



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21d/2021/08.12.23.51-TDI

## MULTISENSOR OPTICAL-SAR APPROACH TO LAND USE AND LAND COVER CHARACTERIZATION IN RORAIMA

Victor Hugo Rohden Prudente

Doctorate Thesis of the Graduate  
Course in remote Sensing, guided  
by Drs. Ieda Del'Arco Sanches, and  
Marcos Adami, approved in August  
18, 2021.

URL of the original document:

<<http://urlib.net/8JMKD3MGP3W34T/458TDR2>>

INPE  
São José dos Campos  
2021

**PUBLISHED BY:**

Instituto Nacional de Pesquisas Espaciais - INPE  
Coordenação de Ensino, Pesquisa e Extensão (COEPE)  
Divisão de Biblioteca (DIBIB)  
CEP 12.227-010  
São José dos Campos - SP - Brasil  
Tel.:(012) 3208-6923/7348  
E-mail: pubtc@inpe.br

**BOARD OF PUBLISHING AND PRESERVATION OF INPE  
INTELLECTUAL PRODUCTION - CEPPII (PORTARIA N°  
176/2018/SEI-INPE):****Chairperson:**

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra  
(CGCT)

**Members:**

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)  
Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e  
Ciência Espaciais (CGCE)  
Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e  
Pesquisas Aplicadas (CGIP)  
Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

**DIGITAL LIBRARY:**

Dr. Gerald Jean Francis Banon  
Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

**DOCUMENT REVIEW:**

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)  
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

**ELECTRONIC EDITING:**

Ivone Martins - Divisão de Biblioteca (DIBIB)  
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21d/2021/08.12.23.51-TDI

## MULTISENSOR OPTICAL-SAR APPROACH TO LAND USE AND LAND COVER CHARACTERIZATION IN RORAIMA

Victor Hugo Rohden Prudente

Doctorate Thesis of the Graduate  
Course in remote Sensing, guided  
by Drs. Ieda Del'Arco Sanches, and  
Marcos Adami, approved in August  
18, 2021.

URL of the original document:

<<http://urlib.net/8JMKD3MGP3W34T/458TDR2>>

INPE  
São José dos Campos  
2021

Cataloging in Publication Data

---

Prudente, Victor Hugo Rohden.

P951m Multisensor optical-sar approach to land use and land cover  
characterization in Roraima / Victor Hugo Rohden Prudente. –  
São José dos Campos : INPE, 2021.

xxv + 144 p. ; (sid.inpe.br/mtc-m21d/2021/08.12.23.51-TDI)

Thesis (Doctorate in Remote Sensing) – Instituto Nacional de  
Pesquisas Espaciais, São José dos Campos, 2021.

Guiding : Drs. Ieda Del'Arco Sanches, and Marcos Adami.

1. Microwave. 2. Cloud cover. 3. Sampling. 4. Tropical areas.  
I.Title.

CDU 528.8:551.576(811.4)

---



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**  
Serviço de Pós-Graduação - SEPGR**DEFESA FINAL DE TESE DE VICTOR HUGO ROHDEN PRUDENTE**  
**BANCA Nº 211/2021, REG 142417/2017**

No dia 18 de agosto de 2021, às 09h00min, por teleconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Doutor em Sensoriamento Remoto. O trabalho precisa da incorporação das correções sugeridas pela Banca Examinadora e revisão final pelo(s) orientador(es).

**Título: “MULTISENSOR OPTICAL-SAR APPROACH TO LAND USE AND LAND COVER CHARACTERIZATION IN RORAIMA”**

**Observações da banca:** As modificações sugeridas pela banca serão implementadas sob a supervisão dos orientadores.

**Membros da banca:**

Dr. Lênio Galvão - Presidente - INPE  
Dra. Ieda Del'Arco Sanches - Orientadora - INPE  
Dr. Marcos Adami - Orientador - INPE  
Dr. Edson Eyji Sano - Membro Externo - Embrapa Cerrado  
Dr. Maristela Ramalho Xaud - Membro Externo - Embrapa



Documento assinado eletronicamente por **Marcos Adami, Pesquisador**, em 19/08/2021, às 14:43 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ieda Del Arco Sanches, Pesquisadora**, em 19/08/2021, às 14:53 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **EDSON EYJI SANO (E), Usuário Externo**, em 19/08/2021, às 15:08 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lênio Soares Galvão, Pesquisador**, em 22/08/2021, às 21:46 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

Documento assinado eletronicamente por **maristela ramalho xaud (E), Usuário Externo**, em



27/08/2021, às 09:41 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site <http://sei.mctic.gov.br/verifica.html>, informando o código verificador **7976863** e o código CRC **D62DD422**.

Referência: Processo nº 01340.005225/2021-05

SEI nº 7976863

*If I have seen further, it is by standing on the shoulders of Giants.*

**Isaac Newton**



*To my parents,  
Marilene and José Adelmo,*



## ACKNOWLEDGMENTS

First, to God for my life. To my family for all the support and incentive during my doctorate. To my parents, Marilene Prudente and José Adelmo Rohden, for incentive me to study all these years. To my sisters, Eulalia and Gilene, and niece Milena, for all the talks and love. To my girlfriend, Eloisa, for all the talks, supports, patience, and help during this doctorate time.

To Dr. Erivelto Mercante, Dr. Jerry Johann, and Dr. Suzana Costa Wrublack, for supervising me in my first steps in this journey. To all GeoLab team (many of them I cannot name all of you here), from the seven years of fellowship, ideas, and talks. To my friends, Jows, for all the support, in special to Willyan Becker.

To Dr. Ieda Sanches, for accepting me in this Doctorate and for all the advices, orientation, support, and help during the period. To Dr. Marcos Adami, for the ideas, the study area proposal, advice, help with the methodology, and all orientation and support. To Dr. Sergii Skakun (the University of Maryland – UMD), for accepting me for the internship program and for all the support and advice during my stay in the USA.

To Embrapa Roraima researchers, in especial to Dr. Haron Xaud and Dr. Maristela Xaud, for the help, and support through TERRAMZ (*Conhecimento Compartilhado para Gestão Territorial Local na Amazônia*) project, during the field mission. Dr. Haron and Dr. Maristela, thank you for teaching me all the about Roraima.

To the Services Coordinated Interface (SCI) team from the European Spatial Agency (ESA) Copernicus, for helping with the Sentinel-1B orbit change, starting to cover the entire Roraima state since September 2018.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the

scholarship, Financial Code 140175/2018-0. Also, to CAPES – PrInt for financed the internship at UMD/USA for one year.

To the National Institute for Space Research (INPE) for the great support. In special, to the Remote Sensing Division staff, for all teachings, talks, coffees, and help. I cannot forget the cleaning ladies, thank you for all your efforts during this time.

To the Department of Geographical Science, from the University of Maryland in the United State of America, for all the support and help during my internship there.

To all my friends at INPE and UMD/USA, for the friendship, laughing, ideas, talks, and adventures. In special, to all my colleges from 2017, friends of room 04/SERE II, and Lucas, Denis, Andeise, and Nildson. Special thanks to Sally, Mohammed, André, and Juliana, from the USA. To all roommates from Beth's house in the USA, thank you very much.

And thanks to many other friends and colleagues that would be impossible to mention here.



