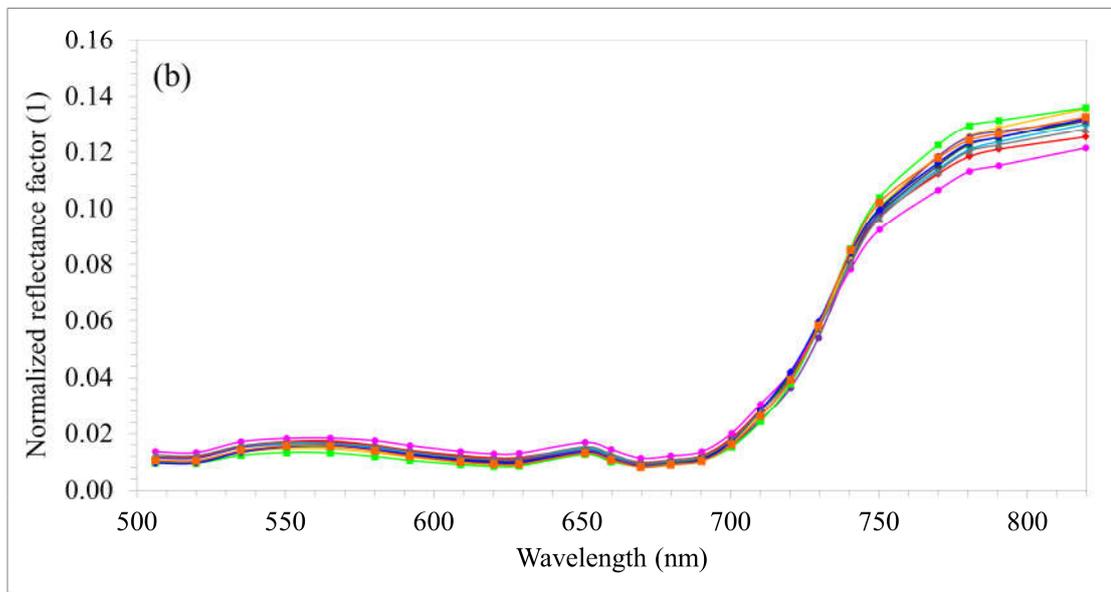
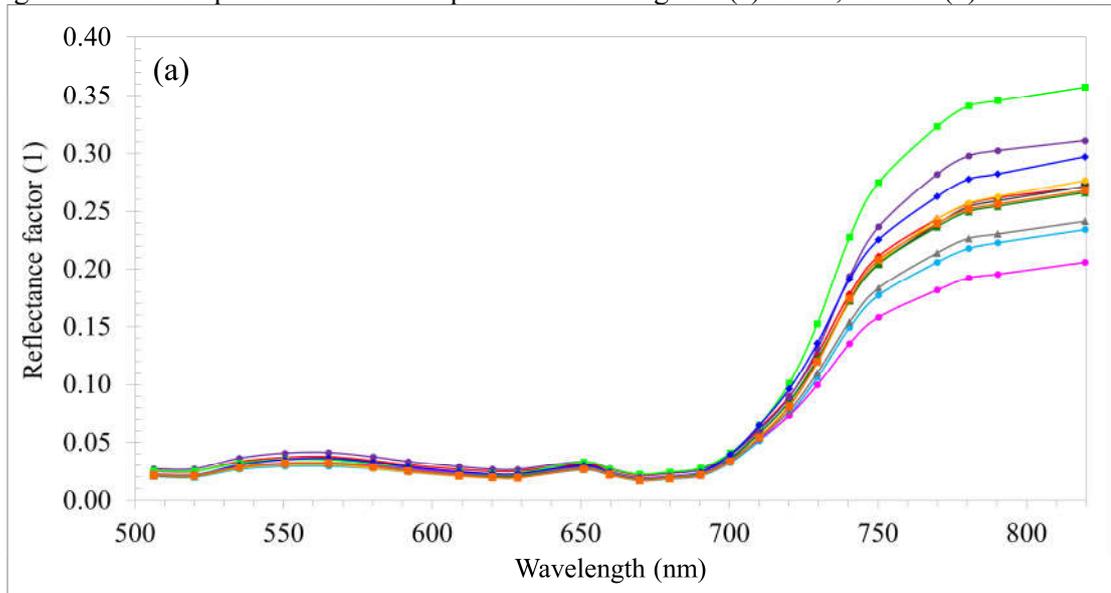
Figure 20 presents the mean reflectance factor spectra (i.e., Mean) and the mean normalized reflectance factor spectra (i.e., MeanNorm). Similar spectral responses are noticed in the VIS region for both Mean and MeanNorm. In the NIR region, the Mean spectra are visually similar between IV, HA, HC, AS and AL and between AG and CL. Despite smaller differences among the MeanNorm spectra, which may lead to higher classification confusion, the spectral variability within the samples of Mean had a higher range (Figure 21). In

ⁱ Results to be published in the ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, and published at Remote Sensing Journal (ISSN 2072-4292). MIYOSHI, G.; IMAI, N.; TOMMASELLI, A.; HONKAVAARA, E. Spectral differences of tree species belonging to Atlantic forest obtained from UAV hyperspectral images. **To be published**, [s. l.], [s.d.].

MIYOSHI, G. T.; IMAI, N. N.; GARCIA TOMMASELLI, A. M.; ANTUNES DE MORAES, M. V.; HONKAVAARA, E. Evaluation of Hyperspectral Multitemporal Information to Improve Tree Species Identification in the Highly Diverse Atlantic Forest. **Remote Sensing**, v. 12, n. 2, 2020. Available at: <<https://www.mdpi.com/2072-4292/12/2/244>>

Figure 21, the range between the minimum and maximum values is visually the same for both the Mean and MeanNorm spectra of all tree species in the VIS part of the electromagnetic spectrum. It is noted that the number of samples of each tree species can affect this range of variation, as observed for SR with 20 samples. However, this behavior was not observed for AL (10 samples) and HC (11 samples). The range variation in the Mean values from the red-edge (700 nm) to near-infrared region (820 nm) had a higher variability when compared with the MeanNorm values, leading to the conclusion that a higher variability may influence classifier performance. Moreover, in Figure 20, an unusual peak may be noticed at the spectral response at 650 nm, probably because this spectral band is located near the edge of the first sensor from the FPI, which acquires information from 647 nm to 900 nm as mentioned in Section 3.2.

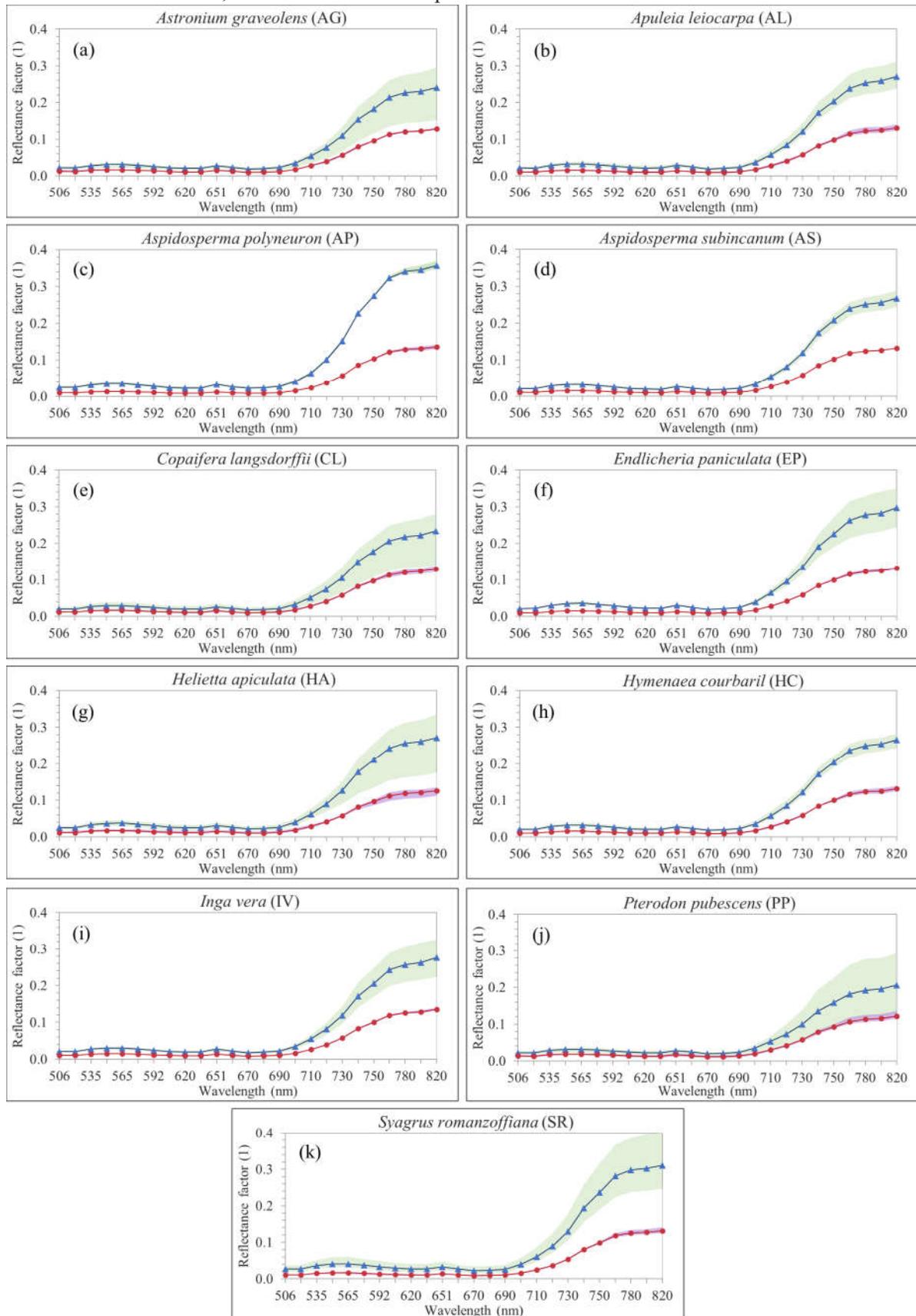
Figure 20 - Mean spectra of each tree species considering the: (a) Mean, and the (b) MeanNorm.



- *Apuleia leiocarpa* (AL) ■ *Aspidosperma polyneuron* (AP) ■ *Astronium graveolens* (AG)
 ■ *Aspidosperma subincanum* (AS) ■ *Copaifera langsdorffii* (CL) ■ *Endlicheria paniculata* (EP)
 ■ *Helietta apiculata* (HA) ■ *Hymenaea courbaril* (HC) ■ *Inga vera* (IV)
 ■ *Pterodon pubescens* (PP) ■ *Syagrus romanzoffiana* (SR)

Source: Adapted from Miyoshi et al. (2020).

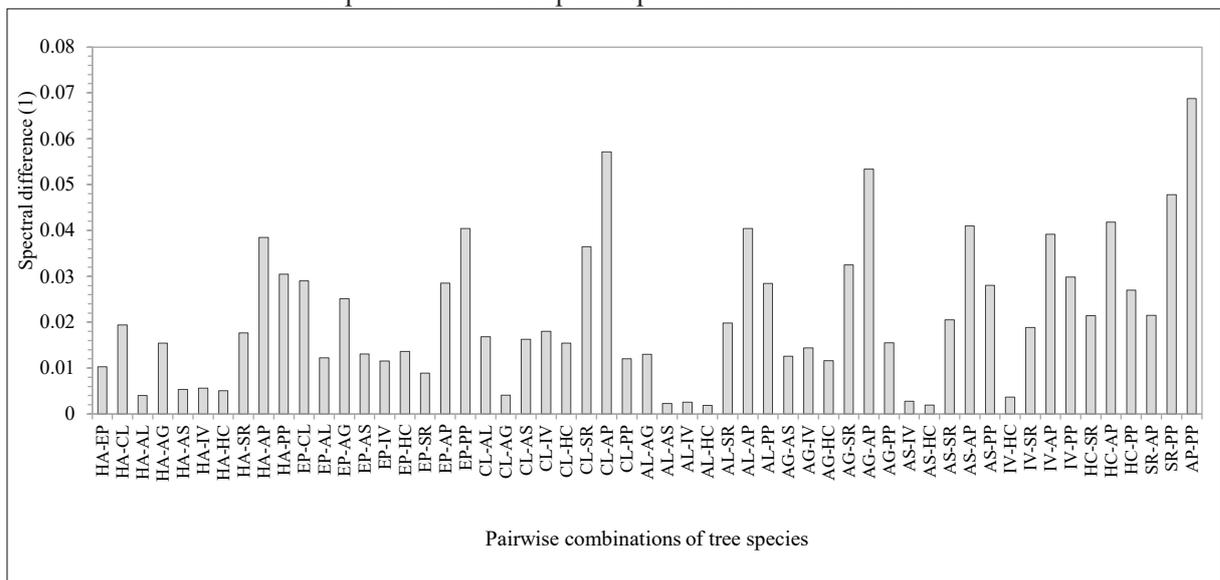
Figure 21 - Values for each tree species considering the mean reflectance factor spectra (Mean) and the mean normalized spectra (MeanNorm). The blue line represents the Mean, the red line represents the MeanNorm, and the shaded area represents the minimum and maximum values.



Source: Adapted from Miyoshi et al. (2020).

The between-species spectral differences calculated using the mean spectra acquired per species, i.e., a unique spectra per class, are presented in Figure 22. As previously mentioned, the spectral distance D was calculated for each pairwise combination. The highest differences are for AP to the other tree species, with an average value of 0.043. This indicates that this tree species presents a higher difference in amplitude when compared with the other tree species, especially from PP where the difference was 0.0688. Further, AP presented a high spectral distance when considering AS even belonging to the same genus. The smallest distance of AP was with SR, where D was 0.0215.

Figure 22 - Spectral differences between the tree species when using the Mean feature. The x-axis represents the tree species pairwise combination¹.



¹Check Table 1 for tree species description.

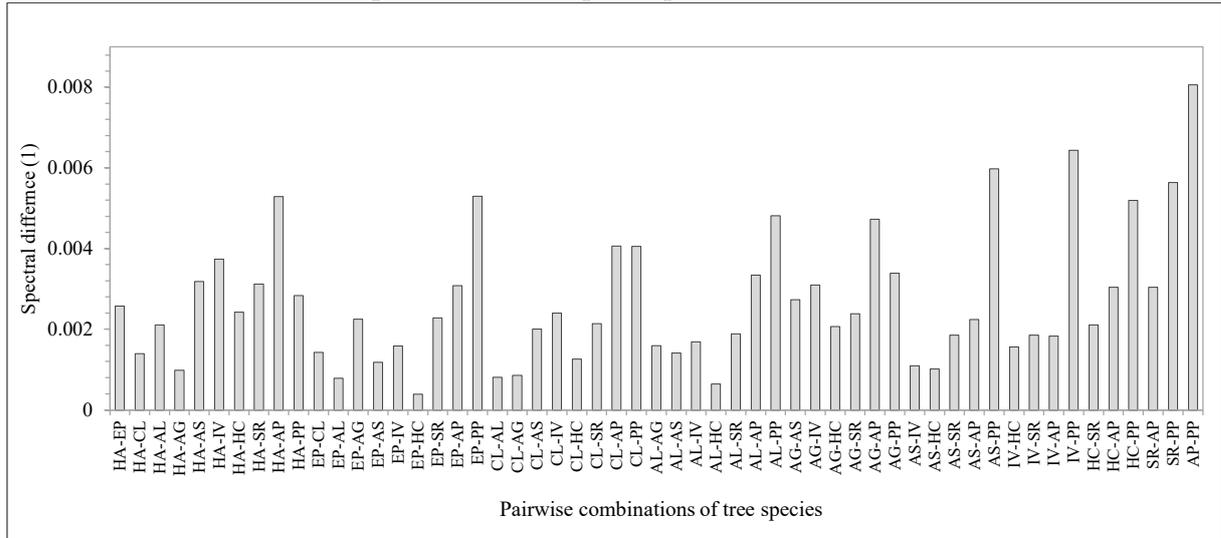
Source: Miyoshi et al. ([s.d.]).

The smallest differences were obtained to the HC with AL and with AS, both with distance equal 0.0019. In sequence, AL with AS (D of 0.0023), IV with AL (D of 0.0026) and, IV with AS (D of 0.0027). From these distances, it can be noticed the challenge when classifying the tree species since they have similar spectra. Moreover, it is observed that the smallest differences are mainly for AL, AS and HC. Even belonging to different botanical genus and families, with different leaf sizes and blossoming, their spectra can be similar. Moreover, 443 of the 1375 tested combinations of the hypothesis test had a p-value lower than 0.05, being the null hypothesis rejected. In other words, only 32% of the observations are likely to belong to different populations, i.e., spectrally different (APPENDIX A).

Considering the MeanNorm spectra, the minimum and maximum spectral differences were 0.0040 and 0.0081 (Figure 23). The highest differences were for PP to the

other tree species, different from the results obtained when using the Mean spectra. The smallest D value was for the pairwise between EP and HC. In fact, HC together with AL and CL presented the smaller values to the other tree species, being possible to infer that confusion among these tree species with the others may occur during the classification process.

Figure 23 - Spectral differences between the tree species when using the MeanNorm feature. The x-axis represents the tree species pairwise combination¹.



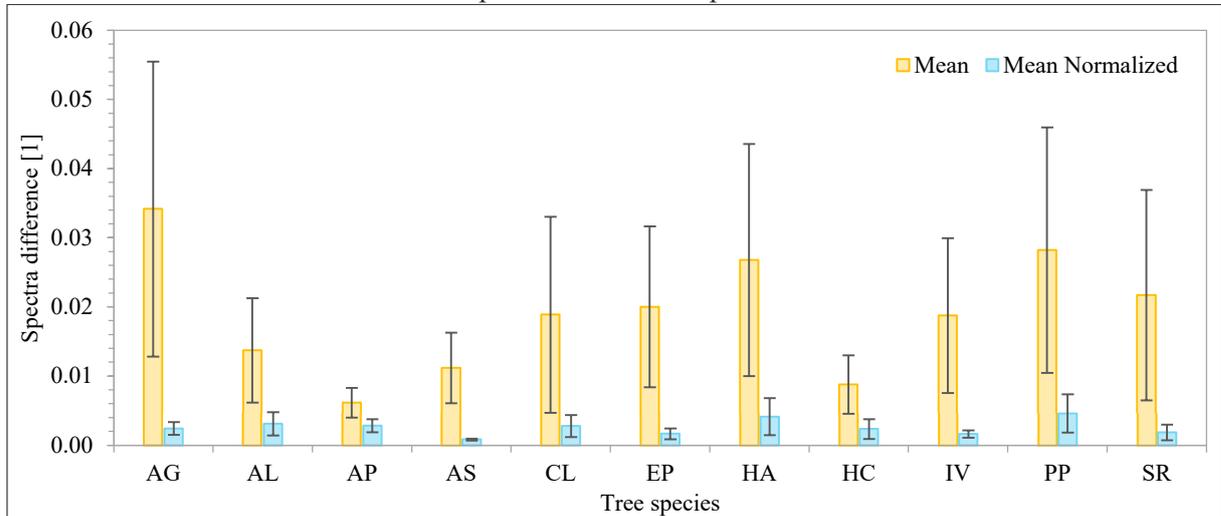
¹Check Table 1 for tree species description.

Source: Miyoshi et al. ([s.d.]).

Results from the hypothesis test indicate that 40% of the observations belong to different populations when using the normalized spectra, i.e., 555 cases of the 1375 tests (APPENDIX A). In this sense, the higher number of rejections of the null hypothesis when using the normalized spectra leads us to the conclusion that this feature could provide better results when classifying the tree species.

For the within-species case, the mean spectral variability of each tree species considering the Mean and the MeanNorm spectra is presented in Figure 24. As expected, the highest differences within-species are for the Mean spectra, indicating that the spectral variability intra species are higher when using the mean spectra than using the normalized values. Before the normalization, the average spectral difference among the samples of AG was 0.0095. After the normalization, the value was reduced to 0.0007 less than twice the original value. Even AP, which presented the smallest difference through the six pairwise combinations, the average difference was reduced from 0.0017 to 0.0008.

Figure 24 - Spectra difference within-species considering the mean spectra and the mean normalized spectra for each sample¹.



¹Check Table 1 for tree species description.

Source: Miyoshi et al. ([s.d.]).

There were no values higher than 0.0046 for the differences within-species when considering the normalized spectra. The smallest value is no longer for AP, but for the AS tree species (D of 0.0008). Indeed, the order of spectral differences has been completely changed. IV was the fifth tree species with the smallest spectral difference within-species, but now it is the second one. Now, HA and PP presented the higher within-species differences with average spectral distances of 0.0041 and 0.0046, respectively. Additionally, comparing the results from Figure 24 with the one obtained in Figure 23 it is noticed that the difference within-species in some cases are higher than the difference between-species as for HA. In this example, the difference within-HA sample was 0.0268 when using the MeanNorm spectra. Pairwise combinations of HA with the other tree species, with exception to PP and AP, presented smaller differences than within spectra differences, i.e., lower than 0.0268. A similar analysis is observed for SR, where the variability among the samples of its species is only lower than the difference of SR with PP, AG, and CL.

From the results obtained in the performed experiments, it can be noticed that the high spectral variability within-tree species can be a challenge when trying to classify the tree species. Possible reasons for these results are related to the different development stages of the trees, since the tree spectra change according to its age, development stage and environment (BUDDENBAUM; SCHLERF; HILL, 2005). Despite of using leaves of trees belonging to the Amazon forest in different ages, Chavana-Bryant et al. (2017) found out that the leaves' spectra were different, supporting our hypothesis about the within-species

variability in tropical environments. Additionally, deciduous trees commonly have an irregular crown shape, which makes the reflectance pattern of tree species more difficult to be recognized (KE; QUACKENBUSH, 2011). Moreover, lower trees can have their spectra affected by neighborhood trees as well as the proportion of sunlit and shadowed pixels.

Finally, considering just the hypothesis tests, it is noticed confusion among the observations. However, the mean normalized spectra presented lower confusion between-species. Secondly, the use of normalized spectra presented lower values of spectral difference D as well as a lower degree of variation. Thus, it can be suggested that the use of normalized spectra decreases the variability of crowns spectral response, caused by different factors such as crown structure and BRDF effects, and it is recommended to be used in classification purposes.

5.2 COMPARISON OF PIXEL-BASED AND REGION-BASED APPROACHESⁱⁱ

This experiment aims to verify which classification approach is more suitable for the tree species identification task and to assess the results from Section 5.1. It is important noting that in this experiment, only the mosaic of hyperspectral images from 2017 was used and the evaluation used the manual delimited polygons because the objective was to verify the best classification approach and not assess the use of multitemporal data to tree species classification.

Results achieved for each experiment are given in Table 8. Analyzing only the use of the pixel-based approach, it is noticed that they achieve similar OA values, 42.858% and 42.922%, respectively for Mean and MeanNorm. Moreover, similar AUCROC values between the Mean and MeanNorm when using the pixel-based approach were achieved for all tree species except for EP. EP is an evergreen tree species with an irregular shape crown (LORENZI, 1992c), which may have caused different shadow densities and consequently have affected the value of AUCROC when using the spectral average (i.e., Mean).

ⁱⁱThis section is an extended version of the paper published at the ISPRS-International Archives of the Photogrammetry. MIYOSHI, G.; IMAI, N.; TOMMASELLI, A.; HONKAVAARA, E. Comparison of Pixel and Region-Based Approaches for Tree Species Mapping in Atlantic Forest Using Hyperspectral Images Acquired by Uav. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. 4213, p. 1875–1880, 2019.