## ABSTRACT

Earth Observation data has an important role in the worldwide Land Use and Land Cover (LULC) mapping process. However, update LULC maps are not fully available to large areas, as in Brazil. The monitoring and mapping of LULC with Optical Remote Sensing (ORS) data, especially in agricultural areas, is highly affected by cloud cover frequency (CCF), mainly in the rainy season. Our first goal was to evaluate the effects of clouds interference in ORS data for agricultural areas in South America, highlighting Roraima state which was selected as the study area for LULC analyses. During the quarter from September to November (P1) and December and February (P2), South America croplands have the higher CCF concentration. These patterns make summer crop monitoring via ORS data very challenging. In Roraima state the cloud-cover is frequent over the entire year, with some increase during May to August, corresponding to the state agriculture period. Few ORS clear observations, limit even the LULC classification process. Besides, existing LULC programs do not cover the entire state yet or do not provide a separation among important classes. For this reason, in our second study, we investigate the possible improvements when SAR (Synthetic Aperture Radar) data is incorporated into the LULC classification process along with ORS data for 2019 in Roraima state. Twenty-nine scenarios, involving a combination of optical (MSI/Sentinel-2) and SAR (SAR/Sentinel-1) based features, as well as times of data acquisition, were considered in this study. Our results show that optical or SAR data used individually are not enough to provide accurate LULC mapping. The best results in terms of overall accuracy (OA) were achieved using metrics of multi-temporal surface reflectance and vegetation index (VI) for optical imagery, and values of backscatter coefficient in different polarizations and their ratios, yielding an OA of  $86.41 \pm 1.74\%$ . Analysis of three periods of data (January to April, May to August, and September to December) used for classification allowed us to identify the optimal period for distinguishing specific classes. Comparing our LULC map with that of MapBiomas, the only LULC map available for 2019 for the whole state, we observed that our method performed better to map annual and perennial crops and water classes. Our methodology provided a more accurate LULC for the Roraima state, and the technique, with some adjustments, was applied to the entire state. In this sense, our last study focuses on comparing different approaches to classify the LULC for Roraima in 2017, 2018, and 2019. We used a two-stage sampling approach, along with temporal metrics, as an alternative to decreasing the data volume and the time-consuming from the traditional roadside LULC classification process. Two-stage sampling process shown potential to be used in LULC area estimation, with similar accuracies to the roadside approach. Analyzing the LULC changes (LULCC), we observed a few changes between 2017 and 2019. The more traditional change was from Forest or Savannas to Pasture, and from Acacias to Annual crops. In general, our study showed that using SAR with ORS data it is possible improvement the LULC classification process to Roraima, bringing information about where and when is happening the changes.

Keywords: microwave, cloud cover, sampling, tropical areas.

# ABORDAGEM MULTISENSOR SAR-ÓPTICA PARA CARACTERIZAR O USO E COBERTURA DA TERRA EM RORAIMA

## RESUMO

Os dados de observação da Terra têm um papel importante no processo de mapeamento mundial do uso e cobertura da terra (LULC). No entanto, os mapas atualizados de LULC não estão totalmente disponíveis em grandes escalas, como no Brasil. O monitoramento e mapeamento de LULC com dados de sensoriamento remoto óptico, especialmente em áreas agrícolas, é altamente afetado pela frequência de cobertura de nuvens, principalmente na estação chuvosa. Nosso primeiro objetivo foi avaliar os efeitos da interferência de nuvens em dados ópticos para áreas agrícolas na América do Sul, em especial, para o estado de Roraima. Nos trimestres de setembro a novembro e de dezembro a fevereiro, as áreas agrícolas da América do Sul têm as maiores concentrações de cobertura de nuvens, tornando desafiador o monitoramento da safra de verão por meio de dados ópticos. Em Roraima, a cobertura de nuvens é frequente durante todo o ano, com maior intensidade durante os meses de maio a agosto que corresponde ao período agrícola do estado. Poucas observações ópticas livres de nuvens limitam o processo de classificação de LULC neste estado. Além disso, os programas de mapeamento e monitoramento de LULC existentes ainda não cobrem todo o estado ou não fornecem uma separação entre classes importantes. Por esse motivo, em nosso segundo estudo, investigamos as possíveis melhorias ao incorporar dados de radar de abertura sintética (SAR) aos dados ópticos no processo de classificação de LULC em Roraima. Foram criados 29 cenários envolvendo dados ópticos do satélite Sentinel-2 MSI e de radar do satélite Sentinel-1 SAR. Nossos resultados mostraram que os dados ópticos e SAR usados individualmente não são suficientes para fornecer um mapeamento preciso de LULC. Os melhores resultados em termos de exatidão global foram alcançados usando métricas de reflectância de superfície multitemporal e índice de vegetação para imagens ópticas e valores de coeficiente de retroespalhamento em diferentes polarizações e razões, resultando em uma exatidão global de  $86,41 \pm 1,74\%$ . A análise de três períodos de dados (janeiro a abril, maio a agosto e setembro a dezembro) permitiu identificar o período ideal para distinguir classes específicas. Comparando nosso mapa de LULC com o do MapBiomass, o único projeto com dados de LULC disponíveis para 2019 para todo o estado, observamos que nosso método teve melhor desempenho para mapear culturas anuais e perenes e classes de água. Nossa metodologia apresentou resultados melhores para o mapeamento de LULC para Roraima. Posteriormente, com alguns ajustes, essa técnica foi aplicada para todo o estado. Nesse sentido, nosso último estudo se concentrou na comparação de diferentes abordagens para classificar classes de LULC de

Roraima em 2017, 2018 e 2019. Usamos uma abordagem de amostragem em dois estágios, juntamente com métricas temporais, como alternativa para diminuir o volume de dados e a demora do processo de classificação tradicional de LULC, que considera o estado como um todo. O processo de amostragem de dois estágios mostrou potencial para ser usado na estimativa de área ocupada por diferentes classes de LULC, com precisões semelhantes à abordagem tradicional. Entre 2017 e 2019, as mudanças mais importantes foram de floresta ou savanas para pastagem, e de formações de Acácias para culturas anuais. De maneira geral, nosso estudo mostrou que, combinando dados SAR com dados ópticos, é possível aprimorar o processo de classificação de LULC para Roraima, trazendo informações sobre, onde e quando estão ocorrendo as mudanças de LULC.

Palavras-chave: micro-ondas, cobertura de nuvens, amostragem, áreas tropicais.

## LIST OF FIGURES

	<u>Pag.</u>
Figure 2.1 – Distance from each case to every to every case, converted into a probability by fitting a normal distribution over the current case. ....	8
Figure 2.2 – The scaled probabilities for each case are stored as a matrix of values.....	9
Figure 2.3 – t-SNE Student’s t distribution fit over the current point instead of a normal distribution. ....	9
Figure 2.4 – Cases are randomly initialized over the new axes (one axis is shown here).....	10
Figure 3.1 – Location of Roraima study state and its 15 municipalities. ....	12
Figure 3.2 – Roraima natural formations according to IBGE. ....	13
Figure 3.3 – Roraima elevation data according to SRTM data. ....	14
Figure 3.4 – Roraima’s soil map.....	15
Figure 3.5 – Area of Annual crops (upper) sowed and Perennial crops (lower) harvested in Roraima state, from 1988 to 2019, in thousands (K) of hectares. ....	17
Figure 3.6 – Roraima’s land use and land cover (LULC) mosaic based on the IBGE initiative of 2018.....	18
Figure 3.7 – Roraima LULC map, Terraclass Amazonia for 2014.....	19
Figure 3.8 – MapBiomias v.5 LULC map for Roraima region for 2019.....	20
Figure 4.1 – Flowchart with the steps of data processing and papers division of this research.....	22
Figure 4.2 – Field data polygons are drawn after collecting roadside data in August and September of 2019 in Roraima. ....	23
Figure 4.3 – Panoramic field photos of land use and land cover classes in the Roraima State obtained during the field campaign in 2019. ....	25
Figure 4.4 – Sentinel-1 (considering 2018 and 2019) and Sentinel-2 (2019) orbits over the Roraima state. ....	27
Figure 4.5 – Spectral and spatial resolutions of the Multispectral Instrument (MSI) sensor on board of the Sentinel-2 satellite. ....	28
Figure 5.1 – Flowchart showing the four methodological steps: process to create the cloud cover frequency (CCF) for South America (SA).....	33
Figure 5.2 – Study area. The colored states represent states/provinces/ departments with more than1% of the South American cropland in 2015 (FAOSTAT, 2018). ....	35
Figure 5.3 – Location of Roraima state locations in SA, including the Sentinel-2 MSI tiles over the state. ....	36
Figure 5.4 – Cloud cover frequency considering MAIAC and MODIS. ....	38