Table 8 - AUCROC values for each tree species identified in each classification experiment as well as its overall accuracy. Shaded cells indicate the highest AUCROC value of each tree species.

| Tree species ¹ | Mean | | MeanNorm | |
|-----------------------------|-------------|--------------|-------------|--------------|
| | Pixel-based | Region-based | Pixel-based | Region-based |
| AL | 0.695 | 0.474 | 0.691 | 0.608 |
| CL | 0.71 | 0.721 | 0.711 | 0.821 |
| EP | 0.722 | 0.812 | 0.788 | 0.818 |
| HA | 0.727 | 0.633 | 0.732 | 0.594 |
| HC | 0.762 | 0.796 | 0.756 | 0.8 |
| IV | 0.798 | 0.729 | 0.8 | 0.627 |
| PP | 0.863 | 0.742 | 0.862 | 0.713 |
| SR | 0.915 | 0.954 | 0.917 | 0.986 |
| Overall accuracy (%) | 42.858 | 38.889 | 42.922 | 55.556 |

¹Check Table 1 for tree species description.

In the pixel-based approach, higher differences in OA values are noted, 38.889% and 55.556%, respectively for Mean and MeanNorm. The use of a region-based approach showed higher differences between the AUCROC values of Mean and MeanNorm, suggesting that the use of normalized spectra is better than when using the mean spectra. In fact, the normalization reduced the illumination differences and improved the identification of most of the tree species. For example, CL and EP had AUCROC values higher than 0.8 (i.e., 0.821 and 0.818, respectively) when using the MeanNorm. However, HA, IV and PP were better identified when using the non-normalized spectra, indicating that the response of shadowed pixels assisted the model to identify these tree species.

Comparing the results obtained in Table 8 it is observed that the use of region-region approach with the MeanNorm produces most of the highest AUCROC values for the tree species identification. Thus, it is possible to conclude that the use of the region-based approach is more appropriate as well as the use of normalized spectra, confirming the findings of Section 5.1. The results are in accordance with the one obtained by the previous study of Miyoshi et al. (2019). Regarding the spectral features, Miyoshi et al. (2019) assumed the pixel values for training and validation whereas we used the normalized pixel values. Additionally, although Clark and Roberts (2012) reported higher accuracy for the majority vote of pixels, Immitzer, Atzberger and Koukal (2012) found out the region-based approach was better. Despite both authors applied RF, it is worth noting their different datasets, not only about imagery data, but also different forest types, tropical rainforest and temperate forest, respectively for Clark and Roberts (2012) and Immitzer, Atzberger and Koukal (2012). Regarding the rainforest studied by Clark and Roberts (2012), it is noted that their forest area

was an old-growth forest located in Costa Rica, not with different successional stages and not being submontane semideciduous seasonal forest.

5.3 CONTRIBUTION OF TEMPORAL DATA TO IDENTIFY THE TREE SPECIESⁱⁱⁱ

The results presented in the previous Sections (5.1 and 5.2) showed that the region-based approach provides better performance in our study area as well as the use of normalized pixel values. Thus, this Section shows an initial assessment of the use of spectral temporal information to improve the tree species identification in the region-based approach.

The classification process was carried out five times with four different imagery datasets (described in Section 3.2): (i) the 2017 (D17); (ii) the 2018 (D18); (iii) the 2019 (D19); (iv) the combination of the 2017, 2018, and 2019 imagery (Dall). For the D17, D18, and D19 datasets, we used the normalized pixel values to extract the spectral features, which are referred to as cases D17_MeanNorm, D18_MeanNorm, and D19_MeanNorm, respectively. Additionally, in the case of the combined dataset (item (iv) in the previously described datasets), the classification was performed using both the normalized and non-normalized values, referred to as Dall_MeanNorm and Dall_Mean, respectively. Table 9 summarizes the number of features used in each case.

Table 9 - The number of spectral features used in each classification investigation.

| Cases | Spectral Data From | | | Number of Features |
|---------------|--------------------|------|------|--------------------|
| | 2017 | 2018 | 2019 | |
| D17_MeanNorm | X | | | 25 |
| D18_MeanNorm | | X | | 25 |
| D19_MeanNorm | | | X | 25 |
| Dall_Mean | X | X | X | 75 |
| Dall_MeanNorm | X | X | X | 75 |

Source: Miyoshi et al. (2020).

As shown in Figure 5 from Section 3.2 despite performing joint geometric processing, there were differences in the spatial position of trees, especially because of the very high spatial resolution of the images. Thus, the spectral features were extracted using the manually delineated ITCs. Similar to the Section 5.2, eight tree species were used (AL, CL, EP, HA, HC, IV, PP, and SR) with the RF method of classification and the use of LOOCV to

ⁱⁱⁱ Paper published at Remote Sensing Journal (ISSN 2072-4292). MIYOSHI, G. T.; IMAI, N. N.; GARCIA TOMMASELLI, A. M.; ANTUNES DE MORAES, M. V.; HONKAVAARA, E. Evaluation of Hyperspectral Multitemporal Information to Improve Tree Species Identification in the Highly Diverse Atlantic Forest. **Remote Sensing**, v. 12, n. 2, 2020.b. Available at: <<https://www.mdpi.com/2072-4292/12/2/244>>

circumvent the unbalanced number of samples. Additionally, the user accuracy and the producer accuracy metrics (LI et al., 2012), and the importance of the features resultant from the RF were included in the analysis of the best classification.

Table 10 provides the AUCROC values after applying the RF with LOOCV to each dataset. AUCROC values varied from 0.313 (AL in the D19_MeanNorm experiment) to 0.999 (SR in the Dall_MeanNorm experiment) showing high variability of values. Considering the highest AUCROC values to each class, it is noted that Dall_MeanNorm presented three classes with the highest and thus, it can be considered the best dataset to identify the tree species.

Table 10 - AUCROC values for each tree species identified in each dataset. AUCROC values are from imagery data of (i) only 2017 (D17_MeanNorm); (ii) only 2018 (D18_MeanNorm); (iii) only 2019 (D19_MeanNorm); (iv) all years and the mean spectral values (Dall_Mean); and (v) all years and the mean normalized values (Dall_MeanNorm).

| Tree Species ¹ | AUCROC | | | | |
|-----------------------------|--------------|--------------|--------------|-----------|---------------|
| | D17_MeanNorm | D18_MeanNorm | D19_MeanNorm | Dall_Mean | Dall_MeanNorm |
| AL | 0.608 | 0.438 | 0.313 | 0.754 | 0.613 |
| CL | 0.821 | 0.678 | 0.517 | 0.742 | 0.768 |
| EP | 0.818 | 0.827 | 0.664 | 0.743 | 0.836 |
| HA | 0.594 | 0.576 | 0.899 | 0.798 | 0.846 |
| HC | 0.800 | 0.809 | 0.847 | 0.699 | 0.847 |
| IV | 0.627 | 0.886 | 0.622 | 0.837 | 0.824 |
| PP | 0.713 | 0.817 | 0.680 | 0.758 | 0.723 |
| SR | 0.986 | 0.997 | 0.915 | 0.936 | 0.999 |
| Overall accuracy (%) | 55.556 | 46.667 | 31.111 | 46.667 | 50 |

¹Check Table 1 for tree species description.

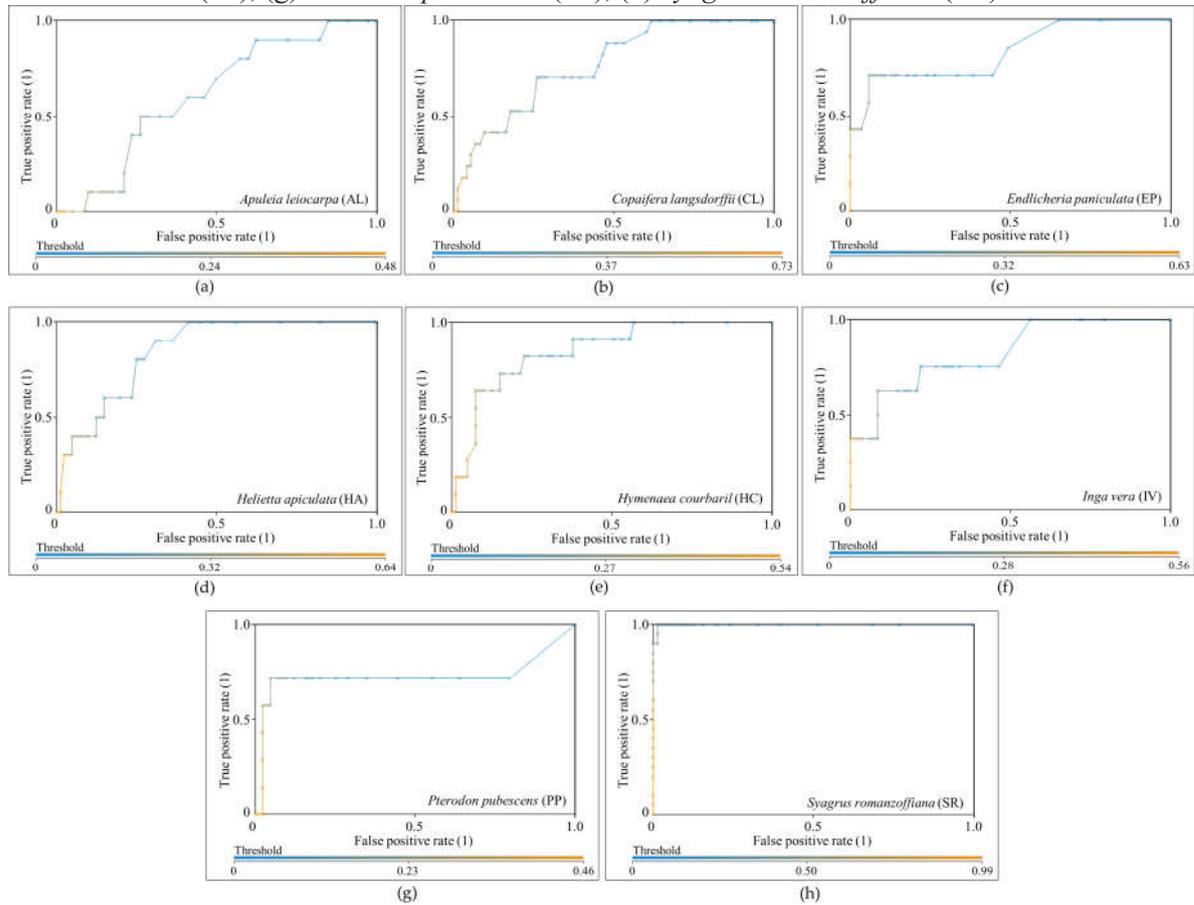
Source: Miyoshi et al. (2020).

Compared with the other datasets, Dall_MeanNorm had the highest AUCROC values for three of the eight tree species, namely, EP, HC, and SR. HA was better modeled in the D19_MeanNorm dataset, with an AUCROC value of 0.899, and it was worst modeled in the D18_MeanNorm dataset (AUCROC of 0.576). In contrast, IV was better and worse identified in D18_MeanNorm and D19_MeanNorm, respectively. Additionally, no significant differences were obtained when using normalized pixels compared with unnormalized ones for this tree species since the AUCROC values were 0.837 for Dall_Mean and 0.824 for Dall_MeanNorm. The identification of CL was similar between Dall_Mean (AUCROC of 0.742) and Dall_MeanNorm (AUCROC of 0.768), and it was best modeled in D17_MeanNorm (AUCROC of 0.821). AL had the lowest AUCROC value in

D19_MeanNorm (0.313), which probably affected its identification in the Dall_MeanNorm dataset, in which its AUCROC was 0.613.

Since Dall_MeanNorm, in general, generated the best results, its ROC curves are shown in Figure 25, and its confusion matrix and user and producer accuracies are presented in Table 11. Figure 25 reveals different threshold values for each tree species, which are related to predictive probabilities (WITTEN; FRANK, 2005). For AL, which had the lowest AUCROC value (0.613), the FPR was higher than 0 (0.088), even when the TPR was equal 0, which indicates that the RF performed poorly in identifying this tree species, as confirmed by the confusion matrix, since none of AL were correctly identified. PP had the second lowest AUCROC value (0.723), and its threshold varied from 0 to 0.46, i.e., similar to AL. As shown in Figure 25 and in the confusion matrix of Table 11, only one tree species was correctly identified, and the TPR was only higher than 0 (TPR of 0.143) when the FPR was 0.024 to a threshold of 0.4. The highest AUCROC value, 0.999 for SR, corresponded to the tree species with the fewest false positives; that is, it was less frequently confused with the other tree species. The ROC curve of SR in Figure 25 shows that a TPR of 1 was obtained when the FPR was 0.014. This fact indicates that samples of this tree species will always be correctly identified with a low degree of confusion to other tree samples. Interestingly, for IV, which was not among the highest AUCROC values, the FPR is equal to 0 until a threshold of 0.44, when the TPR is 0.375. This fact is associated with the confusion matrix of Table 11, which has few false positives for this tree species.

Figure 25 - ROC curves of the identification of each tree species from the application of RF to all imagery datasets (Dall_MeanNorm). (a) *Apuleia leiocarpa* (AL), (b) *Copaifera langsdorffii* (CL), (c) *Endlicheria paniculata* (EP), (d) *Helieta apiculata* (HA), (e) *Hymenaea courbaril* (HC), (f) *Inga vera* (IV), (g) *Pterodon pubescens* (PP), (h) *Syagrus romanzoffiana* (SR).



Source: Miyoshi et al. (2020).

Table 11 - Confusion matrix of the classification of eight tree species and all datasets (Dall_MeanNorm) and its user accuracy and producer accuracy.

| | | Reference ¹ | | | | | | | | |
|-----------------------|----|------------------------|------|------|----|------|------|------|-----|------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | User Accuracy (%) |
| Classified as | AL | 0 | 2 | 0 | 1 | 3 | 0 | 0 | 0 | 0 |
| | CL | 3 | 8 | 2 | 3 | 0 | 2 | 2 | 0 | 40 |
| | EP | 0 | 1 | 4 | 1 | 1 | 0 | 0 | 0 | 57.1 |
| | HA | 2 | 2 | 0 | 3 | 0 | 0 | 2 | 0 | 33.3 |
| | HC | 4 | 2 | 1 | 0 | 6 | 2 | 2 | 0 | 35.3 |
| | IV | 0 | 2 | 0 | 0 | 0 | 3 | 0 | 0 | 60 |
| | PP | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 33.3 |
| | SR | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 20 | 87 |
| Producer accuracy (%) | | 0 | 47.1 | 57.1 | 30 | 54.5 | 37.5 | 14.3 | 100 | Overall accuracy = 50% |

¹Check Table 1 for tree species description.

Source: Miyoshi et al. (2020).

In summary, it was observed that the use of joint spectral normalized features (i.e., Dall_MeanNorm) increased the AUCROC values of three tree species (EP, HC, and SR). In general, when using the mean spectral features together (i.e., Dall_Mean), variations in the

AUCROC values were more apparent compared with the use of spectral information from each dataset separately. The exception in the Dall_Mean results is to AL, whose AUCROC value increased with the use of temporal spectral information without normalization.

All the AUCROC values for SR were higher than 0.9, leading to the conclusion that the identification of this tree species did not depend on multitemporal information or the use of normalized spectra. A similar analysis can be applied to HC, which had similar AUCROC values in all experiments, without counting the normalized spectra. CL had similar AUCROC values in Dall_Mean and Dall_MeanNorm and was better identified in D17_MeanNorm. Therefore, the weather pattern in 2017 can be related to the identification accuracy of CL. There was a higher volume of rain right before the 2017 flight campaign (as seen in Figure 4 from Section 2). Similarly, the weather influenced the detection of other tree species when using a single spectral dataset of that year. The dry weather before the 2018 and 2019 flight campaigns could hinder the ability to identify the AL tree species when using the spectral data of these years.

These three above mentioned tree species have different structures, such as the leaf format, and they have different blossoms and fruit sets (LORENZI, 1992a, 1992b, 1992c). SR is a palm tree with leaves that are 2–3 m in length and spadices that are 80–120 cm in length (LORENZI, 1992a). HC has pinnate leaves and requires sunlight to grow and emerge from the canopy; blooming occurs in the dry season, and fruit appears after 3–4 months (LORENZI, 1992c). AL blooms without leaves, usually in September, and its flowers are white (LORENZI, 1992c). Thus, the use of multitemporal data may have influenced the detection of tree species. Of the previous works in the literature related to the Brazilian Atlantic forest, more specifically the semideciduous forest, the research of Ferreira et al. (2019) is highlighted. They acquired WorldView images during the wet and dry seasons of a well-developed Brazilian semideciduous forest to classify tree species; no improvement in the classification results was observed when using the combined data. On the other hand, Somers and Asner (2014), Deventer, Cho and Mutanga (2017), and Hill et al. (2010) found that tree species classification improved when using multitemporal data because of the different spectral changes in the data.

The previous studies that used multitemporal data acquired datasets from different seasons, and they did not use UAVs or consider a semideciduous forest with different development stages. Deventer, Cho and Mutanga (2017) simulated both WorldView and RapidEye data from the leaf spectra of a subtropical forest in South Africa. Hill et al. (2010)

used the Daedalus 1268 Airborne Thematic Mapper (ATM) sensor to acquire data over a deciduous forest in England. Using WorldView images, Li et al. (2015) studied the multitemporal information of tree species in urban environments. When using UAVs, image acquisition depends on several factors (such as wind conditions since a UAV is a lightweight platform), and there are safety requirements to fulfill. During the spring and summer, when some trees may be blooming, the rainfall is usually higher; for example, summer rain events may occur every day. Although images were acquired in the same season in this study, annual differences in tree phenology provided additional information and enhanced the classification accuracy (Table 10). In this sense, it is noticed the benefit of using multitemporal images even when acquired in the same season. The benefits are due to phenological changes among the tree species. Additionally, it is highlighted the ability of the hyperspectral sensor in acquiring these spectral differences.

The utilization of multitemporal data introduces some challenges to the data processing and classification processes. As shown in Figure 5, there are small differences in tree positions due to tree growth and probably also due to geometric projection characteristics; thus, trees were delineated separately in each dataset. When using structural features, the use of different polygons in the same point cloud might affect the classifier. Furthermore, these variations are challenging when working with very high spatial resolution imagery. Ferreira et al. (2019) used resampled WorldView images at 0.3 and 1.2 m and needed to adjust the polygons of each ITC. Special attention must be paid to the radiometric processing of multitemporal spectral datasets. In this study, the datasets from each year were first processed to ensure that reflectance mosaics were uniform using the radiometric block adjustment, and further normalization of the shadows was shown to be advantageous.

Classification accuracies are always affected by the forest characteristics, the existence of several classes, and dataset characteristics, which should be considered for a reliable comparison of studies. Tuominen et al. (2018) used multisource data to classify 26 different tree species of a Finnish forest into species and genus. They had more than 650 samples and achieved accuracies from 59.9% (when classifying tree species using the RF classifier and DN values of the shortwave infrared range) to 86.9% (when using selected features and the k-NN algorithm to classify the genus). Dalponte et al. (2014) classified three types of trees in a boreal forest with more than 2300 samples and obtained an OA of 93.5% using manually delimited ITCs. Sothe et al. (2019) used hyperspectral imagery with 11 cm

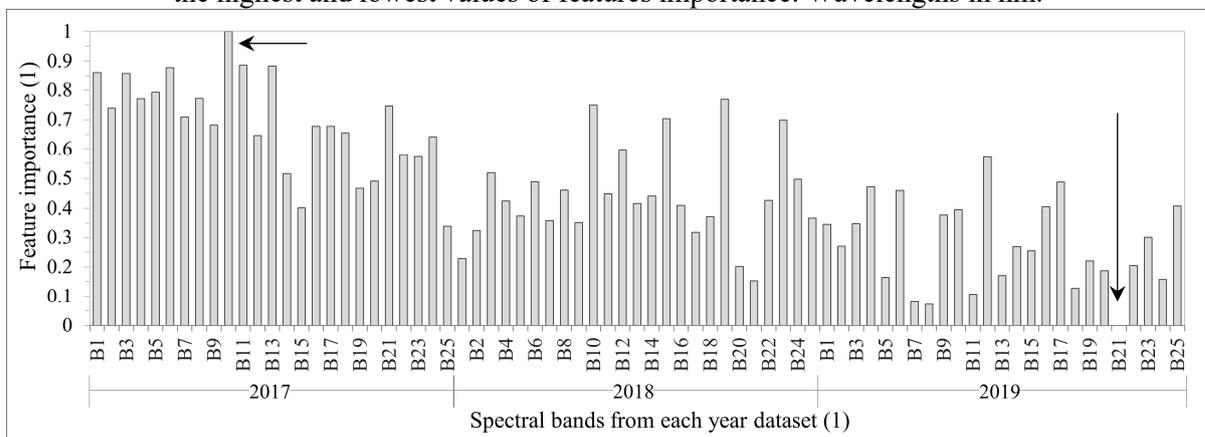
and structural features to classify 12 tree species of a mixed ombrophilous forest, and achieved a maximum OA of 72.4%.

The number of samples affects the classification results and, thus, the analysis results, especially when using an unbalanced number of samples and statistics that consider the OA rather than the class accuracy. Therefore, the use of LOOCV followed by AUCROC analysis is extremely important because AUCROC values are specific to each class. In this study, the number of samples was quite low because of challenges in detecting the training data of a complex forest beside the availability of small areas of forests.

5.3.1 Results of the feature importance

The feature importance in the Dall_MeanNorm dataset, which had the best classification results, is given in Figure 26. The feature importance was scaled from 0 to 1, where 0 represents the least important feature, and 1 represents the most important feature. The least important feature was band 21 in the 2019 dataset, centered at 750.16 nm, and the most important feature was band 10 in the 2017 dataset, centered at 628.73 nm.

Figure 26 - Feature importance when using the RF for 8 tree species and all datasets (Dall_MeanNorm). The x-axis represents the spectral bands in each year. The black arrows point to the highest and lowest values of features importance. Wavelengths in nm.



B1 = 506.22 B2 = 519.94 B3 = 535.09 B4 = 550.39 B5 = 565.10 B6 = 580.16 B7 = 591.90 B8 = 609.00 B9 = 620.22
 B10 = 628.73 B11 = 650.96 B12 = 659.72 B13 = 669.75 B14 = 679.84 B15 = 690.28 B16 = 700.28 B17 = 710.06
 B18 = 720.17 B19 = 729.57 B20 = 740.42 B21 = 750.16 B22 = 769.89 B23 = 780.49 B24 = 790.30 B25 = 819.66

Source: Miyoshi et al. (2020).

In general, the most important features in the 2017 dataset were from the VIS part to the beginning of the red-edge part of the electromagnetic spectrum. In the 2018 dataset, an exception in feature importance may be observed at bands 15, 19, and 23, centered at 690.28 nm, 729.57 nm, and 780.49 nm. These bands were more important than most of the NIR bands in the 2018 dataset. In the 2019 dataset, bands 3, 6, and 12, centered at 535.09 nm, 580.16 nm, and 659.72 nm, respectively, were highlighted because of the peak in the feature

importance value when compared with the other bands from 2019. These bands are in the VIS part of the electromagnetic spectrum; this is related to the leaves' pigment, e.g., chlorophyll and carotenoids, content. As supported by previous research, VIS bands were among the most important features in tree species classification at the crown scale (CLARK; ROBERTS; CLARK, 2005; FERREIRA et al., 2016; MICHEZ et al., 2016; NEVALAINEN et al., 2017). Vegetation spectra are characterized by the peak and absorption in the green and red parts of the electromagnetic spectrum, which helps differentiate tree species.

5.4 ITC DELINEATION^{iv}

This Section shows the results regarding the automatic ITC delineation experiments. The experiments were performed before the use of structural, textural and vegetation index features because it is expected the use of the automatic generated segments in the final experiment. Figure 27 summarizes the experiment, where it is possible to see that SLIC superpixels and watershed methods were applied in the multitemporal data.

Figure 27 - Flowchart of the segmentation experiments.