Figure 5.5 – Crop types distribution based on USGS and FAOSTAT data. Symbols: adapted from GEOGLAM. ....	41
Figure 5.6 – Crop Calendar for the main crops cultivated in SA. Symbols: adapted from GEOGLAM. ....	42
Figure 5.7 – CCF classes over the entire SA for four different periods. ....	45
Figure 5.8 – Spatial and temporal CCF for croplands in each period. ....	47
Figure 5.9 – Detail about the spatial and temporal CCF distribution over the Brazilian states croplands. ....	48
Figure 5.10 – Detail about the spatial and temporal CCF distribution over Ecuador, Colombia, and Venezuela croplands. ....	49
Figure 5.11 – CCF interference for cropland monitoring with ORS data for each period, for Argentina provinces. ....	51
Figure 5.12 – Cloud cover frequency considering MAIAC and MODIS for Roraima state. ....	55
Figure 5.13 – QA60 Cloud cover frequency considering Sentinel-2 MSI optical data over three different periods of 2019 (QP1 – Jan-Apr, QP2 – May-Aug, QP3 – Sep-Dec). ....	56
Figure 5.14 – QA60 Cloud cover frequency considering Sentinel-2 MSI optical data over three different periods of 2017 (QP1 – Jan-Apr, QP2 – May-Aug, QP3 – Sep-Dec). ....	57
Figure 5.15 – Sentinel-2/MSI optical cloud-free cover observation data for 2017 (left map) and 2019 (right map). ....	58
Figure 5.16 – Sentinel-2/MSI optical cloud-free (top images) and Sentinel-1 (bottom images) data observations over 2017 (left images) and 2019(right images). ....	59
Figure 6.1 - Flowchart showing the processing steps. ....	68
Figure 6.2 – Roraima study area, with natural formations (upper-left) and fieldwork detail (down-left). ....	70
Figure 6.3 – Cloud cover frequency map (a) and graph (b) over the study area for MSI data, using quality data from google earth engine considering three periods of 2019 (January to April, May to August, and September to December). ....	72
Figure 6.4 – t-SNE results for each one of the 29 scenarios. ....	77
Figure 6.5 – t-SNE graph for the scenario with the best separation (D3P5). ....	79
Figure 6.6 – Overall Accuracy (OA) of Random Forest (RF) and Multi-layer Perceptron (MLP) classifiers. ....	80
Figure 6.7 – Users (UA) and Producers (PA) accuracies for each class in each scenario considering the MLP classifier. ....	81
Figure 6.8 – Users (UA) and Producers (PA) accuracies for each class in each scenario considering the RF classifier. ....	82

Figure 6.9 – Classification results over the 30 scenarios using MLP classifiers for 2019. ....	84
Figure 6.10 – Classification results over the 30 scenarios using RF classifier . 85	
Figure 6.11 – Best LULC classification scenario (D3P5) map, using the MLP classifier (left), and a 2019 Sentinel-2 RGB True color image, using mean of the band values over the year (right). ....	86
Figure 6.12 – Confusion matrix in terms of validation pixels from polygons derived through field campaign for the best LULC classification scenario (D3P5 – MLP), using field data. ....	87
Figure 6.13 – Confusion matrix in terms of area percentage (OLOFSSON et al., 2014) for the best LULC classification scenario (D3P5 – MLP), using Stratified Random Points. ....	88
Figure 6.14 – Comparison among accuracies from field data and sampling approach. ....	88
Figure 6.15 – Land Use and Land Cover mapping provided by the MapBiomass v5.0 – 2019. ....	89
Figure 6.16 – Map of the differences between our LULC mapping and the LULC provide for MapBiomass v5.0 – 2019 (SOUZA et al., 2020). ....	90
Figure 6.17 – Comparison between our result and the MapBiomass v.5 LULC map. ....	95
Figure 7.1 – Details about the field data mission using a two-stage sampling approach. ....	102
Figure 7.2 – LULC classification map (a) and confusion matrix (b) for 2017 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	105
Figure 7.3 – LULC classification map (a) and confusion matrix (b) for 2018 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	106
Figure 7.4 – LULC classification map (a) and confusion matrix (b) for 2019 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	107
Figure 7.5 – Users (UA) and Producers (PA) Accuracies for 2017 (a), 2018 (b), and 2019 (c) LULC classification, using testing and multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	108
Figure 7.6 – LULC classification (a) and Users (UA) and Producers (PA) Accuracies (b) for the area of the blocks, for 2017, based on multisensor Sentinel-1 and Sentinel-2 data processed in GEE platform using RF classifier. ....	110
Figure 7.7 – LULC classification (a) and Users (UA) and Producers (PA) Accuracies (b) result from blocks to 2018 with multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	111
Figure 7.8 – LULC classification (a) and Users (UA) and Producers (PA) Accuracies (b) result from blocks to 2019 with multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	112

Figure 7.9 – Sample approach LULC classification map (a) and confusion matrix (b) for 2017 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	114
Figure 7.10 – Sample approach LULC classification map (a) and confusion matrix (b) for 2018 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	115
Figure 7.11 – Sample approach LULC classification map (a) and confusion matrix (b) for 2019 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	116
Figure 7.12 – Users (UA) and Producers (PA) Accuracies for 2017 (a), 2018 (b), and 2019 (c) considering the blocks to the LULC classification, using testing and multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier. ....	117
Figure 7.13 – Sankey graph to represent the LULCC based on the roadside field data. ....	118
Figure 7.14 – LULCC classification from 2017 to 2019, using the roadside data. Change details on the right.....	119
Figure 7.15 – Details of LULCC classifications and the two-stage sampling LULC classification from 2017 to 2019. ....	120



## LIST OF TABLES

### Pag.

Table 4.1. Summary of ORS data use feasibility: very likely (+), likely ( $\pm$ ), and unlikely (-) considering the main crops in the main producer states of Brazil and Argentina: Regions, where the crop is inexpressive, are blank. ....	63
Table 5.1. Details about the 29 scenarios formed with different datasets and periods. ....	74
Table 7.1. Area (in km <sup>2</sup> ) for each Roraima LULC class, for 2017, 2018, and 2019, using the roadside data. ....	104
Table 7.2. Estimate area (in km <sup>2</sup> ) for each LULC class for 2017 and 2019 considering the points from the block approach. SE: Standard Error. ....	109
Table 7.3 – Estimate area (in km <sup>2</sup> ) for each LULC class for 2017, 2018, and 2019, after the classification of each block. SE: Standard Error. ....	113
Table 7.4. Area (in km <sup>2</sup> ) for each LULC class for 2017, 2018, and 2019 after the classification in each block. ....	113
Table 7.5. Area (in km <sup>2</sup> ) from Municipal Agricultural Production (PAM) data from IBGE and our roadside and two-stage sampling-based classifications for the Annual Crops class. ....	122



## LIST OF ABBREVIATIONS

<i>CCF</i>	Cloud Cover Frequency
<i>ESA</i>	European Space Agency
<i>ETM+</i>	Enhanced Thematic Mapper Plus
<i>FAO</i>	Food and Agriculture Organization
<i>GEE</i>	Google Earth Engine
<i>GEO</i>	Group on Earth Observation
<i>IBGE</i>	Brazilian Institute of Geography and Statistics
<i>INPE</i>	National Institute for Space Research
<i>ITCZ</i>	Intertropical Convergence Zone
<i>IW</i>	Interferometric Wide
<i>JECAM</i>	Joint Experiment for Crop Assessment and Monitoring
<i>LSWI</i>	Land Surface Water Index
<i>LULC</i>	Land Use and Land Cover
<i>LULCC</i>	Land Use and Land Cover Change
<i>MLP</i>	Multi-Layer Perception
<i>MODIS</i>	Moderate Resolution Imaging Spectroradiometer
<i>MSI</i>	Multispectral Imager
<i>MUX</i>	Multispectral Camera
<i>NASA</i>	National Aeronautics and Space Administration
<i>NDVI</i>	Normalized Difference Vegetation Index
<i>NIR</i>	Near Infrared
<i>OA</i>	Overall Accuracy
<i>OLI</i>	Operational Land Imager

<i>ORS</i>	Optical Remote Sensing
<i>PA</i>	Producer's Accuracy
<i>PAM</i>	Municipal Agricultural Production
<i>PRODES</i>	Deforestation Monitoring in the Brazilian Legal Amazon
<i>RF</i>	Random Forest
<i>RS</i>	Remote sensing
<i>SA</i>	South America
<i>SACS</i>	South Atlantic Convergence Zone
<i>SAR</i>	Synthetic Aperture Radar
<i>SRP</i>	Stratified Random Points
<i>SWIR</i>	Shortwave Infrared
<i>TM</i>	Thematic Mapper
<i>TOA</i>	Top of the Atmosphere
<i>t-SNE</i>	t-distributed Stochastic Neighbor Embedding
<i>UA</i>	User's Accuracy
<i>USGS</i>	United States Geological Survey
<i>VI</i>	Vegetation Index

# CONTENTS

	<u>Pág.</u>
<b>1 INTRODUCTION.....</b>	<b>1</b>
1.1 Objectives.....	2
1.1.1 Specific objectives .....	2
<b>2 THEORETICAL BACKGROUND .....</b>	<b>3</b>
2.1 LULC monitoring.....	3
2.2 Multisensor approach to overcome the limitations of optical remote sensing .....	5
2.3 LULC Classification process.....	6
2.3.1 Classifiers .....	7
2.3.2 Data reducers .....	7
2.3.3 Two-stage sampling.....	10
<b>3 RORAIMA STATE .....</b>	<b>12</b>
3.1 Agriculture expansion.....	16
3.2 Existing LULC mapping programs .....	17
<b>4 MATERIAL AND METHODS .....</b>	<b>21</b>
4.1 Thesis organization.....	21
4.2 Field data .....	22
4.2.1 LULC classes.....	24
4.3 Remote sensing data.....	27
4.3.1 Sentinel-2.....	27
4.3.2 Sentinel-1.....	29
<b>5 IMPACTS OF CLOUD COVER OVER SOUTH AMERICA OPTICAL REMOTE SENSING MONITORING – CROPLANDS AND RORAIMA CASE</b>	<b>30</b>
5.1 Introduction.....	30
5.2 Material and methods.....	32
5.2.1 Study area .....	33
5.2.2 Cloud cover.....	36
5.2.3 Cropland information.....	38
5.2.4 MODIS cloud cover frequency .....	42
5.2.5 QA60 cloud cover frequency.....	43

5.3	Results .....	44
5.3.1	Cloud cover frequency – South America.....	44
5.3.2	Cloud cover frequency – South America croplands .....	46
5.3.3	Cloud cover frequency – Roraima case .....	54
5.4	Discussion .....	59
5.5	Conclusion.....	64
<b>6</b>	<b>MULTISENSOR APPROACH TO LAND USE AND LAND COVER MAPPING IN NORTHERN AMAZON, BRAZIL .....</b>	<b>66</b>
6.1	Introduction.....	66
6.2	Material and methods.....	68
6.2.1	Study area .....	69
6.2.2	Remote sensing data .....	70
6.2.3	Scenarios.....	73
6.2.4	t-Distributed Stochastic Neighbor Embedding (t-SNE).....	74
6.2.5	LULC classification scenarios .....	75
6.3	Results .....	75
6.3.1	Visualization of input features using t-SNE .....	76
6.3.2	Classification.....	79
6.4	Discussion .....	90
6.5	Conclusion.....	96
<b>7</b>	<b>GOOGLE EARTH ENGINE IN A TWO-STAGE SAMPLING MULTISENSOR APPROACH TO ESTIMATE THE LAND USE AND LAND COVER CLASSES OVER RORAIMA STATE.....</b>	<b>98</b>
7.1	Introduction.....	98
7.2	Material and methods.....	99
7.2.1	Study area .....	99
7.2.2	Data .....	99
7.2.3	LULC classification .....	100
7.2.4	Determination of LULC changes .....	103
7.3	Results .....	104
7.3.1	Roadside LULC classification .....	104
7.3.2	Sample-based LULC classification .....	109
7.3.3	LULC changes .....	118
7.4	Discussion .....	120
7.5	Conclusion.....	123

<b>8</b>	<b>OVERALL DISCUSSION.....</b>	<b>124</b>
<b>9</b>	<b>OVERALL CONCLUSION.....</b>	<b>127</b>
	<b>REFERENCES.....</b>	<b>128</b>

## 1. INTRODUCTION

Land use and land cover (LULC) data are important to mitigate environmental impacts associated with food production. The use of optical remote sensing to characterize LULC in tropical regions is challenging since tropical areas have frequent cloud cover (ASNER, 2001; WULDER et al., 2015; MARTINS et al., 2018a) and fragmented landscape types (LU et al., 2012; LAURIN et al., 2013). This difficulty is even more serious for detecting targets with high spatial and temporal dynamics, which is the case of agriculture. Mapping and monitoring LULC changes (LULCC) have economic, social, and environmental interests. Cropland monitoring is important for food security and for environmental impact reduction. Thus, proper agriculture monitoring systems are demanded.

The Roraima state, located under the Intertropical Convergence Zone (ITCZ) influence, has a frequent cloud cover, limiting the use of optical data. Moreover, agriculture is expanding in this state (IBGE, 2018a), mainly over savanna-dominated vegetation called “lavrados”, threatening the environment. This activity needs to be monitored and regulated; otherwise, it can affect the availability of natural resources (e.g., water and soil) and other environmental resources. However, there is a lack of detailed information about LULC and LULCC over the whole state.

One way to overcome this problem is the combined use of multisource remote sensing data, including the Synthetic Aperture Radar (SAR) data which are less affected by atmospheric conditions and cloud coverage than optical data (MOREIRA et al., 2013; WULDER et al., 2015). The optical-SAR approach takes advantage of both well-known interpretation experience of optical data and cloud-free SAR data (MCNAIRN; BRISCO, 2004; ORYNBAIKYZY; GESSNER; CONRAD, 2019; TAVARES et al., 2019; SOARES et al., 2020). The use of two-stage sampling approach can reduce computational demands and cloud limitations.

In this context, this research will test if the combination of SAR and optical in a multisensor timeseries data allows the proper LULC mapping in tropical



regions with frequent cloud cover. Moreover, the use of a two-stage sampling approach is a way to overcome the big data limitations, and analyze the LULC and LUCC, to quantify the possible environmental impacts, in different stages. This research exemplified the difficulty of using only RS optical data for specific periods to provide LULC mapping in a region with high cloud cover conditions throughout the year.

## **1.1 Objectives**

The overall objective of this research is to develop a multisensor and multitemporal approach to map the LULC classes of the Roraima state, Brazil. The specific objectives are:

### **1.1.1 Specific objectives**

- To evaluate the cloud cover frequency of South America territory and the implications of the use of optical data for cropland monitoring, and the use of the result of this evaluation to justify the addition of SAR data in areas with frequent cloud cover, as Roraima.
- To provide a robust methodology to classify the LULC classes of the Roraima state, investigating the benefits of combining SAR and optical data in the LULC mapping process in different periods.
- To evaluate the two-stage sampling to estimate the LULC classes of the Roraima state in 2017, 2018 and 2019, by using multi-sensors data collected along the roadside.

## **2. THEORETICAL BACKGROUND**

This chapter brings a brief theoretical background about LULC monitoring (2.1), multisensor approach with SAR and optical data (2.2), LULC classification process (0). The idea is to provide a general overview of state-of-art that will be further explored in the next chapters.

### **2.1 LULC monitoring**

Mitigating climate change, meanwhile, population growth, and biodiversity loss, is one of the most tricky topics for crop production at the national scale (NELSON et al., 2009; FOLEY et al., 2011). Nowadays, environmentally sustainable food production meanwhile minimizing the environmental impacts is one of the biggest challenges to be faced by the agriculture sector, government, and researchers (SEARCHINGER et al., 2015; ESTES et al., 2016). Part of this food demand is met by the agriculture expansion, resulting in severe environmental impacts (LAMBIN; GEIST; LEPERS, 2003), and contributing to global climatic changes (DENG; ZHAO; YAN, 2013). In this sense, continuous Land Use and Land Cover (LULC) mapping is fundamental to land use management and to understand the environmental effects at local, regional, and global scales (ADAMI et al., 2018; PAVANELLI et al., 2018). Thus, Remote Sensing (RS) technology is widely utilized for synoptic and continuous LULC monitoring, allowing identification of the LULC Changes (LULCC) (WULDER et al., 2015; VELOSO et al., 2017).