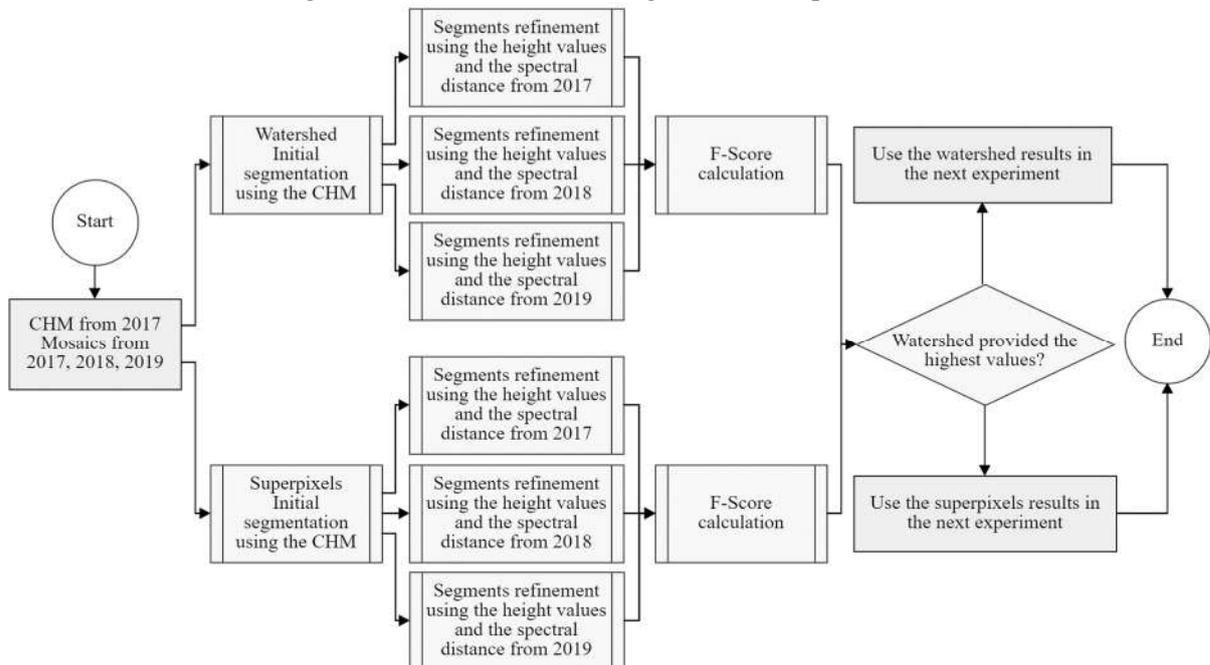


Figure 28 shows an example of the generated superpixels before and after the refinement step whereas Table 12 summarizes the criteria used. All the criteria considered the spectral distance D to overcome the differences of the spatial positions between the different datasets, i.e., acquired in 2017, 2018, and 2019. D value was chosen based on the experiments

^{iv}Paper to be submitted.

performed in Section 5.1 where the mean spectral distance between-species was 0.0027 considering the mean normalized spectra. Besides, the criteria concerning the segments adjacency and height values were empirically chosen. The first criterion considered that the absolute difference between the maximum and mean height of each superpixel should not be higher than 4.5 m; their maximum and mean heights should be greater than 17 m and 15.3 m, respectively; adjacency length between superpixels should be at least 2.2 m, and their absolute difference of maximum heights were no greater than 2.5 m. The second criterion considered 2.5 m as minimum adjacency between superpixels; maximum height between 15.3 and 17 m; and maximum absolute differences equal 1.6 m. The third criterion considered the interval 13.5 to 16 m for the maximum height of superpixels; 2.5 m and 1.6 m as maximum values to the length of adjacency and the absolute difference between maximum heights, respectively; and the difference between maximum and mean heights no greater than 1.5 m.

Figure 28 - Superpixels merged. The red polygons represent the superpixels after they were merged, and the blue lines represent the original superpixels from SLIC.

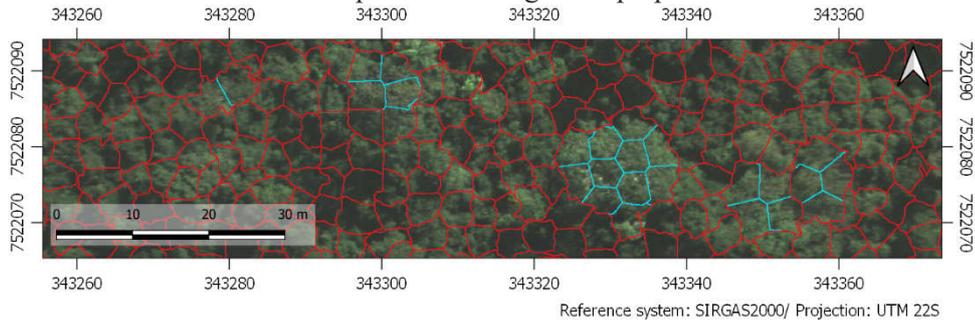


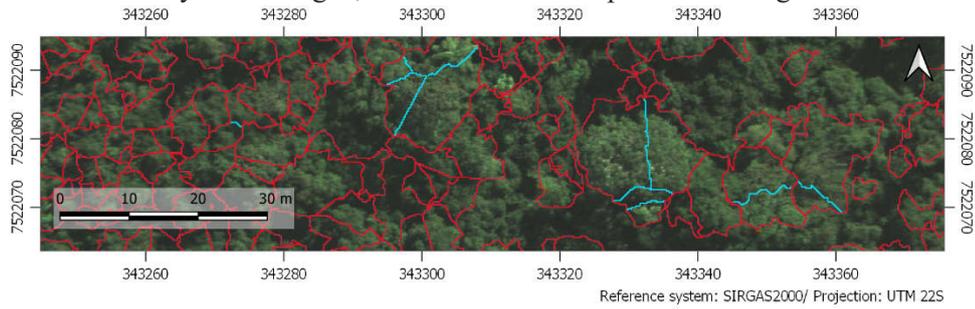
Table 12 - Criteria to merge the segments¹.

| Criteria | Min. adj. (m) | Max. height (m) | Mean height (m) | $ \Delta\text{MaxMean} $ (m) | $ \Delta\text{Max} $ (m) | D (1) |
|----------|---------------|-----------------------------|-----------------|------------------------------|--------------------------|---------------|
| 1 | ≥ 2.2 | ≥ 17.0 | ≥ 15.3 | ≤ 4.5 | ≤ 2.5 | ≤ 0.0027 |
| 2 | ≥ 2.5 | ≥ 15.3 and ≤ 17.0 | - | ≤ 1.6 | ≤ 1.6 | ≤ 0.0027 |
| 3 | ≥ 2.5 | ≥ 13.5 and ≤ 16.0 | - | ≤ 1.5 | ≤ 1 | ≤ 0.0027 |

¹Min. adj.: minimum adjacency between the tested polygons; Max. height: maximum height of the tested polygons; Mean height: mean height of the tested polygons; $|\Delta\text{MaxMean}|$: the absolute difference between the maximum and mean height of the tested polygons; $|\Delta\text{Max}|$: the absolute difference between the maximum height of the tested polygons; D: spectral distance D (Equation (3) from Section 5.1).

As previously mentioned, the refinement step was necessary because of the existence of wider tree crowns and the different datasets. The criteria were the same adopted to merge the superpixels (Table 12) viewing a comparison of the results between superpixels and watershed. Figure 29 shows examples of the segments generated by the watershed algorithm before and after the merge.

Figure 29 - Segments generated by watershed after merging. The red polygons represent the segments after they were merged, and the blue lines represent the original ones.



The results of ITC delineation are summarized in Table 13, being noticed higher accuracy values to the superpixels method for all datasets. F-Score values of superpixels varied from 54.47% to 61.15% whereas for watershed the minimum value was 48.04% and the maximum was 48.8%. Additionally, it is noticed that the highest F-Score values were obtained with the 2017 dataset. A possible reason for this result may be the use of a CHM from 2017 to generate the initial segments. As the tree crowns can vary in the different seasons and years, the use of CHM from different years could provide better results.

Table 13 - Overall accuracies for the individual crown delineation using superpixels and watershed methods. Shaded cells highlight the highest F-Score values for each dataset.

| Dataset from | Method | User accuracy | Producer accuracy | F-Score | Omission error | Commission error |
|--------------|------------|---------------|-------------------|---------|----------------|------------------|
| 2017 | Superpixel | 62.03% | 60.30% | 61.15% | 39.70% | 37.97% |
| | Watershed | 42.83% | 56.34% | 48.67% | 43.66% | 57.17% |
| 2018 | Superpixel | 57.37% | 51.84% | 54.47% | 48.16% | 42.63% |
| | Watershed | 41.93% | 56.22% | 48.04% | 43.78% | 58.07% |
| 2019 | Superpixel | 56.66% | 59.14% | 57.87% | 43.34% | 40.86% |
| | Watershed | 43.42% | 55.8% | 48.84% | 44.2% | 56.58% |

Both, user accuracy and producer accuracy, can be related to the under-segmentation (i.e., when there is more than one ITC in the segment) and over-segmentation (i.e., when the ITC belongs to more than one segment). User accuracy considers the number of true positives and the false positives, being complementary to the commission error and thus, related to the over-segmentation case. In contrast, producer accuracy is complementary to the omission error and considers the number of true positives and the false negatives, so it is related to the under-segmentation. Thus, results Table 13 reveal that both cases of segmentation error occur in all methods and datasets.

Concerning the use of ALS data to delineate the tree crowns, it is noticed that our results are supported by Dalponte et al. (2014) who found 48.5% of their reference trees, but

not by those who achieved OA and F-Score values up to 99% in detecting and delineating the tree crowns (DAI et al., 2018; HU; CHEN; XU, 2017; HYYPPÄ et al., 2001; LI et al., 2012; SILVA et al., 2016; STRÎMBU; STRÎMBU, 2015). In fact, it is necessary to consider that most of the methods developed for tree species delineation have considered different types of forests. Most of the researches were conducted considering high latitude forests, where the coniferous tree types were dominant. These types of trees usually have a cone shape and a circular crown, being difficult to compare these results with ours. Therefore, tree detection and delineation accuracy depend on the type of forest, its characteristics and type of sensor used to acquire the data (CARR; SLYDER, 2018; KE; QUACKENBUSH, 2011; VAUHKONEN et al., 2011).

Considering Brazilian forests, Ferreira et al. (2016) and Wagner et al. (2018) were the pioneer researches to detect ITCs in the Atlantic forest to the best of our knowledge. Ferreira et al. (2016) applied the Jeffreys-Matusita (JM) distance to support the ITC delineation and Wagner et al. (2018) used the rolling ball algorithm and mathematical morphological operations. Wagner et al. (2018) correctly found 79.2% of the reference trees, being 23% of the trees over-segmented and only 0.9% under-segmented. These results are in contrast with our findings and among the reasons, it is possible to cite the differences in the study areas. Both Ferreira et al. (2016) and Wagner et al. (2018) used the Brazilian Atlantic forest located in the central region of São Paulo State. Their and our study areas represent submontane semideciduous forests, however, there are significant differences between the areas. Their area is well-preserved with minimum mean crown size of 20 m², whereas tree crown sizes in our study area are mostly smaller than 20 m² as shown in Figure 8 from Section 3.1. Tree species belonging to our study area have similar heights and smaller tree crowns because of the different stages of development, which makes the ITC delineation challenging.

Mentioning the ITC delineation in multitemporal data, Ferreira et al. (2019) evaluated tree species identification in different seasons using resampled WorldView images. They worked with WorldView images with a spatial resolution of 0.30 m and 1.20 m in the same study area of Ferreira et al. (2016) and Wagner et al. (2018) and showed the need to consider the spatial difference of tree crowns. These finds reinforce the need to consider the spatial difference of trees crowns when using multitemporal data, especially when using very high resolution images, like 10 cm. So, in our case, the use of spectral distance showed to be useful to refine the segments from the different year datasets, since the initial segments were

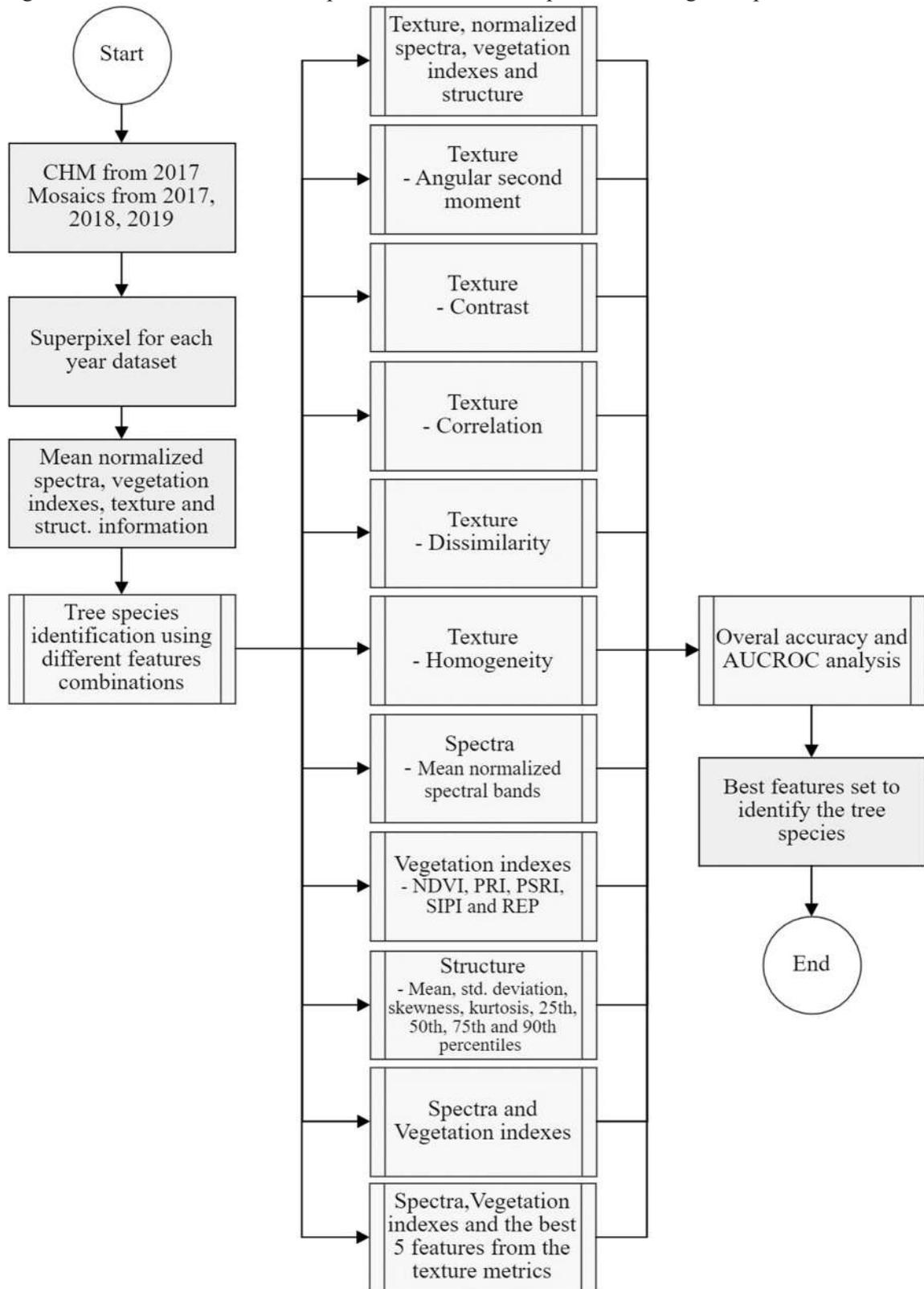
generated with a CHM from 2017. Even not achieving high degrees of F-Score values (i.e., higher than 80%), the superpixel method presented the best results. Further researches considering more mature forests and initial CHMs from its respective date are recommended as well as the use of different spectral distances such as the JM distance.

5.5 TREE SPECIES CLASSIFICATION USING SUPERPIXELS AND DIFFERENT SETS OF MULTITEMPORAL FEATURES^v

Considering the previous results from Sections 5.1 to 5.4 it was noticed the importance of the normalized spectral features, the use of region-based classification method, the contribution of multitemporal spectral information, and the ITC delineation using the superpixel method. Thus, the objective of this Section is to apply all these previous findings to identify the tree species of the initial to more advanced successional stages of our study area using the automatically delineated ITCs. All features were extracted from the superpixels generated in Section 5.4. Figure 30 exemplifies this experiment where it is possible to see that 11 different tree species identification experiments were performed using the multitemporal data and different sets of features.

^vPaper to be submitted.

Figure 30 - Flowchart of the tree species identification experiments using multiple sets of features.



Although RF shows to be insensitive to the number of features used, it can be sensitive to noisy features (SEGAL, 2004). In this sense, different combinations of features and their numbers were evaluated. RF was applied using the Weka software version 3.8.3

(The University of Waikato, Hamilton, New Zealand) (HALL et al., 2009). 500 was the number of decision trees set because a preliminary test showed that using 100 decision trees (default parameter of RF in Weka) most of the attributes would not be used to create the classifying model. Additionally, the use of a higher number of decision trees would not affect the RF results (BELGIU; DRĂGUȚ, 2016; GHOSH; JOSHI, 2014).

The performed experiments considered: (i) all the 473 features described in Table 7; (ii) all mean normalized spectral features from the three datasets; (iii) the vegetation indexes from the tree datasets; (iv) the structural features; (v) the ASM from the three datasets; (vi) the CON from the tree datasets; (vii) the COR from the three datasets; (viii) the DIS from the three datasets; (ix) the HOM from the three datasets; (x) all mean normalized spectral features and the vegetation indexes from the three datasets; (xi) all mean normalized spectral features, the vegetation indexes from the three datasets, and the five most important features from (v) to (ix). Table 14 summarizes each performed experiment, its number of features used and its identification (ID). The results of each classification experiment were assessed using the AUCROC values of each tree species (Section 5.2) together with the OA values.

Table 14 - Classification investigations performed using different features combination and the number of features used in each test.

| ID | Features used | Total number of features |
|--------------|--|--------------------------|
| RF_ALL | All features | 473 |
| RF_ASM | Angular second moment - 25 features/dataset | 75 |
| RF_CON | Contrast - 25 features/dataset | 75 |
| RF_COR | Correlation - 25 features/dataset | 75 |
| RF_DIS | Dissimilarity - 25 features/dataset | 75 |
| RF_HOM | Homogeneity - 25 features/dataset | 75 |
| RF_MNo | Mean normalized spectra - 25 features/dataset | 75 |
| RF_VIs | NDVI, PRI, PSRI, SIPI and REP from each dataset | 15 |
| RF_3D | Mean, standard deviation, skewness, kurtosis, 25th, 50th, 75th e 90th percentiles from the CHM | 8 |
| RF_MNoVIs | RF_MNo and RF_Vis | 90 |
| RF_MNoVIsTXT | RF_MNoVIs with five of the most important features from RF_ASM, RF_CON, RF_COR, RF_DIS, and RF_HOM | 115 |

Table 15 presents the AUCROC values and OA of each classification investigation. Features importance and confusion matrix of each experiment are provided in

Appendix B. From Table 15 it is possible to see that in general, the experiments with the highest number of AUCROC values are for the RF_MNo and RF_VIs. These results can lead us to the conclusion that when using the mean normalized spectra or the vegetation indexes, the classification accuracies are higher. Moreover, considering the OA values and the use of mean normalized spectra, vegetation indexes and the most important textural features, it is confirmed that RF does an indirect feature selection, thus, the textural features did not contribute to the tree species identification. Observing the OA, it is noticed that the highest value is achieved when using only the vegetation indexes. The OA does not consider the different proportions of samples in each class, only the proportion of correctly identified instances, being the analysis of AUCROC a better approach to evaluate the results. As previously mentioned, the AUCROC value is independent of the class distribution (EVANGELISTA et al., 2009; FAWCETT, 2006).

Table 15 - AUCROC values for each tree species identification in each experiment. Shaded cells highlight the highest value of each row.

| Tree species ¹ | RF_A LL | RF_A SM | RF_C ON | RF_C OR | RF_DI S | RF_H OM | RF_M No | RF_VI s | RF_3 D | RF_M NoVIs | RF_M NoVIs TXT ² |
|-----------------------------|------------|------------|------------|------------|------------|------------|------------|------------|-----------|---------------|-----------------------------------|
| AL | 0.334 | 0.509 | 0.437 | 0.268 | 0.439 | 0.575 | 0.330 | 0.391 | 0.490 | 0.329 | 0.308 |
| CL | 0.586 | 0.510 | 0.649 | 0.464 | 0.577 | 0.379 | 0.615 | 0.670 | 0.588 | 0.637 | 0.654 |
| EP | 0.687 | 0.473 | 0.755 | 0.267 | 0.503 | 0.458 | 0.830 | 0.759 | 0.467 | 0.834 | 0.828 |
| HA | 0.778 | 0.414 | 0.627 | 0.411 | 0.499 | 0.393 | 0.808 | 0.701 | 0.628 | 0.786 | 0.807 |
| HC | 0.643 | 0.510 | 0.505 | 0.240 | 0.444 | 0.441 | 0.782 | 0.727 | 0.539 | 0.777 | 0.745 |
| IV | 0.788 | 0.375 | 0.783 | 0.413 | 0.688 | 0.564 | 0.779 | 0.871 | 0.627 | 0.847 | 0.852 |
| PP | 0.546 | 0.657 | 0.730 | 0.503 | 0.628 | 0.733 | 0.682 | 0.703 | 0.441 | 0.690 | 0.684 |
| SR | 0.999 | 0.416 | 0.933 | 0.690 | 0.909 | 0.621 | 0.970 | 0.944 | 0.762 | 0.964 | 0.984 |
| Overall accuracy (%) | 38.9 | 20.0 | 32.2 | 18.9 | 25.6 | 18.9 | 43.3 | 47.8 | 26.7 | 43.3 | 43.3 |

¹Check Table 1 for tree species description.

²The features importance of each tree species classification are in Appendix B.

From the AUCROC values, it is noticed that all models poorly describe the AL tree species. The highest AUCROC value achieved for AL was 0.575 when using the HOM feature of all datasets. For the remaining classifications, the AUCROC values were lower than 0.5 value which according to Evangelista et al. (2009) represents that a classification model is no better than random, i.e., the samples are randomly correctly or incorrectly classified in AL tree species. Besides, bearing the results obtained in Section 5.3, it is noticed that AL also presents the poorest results. Possible reasons that might affect AL identification are the data

acquisition in the driest season, the similarity with other tree species, and the superpixels that delineated the tree species samples. It is important noting the superpixels results did not present a perfect matching with the manually delineated polygons.

Bearing the results for CL, it is noticed the highest AUCROC value when using only the vegetation indexes (AUCROC of 0.67). However, close AUCROC values were achieved for RF_MNoVIsTXT, RF_MNoVIs and RF_CON. Additionally, slightly better results than AL were obtained in all classification experiments. For EP, the highest AUCROC value was 0.834 using the joint normalized spectral features and vegetation indexes (i.e., RF_MNo_VIS), value almost similar when using only the normalized spectral features (AUCROC of 0.83), leading to the conclusion that the use of vegetation indexes is not necessary to correctly identify this tree species. Moreover, these AUCROC values, were similar to the one when using the ITC manually delimited (Section 5.3).

To HA, the highest AUCROC value was 0.808 in the RF_MNo case, almost the same value found to HC (AUCROC of 0.782) which was also better identified in RF_MNo. Further, HA was poorly modeled when using the texture features ASM, COR, DIS, and HOM. Despite these features, CON and the structural features (RF_3D) also poorly modeled HC species, being the lowest AUCROC value when using only the COR as a feature. Considering the results for IV, it is noticed the poorest performance of RF when using ASM (AUCROC of 0.375) or COR (AUCROC of 0.564) as features, but a good performance for this tree species was achieved using the vegetation indexes (RF_VIs; AUCROC of 0.871), similar value to the best result found in Section 5.3.