RS data is essential for agricultural monitoring since they can provide accurate and timely information about crop development (WHITCRAFT; BECKER-RESHEF; JUSTICE, 2015b; VELOSO et al., 2017). Brazil and Argentina Brazil and Argentina are the leaders of agricultural production in South America (ZHONG et al., 2017). Consequently, monitoring this large-scale agriculture is relevant to better understand the relation between food production and its environmental impacts. Due to the large extension and dynamics of cropland areas, satellite-derived information has been used to monitoring croplands at a global scale (FRITZ et al., 2015, 2019), such as the Group on

Earth Observation (GEO) and the Global Agricultural Monitoring Community (GEOGLAM) (WHITCRAFT et al., 2015a; WHITCRAFT; BECKER-RESHEF; JUSTICE, 2015a) and the Joint Experiment for Crop Assessment and Monitoring (JECAM, 2019). To increase the efficiency of the remotely sensed cropland monitoring, a suite of satellite systems is used to understand the spectral response of the croplands during their growing season. Some examples are the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the Terra and Aqua platforms (JUSTICE et al., 2002), the Operational Land Imager (OLI) sensor onboard the Landsat-8 satellite (MARKHAM; STOREY; MORFITT, 2015), the Multispectral Imager (MSI) sensor onboard the Sentinel-2A and Sentinel-2B satellites (DRUSCH et al., 2012) and the Multispectral Camera (MUX) onboard the CBERS-4 satellite (INPE, 2013; MARTINS et al., 2018b). The GEOGLAM group developed a moderate spatial resolution, orbital data integration system for crop monitoring and yield estimation of the most important crops on a global scale (BECKER-RESHEF et al., 2010).

Despite their wide use and importance, the use of optical remote sensing data for crop monitoring is complex and challenging in tropical areas such as in South America (SANCHES et al., 2018b). Climate, socio-economics, and infrastructure are some of the factors that interfere for the crop dynamics in tropical croplands (SANCHES et al., 2018a). Moreover, the cloud cover in South America is influenced by different factors, such as the South Atlantic Convergence Zone (SACZ) (CARVALHO; JONES; LIEBMANN, 2004; SILVA; KOUSKY, 2012), seasonal migration of intertropical convergence zone (COOK, 2009), and the South America low-level jet east from the Andes (VERA et al., 2006). These factors make the acquisition of cloud-free, optical images in tropical regions a big challenge (WHITCRAFT et al., 2015b).

Following the wet season, frequent cloud cover is typically observed in the austral summer. This is because the convergence zone and intense convection oriented from the northwest to southeast in the subtropics near the southeastern coast of Brazil protrudes into the adjoining South Atlantic ocean as SACZ (ZHOU; LAU, 1998).

Optical remote sensing data are traditionally used to identify and characterize LULC. Thus, the cloud-cover limitation affects the LULC mapping and characterization in South America (ASNER, 2001; WULDER et al., 2015; MARTINS et al., 2018a), along with the rapid LULCC (WHITCRAFT et al., 2015b; EBERHARDT et al., 2016). Agriculture are dynamic targets, in which the crop growing phase coincides with the rainy season, are more affected by cloud cover frequency. These difficulties, along with tropical landscapes fragmentations and transitions among vegetation types (LU et al., 2012; LAURIN et al., 2013), make the use of optical data to map LULC classes challenging (SANO et al., 2007; SUGAWARA; RUDORFF; ADAMI, 2008; WHITCRAFT et al., 2015b; EBERHARDT et al., 2016).

## **2.2 Multisensor approach to overcome the limitations of optical remote sensing**

LULC mapping is fundamental to understand the environment effects at local, regional, and global scales (ADAMI et al., 2018; PAVANELLI et al., 2018). In this sense, SAR data, are an option to overcome the limitations of optical remote sensing (WHITCRAFT et al., 2016). SAR sensors are less influenced by cloud cover frequency when compared with optical data (MOREIRA et al., 2013; WULDER et al., 2015). However, SAR signals interact with the surface differently compared with optical data. It depends on surface factors such as the dielectric constant, geometry, topography, and surface roughness. It also depends on the radar image acquisition mode, such as polarization, frequency, and incident angle (STEELE-DUNNE et al., 2017; HARFENMEISTER; SPENGLER; WELTZIEN, 2019). As SAR interpretation is complex, identifying vegetation types is not an easy task (TAVARES et al., 2019). In this sense, SAR use is more common only to overcome cloud limitations in specific tasks, such as in LULC classification. It has been less widespread than optical data (SOARES et al., 2020).

Thus, methods that integrate optical and SAR data have been explored in LULC studies (INGLADA et al., 2016; CLERICI; VALBUENA CALDERÓN;

POSADA, 2017; TORBICK et al., 2017a; REICHE et al., 2018; VAN TRICHT et al., 2018). Multisensor data overcome SAR interpretation with optical data and the cloud limitation with SAR data (MCNAIRN; BRISCO, 2004; ORYNBAIKYZY; GESSNER; CONRAD, 2019; TAVARES et al., 2019; SOARES et al., 2020). Sentinel-1 and Sentinel-2 data from the European Space Agency (ESA) are two of the most useful satellite data to provide LULC data (CHATZIANTONIOU; PSOMIADIS; PETROPOULOS, 2017; STEINHAUSEN et al., 2018; IENCO et al., 2019; TAVARES et al., 2019). Sentinel-2 has a constellation of two satellites (Sentinel-2A and Sentinel-2B). Sentinel-2A was launched in November, 2015 while the Sentinel-2B was launched in June, 2017, operating in the optical spectral region (DRUSCH et al., 2012). Sentinel-1A was launched in April, 2014 and Sentinel-1B was launched in April, 2016 (TORRES et al., 2012; ESA, 2017). They are SAR sensors that operate in the C-band (wavelength of ~5.6 cm; frequency of 5250–5570 MHz). The main image acquisition mode is the Interferometric Wide (IW) swath with dual-polarization (VV and VH polarizations) (TORRES et al., 2012). However, the methods to integrate the optical and SAR multisensor data are specific and complex (WULDER et al., 2015). The multisensor SAR-optical approach provides a high dimensionality dataset, demanding more computational power in the LULC classification process.

### **2.3 LULC classification process**

LULC classification process using a multisensor SAR-Optical data is time and computational consuming. Handling the high dimensionality data in the SAR-optical approach is challenging. Techniques to reduce data collection and to visualize class separation are useful to reduce the dimensionality problem. Cloud processing, which has higher computational power and does not demand downloading the remote sensing data, is an advance to the multisensor data processing capability (GORELICK et al., 2017; PALAZZO et al., 2018; ZHANG et al., 2018a).

### **2.3.1 Classifiers**

To the LULC classification process in these large datasets, traditional parametric classifiers do not produce good results (CLERICI; VALBUENA CALDERÓN; POSADA, 2017). Thus, it is necessary to use non-parametric classifiers (LU et al., 2012). On SAR-optical approach, Random Forest (RF) (RODRIGUEZ-GALIANO et al., 2012; INGLADA et al., 2016; TORBICK et al., 2016, 2017a, 2017b; CLERICI; VALBUENA CALDERÓN; POSADA, 2017; ZHOU et al., 2017; PAVANELLI et al., 2018) and Multilayer Perceptron (MLP) (SKAKUN et al., 2016; CAMARGO et al., 2019) are two of the most commonly non-parametric classifiers used to produce LULC classification (LU et al., 2012; CLERICI; VALBUENA CALDERÓN; POSADA, 2017).

The RF classifier provides robustness and capability of holding a high number of variables (JHONNERIE et al., 2015; DINIZ; GAMA; ADAMI, 2020) and high data dimensionality (TORBICK et al., 2017b). The MLP is a feed-forward artificial neural network trained by the backpropagation method and designed to map a set of input vectors to a set of output vectors (SKAKUN et al., 2016; CAMARGO et al., 2019).