The identification of PP was better when using the HOM textural feature, achieving an AUCROC value of 0.733. The same tree species was poorly modeled using the structural features (RF_3D; AUCROC of 0.441), and the COR features (RF_COR; AUCROC of 0.503). Last, but not least, interesting results were found to SR, which achieved AUCROC values higher than 0.9 to RF_ALL, RF_CON, RF_DIS, RF_MNo, RF_VIs, RF_MNoVIs, and RF_MNoVISTXT. The textural features ASM, HOM and COR produced AUCROC values between 0.416 and 0.69. From the AUCROC value using the structural features (RF_3D, AUCROC of 0.762) it can be addressed that the structural features were not suitable to identify the SR tree species or even that the superpixels could not correctly represent the star shape of SR thus, affecting its structural features.

Comparing our results with previous ones from the literature, different findings are observed. Nevalainen et al. (2017) and Tuominen et al. (2018) classified tree species

belonging to Finnish forests using hyperspectral images acquired with different models of Rikola hyperspectral camera. Besides the use of hyperspectral information, Nevalainen et al. (2017) used structural features extracted from a CHM of the area and performed its classification experiments using selected features considering the mean spectra and the mean normalized spectra. With the RF classifier, the authors achieved accuracies up to 94.9%. When using the non-normalized spectra, they noticed that the structural features were among the most important features, which differs from our findings. As it is possible to see in Appendix B and in Table 15 the use of structural features did not present the best results. Potential reasons can be the different forest areas and tree structures. Figure 9 in Section 3.1 shows that only a few tree species stood out from the mean height, which hinder the tree species identification using only this attribute. Nevertheless, their findings support our results considering the use of normalized spectra, which presented the best results. Tuominen et al. (2018) also applied the spectral features from the hyperspectral images and structural features from the CHM. Even though they did not use normalized spectral features, they found better tree species identification results for RF application using the structural features and visible to near infrared (VNIR) and SWIR spectral features (accuracy of 0.823). Unfortunately, SWIR data is not available in our dataset as well as in the Rikola camera used in our experiments.

Regarding the use of GLCM texture features, Dian, Li and Pang (2014) used airborne images captured with the Compact Airborne Spectrographic Image (CASI) 1500 hyperspectral imager to identify five tree species in a Chinese forest. As classifiers, the SVM-RBF achieved the best results (OA of 83.4%), using textural features and features derived from the minimum noise fraction (MNF). Without the use of textural features, they achieved OA of 83.24%, showing no great improvement when textural features were used, which support our results. Kim, Madden and Warner (2009) classified different forest types in the United States of America using Ikonos satellite imagery (GSD of 4 m). When using textural features solely, they achieved OA varying from 60% to 79.3% and showed that the use of textural features improved accuracy by only 0.3% when compared with the results using spectral features. Using both features together, the OA decreased confirming our results presented in Table 15.

When using automatic segments derived from the mean-shift algorithm and VNIR spectral bands, first derivative, texture, vegetation indexes, and principal components, Maschler, Atzberger and Immitzer (2018) achieved the OA of 89.4% when classifying 13 tree species from temperate Austrian forest. Using only the VNIR spectral bands and the textural

features, with manually delimited polygons, the OA decreased to 59.4%. The use of VNIR spectral bands with vegetation indexes produced the OA of 75.5% leading us to the conclusion that vegetation indexes were better features to identify tree species than textural features. This is consistent with our results.

Considering Brazilian forests, Sothe et al. (2020) showed that the GLCM textural features were not among the most important features, which our results also showed. When using the combination of VNIR spectral bands, vegetation indexes, MNF, and GLCM features, they achieved the OA of 59.17% with RF, whereas the best OA was 66.06% when using only VNIR bands and structural features from the CHM. This is not in accordance with our results, where the use of structural features did not improve tree species identification. However, this could be explained by the differences in the study areas. Although, both areas belong to the Brazilian Atlantic forest biome, Sothe et al. (2020) study area belongs to the mixed ombrophilous forest whereas our study area is semideciduous seasonal forest. Additionally, regarding the use of vegetation indexes, Ferreira et al. (2016) showed that their utilization improved the tree species identification of semideciduous seasonal forest located 110 km northwest of São Paulo city, in São Paulo State.

In summary, when evaluating the impact of using the different types of features, it can be noticed that the use of textural features as well as the use of structural features did not improve the tree species classification accuracy in our study. In fact, the tree species identification had the accuracy decreased when using those features, especially the COR feature. The possible reason for the poor performance of textural features include the irregular tree crowns, such as, SR having a star shape, CL with a dense globose crown, AL with an umbrella shape crown or EP with an irregular crown, not counting the high spatial resolution of the used images (i.e., 10 cm) and the use of a 5×5 window to calculate the GLCM. Additionally, the varying development stages of trees and their similar heights, and the very high spatial resolution of images caused further challenges for the tree species identification. Thus, the textural features could be considered as noisy information in our study. Concerning the structural features, it was observed that the similarity of the tree heights was the reason why the accuracy of the tree species identification did not improve when using this feature (Figure 9 from Section 3.1). Despite some studies showing the usefulness of structural attributes, in our study area they did not improve the tree species identification. On the other hand, the vegetation indexes provided a clear improvement in the species identification.

It is also important to highlight the use of superpixels as ITCs in the results classification. Despite the F-Score values achieved in Section 5.4 (61.15%), the use of superpixels provides similar AUCROC values in the identification when compared with the AUCROC values achieved using the manually delimited polygons from Section 5.3. Thus, the over-segmented trees and the superpixels nearest to the centroids of the true polygons were suitable to provide a spectral signature to model the different tree species.

6 CONCLUSION

The objective of this doctoral dissertation was to develop a methodology to improve the tree species identification and to evaluate whether the multitemporal information could improve the tree species identification. Hyperspectral images were acquired by Rikola camera onboard an unmanned aerial vehicle (UAV) over an area of the Brazilian Atlantic forest having great species diversity and different successional stages. Further objectives were the evaluation of spectral differences, the automatic ITC delineation and the combination of different temporal features to the classification task.

The use of mean normalized spectral features showed a better performance than the non-normalized features in classifying tree species. Even applying the radiometric block adjustment, the pixel normalization indeed reduced the differences in shadowed and sunlit pixels and thus, increasing the tree species separability. Radiometric block adjustment was equally important and highlighted. Different cloud covering density affects the spectral response of samples from the same tree species because the incident light is different and the method to acquire the spectral response of the images is the empirical line method. The importance of the radiometric block adjustment should be emphasized because the high spatial resolution images show detailed information of the tree crown and are subject to the anisotropy effects when not properly corrected.

Furthermore, the region-based approach presented the best results when compared with the pixel-based approach. Temporal spectral information improved the performance of the random forest classifier for three of the eight tree species analyzed, indicating that better accuracy could be obtained when using temporal spectral information. Separated analysis of single-date datasets showed that the weather pattern directly influenced the classification performance of some of the tree species. The analysis of datasets from several years of the same season showed that differences in weather conditions in different years resulted in some changes in the species spectra and these changes were useful for differentiating some of the selected tree species.

Automatic ITC delineation was shown to be a highly complex task. The lack of a standard tree shape, the high forest density, its different development stages, and the similarity of heights directly affected the automatic ITC delineation, are weakness in all techniques for tree species identification. Considering the *Syagrus romanzoffiana*, this task is even more challenging. Its regular shape requires smaller superpixels, but it may cause the over-segmentation of wider crowns. Both assessed methods did not achieve an F-Score value

higher than 70%. However, the superpixels application provided similar AUCROC values when compared with the use of manually delineated polygons.

The knowledge of the different tree heights was essential as well as the use of the spectral information. The use of spectral differences was crucial to deal with the different spatial positions of the ITC over the years. Concerning the spatial position of the trees, the initial EOP information from the camera GPS was important to geometrically produce the mosaic of hyperspectral images. There is a challenge to introduce GCPs inside of the forest because of its high density. Furthermore, even though the georeferencing of the three datasets was carried out in a single process there appeared small geometric differences as expected.

Weather conditions directly affect the tree species bloom or defoliation because some species were better identified when using all temporal data, such as *Hymenaea Courbaril* and *Inga vera*. Further, the use of vegetation indexes is of fundamental importance. They were shown to provide similar results as the use of normalized features. The use of textural features was shown not to be relevant in our study area due to the high spatial resolution of the images, which might result in the textural features to be noisy and thus, not producing the best results. A similar analysis is applied to the use of structural features because the similar tree heights did not improve the tree species identification. Finally, despite the RF appeared to be insensitive to the number of attributes, the results showed its sensitivity to noisy features, as pointed out by other researches also. When using all textural, spectral, vegetation indexes and structural features, the results were worse than when using only the spectral features or the vegetation indexes.

To the best of our knowledge, this is the first work to use hyperspectral UAV images acquired over several years to classify the highly diverse Atlantic Forest. Improvements should be applied regarding the number of samples per class and the seasonality for data acquisition. For some species, finding a higher number of tree samples is quite challenging, such as for *Aspidosperma polyneuron* which only had three individual samples identified in the field and was removed from the classification experiments.

6.1 CONTRIBUTIONS AND FUTURE WORKS RECOMMENDATION

As final remarks of this doctoral dissertation, it is highlighted the use of temporal information for tree species identification. Despite the images were not acquired in different seasons, it was possible to improve the identification of at least three tree species. The use of

an innovative lightweight hyperspectral sensor acquiring information from the VIS to the NIR over a small fragment of the Brazilian Atlantic forest in one of the novelties of this research. The multitemporal data analysis was a very challenging task because it involved not only the data acquisition, but the understanding of how to process and analyze all data together as well as the comprehension of forest components and behavior over the years. Another point to be reminded is the use of an area still not well-developed with similar tree heights surrounded by crops. Moreover, it was a protected area inside an ecological station, being required authorization from the environmental agencies to collect the data. It is worthy of mention the need for suitable forest management even when protected by laws. Therefore, the reported results are of great importance to decision-makers and can be used as key information to monitor this fragment.

Bearing the recommendations, there is the use of a higher number of samples and tree species. Despite being a small fragment, a higher number of samples per tree species and the use of a higher number of classes can improve the monitoring task of this forest. The lack of samples could affect the classification results because of the unbalanced number of samples. Image acquisition in different seasons is another recommendation. Images acquired during Spring, Summer or Autumn can show higher discrepancies in the ITCs because of the different aspects of the soil moisture, weather and pigment content in each ITC. The use of a higher number of tree characteristics in different seasons could improve not only the tree species identification but also follow its evolution, consequently providing information to monitor the degree of forest restoration and conservation.

The employment of recent deep learning approaches is encouraged. They are an emergent approach from the machine learning field being the state of the art of the classification methods in Remote Sensing. When using deep learning approaches, increasing the number of samples will be of higher importance, since the application of these algorithms requires a larger number of samples to properly model the classifiers. In the case of our study area it is possible to identify the *Syagrus romanzoffiana* because, during fieldworks and image interpretation, hundreds of samples were recognized. Nevertheless, the use of different machine learning algorithms is also suggested such as the SVM and the Multilayer perceptron.

Considering the assessed features, there is the recommendation to apply different criteria to calculate the features importance and their application in classification experiments. Regarding the textural features, there is the use of different window sizes, the use of non-

normalized pixel values, the use of different spatial resolution imagery and the use of other textural features not used in this doctoral dissertation. Regarding the vegetation indexes, different vegetation indexes assessment is encouraged. Hence, there is a recommendation to evaluate multispectral images in the multitemporal form.

Finally, as a final recommendation, there is the application of the developed methodology in well-developed areas or even in different forest areas, such as the remaining types of Atlantic Forest, the Amazon forest or the Northern forests.

REFERENCES

- AASEN, H.; HONKAVAARA, E.; LUCIEER, A.; ZARCO-TEJADA, P. J. Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows. **Remote Sensing**, v. 10, n. 7, 2018. Available at: <<https://www.mdpi.com/2072-4292/10/7/1091>>
- ACHANTA, R.; SHAJI, A.; SMITH, K.; LUCCHI, A.; FUA, P.; SÜSSTRUNK, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 34, n. 11, p. 2274–2282, 2012.
- ALEXANDER, C.; KORSTJENS, A. H.; HANKINSON, E.; USHER, G.; HARRISON, N.; NOWAK, M. G.; ABDULLAH, A.; WICH, S. A.; HILL, R. A. Locating emergent trees in a tropical rainforest using data from an Unmanned Aerial Vehicle (UAV). **International Journal of Applied Earth Observation and Geoinformation**, v. 72, p. 86–90, 2018.
- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; DE MORAES GONÇALVES, J. L.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.
- AXELSSON, P. DEM generation from laser scanner data using adaptive TIN models. **International archives of photogrammetry and remote sensing**, v. 33, n. 4, p. 110–117, 2000.
- BALDECK, C. A.; ASNER, G. P.; MARTIN, R. E.; ANDERSON, C. B.; KNAPP, D. E.; KELLNER, J. R.; WRIGHT, S. J. Operational Tree Species Mapping in a Diverse Tropical Forest with Airborne Imaging Spectroscopy. **PLOS ONE**, v. 10, n. 7, p. 1–21, 2015.
- BELGIU, M.; DRĂGUȚ, L. Random forest in remote sensing: A review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 114, p. 24–31, 2016.
- BERVEGLIERI, A.; IMAI, N. N.; TOMMASELLI, A. M. G.; CASAGRANDE, B.; HONKAVAARA, E. Successional stages and their evolution in tropical forests using multi-temporal photogrammetric surface models and superpixels. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 146, p. 548–558, 2018.
- BERVEGLIERI, A.; TOMMASELLI, A. M. G.; IMAI, N. N.; RIBEIRO, E. A. W.; GUIMARÃES, R. B.; HONKAVAARA, E. Identification of Successional Stages and Cover Changes of Tropical Forest Based on Digital Surface Model Analysis. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 9, n. 12, p. 5385–5397, 2016.
- BEUCHER, S.; MEYER, F. The morphological approach to segmentation: the watershed transformation. **Mathematical morphology in image processing**, v. 34, p. 433–481, 1993.
- BRADLEY, A. P.; DUIN, R. P. W.; PACLIK, P.; LANDGREBE, T. C. W.; BRADLEY, A. P.; DUIN, R. P. W.; PACLIK, P.; LANDGREBE, T. C. W. Precision-recall operating characteristic (P-ROC) curves in imprecise environments. *In: 18TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR'06)*, 2006.
- BRASIL. 11.428. LEI N° 11.428, DE 22 DE DEZEMBRO DE 2006. 22 Dez. 2006.

- BREIMAN, L. Random forests. **Machine learning**, v. 45, n. 1, p. 5–32, 2001.
- BROVELLI, M. A.; CRESPI, M.; FRATARCANGELI, F.; GIANNONE, F.; REALINI, E. Accuracy assessment of high resolution satellite imagery orientation by leave-one-out method. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 63, n. 4, p. 427–440, 2008.
- BUDDENBAUM, H.; SCHLERF, M.; HILL, J. Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. **International Journal of Remote Sensing**, v. 26, n. 24, p. 5453–5465, 2005.
- CARR, J. C.; SLYDER, J. B. Individual tree segmentation from a leaf-off photogrammetric point cloud. **International Journal of Remote Sensing**, v. 39, n. 15–16, p. 5195–5210, 2018.
- CASTRO-ESAU, K. L.; SÁNCHEZ-AZOFEIFA, G. A.; RIVARD, B.; WRIGHT, S. J.; QUESADA, M. Variability in leaf optical properties of Mesoamerican trees and the potential for species classification. **American Journal of Botany**, v. 93, n. 4, p. 517–530, 2006.
- CHAVANA-BRYANT, C.; MALHI, Y.; WU, J.; ASNER, G. P.; ANASTASIOU, A.; ENQUIST, B. J.; COSIO CARAVASI, E. G.; DOUGHTY, C. E.; SALESKA, S. R.; MARTIN, R. E.; GERARD, F. F. Leaf aging of Amazonian canopy trees as revealed by spectral and physiochemical measurements. **New Phytologist**, v. 214, n. 3, p. 1049–1063, 2017.
- CHEN, C.; LIAW, A.; BREIMAN, L.; OTHERS. Using random forest to learn imbalanced data. **University of California, Berkeley**, v. 110, n. 1–12, p. 24, 2004.
- CLARK, M. L.; ROBERTS, D. A. Species-Level Differences in Hyperspectral Metrics among Tropical Rainforest Trees as Determined by a Tree-Based Classifier. **Remote Sensing**, v. 4, n. 6, p. 1820–1855, 2012.
- CLARK, M. L.; ROBERTS, D. A.; CLARK, D. B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. **Remote Sensing of Environment**, v. 96, n. 3, p. 375–398, 2005.
- COLGAN, M. S.; BALDECK, C. A.; FÉRET, J.-B.; ASNER, G. P. Mapping Savanna Tree Species at Ecosystem Scales Using Support Vector Machine Classification and BRDF Correction on Airborne Hyperspectral and LiDAR Data. **Remote Sensing**, v. 4, n. 11, p. 3462–3480, 2012.
- COLOMINA, I.; MOLINA, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 92, p. 79–97, 2014.
- CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v. 37, n. 1, p. 35–46, 1991.
- CUBERT. **UAV mapping with FirefLEYE - Cubert**. [s.d.]. Available at: <<https://cubert-gmbh.com/applications/uav-mapping-with-firefleye/>>. Accessed on 25 out. 2019.
- D'OLIVEIRA, M. V. N.; ALVARADO, E. C.; SANTOS, J. C.; CARVALHO, J. A. Forest natural regeneration and biomass production after slash and burn in a seasonally dry forest in

the Southern Brazilian Amazon. **Forest Ecology and Management**, v. 261, n. 9, p. 1490–1498, 2011.

DAI, W.; YANG, B.; DONG, Z.; SHAKER, A. A new method for 3D individual tree extraction using multispectral airborne LiDAR point clouds. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 144, p. 400–411, 2018.

DALPONTE, M.; FRIZZERA, L.; GIANELLE, D. Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. **PeerJ**, v. 6, p. e6227, 2019.

DALPONTE, M.; ØRKA, H. O.; ENE, L. T.; GOBAKKEN, T.; NÆSSET, E. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. **Remote Sensing of Environment**, v. 140, p. 306–317, 2014.

DEVENTER, H. Van; CHO, M. A.; MUTANGA, O. Improving the classification of six evergreen subtropical tree species with multi-season data from leaf spectra simulated to WorldView-2 and RapidEye. **International Journal of Remote Sensing**, v. 38, n. 17, p. 4804–4830, 2017.

DIAN, Y.; LI, Z.; PANG, Y. Spectral and Texture Features Combined for Forest Tree species Classification with Airborne Hyperspectral Imagery. **Journal of the Indian Society of Remote Sensing**, v. 43, n. 1, p. 101–107, 2014.

EVANGELISTA, P. H.; STOHLGREN, T. J.; MORISETTE, J. T.; KUMAR, S. Mapping Invasive Tamarisk (Tamarix): A Comparison of Single-Scene and Time-Series Analyses of Remotely Sensed Data. **Remote Sensing**, v. 1, n. 3, p. 519–533, 2009.

EYSN, L.; HOLLAUS, M.; LINDBERG, E.; BERGER, F.; MONNET, J.-M.; DALPONTE, M.; KOBAL, M.; PELLEGRINI, M.; LINGUA, E.; MONGUS, D.; PFEIFER, N. A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space. **Forests**, v. 6, n. 5, p. 1721–1747, 2015.

FAN, J.; UPADHYE, S.; WORSTER, A. Understanding receiver operating characteristic (ROC) curves. **Canadian Journal of Emergency Medicine**, v. 8, n. 1, p. 19–20, 2006.

FARQUAD, M. A. H.; BOSE, I. Preprocessing unbalanced data using support vector machine. **Decision Support Systems**, v. 53, n. 1, p. 226–233, 2012.

FASSNACHT, F. E.; LATIFI, H.; STEREŃCZAK, K.; MODZELEWSKA, A.; LEFSKY, M.; WASER, L. T.; STRAUB, C.; GHOSH, A. Review of studies on tree species classification from remotely sensed data. **Remote Sensing of Environment**, v. 186, p. 64–87, 2016.

FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861–874, 2006.

FERET, J.; ASNER, G. P. Tree Species Discrimination in Tropical Forests Using Airborne Imaging Spectroscopy. **IEEE Transactions on Geoscience and Remote Sensing**, v. 51, n. 1, p. 73–84, 2013.

FERREIRA, M. P.; WAGNER, F. H.; ARAGÃO, L. E. O. C.; SHIMABUKURO, Y. E.; FILHO, C. R. de S. Tree species classification in tropical forests using visible to shortwave

infrared WorldView-3 images and texture analysis. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 149, p. 119–131, 2019.

FERREIRA, M. P.; ZORTEA, M.; ZANOTTA, D. C.; SHIMABUKURO, Y. E.; FILHO, C. R. de S. Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. **Remote Sensing of Environment**, v. 179, p. 66–78, 2016.

GAMON, J. A.; PEÑUELAS, J.; FIELD, C. B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. **Remote Sensing of Environment**, v. 41, n. 1, p. 35–44, 1992.

GHOSH, A.; JOSHI, P. K. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. **International Journal of Applied Earth Observation and Geoinformation**, v. 26, p. 298–311, 2014.

GUYOT, G.; BARET, F. Utilisation de la haute resolution spectrale pour suivre l'état des couverts vegetaux. *In*: **SPECTRAL SIGNATURES OF OBJECTS IN REMOTE SENSING**, 1988.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA data mining software: an update. **SIGKDD Explorations**, v. 11, n. 1, p. 10–18, 2009.

HALL-BEYER, M. GLCM Texture: A Tutorial v. 1.0 through 2.7. [s. l.], 2017.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural Features for Image Classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. SMC-3, n. 6, p. 610–621, 1973.

HEINZEL, J.; KOCH, B. Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. **International Journal of Applied Earth Observation and Geoinformation**, v. 18, p. 101–110, 2012.

HILL, R. A.; WILSON, A. K.; GEORGE, M.; HINSLEY, S. A. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. **Applied Vegetation Science**, v. 13, n. 1, p. 86–99, 2010.

HONKAVAARA, E.; KHORAMSHAHI, E. Radiometric Correction of Close-Range Spectral Image Blocks Captured Using an Unmanned Aerial Vehicle with a Radiometric Block Adjustment. **Remote Sensing**, v. 10, n. 2, 2018. Available at: <<https://www.mdpi.com/2072-4292/10/2/256>>

HONKAVAARA, E.; ROSNELL, T.; OLIVEIRA, R.; TOMMASELLI, A. Band registration of tuneable frame format hyperspectral UAV imagers in complex scenes. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 134, p. 96–109, 2017.

HONKAVAARA, E.; SAARI, H.; KAIVOSOJA, J.; PÖLÖNEN, I.; HAKALA, T.; LITKEY, P.; MÄKYNEN, J.; PESONEN, L. Processing and Assessment of Spectrometric, Stereoscopic Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture. **Remote Sensing**, v. 5, n. 10, p. 5006–5039, 2013.

HU, X.; CHEN, W.; XU, W. Adaptive Mean Shift-Based Identification of Individual Trees Using Airborne LiDAR Data. **Remote Sensing**, v. 9, n. 2, 2017. Available at: <<https://www.mdpi.com/2072-4292/9/2/148>>

HYYPPÄ, J.; KELLE, O.; LEHIKONEN, M.; INKINEN, M. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 5, p. 969–975, 2001.

IBGE. Manual técnico da vegetação brasileira. **Manuais técnicos em geociências**, v. 1, 2012.

IMMITZER, M.; ATZBERGER, C.; KOUKAL, T. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. **Remote Sensing**, v. 4, n. 9, p. 2661–2693, 2012.

IMMITZER, M.; NEUWIRTH, M.; BÖCK, S.; BRENNER, H.; VUOLO, F.; ATZBERGER, C. Optimal Input Features for Tree Species Classification in Central Europe Based on Multi-Temporal Sentinel-2 Data. **Remote Sensing**, v. 11, n. 22, 2019. Available at: <<https://www.mdpi.com/2072-4292/11/22/2599>>

INMET - INSTITUTO NACIONAL DE METEOROLOGIA. **Estações Automáticas - Gráficos**. 2019. Available at: <http://www.inmet.gov.br/portal/index.php?r=home/page&page=rede_estacoes_auto_graf>. Accessed on 8 Nov. 2019.

ISENBURG, M. **LAStools - efficient LiDAR processing software**. 141017, unlicensed. [s.l.: s.n.]. Available at: <<http://rapidlasso.com/LAStoolss>>

ISENBURG, M. **Readme file of lasnoise tool**. [s.d.]. Available at: <http://lastools.org/download/lasnoise_README.txt>.

ISI WEB-OF-SCIENCE. **ISI-Web-Of-Science Search**. [s.d.]. Available at: <<https://bit.ly/2LeMWnO>>. Accessed on 11 Dez. 2018.

JENSEN, J. R. **Remote Sensing of the Environment: An Earth Resource Perspective**. [s.l.]: Pearson Prentice Hall, 2007.

KAARTINEN, H.; HYYPPÄ, J.; YU, X.; VASTARANTA, M.; HYYPPÄ, H.; KUKKO, A.; HOLOPAINEN, M.; HEIPKE, C.; HIRSCHMUGL, M.; MORSDORF, F.; NÆSSET, E.; PITKÄNEN, J.; POPESCU, S.; SOLBERG, S.; WOLF, B. M.; WU, J.-C. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. **Remote Sensing**, v. 4, n. 4, p. 950–974, 2012.

KANG, K. K.; HOEKSTRA, M.; FOROUTAN, M.; CHEGOONIAN, A. M.; ZOLFAGHARI, K.; DUGUAY, C. R. Operating Procedures and Calibration of a Hyperspectral Sensor Onboard a Remotely Piloted Aircraft System For Water and Agriculture Monitoring. *In: IGARSS 2019 - 2019 IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM*, 2019.

KARASIAK, N.; DEJOUX, J.-F.; FAUVEL, M.; WILLM, J.; MONTEIL, C.; SHEEREN, D. Statistical Stability and Spatial Instability in Mapping Forest Tree Species by Comparing 9 Years of Satellite Image Time Series. **Remote Sensing**, v. 11, n. 21, 2019. Available at: <<https://www.mdpi.com/2072-4292/11/21/2512>>

KE, Y.; QUACKENBUSH, L. J. A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. **International Journal of Remote Sensing**, v. 32, n. 17, p. 4725–4747, 2011.

KEY, T.; WARNER, T. A.; MCGRAW, J. B.; FAJVAN, M. A. A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest. **Remote Sensing of Environment**, v. 75, n. 1, p. 100–112, 2001.

KIM, M.; MADDEN, M.; WARNER, T. A. Forest Type Mapping using Object-specific Texture Measures from Multispectral Ikonos Imagery. **Photogrammetric Engineering & Remote Sensing**, v. 75, n. 7, p. 819–829, 2009.

KUNERT, N.; APARECIDO, L. M. T.; WOLFF, S.; HIGUCHI, N.; SANTOS, J. Dos; ARAUJO, A. C. De; TRUMBORE, S. A revised hydrological model for the Central Amazon: The importance of emergent canopy trees in the forest water budget. **Agricultural and Forest Meteorology**, v. 239, p. 47–57, 2017.

LARY, D. J.; ALAVI, A. H.; GANDOMI, A. H.; WALKER, A. L. Machine learning in geosciences and remote sensing. **Geoscience Frontiers**, v. 7, n. 1, p. 3–10, 2016.

LI, D.; KE, Y.; GONG, H.; LI, X. Object-Based Urban Tree Species Classification Using Bi-Temporal WorldView-2 and WorldView-3 Images. **Remote Sensing**, v. 7, n. 12, p. 16917–16937, 2015.

LI, P.; XIAO, X. Multispectral image segmentation by a multichannel watershed-based approach. **International Journal of Remote Sensing**, v. 28, n. 19, p. 4429–4452, 2007.

LI, W.; GUO, Q.; JAKUBOWSKI, M. K.; KELLY, M. A New Method for Segmenting Individual Trees from the Lidar Point Cloud. **Photogrammetric Engineering & Remote Sensing**, v. 78, n. 1, p. 75–84, 2012.

LIESENBERG, V. **Análise multi-angulaar de fitofisionomias do bioma Cerrado com dados MISR/TERRA**. 2005. Dissertação de Mestrado - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, SP, Brazil, 2005. Available at: <<http://mtc-m12.sid.inpe.br/col/sid.inpe.br/iris@1913/2005/08.03.19.56/doc/publicacao.pdf>>

LIMA, R. B.; BUFALINO, L.; ALVES JUNIOR, F. T.; SILVA, J. A. A. Da; FERREIRA, R. L. C. Diameter distribution in a Brazilian tropical dry forest domain: predictions for the stand and species. **Anais da Academia Brasileira de Ciências**, v. 89, n. 2, p. 1189–1203, 2017.

LIRA, P. K.; TAMBOSI, L. R.; EWERS, R. M.; METZGER, J. P. Land-use and land-cover change in Atlantic Forest landscapes. **Forest Ecology and Management**, v. 278, p. 80–89, 2012.

LORENZI, H. Árvores brasileiras, vol1, ed1. **Plantarum, Nova Odessa**, v. 1, n. 1, 1992. a.

LORENZI, H. Árvores brasileiras, vol1, ed5. **Plantarum, Nova Odessa**, v. 1, n. 5, 1992. b.

LORENZI, H. Árvores brasileiras, vol2, ed1. **Plantarum, Nova Odessa**, v. 2, n. 1, 1992. c.

MANN, H. B.; WHITNEY, D. R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. **The Annals of Mathematical Statistics**, v. 18, n. 1, p. 50–60, 1947.

MASCHLER, J.; ATZBERGER, C.; IMMITZER, M. Individual Tree Crown Segmentation and Classification of 13 Tree Species Using Airborne Hyperspectral Data. **Remote Sensing**, v. 10, n. 8, 2018. Available at: <<https://www.mdpi.com/2072-4292/10/8/1218>>

MATSUKI, T.; YOKOYA, N.; IWASAKI, A. Hyperspectral Tree Species Classification of Japanese Complex Mixed Forest With the Aid of Lidar Data. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 8, n. 5, p. 2177–2187, 2015.

MELGANI, F.; BRUZZONE, L. Classification of hyperspectral remote sensing images with support vector machines. **IEEE Transactions on Geoscience and Remote Sensing**, v. 42, n. 8, p. 1778–1790, 2004.

MENZE, B. H.; KELM, B. M.; MASUCH, R.; HIMMELREICH, U.; BACHERT, P.; PETRICH, W.; HAMPRECHT, F. A. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. **BMC Bioinformatics**, v. 10, n. 1, p. 213, 2009.

MERZLYAK, M. N.; GITELSON, A. A.; CHIVKUNOVA, O. B.; RAKITIN, V. YU. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. **Physiologia Plantarum**, v. 106, n. 1, p. 135–141, 1999.

MICASENSE. **RedEdge-MX** - **MicaSense**. [s.d.]. Available at: <<https://www.micasense.com/rededge-mx>>. Accessed on 26 Oct. 2019.

MICHEZ, A.; PIÉGAY, H.; LISEIN, J.; CLAESSENS, H.; LEJEUNE, P. Classification of riparian forest species and health condition using multi-temporal and hyperspatial imagery from unmanned aerial system. **Environmental monitoring and assessment**, v. 188, n. 3, p. 146, 2016.

MIYOSHI, G.; IMAI, N.; TOMMASELLI, A.; HONKAVAARA, E. Comparison of Pixel and Region-Based Approaches for Tree Species Mapping in Atlantic Forest Using Hyperspectral Images Acquired by Uav. **ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 4213, p. 1875–1880, 2019.

MIYOSHI, G.; IMAI, N.; TOMMASELLI, A.; HONKAVAARA, E. Spectral differences of tree species belonging to Atlantic forest obtained from UAV hyperspectral images. **To be published**, [s.d.].

MIYOSHI, G. T.; IMAI, N. N.; GARCIA TOMMASELLI, A. M.; ANTUNES DE MORAES, M. V.; HONKAVAARA, E. Evaluation of Hyperspectral Multitemporal Information to Improve Tree Species Identification in the Highly Diverse Atlantic Forest. **Remote Sensing**, v. 12, n. 2, 2020. Available at: <<https://www.mdpi.com/2072-4292/12/2/244>>

MIYOSHI, G. T.; IMAI, N. N.; TOMMASELLI, A. M. G.; HONKAVAARA, E.; NÄSI, R.; MORIYA, É. A. S. Radiometric block adjustment of hyperspectral image blocks in the

Brazilian environment. **International Journal of Remote Sensing**, v. 39, n. 15–16, p. 4910–4930, 2018.

NACHAR, N. The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. **Tutorials in Quantitative Methods for Psychology**, v. 4, n. 1, p. 13–20, 2008.

NÄSI, R.; HONKAVAARA, E.; LYYTIKÄINEN-SAARENMAA, P.; BLOMQVIST, M.; LITKEY, P.; HAKALA, T.; VILJANEN, N.; KANTOLA, T.; TANHUANPÄÄ, T.; HOLOPAINEN, M. Using UAV-Based Photogrammetry and Hyperspectral Imaging for Mapping Bark Beetle Damage at Tree-Level. **Remote Sensing**, v. 7, n. 11, p. 15467–15493, 2015.

NEVALAINEN, O.; HONKAVAARA, E.; TUOMINEN, S.; VILJANEN, N.; HAKALA, T.; YU, X.; HYYPPÄ, J.; SAARI, H.; PÖLÖNEN, I.; IMAI, N. N.; TOMMASSELLI, A. M. G. Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. **Remote Sensing**, v. 9, n. 3, 2017. Available at: <<https://www.mdpi.com/2072-4292/9/3/185>>

NUVEM UAV. **Batmap**. [s.d.]. Available at: <<http://nuvemuav.com/batmap>>. Accessed on 5 Dez. 2019.

OLIVEIRA, L. M. De; GALVÃO, L. S.; PONZONI, F. J. Topographic effects on the determination of hyperspectral vegetation indices: a case study in southeastern Brazil. **Geocarto International**, v. 0, n. 0, p. 1–18, 2019.

OLIVEIRA, R. A. De; TOMMASSELLI, A. M. G.; HONKAVAARA, E. Geometric Calibration of a Hyperspectral Frame Camera. **The Photogrammetric Record**, v. 31, n. 155, p. 325–347, 2016.

OTERO, V.; KERCHOVE, R. V. D.; SATYANARAYANA, B.; MARTÍNEZ-ESPINOSA, C.; FISOL, M. A. B.; IBRAHIM, M. R. B.; SULONG, I.; MOHD-LOKMAN, H.; LUCAS, R.; DAHDOUH-GUEBAS, F. Managing mangrove forests from the sky: Forest inventory using field data and Unmanned Aerial Vehicle (UAV) imagery in the Matang Mangrove Forest Reserve, peninsular Malaysia. **Forest Ecology and Management**, v. 411, p. 35–45, 2018.

PANEQUE-GÁLVEZ, J.; MCCALL, M. K.; NAPOLETANO, B. M.; WICH, S. A.; KOH, L. P. Small Drones for Community-Based Forest Monitoring: An Assessment of Their Feasibility and Potential in Tropical Areas. **Forests**, v. 5, n. 6, p. 1481–1507, 2014.

PENUELAS, J.; BARET, F.; FILELLA, I. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. **Photosynthetica**, v. 31, n. 2, p. 221–230, 1995.

PLOWRIGHT, A. **ForestTools: Analyzing Remotely Sensed Forest Data**. 0.2.0. [s.l.: s.n.]. Available at: <<https://github.com/andrew-plowright/ForestTools>>

PONZONI, F. J.; KUPLICH, T. M.; SHIMABUKURO, Y. E. **Sensoriamento remoto no estudo da vegetação**. [s.l.]: Oficina de Textos, 2012. Available at: <<https://books.google.com.br/books?id=wcPJNAEACAAJ>>

POPESCU, D.; STOICAN, F.; STAMATESCU, G.; ICHIM, L.; DRAGANA, C. Advanced UAV–WSN System for Intelligent Monitoring in Precision Agriculture. **Sensors**, v. 20, n. 3, 2020. Available at: <<https://www.mdpi.com/1424-8220/20/3/817>>

POPESCU, S. C.; WYNNE, R. H. Seeing the Trees in the Forest. **Photogrammetric Engineering & Remote Sensing**, v. 70, n. 5, p. 589–604, 2004.

PRICE, J. C. How unique are spectral signatures? **Remote Sensing of Environment**, v. 49, n. 3, p. 181–186, 1994.

QGIS DEVELOPMENT TEAM. **QGIS Geographic Information System**. [s.l.]: Open Source Geospatial Foundation, 2009. Available at: <<http://qgis.org>>

RIBEIRO, M. C.; METZGER, J. P.; MARTENSEN, A. C.; PONZONI, F. J.; HIROTA, M. M. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. **Biological Conservation**, v. 142, n. 6, p. 1141–1153, 2009.

RICHARDS, J. A.; JIA, X. Interpretation of hyperspectral image data. **Remote Sensing Digital Image Analysis: An Introduction**, p. 359–388, 2006.

ROUSE, J.; HAAS, R.; SCHELL, J.; DEERING, D. Monitoring vegetation systems in the Great Plains with ERTS. **NASA special publication**, v. 351, p. 309, 1974.

SANCHEZ-AZOFEIFA, A.; ANTONIO GUZMÁN, J.; CAMPOS, C. A.; CASTRO, S.; GARCIA-MILLAN, V.; NIGHTINGALE, J.; RANKINE, C. Twenty-first century remote sensing technologies are revolutionizing the study of tropical forests. **Biotropica**, v. 49, n. 5, p. 604–619, 2017.

SANTOS, J. A.; FARIA, F.; CALUMBY, R.; TORRES, R. d S.; LAMPARELLI, R. A. C. A Genetic Programming approach for coffee crop recognition. *In: 2010 IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM*, 2010.

SEGAL, M. R. **Machine learning benchmarks and random forest regression**. [s. l.], 2004.

SENOP. **Datasheet**. [s.d.]. Available at: <https://senop.fi/files/rikola/pdf/Hyperspectral+Camera_Datasheet.pdf>. Accessed on 25 Oct. 2019.

SHAKHATREH, H.; SAWALMEH, A. H.; AL-FUQAHA, A.; DOU, Z.; ALMAITA, E.; KHALIL, I.; OTHMAN, N. S.; KHREISHAH, A.; GUIZANI, M. Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges. **IEEE Access**, v. 7, p. 48572–48634, 2019.

SILVA, C. A.; HUDAK, A. T.; VIERLING, L. A.; LOUDERMILK, E. L.; O'BRIEN, J.; HIERS, J. K.; JACK, S. B.; GONZALEZ-BENECKE, C. A.; LEE, H.; FALKOWSKI, M. J.; KHOSRAVIPOUR, A. Imputation of Individual Longleaf Pine (*Pinus palustris* Mill.) Tree Attributes from Field and LiDAR Data. **Canadian journal of remote sensing**, v. 42, n. 5, p. 554-573, 2016.

SILVA, F. R. Da; BEGNINI, R. M.; LOPES, B. C.; CASTELLANI, T. T. Seed dispersal and predation in the palm *Syagrus romanzoffiana* on two islands with different faunal richness,

southern Brazil. **Studies on Neotropical Fauna and Environment**, v. 46, n. 3, p. 163–171, 2011.

SIMS, D. A.; GAMON, J. A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. **Remote Sensing of Environment**, v. 81, n. 2, p. 337–354, 2002.

SMITH, G. M.; MILTON, E. J. The use of the empirical line method to calibrate remotely sensed data to reflectance. **International Journal of Remote Sensing**, v. 20, n. 13, p. 2653–2662, 1999.

SOMERS, B.; ASNER, G. P. Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: Wavelength adaptive spectral mixture analysis. **International Journal of Applied Earth Observation and Geoinformation**, v. 31, p. 57–66, 2014.

SOTHE, C.; ALMEIDA, C. M. D.; SCHIMALSKI, M. B.; ROSA, L. E. C. L.; CASTRO, J. D. B.; FEITOSA, R. Q.; DALPONTE, M.; LIMA, C. L.; LIESENBERG, V.; MIYOSHI, G. T.; TOMMASELLI, A. M. G. Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. **GIScience & Remote Sensing**, v. 0, n. 0, p. 1–26, 2020.

SOTHE, C.; DALPONTE, M.; ALMEIDA, C. M. De; SCHIMALSKI, M. B.; LIMA, C. L.; LIESENBERG, V.; MIYOSHI, G. T.; TOMMASELLI, A. M. G. Tree Species Classification in a Highly Diverse Subtropical Forest Integrating UAV-Based Photogrammetric Point Cloud and Hyperspectral Data. **Remote Sensing**, v. 11, n. 11, 2019. Available at: <<https://www.mdpi.com/2072-4292/11/11/1338>>

STRÎMBU, V. F.; STRÎMBU, B. M. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 104, p. 30–43, 2015.

TANHUANPÄÄ, T.; VASTARANTA, M.; KANKARE, V.; HOLOPAINEN, M.; HYYPPÄ, J.; HYYPPÄ, H.; ALHO, P.; RAISIO, J. Mapping of urban roadside trees – A case study in the tree register update process in Helsinki City. **Urban Forestry & Urban Greening**, v. 13, n. 3, p. 562–570, 2014.

TARABALKA, Y.; CHANUSSOT, J.; BENEDIKTSSON, J. A.; ANGULO, J.; FAUVEL, M. Segmentation and Classification of Hyperspectral Data using Watershed. In: **IGARSS 2008 - 2008 IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM**, 2008.

TUOMINEN, S.; NÄSI, R.; HONKAVAARA, E.; BALAZS, A.; HAKALA, T.; VILJANEN, N.; PÖLÖNEN, I.; SAARI, H.; OJANEN, H. Assessment of Classifiers and Remote Sensing Features of Hyperspectral Imagery and Stereo-Photogrammetric Point Clouds for Recognition of Tree Species in a Forest Area of High Species Diversity. **Remote Sensing**, v. 10, n. 5, 2018. Available at: <<https://www.mdpi.com/2072-4292/10/5/714>>

VAN DER WALT, S.; SCHÖNBERGER, J. L.; NUNEZ-IGLESIAS, J.; BOULOGNE, F.; WARNER, J. D.; YAGER, N.; GOUILLART, E.; YU, T.; CONTRIBUTORS, The Scikit-image. Scikit-image: image processing in Python. **PeerJ**, v. 2, p. e453, 2014.

VAUHKONEN, J.; ENE, L.; GUPTA, S.; HEINZEL, J.; HOLMGREN, J.; PITKÄNEN, J.; SOLBERG, S.; WANG, Y.; WEINACKER, H.; HAUGLIN, K. M.; LIEN, V.; PACKALÉN, P.; GOBAKKEN, T.; KOCH, B.; NÆSSET, E.; TOKOLA, T.; MALTAMO, M. Comparative testing of single-tree detection algorithms under different types of forest. **Forestry: An International Journal of Forest Research**, v. 85, n. 1, p. 27–40, 2011.

VIRTANEN, P.; GOMMERS, R.; OLIPHANT, T. E.; HABERLAND, M.; REDDY, T.; COURNAPEAU, D.; BUROVSKI, E.; PETERSON, P.; WECKESSER, W.; BRIGHT, J.; VAN DER WALT, S. J.; BRETT, M.; WILSON, J.; JARROD MILLMAN, K.; MAYOROV, N.; NELSON, A. R. J.; JONES, E.; KERN, R.; LARSON, E.; CAREY, C.; POLAT, İ.; FENG, Y.; MOORE, E. W.; VAND ERPLAS, J.; LAXALDE, D.; PERKTOLD, J.; CIMRMAN, R.; HENRIKSEN, I.; QUINTERO, E. A.; HARRIS, C. R.; ARCHIBALD, A. M.; RIBEIRO, A. H.; PEDREGOSA, F.; VAN MULBREGT, P.; CONTRIBUTORS, S. I. 0. SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python. **arXiv e-prints**, p. arXiv:1907.10121, 2019.

WAGNER, F. H.; FERREIRA, M. P.; SANCHEZ, A.; HIRYE, M. C. M.; ZORTEA, M.; GLOOR, E.; PHILLIPS, O. L.; FILHO, C. R. de S.; SHIMABUKURO, Y. E.; ARAGÃO, L. E. O. C. Individual tree crown delineation in a highly diverse tropical forest using very high resolution satellite images. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 145, p. 362–377, 2018.

WANG, Q.; NI-MEISTER, W. Forest Canopy Height and Gaps from Multiangular BRDF, Assessed with Airborne LiDAR Data (Short Title: Vegetation Structure from LiDAR and Multiangular Data). **Remote Sensing**, v. 11, n. 21, 2019. Available at: <<https://www.mdpi.com/2072-4292/11/21/2566>>

WILCOXON, F. **Individual comparisons by ranking methods**. **Biom Bull 1: 80–83**, 1945.

WITTEN, I. H.; FRANK, E. **Data Mining: Practical Machine Learning Tools and Techniques**. 2nd. ed. [s.l.]: Morgan Kaufmann, 2005.

ZHANG, C.; CHEN, K.; LIU, Y.; KOVACS, J. M.; FLORES-VERDUGO, F.; DE SANTIAGO, F. J. F. Spectral response to varying levels of leaf pigments collected from a degraded mangrove forest. **Journal of Applied Remote Sensing**, v. 6, n. 1, p. 063501, 2012.

ZHANG, J.; RIVARD, B.; SÁNCHEZ-AZOFEIFA, A.; CASTRO-ESAU, K. Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. **Remote Sensing of Environment**, v. 105, n. 2, p. 129–141, 2006.