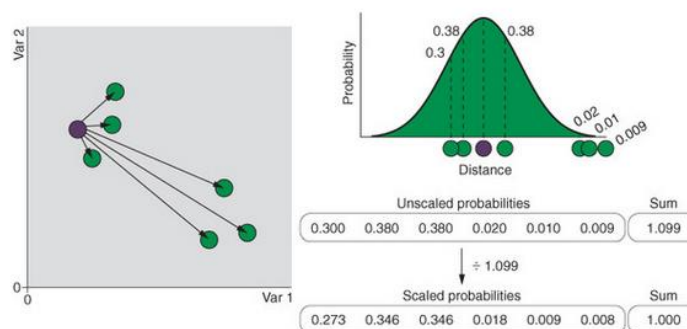
### **2.3.2 Data reducers**

The t-distributed Stochastic Neighbor Embedding (t-SNE), unsupervised technique, uses a heavy-tailed t-distribution to describe data similarity in two or three-dimensional maps (VAN DER MAATEN; HINTON, 2008; SONG et al., 2019). t-SNE converts the high-dimensional Euclidean distances between datapoints into a conditional probability, to keep close into a low-dimensional space the points that are near high-dimensional space (VIOLANTE, 2018; SONG et al., 2019). Different from a Principal Component Analysis (PCA), where seeks to maximize the variance and preserve large pairwise distance, t-SNE preserve only small pairwise distances, keeping the very similar data points close together in a lower-dimensional space (local similarities) (VIOLANTE, 2018; DEY et al., 2020; KHANDELWAL, 2020).

The t-SNE measures the distance among each single point, transformed into a probability by fitting a normal distribution over the current point (Figure 2.1). The results for each point are stored as a matrix of values (Figure 2.2). Converting distances in the lower dimensional representation into probabilities, t-SNE fits a Student's t distribution over the current case (Figure 2.3). The t-Sne job now is to "suffle" the data points around these new axes (one example in Figure 2.4). The probability matrix is computed for this axis, and the cases are shuffled around to make this matrix resemble the original, high-dimensional matrix. During shuffling, cases are attracted toward cases that are similar to them (lines with circles) and repulsed away from cases that are dissimilar (lines with triangles) (RHYS, 2020).

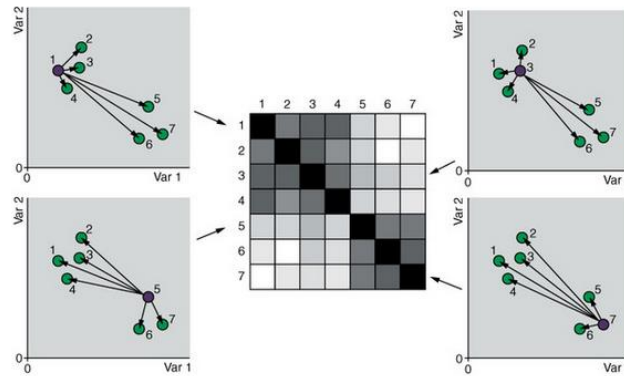
The t-SNE approach helps to understand if the highly dimensional dataset can or cannot provide visual separability among the classes (VAN DER MAATEN; HINTON, 2008; MAATEN, 2014; SONG et al., 2019). However, the t-SNE technique is not commonly used for data to visualize the data in the remote sensing field. Dey et al. (2020) used t-SNE to visualize the separation among different crop classes using polarimetric SAR data. Martins et al. (2020) also used t-SNE to visualize the features learned by deep neural network in each hidden layer.

Figure 2.1 – Distance from each case to every to every case, converted into a probability by fitting a normal distribution over the current case.



These probabilities are scaled by dividing them by their sum, so that they add to 1.  
Source: Rhys (2020).

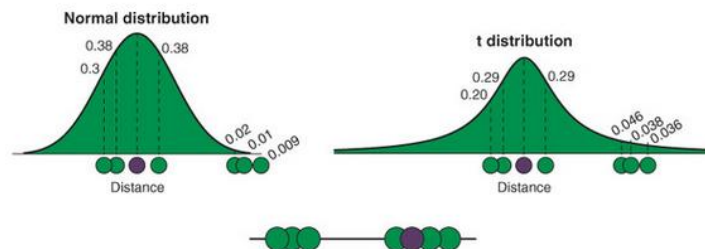
Figure 2.2 – The scaled probabilities for each case are stored as a matrix of values.



This is visualized here as heat map: the closer two cases are, the darker the box is that represents their distance in the heatmap.

Source: Rhys (2020).

Figure 2.3 – t-SNE Student's t distribution fit over the current point instead of a normal distribution.

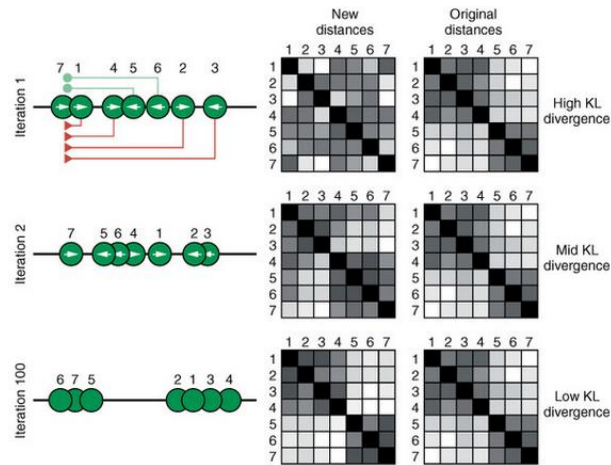


The Student's t distribution has a longer tail, meaning dissimilar cases are pushed further away to achieve the same probability as in the high-dimensional representation.

Source: Rhys (2020).



Figure 2.4 – Cases are randomly initialized over the new axes (one axis is shown here).



Source: Rhys (2020).

The use of different metrics to explore the temporal variation from remote sensing data is a way to decrease the data dimensionality. This statical transformation of time-series images is called multi-temporal metrics (SONG et al., 2017), which is useful for LULC characterization (BECKER et al., 2021). Metrics, as standard derivation, median, mean, minimum, maximum, and variance, among others, are a statical derivate of a time series imagery (ZALLES et al., 2021), representing key characterize of LULC classes in time for each pixel (KING et al., 2017; SONG et al., 2017; BECKER et al., 2021).

### 2.3.3 Two-stage sampling

Even using metrics and cloud processing techniques, classifying large areas based on moderate spatial resolution data (i.e., 30 m) could be a hard task, and sometimes, inviable. Stratified block processing can save time to provide accurate information with less computational effort (KING et al., 2017; SONG et al., 2017). This approach consists in partitioned the area of interest in blocks to randomly select some blocks in a two-stage sample approach (SONG et al., 2017; KRYLOV et al., 2019). Inside each block is randomly sample points to represent the internal block variation. In this way, it is possible the integration of map-making and sample-based process (SONG et al., 2017).