ZHANG, S.; LI, S.; FU, W.; FANG, L. Multiscale Superpixel-Based Sparse Representation for Hyperspectral Image Classification. **Remote Sensing**, v. 9, n. 2, 2017. Available at: <<https://www.mdpi.com/2072-4292/9/2/139>>

APPENDIX A – WILCOXON-MANN-WHITNEY RESULTS

From the spectral response of each tree species recognized in the field, the mean reflectance factor spectra were obtained as well as the mean normalized spectra. In sequence, the spectral differences between-species (inter-classes) were calculated and the Wilcoxon-Mann-Whitney test was applied. Tables 1 to 4 show the p-value of the Wilcoxon-Mann-Whitney test to each pairwise combination and spectral band considering the mean reflectance factor and the mean normalized reflectance factor. The test was applied with a confidence level of 5% ($\alpha = 5\%$) indicating that when the p-value is lower than 0.05 the null hypothesis is rejected, i.e., it can be inferred that there is no spectral difference between the pairwise tested.

| | | | | | | | | | | | | | Continuation |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------|
| Pairwise combination | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 | B13 |
| CO - PE | 0.020 | 0.020 | 0.026 | 0.044 | 0.072 | 0.112 | 0.112 | 0.138 | 0.112 | 0.138 | 0.015 | 0.026 | 0.020 |
| CO - SU | 0.182 | 0.182 | 0.253 | 0.374 | 0.341 | 0.204 | 0.204 | 0.127 | 0.144 | 0.127 | 0.374 | 0.127 | 0.099 |
| GA - GR | 0.944 | 0.832 | 0.944 | 0.944 | 0.944 | 0.944 | 0.944 | 0.944 | 0.724 | 0.944 | 0.944 | 0.944 | 0.944 |
| GA - GT | 0.832 | 0.832 | 0.944 | 0.944 | 0.944 | 0.724 | 0.525 | 0.289 | 0.437 | 0.289 | 0.077 | 0.179 | 0.179 |
| GA - IN | 0.450 | 0.230 | 0.267 | 0.307 | 0.307 | 0.143 | 0.168 | 0.143 | 0.120 | 0.168 | 0.083 | 0.120 | 0.230 |
| GA - JA | 0.379 | 0.307 | 0.550 | 0.860 | 0.860 | 0.699 | 0.597 | 0.699 | 0.550 | 0.460 | 0.193 | 0.307 | 0.275 |
| GA - JE | 0.001 | 0.001 | 0.000 | 0.000 | 0.001 | 0.001 | 0.002 | 0.006 | 0.011 | 0.011 | 0.037 | 0.029 | 0.041 |
| GA - PE | 0.035 | 0.108 | 0.151 | 0.205 | 0.272 | 0.554 | 0.554 | 0.554 | 0.673 | 0.673 | 0.035 | 0.272 | 0.108 |
| GA - SU | 0.407 | 0.464 | 0.807 | 0.884 | 0.961 | 0.733 | 0.733 | 0.661 | 0.733 | 0.464 | 0.961 | 0.807 | 0.591 |
| GR - GT | 0.665 | 0.665 | 0.665 | 0.665 | 0.885 | 0.665 | 0.665 | 0.471 | 0.471 | 0.471 | 0.665 | 0.471 | 0.471 |
| GR - IN | 0.350 | 0.270 | 0.270 | 0.270 | 0.350 | 0.270 | 0.203 | 0.203 | 0.203 | 0.203 | 0.552 | 0.270 | 0.350 |
| GR - JA | 0.396 | 0.396 | 0.648 | 0.744 | 0.845 | 0.744 | 0.744 | 0.744 | 0.557 | 0.648 | 0.744 | 0.557 | 0.473 |
| GR - JE | 0.022 | 0.040 | 0.018 | 0.018 | 0.027 | 0.040 | 0.069 | 0.081 | 0.096 | 0.096 | 0.081 | 0.131 | 0.261 |
| GR - PE | 0.216 | 0.216 | 0.377 | 0.596 | 0.596 | 0.860 | 0.860 | 0.860 | 0.860 | 0.860 | 0.216 | 0.377 | 0.377 |
| GR - SU | 0.508 | 0.508 | 0.925 | 0.925 | 0.925 | 0.925 | 0.637 | 0.395 | 0.508 | 0.395 | 0.925 | 0.777 | 0.637 |
| GT - IN | 0.671 | 0.671 | 0.350 | 0.203 | 0.270 | 0.552 | 0.671 | 0.799 | 0.932 | 0.932 | 0.552 | 0.445 | 0.350 |
| GT - JA | 0.744 | 0.648 | 0.845 | 0.744 | 0.948 | 0.948 | 0.845 | 0.744 | 0.557 | 0.328 | 0.648 | 0.648 | 0.744 |
| GT - JE | 0.008 | 0.010 | 0.008 | 0.012 | 0.012 | 0.015 | 0.012 | 0.012 | 0.012 | 0.012 | 0.018 | 0.012 | 0.018 |
| GT - PE | 0.112 | 0.112 | 0.112 | 0.377 | 0.377 | 0.377 | 0.377 | 0.216 | 0.216 | 0.216 | 0.052 | 0.052 | 0.052 |
| GT - SU | 0.395 | 0.508 | 0.777 | 0.925 | 0.925 | 0.508 | 0.508 | 0.395 | 0.299 | 0.219 | 0.508 | 0.299 | 0.219 |
| IN - JA | 0.967 | 0.967 | 0.592 | 0.386 | 0.342 | 0.265 | 0.342 | 0.302 | 0.342 | 0.483 | 0.967 | 0.967 | 0.901 |
| IN - JE | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.003 | 0.002 | 0.007 |
| IN - PE | 0.019 | 0.019 | 0.032 | 0.083 | 0.126 | 0.053 | 0.083 | 0.053 | 0.053 | 0.053 | 0.019 | 0.032 | 0.032 |
| IN - SU | 0.224 | 0.183 | 0.385 | 0.603 | 0.685 | 0.272 | 0.183 | 0.183 | 0.183 | 0.148 | 0.385 | 0.183 | 0.183 |
| JA - JE | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.001 | 0.002 | 0.001 | 0.002 | 0.002 | 0.002 |
| JA - PE | 0.029 | 0.029 | 0.087 | 0.161 | 0.276 | 0.276 | 0.276 | 0.213 | 0.213 | 0.276 | 0.029 | 0.043 | 0.029 |

Continuation

| Pairwise combination | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 | B13 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| JA - SU | 0.239 | 0.277 | 0.587 | 1.000 | 1.000 | 0.587 | 0.469 | 0.469 | 0.415 | 0.319 | 0.587 | 0.277 | 0.205 |
| JE - PE | 0.438 | 0.338 | 0.157 | 0.075 | 0.110 | 0.091 | 0.110 | 0.110 | 0.218 | 0.294 | 0.964 | 0.964 | 0.964 |
| JE - SU | 0.014 | 0.016 | 0.004 | 0.003 | 0.004 | 0.014 | 0.022 | 0.043 | 0.064 | 0.092 | 0.029 | 0.128 | 0.257 |
| PE - SU | 0.255 | 0.362 | 0.255 | 0.255 | 0.494 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.172 | 0.362 | 0.362 |

Table 2 - P-values of the Wilcoxon-Mann-Whitney test applied to spectral bands 14 to 25 in each pairwise combination when using the mean reflectance factor spectra. Shaded values represent the one where the test was rejected.

| Pairwise combination | B14 | B15 | B16 | B17 | B18 | B19 | B20 | B21 | B22 | B23 | B24 | B25 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| AM - CA | 0.036 | 0.071 | 0.961 | 0.591 | 0.354 | 0.306 | 0.526 | 0.661 | 0.464 | 0.464 | 0.526 | 0.354 |
| AM - CO | 0.002 | 0.002 | 0.002 | 0.005 | 0.008 | 0.008 | 0.008 | 0.011 | 0.022 | 0.022 | 0.026 | 0.042 |
| AM - GA | 0.064 | 0.121 | 0.141 | 0.345 | 0.385 | 0.473 | 0.521 | 0.473 | 0.571 | 0.678 | 0.791 | 0.850 |
| AM - GR | 0.289 | 0.289 | 0.229 | 0.229 | 0.358 | 0.358 | 0.437 | 0.289 | 0.358 | 0.358 | 0.358 | 0.437 |
| AM - GT | 0.028 | 0.028 | 0.040 | 0.104 | 0.229 | 0.437 | 0.621 | 0.832 | 0.724 | 0.724 | 0.724 | 0.944 |
| AM - IN | 0.011 | 0.024 | 0.015 | 0.083 | 0.230 | 0.351 | 0.625 | 0.625 | 0.824 | 0.894 | 0.965 | 0.965 |
| AM - JA | 0.008 | 0.022 | 0.073 | 0.245 | 0.379 | 0.418 | 0.550 | 0.504 | 0.504 | 0.550 | 0.699 | 0.916 |
| AM - JE | 0.947 | 0.912 | 0.582 | 0.495 | 0.775 | 0.843 | 0.281 | 0.153 | 0.045 | 0.029 | 0.062 | 0.068 |
| AM - PE | 0.447 | 0.673 | 0.933 | 0.800 | 0.205 | 0.023 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 |
| AM - SU | 0.262 | 0.262 | 0.188 | 0.130 | 0.057 | 0.036 | 0.022 | 0.022 | 0.022 | 0.022 | 0.017 | 0.017 |
| CA - CO | 0.086 | 0.019 | 0.005 | 0.002 | 0.002 | 0.002 | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| CA - GA | 0.807 | 0.733 | 0.188 | 0.071 | 0.028 | 0.071 | 0.107 | 0.130 | 0.130 | 0.157 | 0.130 | 0.130 |
| CA - GR | 0.777 | 0.777 | 0.395 | 0.219 | 0.219 | 0.219 | 0.156 | 0.219 | 0.219 | 0.156 | 0.219 | 0.156 |
| CA - GT | 0.219 | 0.047 | 0.030 | 0.030 | 0.047 | 0.156 | 0.219 | 0.299 | 0.219 | 0.156 | 0.219 | 0.156 |
| CA - IN | 0.224 | 0.093 | 0.024 | 0.024 | 0.043 | 0.118 | 0.183 | 0.183 | 0.272 | 0.325 | 0.325 | 0.325 |
| CA - JA | 0.147 | 0.124 | 0.070 | 0.046 | 0.046 | 0.103 | 0.057 | 0.085 | 0.070 | 0.057 | 0.070 | 0.037 |
| CA - JE | 0.092 | 0.257 | 0.678 | 0.175 | 0.143 | 0.455 | 0.978 | 0.455 | 0.175 | 0.194 | 0.175 | 0.489 |
| CA - PE | 0.040 | 0.040 | 0.820 | 1.000 | 0.649 | 0.172 | 0.040 | 0.023 | 0.023 | 0.023 | 0.040 | 0.040 |

| Continuation | | | | | | | | | | | | |
|-----------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Pairwise combination | B14 | B15 | B16 | B17 | B18 | B19 | B20 | B21 | B22 | B23 | B24 | B25 |
| CA - SU | 0.443 | 0.898 | 0.160 | 0.055 | 0.022 | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 | 0.007 |
| CO - GA | 0.093 | 0.083 | 0.075 | 0.037 | 0.013 | 0.007 | 0.006 | 0.006 | 0.006 | 0.005 | 0.004 | 0.005 |
| CO - GR | 0.395 | 0.347 | 0.303 | 0.303 | 0.347 | 0.347 | 0.395 | 0.447 | 0.447 | 0.447 | 0.447 | 0.447 |
| CO - GT | 0.754 | 0.687 | 0.622 | 0.347 | 0.194 | 0.098 | 0.028 | 0.035 | 0.035 | 0.035 | 0.054 | 0.066 |
| CO - IN | 0.398 | 0.336 | 0.398 | 0.256 | 0.190 | 0.086 | 0.034 | 0.025 | 0.013 | 0.013 | 0.013 | 0.013 |
| CO - JA | 0.371 | 0.221 | 0.100 | 0.039 | 0.010 | 0.003 | 0.003 | 0.003 | 0.004 | 0.003 | 0.006 | 0.008 |
| CO - JE | 0.001 | 0.001 | 0.002 | 0.003 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| CO - PE | 0.020 | 0.015 | 0.020 | 0.015 | 0.011 | 0.008 | 0.008 | 0.008 | 0.008 | 0.008 | 0.008 | 0.008 |
| CO - SU | 0.127 | 0.253 | 0.446 | 0.657 | 0.849 | 0.409 | 0.162 | 0.099 | 0.086 | 0.075 | 0.066 | 0.086 |
| GA - GR | 0.944 | 0.944 | 0.944 | 0.944 | 0.944 | 0.832 | 0.724 | 0.724 | 0.724 | 0.525 | 0.437 | 0.437 |
| GA - GT | 0.229 | 0.229 | 0.289 | 0.437 | 0.832 | 0.944 | 0.724 | 0.621 | 0.944 | 0.944 | 0.944 | 0.944 |
| GA - IN | 0.307 | 0.505 | 0.351 | 0.399 | 0.505 | 0.824 | 0.965 | 0.965 | 0.824 | 0.894 | 0.965 | 0.689 |
| GA - JA | 0.218 | 0.504 | 0.805 | 0.972 | 0.916 | 0.916 | 0.860 | 0.751 | 0.860 | 0.972 | 0.860 | 0.805 |
| GA - JE | 0.099 | 0.108 | 0.244 | 0.524 | 0.441 | 0.180 | 0.024 | 0.003 | 0.002 | 0.003 | 0.003 | 0.007 |
| GA - PE | 0.076 | 0.076 | 0.151 | 0.205 | 0.052 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 |
| GA - SU | 0.661 | 0.884 | 0.961 | 0.306 | 0.157 | 0.017 | 0.013 | 0.013 | 0.010 | 0.010 | 0.010 | 0.010 |
| GR - GT | 0.471 | 0.471 | 0.471 | 0.665 | 0.665 | 0.885 | 0.665 | 0.471 | 0.885 | 0.885 | 0.885 | 0.885 |
| GR - IN | 0.445 | 0.445 | 0.552 | 0.671 | 0.932 | 0.799 | 0.552 | 0.552 | 0.552 | 0.445 | 0.350 | 0.350 |
| GR - JA | 0.473 | 0.557 | 0.845 | 0.948 | 0.948 | 0.744 | 0.648 | 0.557 | 0.744 | 0.744 | 0.648 | 0.744 |
| GR - JE | 0.373 | 0.416 | 0.373 | 0.510 | 0.510 | 0.201 | 0.069 | 0.040 | 0.018 | 0.018 | 0.010 | 0.027 |
| GR - PE | 0.216 | 0.377 | 0.377 | 0.112 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 |
| GR - SU | 0.925 | 0.925 | 0.925 | 0.925 | 0.637 | 0.395 | 0.395 | 0.395 | 0.299 | 0.395 | 0.299 | 0.299 |
| GT - IN | 0.445 | 0.552 | 0.932 | 0.932 | 0.932 | 0.799 | 0.799 | 0.932 | 0.932 | 0.799 | 0.799 | 0.671 |
| GT - JA | 0.744 | 0.744 | 0.648 | 0.648 | 0.744 | 0.948 | 0.948 | 0.648 | 0.648 | 0.845 | 0.948 | 0.744 |
| GT - JE | 0.040 | 0.048 | 0.048 | 0.201 | 0.261 | 0.230 | 0.112 | 0.081 | 0.018 | 0.018 | 0.022 | 0.040 |

Continuation

| Pairwise combination | B14 | B15 | B16 | B17 | B18 | B19 | B20 | B21 | B22 | B23 | B24 | B25 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| GT - PE | 0.052 | 0.052 | 0.112 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 |
| GT - SU | 0.299 | 0.395 | 0.508 | 0.925 | 0.395 | 0.108 | 0.073 | 0.073 | 0.073 | 0.073 | 0.073 | 0.073 |
| IN - JA | 0.592 | 0.837 | 0.592 | 0.386 | 0.536 | 0.710 | 0.773 | 0.967 | 0.710 | 0.710 | 0.592 | 0.483 |
| IN - JE | 0.014 | 0.027 | 0.045 | 0.109 | 0.213 | 0.134 | 0.031 | 0.027 | 0.016 | 0.016 | 0.024 | 0.045 |
| IN - PE | 0.032 | 0.032 | 0.053 | 0.083 | 0.032 | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 |
| IN - SU | 0.224 | 0.272 | 0.603 | 0.862 | 0.385 | 0.073 | 0.032 | 0.018 | 0.013 | 0.013 | 0.013 | 0.024 |
| JA - JE | 0.003 | 0.016 | 0.143 | 0.470 | 0.421 | 0.239 | 0.010 | 0.002 | 0.000 | 0.000 | 0.000 | 0.002 |
| JA - PE | 0.029 | 0.029 | 0.087 | 0.161 | 0.029 | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 | 0.013 |
| JA - SU | 0.277 | 0.415 | 1.000 | 0.319 | 0.124 | 0.030 | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 |
| JE - PE | 0.494 | 0.494 | 0.616 | 0.254 | 0.061 | 0.032 | 0.032 | 0.032 | 0.040 | 0.050 | 0.050 | 0.050 |
| JE - SU | 0.306 | 0.361 | 0.257 | 0.143 | 0.025 | 0.005 | 0.001 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| PE - SU | 0.172 | 0.255 | 0.362 | 0.111 | 0.040 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 |

Table 3 - P-values of the Wilcoxon-Mann-Whitney test applied to spectral bands 1 to 13 in each pairwise combination when using the normalized mean reflectance spectra. Shaded values represent the one where the test was rejected.

| Pairwise combination | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 | B13 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| AM - CA | 0.005 | 0.005 | 0.005 | 0.088 | 0.157 | 0.188 | 0.107 | 0.071 | 0.057 | 0.036 | 0.045 | 0.022 | 0.007 |
| AM - CO | 0.861 | 0.900 | 0.782 | 0.238 | 0.139 | 0.407 | 0.380 | 0.200 | 0.200 | 0.183 | 0.980 | 0.900 | 0.280 |
| AM - GA | 0.186 | 0.104 | 0.054 | 0.031 | 0.038 | 0.121 | 0.241 | 0.241 | 0.212 | 0.186 | 0.970 | 0.623 | 0.473 |
| AM - GR | 0.621 | 0.621 | 0.832 | 0.944 | 0.621 | 0.944 | 0.944 | 0.525 | 0.944 | 0.944 | 0.358 | 0.437 | 0.832 |
| AM - GT | 0.358 | 0.179 | 0.289 | 0.358 | 0.229 | 0.138 | 0.056 | 0.040 | 0.028 | 0.040 | 0.009 | 0.020 | 0.028 |
| AM - IN | 0.083 | 0.030 | 0.011 | 0.004 | 0.004 | 0.005 | 0.011 | 0.004 | 0.005 | 0.009 | 0.046 | 0.046 | 0.024 |
| AM - JA | 0.027 | 0.010 | 0.032 | 0.032 | 0.032 | 0.053 | 0.045 | 0.032 | 0.032 | 0.027 | 0.130 | 0.062 | 0.012 |
| AM - JE | 0.947 | 0.878 | 0.878 | 0.878 | 0.644 | 0.708 | 0.775 | 0.441 | 0.322 | 0.244 | 0.344 | 0.194 | 0.068 |

| | | | | | | | | | | | | | Continuation |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------|
| Pairwise combination | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 | B13 |
| AM - PE | 0.023 | 0.035 | 0.014 | 0.014 | 0.014 | 0.014 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.035 | 0.076 |
| AM - SU | 0.130 | 0.130 | 0.130 | 0.223 | 0.306 | 0.157 | 0.157 | 0.223 | 0.223 | 0.223 | 0.157 | 0.157 | 0.223 |
| CA - CO | 0.002 | 0.002 | 0.002 | 0.099 | 0.703 | 0.485 | 0.310 | 0.525 | 0.310 | 0.228 | 0.049 | 0.009 | 0.086 |
| CA - GA | 0.057 | 0.107 | 0.306 | 0.961 | 0.354 | 0.961 | 0.961 | 0.733 | 0.464 | 0.354 | 0.130 | 0.057 | 0.188 |
| CA - GR | 0.018 | 0.011 | 0.018 | 0.073 | 0.108 | 0.073 | 0.030 | 0.030 | 0.030 | 0.011 | 0.030 | 0.018 | 0.018 |
| CA - GT | 0.108 | 0.108 | 0.156 | 0.219 | 0.637 | 0.299 | 0.219 | 0.108 | 0.156 | 0.108 | 0.156 | 0.925 | 0.777 |
| CA - IN | 0.148 | 0.183 | 0.954 | 0.073 | 0.009 | 0.024 | 0.056 | 0.018 | 0.118 | 0.148 | 0.862 | 0.772 | 0.862 |
| CA - JA | 0.037 | 0.174 | 0.205 | 0.928 | 0.526 | 0.856 | 0.928 | 0.928 | 1.000 | 0.928 | 0.365 | 0.277 | 0.365 |
| CA - JE | 0.001 | 0.001 | 0.001 | 0.002 | 0.006 | 0.012 | 0.010 | 0.019 | 0.025 | 0.025 | 0.064 | 0.033 | 0.064 |
| CA - PE | 0.820 | 1.000 | 0.068 | 0.023 | 0.023 | 0.023 | 0.023 | 0.068 | 0.068 | 0.068 | 0.172 | 0.362 | 1.000 |
| CA - SU | 0.003 | 0.005 | 0.005 | 0.005 | 0.011 | 0.015 | 0.015 | 0.022 | 0.041 | 0.030 | 0.030 | 0.011 | 0.030 |
| CO - GA | 0.060 | 0.060 | 0.047 | 0.075 | 0.167 | 0.353 | 0.380 | 0.670 | 0.633 | 0.782 | 0.436 | 0.436 | 0.861 |
| CO - GR | 0.447 | 0.303 | 0.502 | 0.303 | 0.395 | 0.303 | 0.395 | 0.347 | 0.227 | 0.347 | 0.447 | 0.303 | 0.263 |
| CO - GT | 0.140 | 0.140 | 0.227 | 0.687 | 0.447 | 0.263 | 0.140 | 0.117 | 0.098 | 0.066 | 0.011 | 0.014 | 0.140 |
| CO - IN | 0.013 | 0.011 | 0.008 | 0.003 | 0.004 | 0.004 | 0.011 | 0.008 | 0.010 | 0.034 | 0.025 | 0.025 | 0.137 |
| CO - JA | 0.000 | 0.001 | 0.013 | 0.082 | 0.173 | 0.221 | 0.240 | 0.452 | 0.397 | 0.397 | 0.121 | 0.060 | 0.221 |
| CO - JE | 0.891 | 0.552 | 0.369 | 0.110 | 0.046 | 0.241 | 0.266 | 0.185 | 0.337 | 0.455 | 0.253 | 0.148 | 0.749 |
| CO - PE | 0.008 | 0.015 | 0.008 | 0.011 | 0.008 | 0.008 | 0.011 | 0.026 | 0.034 | 0.044 | 0.034 | 0.020 | 0.244 |
| CO - SU | 0.042 | 0.036 | 0.016 | 0.011 | 0.006 | 0.009 | 0.009 | 0.016 | 0.031 | 0.026 | 0.042 | 0.031 | 0.031 |
| GA - GR | 0.077 | 0.040 | 0.040 | 0.077 | 0.179 | 0.179 | 0.229 | 0.437 | 0.437 | 0.437 | 0.437 | 0.289 | 0.437 |
| GA - GT | 0.944 | 0.944 | 0.525 | 0.358 | 0.525 | 0.832 | 0.437 | 0.229 | 0.179 | 0.104 | 0.104 | 0.138 | 0.179 |
| GA - IN | 0.689 | 0.450 | 0.230 | 0.100 | 0.143 | 0.056 | 0.120 | 0.037 | 0.083 | 0.143 | 0.100 | 0.198 | 0.267 |
| GA - JA | 0.504 | 0.597 | 0.860 | 0.972 | 0.972 | 0.751 | 0.647 | 0.647 | 0.597 | 0.597 | 0.218 | 0.245 | 0.307 |
| GA - JE | 0.024 | 0.011 | 0.003 | 0.003 | 0.013 | 0.033 | 0.108 | 0.226 | 0.468 | 0.644 | 0.301 | 0.281 | 0.582 |
| GA - PE | 0.151 | 0.108 | 0.035 | 0.035 | 0.023 | 0.023 | 0.023 | 0.076 | 0.151 | 0.151 | 0.052 | 0.108 | 0.272 |

| | | | | | | | | | | | | | Continuation |
|-----------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|--------------|
| Pairwise combination | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | B10 | B11 | B12 | B13 |
| GA - SU | 0.010 | 0.013 | 0.004 | 0.004 | 0.005 | 0.007 | 0.013 | 0.022 | 0.045 | 0.036 | 0.045 | 0.036 | 0.057 |
| GR - GT | 0.061 | 0.030 | 0.194 | 0.194 | 0.112 | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 | 0.030 |
| GR - IN | 0.051 | 0.034 | 0.034 | 0.022 | 0.014 | 0.014 | 0.014 | 0.009 | 0.009 | 0.014 | 0.022 | 0.034 | 0.034 |
| GR - JA | 0.011 | 0.005 | 0.031 | 0.078 | 0.103 | 0.058 | 0.058 | 0.078 | 0.022 | 0.043 | 0.016 | 0.011 | 0.022 |
| GR - JE | 0.727 | 0.847 | 0.786 | 0.670 | 0.510 | 0.969 | 0.847 | 0.908 | 0.561 | 0.561 | 0.040 | 0.027 | 0.069 |
| GR - PE | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.112 |
| GR - SU | 0.637 | 0.637 | 0.395 | 0.219 | 0.108 | 0.219 | 0.219 | 0.299 | 0.299 | 0.299 | 0.508 | 0.395 | 0.219 |
| GT - IN | 0.552 | 0.445 | 0.149 | 0.075 | 0.051 | 0.203 | 0.350 | 0.552 | 0.799 | 0.932 | 0.671 | 0.552 | 0.671 |
| GT - JA | 0.215 | 0.473 | 0.744 | 0.557 | 0.845 | 0.845 | 0.396 | 0.170 | 0.103 | 0.133 | 0.078 | 0.215 | 0.215 |
| GT - JE | 0.027 | 0.018 | 0.015 | 0.012 | 0.018 | 0.010 | 0.010 | 0.006 | 0.015 | 0.015 | 0.012 | 0.040 | 0.048 |
| GT - PE | 0.112 | 0.112 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.216 | 0.377 | 0.377 | 0.596 | 0.596 | 0.860 |
| GT - SU | 0.047 | 0.030 | 0.011 | 0.018 | 0.030 | 0.030 | 0.011 | 0.030 | 0.073 | 0.030 | 0.018 | 0.018 | 0.073 |
| IN - JA | 0.967 | 0.967 | 0.342 | 0.091 | 0.076 | 0.063 | 0.127 | 0.052 | 0.076 | 0.107 | 0.386 | 0.483 | 0.773 |
| IN - JE | 0.012 | 0.003 | 0.002 | 0.000 | 0.000 | 0.001 | 0.001 | 0.001 | 0.002 | 0.004 | 0.098 | 0.088 | 0.162 |
| IN - PE | 0.358 | 0.475 | 0.126 | 0.126 | 0.083 | 0.126 | 0.083 | 0.185 | 0.262 | 0.262 | 0.185 | 0.185 | 0.919 |
| IN - SU | 0.013 | 0.007 | 0.003 | 0.003 | 0.002 | 0.002 | 0.003 | 0.005 | 0.009 | 0.013 | 0.018 | 0.018 | 0.043 |
| JA - JE | 0.001 | 0.000 | 0.000 | 0.001 | 0.004 | 0.004 | 0.005 | 0.006 | 0.027 | 0.024 | 0.312 | 0.193 | 0.103 |
| JA - PE | 0.161 | 0.350 | 0.029 | 0.029 | 0.029 | 0.029 | 0.029 | 0.043 | 0.062 | 0.120 | 0.043 | 0.062 | 0.876 |
| JA - SU | 0.004 | 0.002 | 0.004 | 0.003 | 0.004 | 0.005 | 0.009 | 0.019 | 0.019 | 0.019 | 0.030 | 0.024 | 0.019 |
| JE - PE | 0.012 | 0.016 | 0.009 | 0.009 | 0.009 | 0.009 | 0.009 | 0.009 | 0.016 | 0.016 | 0.020 | 0.025 | 0.338 |
| JE - SU | 0.033 | 0.056 | 0.050 | 0.064 | 0.056 | 0.043 | 0.038 | 0.043 | 0.043 | 0.043 | 0.029 | 0.019 | 0.033 |
| PE - SU | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.023 | 0.040 | 0.040 | 0.040 | 0.040 | 0.111 |

Table 4 - P-values of the Wilcoxon-Mann-Whitney test applied to spectral bands 14 to 25 in each pairwise combination when using the normalized mean reflectance factor spectra. Shaded values represent the one where the test was rejected.

| Pairwise combination | B14 | B15 | B16 | B17 | B18 | B19 | B20 | B21 | B22 | B23 | B24 | B25 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| AM - CA | 0.010 | 0.028 | 0.188 | 0.961 | 0.354 | 0.013 | 0.010 | 0.188 | 0.188 | 0.088 | 0.130 | 0.088 |
| AM - CO | 0.498 | 0.238 | 0.530 | 0.633 | 0.436 | 0.744 | 0.861 | 0.940 | 0.530 | 0.353 | 0.564 | 0.139 |
| AM - GA | 0.473 | 0.308 | 0.521 | 0.473 | 0.734 | 0.678 | 0.385 | 0.571 | 0.521 | 0.308 | 0.345 | 0.064 |
| AM - GR | 0.724 | 0.832 | 0.832 | 0.724 | 0.525 | 0.944 | 0.289 | 0.179 | 0.832 | 0.621 | 0.621 | 0.437 |
| AM - GT | 0.028 | 0.028 | 0.028 | 0.077 | 0.229 | 0.621 | 0.013 | 0.006 | 0.056 | 0.056 | 0.289 | 0.138 |
| AM - IN | 0.037 | 0.037 | 0.030 | 0.046 | 0.083 | 0.756 | 0.307 | 0.083 | 0.019 | 0.009 | 0.019 | 0.003 |
| AM - JA | 0.012 | 0.053 | 0.245 | 0.460 | 0.972 | 0.038 | 0.004 | 0.045 | 0.170 | 0.098 | 0.218 | 0.085 |
| AM - JE | 0.010 | 0.007 | 0.000 | 0.000 | 0.000 | 0.000 | 0.019 | 0.468 | 0.002 | 0.003 | 0.017 | 0.041 |
| AM - PE | 0.076 | 0.076 | 0.023 | 0.023 | 0.052 | 0.800 | 0.014 | 0.014 | 0.023 | 0.023 | 0.052 | 0.052 |
| AM - SU | 0.262 | 0.223 | 0.107 | 0.088 | 0.464 | 0.807 | 0.130 | 0.130 | 0.188 | 0.188 | 0.157 | 0.306 |
| CA - CO | 0.036 | 0.162 | 0.611 | 0.657 | 0.066 | 0.008 | 0.004 | 0.204 | 0.310 | 0.204 | 0.657 | 0.374 |
| CA - GA | 0.157 | 0.262 | 0.884 | 0.591 | 0.262 | 0.107 | 0.157 | 0.661 | 0.733 | 0.661 | 0.807 | 0.884 |
| CA - GR | 0.018 | 0.018 | 0.508 | 0.395 | 0.073 | 0.030 | 0.011 | 0.011 | 0.073 | 0.073 | 0.219 | 0.047 |
| CA - GT | 0.925 | 0.395 | 0.047 | 0.011 | 0.011 | 0.108 | 0.156 | 0.073 | 0.156 | 0.777 | 0.219 | 0.925 |
| CA - IN | 0.524 | 0.954 | 0.073 | 0.002 | 0.002 | 0.009 | 0.118 | 0.325 | 0.056 | 0.056 | 0.032 | 0.018 |
| CA - JA | 0.277 | 0.319 | 0.856 | 0.239 | 0.070 | 0.365 | 0.717 | 0.319 | 1.000 | 0.786 | 0.928 | 0.651 |
| CA - JE | 0.175 | 0.847 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.422 | 0.081 | 0.128 | 0.103 | 0.306 |
| CA - PE | 1.000 | 0.820 | 0.040 | 0.023 | 0.023 | 0.023 | 0.068 | 0.040 | 0.023 | 0.068 | 0.068 | 0.362 |
| CA - SU | 0.030 | 0.030 | 0.030 | 0.055 | 0.898 | 0.030 | 0.007 | 0.030 | 0.022 | 0.022 | 0.030 | 0.041 |
| CO - GA | 0.782 | 0.861 | 0.821 | 0.707 | 0.980 | 0.633 | 0.328 | 0.530 | 0.633 | 0.467 | 0.328 | 0.530 |
| CO - GR | 0.303 | 0.263 | 0.823 | 0.964 | 0.754 | 0.754 | 0.263 | 0.347 | 0.561 | 0.622 | 0.502 | 0.395 |
| CO - GT | 0.117 | 0.054 | 0.022 | 0.035 | 0.117 | 0.893 | 0.005 | 0.003 | 0.140 | 0.165 | 0.263 | 0.347 |
| CO - IN | 0.109 | 0.232 | 0.039 | 0.025 | 0.051 | 0.308 | 0.154 | 0.076 | 0.008 | 0.006 | 0.013 | 0.004 |
| CO - JA | 0.132 | 0.371 | 0.424 | 0.707 | 0.814 | 0.027 | 0.000 | 0.021 | 0.301 | 0.347 | 0.347 | 0.638 |
| CO - JE | 0.148 | 0.075 | 0.000 | 0.000 | 0.000 | 0.000 | 0.011 | 0.307 | 0.002 | 0.002 | 0.012 | 0.726 |

Continuation

| Pairwise combination | B14 | B15 | B16 | B17 | B18 | B19 | B20 | B21 | B22 | B23 | B24 | B25 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| CO - PE | 0.244 | 0.169 | 0.026 | 0.011 | 0.044 | 0.672 | 0.008 | 0.008 | 0.008 | 0.015 | 0.072 | 0.138 |
| CO - SU | 0.031 | 0.026 | 0.016 | 0.008 | 0.112 | 0.751 | 0.031 | 0.016 | 0.013 | 0.016 | 0.013 | 0.016 |
| GA - GR | 0.229 | 0.358 | 0.525 | 0.944 | 0.621 | 0.437 | 0.289 | 0.229 | 0.525 | 0.437 | 0.229 | 0.289 |
| GA - GT | 0.289 | 0.138 | 0.056 | 0.179 | 0.229 | 0.944 | 0.056 | 0.009 | 0.179 | 0.437 | 0.832 | 0.621 |
| GA - IN | 0.307 | 0.307 | 0.100 | 0.083 | 0.143 | 0.399 | 0.689 | 0.198 | 0.056 | 0.069 | 0.120 | 0.030 |
| GA - JA | 0.218 | 0.379 | 0.972 | 0.972 | 0.916 | 0.218 | 0.170 | 0.113 | 0.699 | 0.805 | 0.972 | 0.972 |
| GA - JE | 0.180 | 0.166 | 0.004 | 0.001 | 0.000 | 0.000 | 0.002 | 0.912 | 0.026 | 0.068 | 0.344 | 0.843 |
| GA - PE | 0.205 | 0.205 | 0.052 | 0.035 | 0.108 | 0.554 | 0.035 | 0.014 | 0.035 | 0.052 | 0.205 | 0.272 |
| GA - SU | 0.057 | 0.045 | 0.022 | 0.028 | 0.306 | 0.464 | 0.036 | 0.022 | 0.028 | 0.028 | 0.017 | 0.022 |
| GR - GT | 0.030 | 0.030 | 0.030 | 0.030 | 0.312 | 0.471 | 0.030 | 0.030 | 0.030 | 0.061 | 0.030 | 0.030 |
| GR - IN | 0.034 | 0.051 | 0.022 | 0.014 | 0.552 | 0.932 | 0.107 | 0.014 | 0.009 | 0.009 | 0.009 | 0.009 |
| GR - JA | 0.022 | 0.043 | 0.215 | 0.948 | 0.396 | 0.103 | 0.005 | 0.016 | 0.103 | 0.215 | 0.215 | 0.215 |
| GR - JE | 0.022 | 0.010 | 0.004 | 0.002 | 0.033 | 0.081 | 0.416 | 0.027 | 0.008 | 0.006 | 0.012 | 0.112 |
| GR - PE | 0.112 | 0.112 | 0.052 | 0.052 | 0.377 | 0.860 | 0.052 | 0.052 | 0.052 | 0.052 | 0.052 | 0.216 |
| GR - SU | 0.219 | 0.219 | 0.108 | 0.108 | 0.219 | 0.925 | 0.637 | 0.299 | 0.156 | 0.219 | 0.156 | 0.156 |
| GT - IN | 0.552 | 0.552 | 0.350 | 0.270 | 0.799 | 0.270 | 0.107 | 0.075 | 0.445 | 0.203 | 0.107 | 0.034 |
| GT - JA | 0.215 | 0.215 | 0.078 | 0.058 | 0.043 | 0.473 | 0.043 | 0.043 | 0.078 | 0.396 | 0.396 | 0.557 |
| GT - JE | 0.201 | 0.230 | 0.727 | 0.027 | 0.002 | 0.002 | 0.002 | 0.002 | 0.561 | 0.261 | 0.615 | 0.261 |
| GT - PE | 0.860 | 0.596 | 0.596 | 0.112 | 0.216 | 0.596 | 0.860 | 0.052 | 0.052 | 0.052 | 0.377 | 0.596 |
| GT - SU | 0.073 | 0.073 | 0.030 | 0.018 | 0.156 | 0.925 | 0.018 | 0.011 | 0.030 | 0.030 | 0.073 | 0.073 |
| IN - JA | 0.901 | 0.901 | 0.091 | 0.019 | 0.023 | 0.023 | 0.063 | 0.967 | 0.063 | 0.063 | 0.063 | 0.052 |
| IN - JE | 0.666 | 0.819 | 0.109 | 0.002 | 0.000 | 0.000 | 0.001 | 0.079 | 0.939 | 0.703 | 0.232 | 0.003 |
| IN - PE | 0.760 | 0.760 | 0.083 | 0.032 | 0.262 | 0.919 | 0.053 | 0.032 | 0.053 | 0.126 | 0.475 | 0.919 |
| IN - SU | 0.043 | 0.018 | 0.009 | 0.009 | 0.056 | 0.772 | 0.032 | 0.009 | 0.005 | 0.005 | 0.007 | 0.005 |
| JA - JE | 0.522 | 0.496 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 | 0.034 | 0.041 | 0.041 | 0.112 | 0.853 |
| JA - PE | 0.876 | 0.640 | 0.062 | 0.020 | 0.029 | 0.087 | 0.029 | 0.020 | 0.043 | 0.043 | 0.161 | 0.276 |
| JA - SU | 0.015 | 0.015 | 0.019 | 0.030 | 0.651 | 0.103 | 0.003 | 0.005 | 0.019 | 0.019 | 0.019 | 0.019 |
| JE - PE | 0.681 | 0.749 | 0.294 | 0.494 | 0.091 | 0.009 | 0.007 | 0.007 | 0.040 | 0.061 | 0.254 | 0.157 |
| JE - SU | 0.016 | 0.005 | 0.001 | 0.000 | 0.000 | 0.002 | 0.115 | 0.009 | 0.001 | 0.001 | 0.002 | 0.010 |
| PE - SU | 0.172 | 0.068 | 0.040 | 0.023 | 0.111 | 1.000 | 0.023 | 0.023 | 0.023 | 0.023 | 0.040 | 0.040 |

APPENDIX B – FEATURES IMPORTANCE AND CONFUSION MATRIXES WHEN USING DIFFERENT FEATURES COMBINATION

Table 1 to Table 33 show the features importance and its confusion matrix for all classification tests performed in Section 5.5. The feature importance of each test was scaled from 0 to 1, where 0 represents the least important feature, and 1 represents the most important feature. In the Tables, each variable represents:

| | |
|--|--|
| B1: spectral band centered at 506.22 nm | HOM: homogeneity |
| B2: spectral band centered at 519.94 nm | NDVI: normalized difference vegetation index |
| B3: spectral band centered at 535.09 nm | PRI: photochemical reflectance index |
| B4: spectral band centered at 550.39 nm | PSRI: plant senescence reflectance index |
| B5: spectral band centered at 565.10 nm | REP: red-edge position |
| B6: spectral band centered at 580.16 nm | SIPI: structure insensitive pigment reflectance index |
| B7: spectral band centered at 591.90 nm | avg: mean height |
| B8: spectral band centered at 609.00 nm | std: standard deviation of height |
| B9: spectral band centered at 620.22 nm | kur: kurtosis of height |
| B10: spectral band centered at 628.73 nm | ske: skewness of height |
| B11: spectral band centered at 650.96 nm | p25, p50, p75, p90: 25 th , 50 th , 75 th , and 90 th percentile of height |
| B12: spectral band centered at 659.72 nm | |
| B13: spectral band centered at 669.75 nm | |
| B14: spectral band centered at 679.84 nm | |
| B15: spectral band centered at 690.28 nm | |
| B16: spectral band centered at 700.28 nm | |
| B17: spectral band centered at 710.06 nm | |
| B18: spectral band centered at 720.17 nm | |
| B19: spectral band centered at 729.57 nm | |
| B20: spectral band centered at 740.42 nm | |
| B21: spectral band centered at 750.16 nm | |
| B22: spectral band centered at 769.89 nm | |
| B23: spectral band centered at 780.49 nm | |
| B24: spectral band centered at 790.30 nm | |
| B25: spectral band centered at 819.66 nm | |
| MNo: mean normalized spectra | |
| ASM: angular second moment | |
| CON: contrast | |
| COR: correlation | |
| DIS: dissimilarity | |

Table 1 - Feature importance when using all vegetation indexes, spectral, textural, and structural features.

| | ASM | | | COM | | | COR | | | DIS | | | HOM | | | Mno | | |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 |
| B1 | 0.890 | 0.716 | 0.695 | 0.810 | 0.752 | 0.737 | 0.823 | 0.715 | 0.795 | 0.752 | 0.750 | 0.735 | 0.752 | 0.675 | 0.558 | 0.715 | 0.701 | 0.697 |
| B2 | 0.860 | 0.727 | 0.852 | 0.776 | 0.757 | 0.671 | 0.791 | 0.741 | 0.714 | 0.777 | 0.705 | 0.675 | 0.801 | 0.974 | 0.550 | 0.697 | 0.620 | 0.726 |
| B3 | 0.909 | 0.829 | 0.636 | 0.805 | 0.800 | 0.617 | 0.887 | 0.728 | 0.479 | 0.738 | 0.654 | 0.646 | 0.773 | 0.690 | 0.768 | 0.759 | 0.662 | 0.665 |
| B4 | 0.882 | 0.879 | 0.574 | 0.759 | 0.718 | 0.626 | 0.775 | 0.789 | 0.599 | 0.834 | 0.706 | 0.734 | 0.854 | 0.678 | 0.803 | 0.792 | 0.667 | 0.664 |
| B5 | 0.888 | 0.639 | 0.903 | 0.804 | 0.741 | 0.690 | 0.882 | 0.714 | 0.681 | 0.821 | 0.776 | 0.584 | 0.824 | 0.758 | 0.000 | 0.758 | 0.655 | 0.679 |
| B6 | 0.938 | 0.763 | 0.747 | 0.814 | 0.687 | 0.733 | 0.807 | 0.785 | 0.665 | 0.796 | 0.670 | 0.669 | 0.696 | 0.798 | 0.414 | 0.731 | 0.643 | 0.682 |
| B7 | 0.925 | 0.770 | 0.829 | 0.824 | 0.817 | 0.620 | 0.804 | 0.749 | 0.674 | 0.758 | 0.661 | 0.640 | 0.793 | 0.813 | 0.829 | 0.774 | 0.734 | 0.616 |
| B8 | 0.841 | 0.728 | 0.587 | 0.803 | 0.741 | 0.671 | 0.811 | 0.665 | 0.626 | 0.816 | 0.727 | 0.665 | 0.806 | 0.833 | 0.768 | 0.765 | 0.683 | 0.784 |
| B9 | 0.866 | 0.835 | 0.648 | 0.808 | 0.749 | 0.708 | 0.809 | 0.736 | 0.627 | 0.797 | 0.797 | 0.698 | 0.671 | 0.673 | 0.676 | 0.758 | 0.672 | 0.727 |
| B10 | 0.839 | 0.799 | 0.754 | 0.811 | 0.752 | 0.649 | 0.807 | 0.742 | 0.747 | 0.740 | 0.767 | 0.648 | 0.893 | 0.608 | 0.479 | 0.742 | 0.727 | 0.662 |
| B11 | 0.882 | 0.772 | 0.728 | 0.828 | 0.833 | 0.678 | 0.731 | 0.736 | 0.759 | 0.773 | 0.784 | 0.647 | 0.837 | 0.568 | 0.567 | 0.672 | 0.705 | 0.767 |
| B12 | 0.884 | 0.733 | 0.605 | 0.849 | 0.763 | 0.676 | 0.785 | 0.801 | 0.561 | 0.839 | 0.627 | 0.665 | 0.802 | 0.611 | 0.839 | 0.733 | 0.698 | 0.672 |
| B13 | 0.837 | 0.679 | 0.734 | 0.832 | 0.735 | 0.705 | 0.849 | 0.777 | 0.650 | 0.795 | 0.678 | 0.603 | 0.725 | 0.734 | 0.657 | 0.775 | 0.658 | 0.613 |
| B14 | 0.870 | 0.702 | 0.689 | 0.877 | 0.700 | 0.689 | 0.827 | 0.840 | 0.916 | 0.769 | 0.714 | 0.693 | 0.746 | 0.658 | 0.545 | 0.748 | 0.679 | 0.674 |
| B15 | 0.874 | 0.688 | 0.873 | 0.893 | 0.843 | 0.765 | 0.853 | 0.692 | 0.688 | 0.805 | 0.688 | 0.799 | 0.795 | 0.646 | 0.540 | 0.705 | 0.690 | 0.688 |
| B16 | 0.898 | 0.699 | 0.718 | 0.843 | 0.656 | 0.622 | 0.841 | 0.780 | 0.298 | 0.828 | 0.685 | 0.649 | 0.803 | 0.697 | 0.711 | 0.738 | 0.736 | 0.622 |
| B17 | 0.939 | 0.723 | 0.606 | 0.831 | 0.760 | 0.793 | 0.788 | 1.000 | 0.522 | 0.760 | 0.726 | 0.745 | 0.818 | 0.756 | 0.985 | 0.742 | 0.713 | 0.692 |
| B18 | 0.864 | 0.694 | 0.818 | 0.829 | 0.778 | 0.708 | 0.824 | 0.727 | 0.620 | 0.828 | 0.721 | 0.596 | 0.760 | 0.683 | 0.739 | 0.776 | 0.733 | 0.681 |
| B19 | 0.869 | 0.887 | 0.717 | 0.770 | 0.685 | 0.691 | 0.841 | 0.773 | 0.799 | 0.767 | 0.749 | 0.738 | 0.824 | 0.646 | 0.662 | 0.755 | 0.717 | 0.728 |
| B20 | 0.916 | 0.719 | 0.766 | 0.819 | 0.745 | 0.622 | 0.837 | 0.750 | 0.630 | 0.777 | 0.720 | 0.742 | 0.838 | 0.700 | 0.790 | 0.679 | 0.697 | 0.735 |
| B21 | 0.850 | 0.638 | 0.771 | 0.751 | 0.761 | 0.724 | 0.805 | 0.657 | 0.586 | 0.692 | 0.730 | 0.655 | 0.837 | 0.798 | 0.652 | 0.766 | 0.729 | 0.638 |
| B22 | 0.913 | 0.769 | 0.740 | 0.768 | 0.699 | 0.672 | 0.760 | 0.725 | 0.622 | 0.755 | 0.644 | 0.699 | 0.801 | 0.794 | 0.925 | 0.721 | 0.602 | 0.698 |
| B23 | 0.893 | 0.672 | 0.930 | 0.783 | 0.731 | 0.761 | 0.737 | 0.695 | 0.676 | 0.753 | 0.711 | 0.661 | 0.806 | 0.704 | 0.813 | 0.712 | 0.570 | 0.633 |
| B24 | 0.844 | 0.633 | 0.658 | 0.814 | 0.664 | 0.822 | 0.814 | 0.773 | 0.643 | 0.741 | 0.690 | 0.678 | 0.782 | 0.939 | 0.541 | 0.701 | 0.680 | 0.691 |
| B25 | 0.883 | 0.602 | 0.549 | 0.806 | 0.734 | 0.615 | 0.763 | 0.800 | 0.713 | 0.834 | 0.796 | 0.720 | 0.919 | 0.843 | 0.000 | 0.755 | 0.631 | 0.666 |

Continuation

| NDVI | | | PRI | | | PSRI | | | REP | | | SIPI | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--|
| 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | |
| 0.727 | 0.719 | 0.651 | 0.634 | 0.601 | 0.725 | 0.774 | 0.668 | 0.694 | 0.656 | 0.738 | 0.660 | 0.631 | 0.663 | 0.692 | |
| avg | | std | | kur | | ske | | p25 | | p50 | | p75 | | p90 | |
| 0.853 | | 0.809 | | 0.876 | | 0.832 | | 0.870 | | 0.841 | | 0.865 | | 0.836 | |