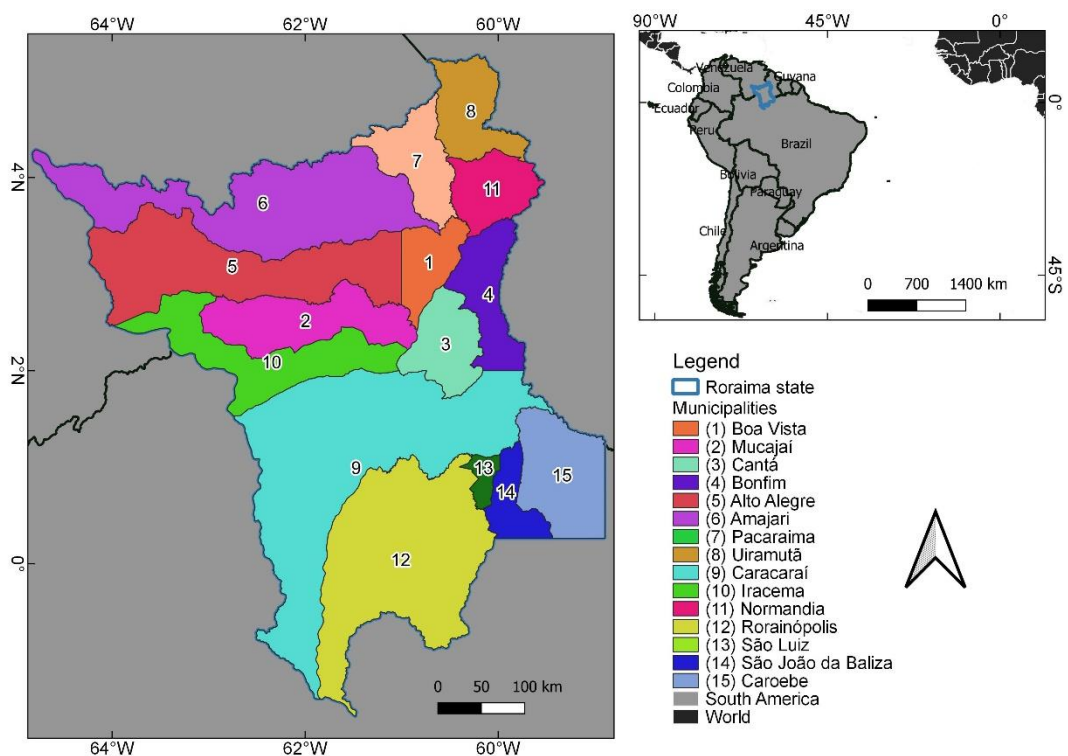
As it is visited only the selected blocks, this reduces the travel time and cost to field sample location (KRYLOV et al., 2019). The use of two-stage sampling methods allows the area to be estimated and it is a cost-effective way to obtain LULC information. This approach is used for crop classification estimation (KING et al., 2017; SONG et al., 2017, 2021a, 2021b), and tree cover loss (KRYLOV et al., 2019).

Using cloud processing (GORELICK et al., 2017) along metrics to explore the temporal variation from the remote sensing data (SONG et al., 2017), and sample-based methods (KING et al., 2017; SONG et al., 2017) are a viable way to provide LULC information as quick as possible.

### 3. RORAIMA STATE

Roraima state is localized in the North region of Brazil, near the states of Amazonas and Pará, Venezuela, and Guiana (Figure 3.1). It is the Brazilian state with the lowest number of inhabitants: around 606,000 inhabitants, 332,000 living in Boa Vista, the capital of the state. This state occupies an area of 224,300 km<sup>2</sup> and has 15 municipalities (IBGE, 2018b).

Figure 3.1 – Location of Roraima study state and its 15 municipalities.

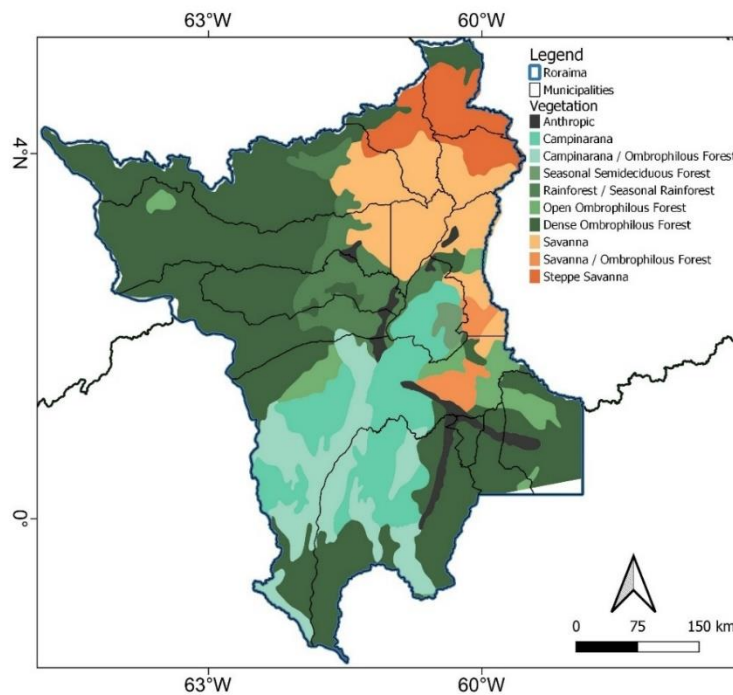


Source: Author's production.

There are three groups of natural vegetation formations in Roraima: rainforest, campina-campinarana, and savannas, also called “lavrados” (BARBOSA; KEIZER; PINTO, 2010), as shown in Figure 3.1 (IBGE, 2012). Rainforest covers the main part of the center-southern part of the state. It is divided in Seasonal Semideciduous Forest, Seasonal Forest, Open, and Dense Ombrophilous Forest. Campina-campinarana is formed by campinas (small shrubs) and campinarana, which are mostly found in the center-southern part of

the state, surrounded by the rainforest. Savannas are formed by savanna and steppe savanna and are present mostly in the northeast part of the state. This class is predominated by grasslands with few shrubs (BARBOSA; KEIZER; PINTO, 2010; PAVANELLI et al., 2018). Some regions show ecologic tension, where the transition between two or more types of vegetation occurs, as Savannas and Forest, and Campinarana and Forest (BARBOSA; BACELAR-LIMA, 2008; BARBOSA; KEIZER; PINTO, 2010; IBGE, 2012).

Figure 3.2 – Roraima natural formations according to IBGE.

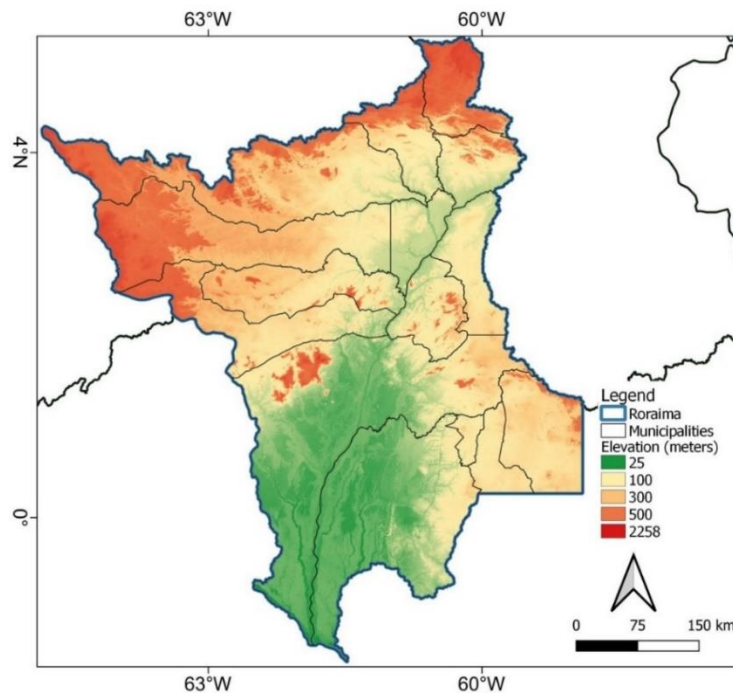


Source: Adapted from Barcelar-Lima (2008) and IBGE (2012).

The elevation (Figure 3.3) ranges from 30 meters in the Negro River to 2000 meters in the Roraima mountain (BARBOSA; BACELAR-LIMA, 2008). This altitude range act as a natural barrier in the north of the Roraima, blocking the moisture brought by the ocean winds along the ITCZ. This generates a precipitation gradient and a persistent cloud coverage in the state, which greatly limits the use of optical satellite images. Barni et al. (2020a) describe that Roraima has two well-defined climatic seasons, but it is different in the areas of

the state located in the north from the areas in the southern hemisphere. Most of Roraima territory has rainy seasons concentrated between April to September, with a peak in June. The dry period starts in October and go until March (MORAIS; CARVALHO, 2015). In the lavrados, the small lakes fill and connect in the rainy period. The mean annual temperature is 28°C (PAVANELLI et al., 2018).

Figure 3.3 – Roraima elevation data according to SRTM data.