Table 2 - Feature importance when using the angular second moment textural feature

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|
| B1 | 0.867 | 0.429 | 0.232 | B6 | 0.889 | 0.434 | 0.262 | B11 | 0.879 | 0.332 | 0.166 | B16 | 0.885 | 0.304 | 0.238 | B21 | 0.880 | 0.315 | 0.102 |
| B2 | 0.874 | 0.391 | 0.301 | B7 | 0.888 | 0.295 | 0.153 | B12 | 0.941 | 0.492 | 0.311 | B17 | 0.765 | 0.129 | 0.187 | B22 | 0.923 | 0.266 | 0.000 |
| B3 | 0.940 | 0.451 | 0.272 | B8 | 0.856 | 0.487 | 0.214 | B13 | 0.842 | 0.532 | 0.456 | B18 | 0.801 | 0.391 | 0.210 | B23 | 0.940 | 0.141 | 0.085 |
| B4 | 0.899 | 0.427 | 0.213 | B9 | 0.886 | 0.442 | 0.330 | B14 | 0.844 | 0.500 | 0.495 | B19 | 0.951 | 0.429 | 0.168 | B24 | 1.000 | 0.275 | 0.168 |
| B5 | 0.894 | 0.482 | 0.242 | B10 | 0.875 | 0.519 | 0.143 | B15 | 0.869 | 0.432 | 0.481 | B20 | 0.938 | 0.330 | 0.140 | B25 | 0.803 | 0.193 | 0.154 |

Table 3 - Feature importance when using the contrast textural feature

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|
| B1 | 0.852 | 0.761 | 0.339 | B6 | 0.838 | 0.437 | 0.105 | B11 | 1.000 | 0.593 | 0.236 | B16 | 0.827 | 0.348 | 0.214 | B21 | 0.484 | 0.452 | 0.462 |
| B2 | 0.917 | 0.558 | 0.352 | B7 | 0.872 | 0.449 | 0.316 | B12 | 0.921 | 0.449 | 0.335 | B17 | 0.674 | 0.236 | 0.523 | B22 | 0.334 | 0.361 | 0.074 |
| B3 | 0.878 | 0.582 | 0.240 | B8 | 0.963 | 0.492 | 0.274 | B13 | 0.949 | 0.440 | 0.274 | B18 | 0.679 | 0.483 | 0.256 | B23 | 0.479 | 0.325 | 0.000 |
| B4 | 0.841 | 0.699 | 0.227 | B9 | 0.839 | 0.569 | 0.246 | B14 | 0.911 | 0.524 | 0.315 | B19 | 0.602 | 0.319 | 0.300 | B24 | 0.522 | 0.308 | 0.308 |
| B5 | 0.839 | 0.587 | 0.292 | B10 | 0.970 | 0.486 | 0.281 | B15 | 0.939 | 0.604 | 0.218 | B20 | 0.505 | 0.408 | 0.443 | B25 | 0.488 | 0.407 | 0.203 |

Table 4 - Feature importance when using the correlation textural feature

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|
| B1 | 0.938 | 0.675 | 0.318 | B6 | 0.934 | 0.610 | 0.393 | B11 | 0.941 | 0.357 | 0.247 | B16 | 0.844 | 0.430 | 0.131 | B21 | 0.487 | 0.254 | 0.203 |
| B2 | 0.880 | 0.652 | 0.393 | B7 | 0.938 | 0.561 | 0.270 | B12 | 0.817 | 0.587 | 0.200 | B17 | 0.681 | 0.523 | 0.225 | B22 | 0.358 | 0.167 | 0.161 |
| B3 | 1.000 | 0.642 | 0.356 | B8 | 0.907 | 0.472 | 0.171 | B13 | 0.692 | 0.590 | 0.108 | B18 | 0.724 | 0.252 | 0.016 | B23 | 0.449 | 0.102 | 0.119 |
| B4 | 0.928 | 0.601 | 0.275 | B9 | 0.942 | 0.528 | 0.095 | B14 | 0.870 | 0.331 | 0.265 | B19 | 0.807 | 0.205 | 0.236 | B24 | 0.370 | 0.141 | 0.047 |
| B5 | 0.929 | 0.655 | 0.206 | B10 | 0.777 | 0.396 | 0.144 | B15 | 0.917 | 0.213 | 0.191 | B20 | 0.739 | 0.188 | 0.280 | B25 | 0.385 | 0.000 | 0.046 |

Table 5 - Feature importance when using the dissimilarity textural feature

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|
| B1 | 0.947 | 0.630 | 0.175 | B6 | 0.922 | 0.588 | 0.077 | B11 | 0.961 | 0.596 | 0.266 | B16 | 0.675 | 0.279 | 0.234 | B21 | 0.428 | 0.373 | 0.193 |
| B2 | 0.915 | 0.657 | 0.060 | B7 | 0.929 | 0.440 | 0.025 | B12 | 0.887 | 0.432 | 0.325 | B17 | 0.703 | 0.301 | 0.277 | B22 | 0.457 | 0.389 | 0.304 |
| B3 | 0.903 | 0.706 | 0.250 | B8 | 0.949 | 0.434 | 0.091 | B13 | 0.873 | 0.517 | 0.074 | B18 | 0.665 | 0.216 | 0.323 | B23 | 0.349 | 0.365 | 0.128 |
| B4 | 0.912 | 0.660 | 0.216 | B9 | 0.885 | 0.557 | 0.212 | B14 | 1.000 | 0.412 | 0.216 | B19 | 0.712 | 0.429 | 0.332 | B24 | 0.443 | 0.254 | 0.000 |
| B5 | 0.819 | 0.490 | 0.335 | B10 | 0.990 | 0.436 | 0.222 | B15 | 0.992 | 0.302 | 0.136 | B20 | 0.495 | 0.246 | 0.260 | B25 | 0.381 | 0.485 | 0.075 |

Table 6 - Feature importance when using the homogeneity textural feature

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|
| B1 | 0.945 | 0.529 | 0.236 | B6 | 0.944 | 0.384 | 0.200 | B11 | 0.802 | 0.344 | 0.292 | B16 | 0.709 | 0.318 | 0.290 | B21 | 0.741 | 0.473 | 0.077 |
| B2 | 1.000 | 0.478 | 0.303 | B7 | 0.921 | 0.482 | 0.482 | B12 | 0.849 | 0.485 | 0.395 | B17 | 0.740 | 0.323 | 0.244 | B22 | 0.799 | 0.214 | 0.155 |
| B3 | 0.885 | 0.506 | 0.385 | B8 | 0.896 | 0.351 | 0.323 | B13 | 0.825 | 0.360 | 0.304 | B18 | 0.692 | 0.295 | 0.112 | B23 | 0.786 | 0.332 | 0.265 |
| B4 | 0.992 | 0.415 | 0.402 | B9 | 0.878 | 0.390 | 0.255 | B14 | 0.755 | 0.489 | 0.285 | B19 | 0.579 | 0.352 | 0.101 | B24 | 0.593 | 0.345 | 0.152 |
| B5 | 0.915 | 0.407 | 0.487 | B10 | 0.913 | 0.366 | 0.395 | B15 | 0.805 | 0.405 | 0.364 | B20 | 0.797 | 0.130 | 0.165 | B25 | 0.758 | 0.335 | 0.000 |

Table 7 - Feature importance when using the mean normalized spectral features

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|-----|-------|-------|-------|
| B1 | 0.950 | 0.563 | 0.199 | B6 | 0.918 | 0.465 | 0.000 | B11 | 0.691 | 0.391 | 0.491 | B16 | 0.637 | 0.304 | 0.104 | B21 | 0.697 | 0.156 | 0.066 |
| B2 | 0.984 | 0.361 | 0.414 | B7 | 0.928 | 0.377 | 0.111 | B12 | 0.848 | 0.143 | 0.101 | B17 | 0.645 | 0.200 | 0.114 | B22 | 0.521 | 0.012 | 0.200 |
| B3 | 0.890 | 0.427 | 0.101 | B8 | 1.000 | 0.615 | 0.104 | B13 | 0.888 | 0.186 | 0.080 | B18 | 0.709 | 0.063 | 0.175 | B23 | 0.418 | 0.198 | 0.022 |
| B4 | 0.924 | 0.428 | 0.158 | B9 | 0.946 | 0.382 | 0.202 | B14 | 0.807 | 0.342 | 0.032 | B19 | 0.517 | 0.271 | 0.193 | B24 | 0.415 | 0.063 | 0.097 |
| B5 | 0.890 | 0.609 | 0.244 | B10 | 0.948 | 0.256 | 0.300 | B15 | 0.880 | 0.633 | 0.110 | B20 | 0.474 | 0.273 | 0.175 | B25 | 0.481 | 0.123 | 0.398 |

Table 8 - Feature importance when using the vegetation indexes

| NDVI | | | PRI | | | PSRI | | | REP | | | SIPI | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 | 2017 | 2018 | 2019 |
| 1.000 | 0.666 | 0.143 | 0.863 | 0.405 | 0.123 | 0.831 | 0.390 | 0.129 | 0.709 | 0.383 | 0.000 | 0.758 | 0.101 | 0.115 |

Table 9 - Feature importance when using the structural features

| avg | std | kur | ske | p25 | p50 | p75 | p90 |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 1.000 | 0.869 | 0.632 | 0.805 | 0.422 | 0.340 | 0.219 | 0.000 |

Table 10 - Feature importance when using the mean normalized spectra and the vegetation indexes

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|-------|-------|-------|
| B1 | 1.000 | 0.408 | 0.224 | B6 | 0.982 | 0.486 | 0.260 | B11 | 0.785 | 0.437 | 0.469 | B16 | 0.662 | 0.409 | 0.344 | B21 | 0.683 | 0.309 | 0.324 |
| B2 | 0.999 | 0.550 | 0.451 | B7 | 0.966 | 0.482 | 0.237 | B12 | 0.819 | 0.488 | 0.288 | B17 | 0.640 | 0.231 | 0.306 | B22 | 0.654 | 0.178 | 0.013 |
| B3 | 0.863 | 0.503 | 0.105 | B8 | 0.855 | 0.436 | 0.036 | B13 | 0.938 | 0.455 | 0.263 | B18 | 0.574 | 0.192 | 0.145 | B23 | 0.697 | 0.288 | 0.341 |
| B4 | 0.818 | 0.586 | 0.273 | B9 | 0.970 | 0.377 | 0.189 | B14 | 0.851 | 0.577 | 0.000 | B19 | 0.604 | 0.478 | 0.206 | B24 | 0.482 | 0.157 | 0.315 |
| B5 | 0.952 | 0.532 | 0.280 | B10 | 0.883 | 0.497 | 0.175 | B15 | 0.738 | 0.402 | 0.188 | B20 | 0.517 | 0.196 | 0.117 | B25 | 0.703 | 0.122 | 0.287 |
| | NDVI | | | PRI | | | PSRI | | | REP | | | SIPI | | | | | | |
| | 0.260 | 0.343 | 0.003 | 0.340 | 0.069 | 0.711 | 0.048 | 0.148 | 0.556 | 0.267 | 0.445 | 0.320 | 0.176 | 0.109 | 0.410 | | | | |

Table 11 - Feature importance when using the mean normalized spectra, the vegetation indexes and the five most important textural features from the previous classifications

| | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 | | 2017 | 2018 | 2019 |
|------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|---------------|---------------|---------------|-------|-------|-----|-------------|-------|-------|
| B1 | 0.923 | 0.751 | 0.335 | B6 | 0.843 | 0.543 | 0.221 | B11 | 0.868 | 0.422 | 0.262 | B16 | 0.645 | 0.550 | 0.042 | B21 | 0.777 | 0.580 | 0.105 |
| B2 | 0.706 | 0.753 | 0.299 | B7 | 0.984 | 0.499 | 0.313 | B12 | 0.782 | 0.604 | 0.397 | B17 | 0.738 | 0.313 | 0.096 | B22 | 0.559 | 0.179 | 0.000 |
| B3 | 0.849 | 0.626 | 0.211 | B8 | 0.835 | 0.492 | 0.058 | B13 | 0.760 | 0.717 | 0.124 | B18 | 0.715 | 0.334 | 0.143 | B23 | 0.607 | 0.271 | 0.046 |
| B4 | 0.800 | 0.449 | 0.266 | B9 | 0.836 | 0.457 | 0.307 | B14 | 0.835 | 0.584 | 0.109 | B19 | 0.548 | 0.467 | 0.296 | B24 | 0.482 | 0.522 | 0.301 |
| B5 | 0.796 | 0.693 | 0.169 | B10 | 1.000 | 0.730 | 0.373 | B15 | 0.765 | 0.596 | 0.130 | B20 | 0.639 | 0.247 | 0.442 | B25 | 0.681 | 0.309 | 0.140 |
| | NDVI | | | | PRI | | | | PSRI | | | | REP | | | | SIPI | | |
| | 0.397 | 0.490 | 0.041 | 0.244 | 0.157 | 0.207 | 0.099 | 0.274 | 0.416 | 0.329 | 0.405 | 0.006 | 0.326 | 0.274 | 0.199 | | | | |
| | ASM B3 | ASM B12 | ASM B13 | ASM B19 | ASM B24 | CON B8 | CON B1 | CON B11 | CON B13 | CON B15 | COR B1 | COR B3 | COR B7 | | | | | | |
| 2017 | 0.576 | 0.218 | 0.112 | 0.093 | 0.127 | 0.132 | 0.304 | 0.096 | 0.356 | 0.046 | 0.063 | 0.185 | 0.341 | | | | | | |
| | COR B9 | COR B11 | DIS B8 | DIS B1 | DIS B11 | DIS B14 | DIS B15 | HOM B1 | HOM B2 | HOM B4 | HOM B6 | HOM B7 | | | | | | | |
| | 0.421 | 0.269 | 0.266 | 0.360 | 0.295 | 0.425 | 0.007 | 0.749 | 0.193 | 0.197 | 0.207 | 0.163 | | | | | | | |

Table 12 - Confusion matrix when using all vegetation indexes, spectral, textural, and structural features

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|------|----|------|------|----|-----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 1 | 0 | 1 | 2 | 1 | 0 | 0 | 0 |
| | CL | 3 | 7 | 3 | 2 | 3 | 4 | 5 | 0 | 25.9 |
| | EP | 1 | 1 | 3 | 1 | 2 | 1 | 1 | 0 | 30 |
| | HA | 2 | 2 | 1 | 2 | 0 | 0 | 1 | 0 | 25 |
| | HC | 2 | 4 | 0 | 0 | 2 | 1 | 0 | 0 | 22.2 |
| | IV | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 25 |
| | PP | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | SR | 0 | 1 | 0 | 3 | 1 | 0 | 0 | 20 | 80 |
| Producer accuracy (%) | | 0 | 41.2 | 42.9 | 20 | 18.2 | 12.5 | 0 | 100 | Overall accuracy = 38.89% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 13 - Confusion matrix when using the angular second moment textural feature

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|------|----|------|----|------|----|------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 3 | 0 | 0 | 2 | 0 | 0 | 2 | 0 |
| | CL | 1 | 5 | 1 | 0 | 1 | 1 | 1 | 1 | 45.5 |
| | EP | 0 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 40 |
| | HA | 0 | 2 | 0 | 0 | 1 | 1 | 1 | 7 | 0 |
| | HC | 4 | 0 | 1 | 0 | 2 | 2 | 0 | 0 | 22.2 |
| | IV | 1 | 1 | 1 | 1 | 3 | 2 | 0 | 1 | 20 |
| | PP | 0 | 1 | 0 | 2 | 0 | 0 | 1 | 2 | 16.7 |
| | SR | 4 | 3 | 2 | 7 | 2 | 2 | 4 | 6 | 20 |
| Producer accuracy (%) | | 0 | 29.4 | 28.6 | 0 | 18.2 | 25 | 14.3 | 30 | Overall accuracy = 20% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 14 - Confusion matrix when using the contrast textural feature

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|----|------|------|----|------|------|------|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 3 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| | CL | 4 | 8 | 2 | 1 | 6 | 2 | 5 | 2 | 26.7 |
| | EP | 2 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 20 |
| | HA | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 20 |
| | HC | 1 | 3 | 2 | 1 | 0 | 2 | 0 | 1 | 0 |
| | IV | 1 | 2 | 0 | 0 | 1 | 2 | 0 | 0 | 33.3 |
| | PP | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 25 |
| | SR | 1 | 0 | 1 | 3 | 2 | 0 | 0 | 16 | 69.6 |
| Producer accuracy (%) | | 0 | 34 | 16.7 | 13.3 | 0 | 28.6 | 18.2 | 74.4 | Overall accuracy = 32.22% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 15 - Confusion matrix when using the correlation textural feature

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|----|----|----|----|------|----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 2 | 0 |
| | CL | 3 | 7 | 3 | 4 | 4 | 3 | 1 | 3 | 25 |
| | EP | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 |
| | HA | 0 | 1 | 1 | 0 | 1 | 1 | 2 | 3 | 0 |
| | HC | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | IV | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 1 | 0 |
| | PP | 1 | 1 | 1 | 0 | 1 | 2 | 1 | 0 | 14.3 |
| | SR | 5 | 4 | 2 | 5 | 4 | 2 | 0 | 9 | 29 |
| Producer accuracy (%) | | 0 | 41.2 | 0 | 0 | 0 | 0 | 14.3 | 45 | Overall accuracy = 18.89% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helieta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 16 - Confusion matrix when using the dissimilarity textural feature

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|----|----|----|------|----|----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 1 | 0 | 0 | 1 | 2 | 1 | 0 | 0 |
| | CL | 5 | 7 | 1 | 2 | 6 | 3 | 5 | 4 | 21.2 |
| | EP | 1 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | 0 |
| | HA | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 25 |
| | HC | 1 | 4 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| | IV | 1 | 2 | 1 | 0 | 1 | 1 | 0 | 0 | 16.7 |
| | PP | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | SR | 1 | 2 | 3 | 3 | 3 | 0 | 0 | 14 | 53.8 |
| Producer accuracy (%) | | 0 | 41.2 | 0 | 10 | 0 | 12.5 | 0 | 70 | Overall accuracy = 25.56% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helieta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 17 - Confusion matrix when using the homogeneity textural feature

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|-----|------|----|-----|------|------|----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 1 | 0 | 0 | 3 | 0 | 1 | 2 | 0 |
| | CL | 1 | 1 | 2 | 4 | 0 | 3 | 1 | 1 | 7.7 |
| | EP | 0 | 2 | 2 | 0 | 1 | 0 | 0 | 0 | 40 |
| | HA | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 3 | 0 |
| | HC | 4 | 0 | 1 | 1 | 1 | 2 | 1 | 2 | 8.3 |
| | IV | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 2 | 14.3 |
| | PP | 0 | 1 | 0 | 1 | 1 | 0 | 2 | 0 | 40 |
| | SR | 5 | 8 | 1 | 4 | 4 | 1 | 2 | 10 | 28.6 |
| Producer accuracy (%) | | 0 | 5.9 | 28.6 | 0 | 9.1 | 12.5 | 28.6 | 50 | Overall accuracy = 18.89% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helieta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 18 - Confusion matrix when using the mean normalized spectral features

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|------|----|------|------|----|----|--------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 20 |
| | CL | 2 | 9 | 2 | 2 | 4 | 3 | 2 | 1 | 36 |
| | EP | 0 | 1 | 3 | 0 | 1 | 1 | 0 | 0 | 50 |
| | HA | 2 | 0 | 0 | 3 | 0 | 1 | 2 | 0 | 37.5 |
| | HC | 2 | 2 | 2 | 0 | 4 | 0 | 1 | 1 | 33.3 |
| | IV | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 25 |
| | PP | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | SR | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 18 | 69.2 |
| Producer accuracy (%) | | 10 | 52.9 | 42.9 | 30 | 36.4 | 12.5 | 0 | 90 | Overall accuracy = 43.3% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 19 - Confusion matrix when using the vegetation indexes

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|------|----|------|------|------|----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | CL | 6 | 7 | 2 | 6 | 4 | 1 | 1 | 0 | 25.9 |
| | EP | 1 | 0 | 4 | 0 | 1 | 0 | 0 | 0 | 66.7 |
| | HA | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 33.3 |
| | HC | 0 | 5 | 1 | 1 | 4 | 0 | 1 | 0 | 33.3 |
| | IV | 1 | 0 | 0 | 0 | 1 | 5 | 0 | 2 | 55.6 |
| | PP | 1 | 0 | 0 | 1 | 0 | 0 | 4 | 0 | 66.7 |
| | SR | 1 | 0 | 0 | 1 | 1 | 2 | 1 | 18 | 75 |
| Producer accuracy (%) | | 0 | 41.2 | 57.1 | 10 | 36.4 | 62.5 | 57.1 | 90 | Overall accuracy = 47.78% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 20 - Confusion matrix when using the structural features

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|----|----|------|----|------|----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 1 | 0 | 1 | 0 | 2 | 0 | 1 | 0 | 20 |
| | CL | 0 | 5 | 2 | 3 | 1 | 2 | 0 | 3 | 31.3 |
| | EP | 3 | 1 | 0 | 0 | 0 | 5 | 0 | 1 | 0 |
| | HA | 1 | 2 | 0 | 2 | 2 | 0 | 0 | 3 | 20 |
| | HC | 2 | 2 | 0 | 3 | 3 | 0 | 0 | 1 | 27.3 |
| | IV | 1 | 2 | 3 | 0 | 0 | 0 | 1 | 1 | 0 |
| | PP | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 50 |
| | SR | 1 | 4 | 1 | 2 | 3 | 1 | 3 | 11 | 42.3 |
| Producer accuracy (%) | | 10 | 29.4 | 0 | 20 | 27.3 | 0 | 28.6 | 55 | Overall accuracy = 26.67% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 21 - Confusion matrix when using the mean normalized spectra and the vegetation indexes

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|------|----|------|------|----|----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| | CL | 4 | 8 | 2 | 1 | 4 | 3 | 2 | 0 | 33.3 |
| | EP | 0 | 1 | 3 | 0 | 1 | 1 | 0 | 0 | 50 |
| | HA | 0 | 2 | 0 | 5 | 0 | 1 | 2 | 0 | 50 |
| | HC | 3 | 3 | 1 | 0 | 4 | 0 | 1 | 1 | 30.8 |
| | IV | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 16.7 |
| | PP | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | SR | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 18 | 72 |
| Producer accuracy (%) | | 0 | 47.1 | 42.9 | 50 | 36.4 | 12.5 | 0 | 90 | Overall accuracy = 43.33% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.

Table 22 - Confusion matrix when using the mean normalized spectra, the vegetation indexes and the five most important textural features from the previous classifications

| | | Reference ¹ | | | | | | | | User Accuracy (%) |
|-----------------------|----|------------------------|------|------|----|------|------|----|-----|---------------------------|
| | | AL | CL | EP | HA | HC | IV | PP | SR | |
| Classified as | AL | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| | CL | 4 | 8 | 2 | 1 | 4 | 4 | 1 | 0 | 33.3 |
| | EP | 0 | 1 | 3 | 0 | 2 | 1 | 0 | 0 | 42.9 |
| | HA | 1 | 2 | 0 | 4 | 0 | 1 | 3 | 0 | 36.4 |
| | HC | 2 | 4 | 2 | 0 | 3 | 0 | 2 | 0 | 23.1 |
| | IV | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 25 |
| | PP | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | SR | 1 | 2 | 0 | 1 | 1 | 1 | 0 | 20 | 76.9 |
| Producer accuracy (%) | | 0 | 47.1 | 42.9 | 20 | 27.3 | 12.5 | 0 | 100 | Overall accuracy = 43.33% |

AL: *Apuleia leiocarpa*; CL: *Copaifera langsdorffii*; EP: *Endlicheria paniculata*; HA: *Helietta apiculata*; HC: *Hymenaea courbaril*; IV: *Inga vera*; PP: *Pterodon pubescens*; SR: *Syagrus romanzoffiana*.