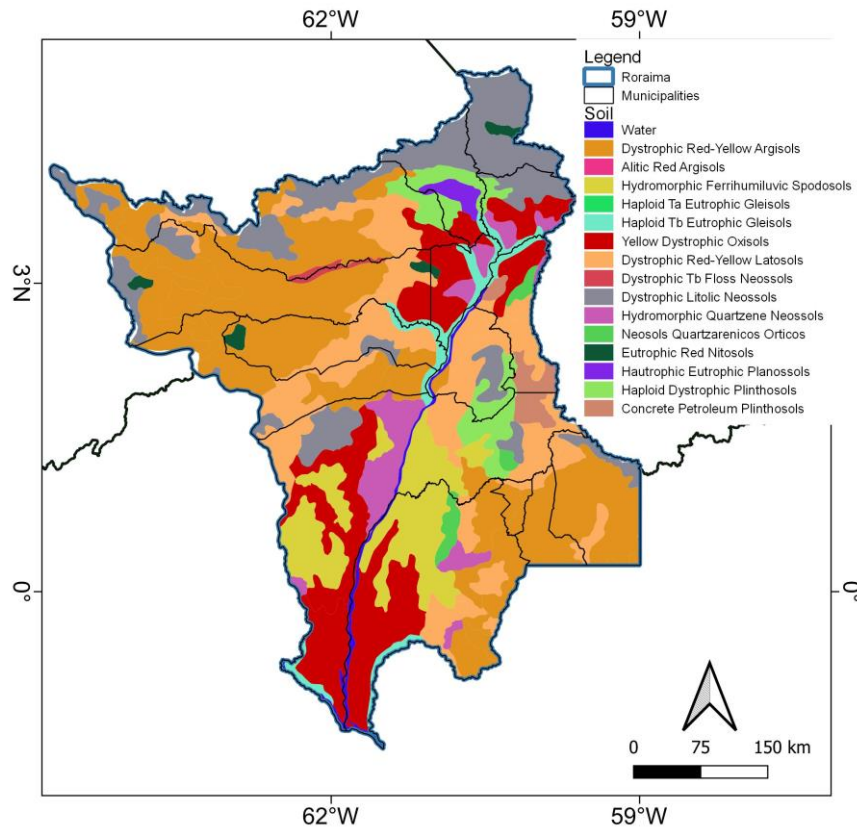


Source: Author's production.

Roraima soils are formed mainly by Argisols (33%), Latosols (22%) and Plinthosols (14%) groups (MELO et al., 2010). At the center-southern part of Roraima, we mostly find the dystrophic Red-Yellow Latosols and the dystrophic Red-Yellow Argisols under the rainforests and savannas. Dystropheric Litolic Neossols is present in the northern and center of Roraima, with mountains and outcrops from the Escudo das Guianas. Hydromorphic Ferrihumiluvic Spodosols and Yellow Dystrophic Oxisols are present in the campinarana region in the

center-southern part of Roraima (BARBOSA; BACELAR-LIMA, 2008; SANTOS et al., 2011).

Figure 3.4 – Roraima’s soil map.



Source: Adapted from Santos et al. (2011).

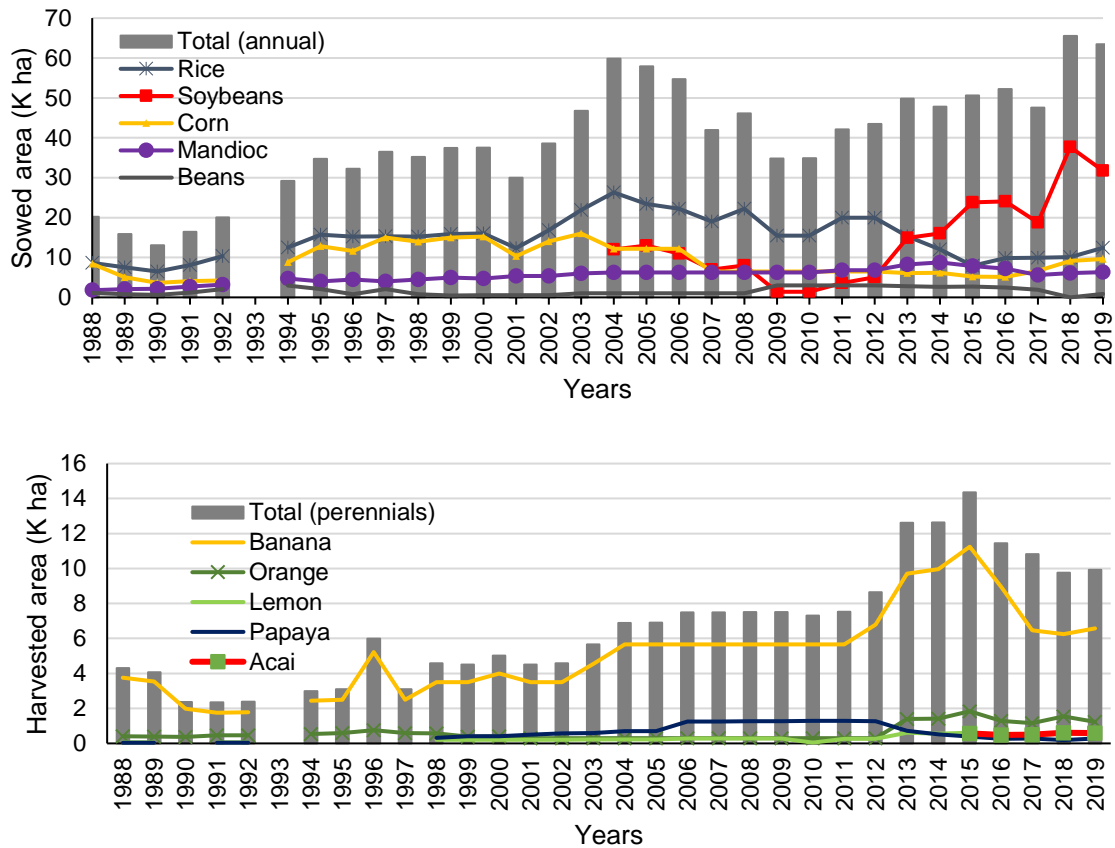
The combination of precipitations and altitude gradients allowed the formations of a mosaic of open field and forest system in the lavrados. Moreover, differences between altitudes and precipitation, located between the Orinoco and Essequibo Amazonas watersheds, and geologic basalt residue and depressions in the land, allowed in low and middle altitude (< 600 m), the emergence of a system of perennial and seasonal lakes. These are connected in the rainy season, characterizing the importance of these regions to the biodiversity and water resource conservations (BARBOSA et al., 2007).

### 3.1 Agriculture expansion

The combination among adequate climatic, altitude, and water availability factors, low land prices and government subsidies, has encouraged agriculture and livestock in Roraima (CARVALHO; MUSTIN, 2017). The crop calendar, with harvest during the off-season for the other Brazilian states (April-September), favors better prices and facilitates production chain logistics. As a consequence, agriculture in Roraima is expanding, mainly over the lavrados, where the adoption of drainage channels to use water from the lakes is frequent. The conversion of areas with Acacia formation in soybean plantations is also common in lavrados. However, crop areas are not significant in Roraima yet, representing less than 1% of the territory, not being mapped by other LULC studies (SONG et al., 2021a). This activity needs to be monitored and regulated; otherwise, it can affect the availability of natural resources (e.g. water and soil) (CARVALHO; MUSTIN, 2017).

According to the Municipal Agricultural Production (PAM) data, from the Brazilian Institute of Geography and Statistics (IBGE, 2018a), agriculture expansion was intensified after 2010, mainly related to the soybean crop expansion. Other annual crops present in Roraima are corn and cassava, cultivated mainly in small areas, and beans. The area cultivated with rice decreased during this time due to the demarcation of the Raposa Serra do Sol indigenous land. In the southern part of Roraima, we find cultivations of banana, orange, and palm oil. Cattle ranching is the main economic activity in this region.

Figure 3.5 – Area of Annual crops (upper) sowed and Perennial crops (lower) harvested in Roraima state, from 1988 to 2019, in thousands (K) of hectares.



Source: IBGE (2018a).

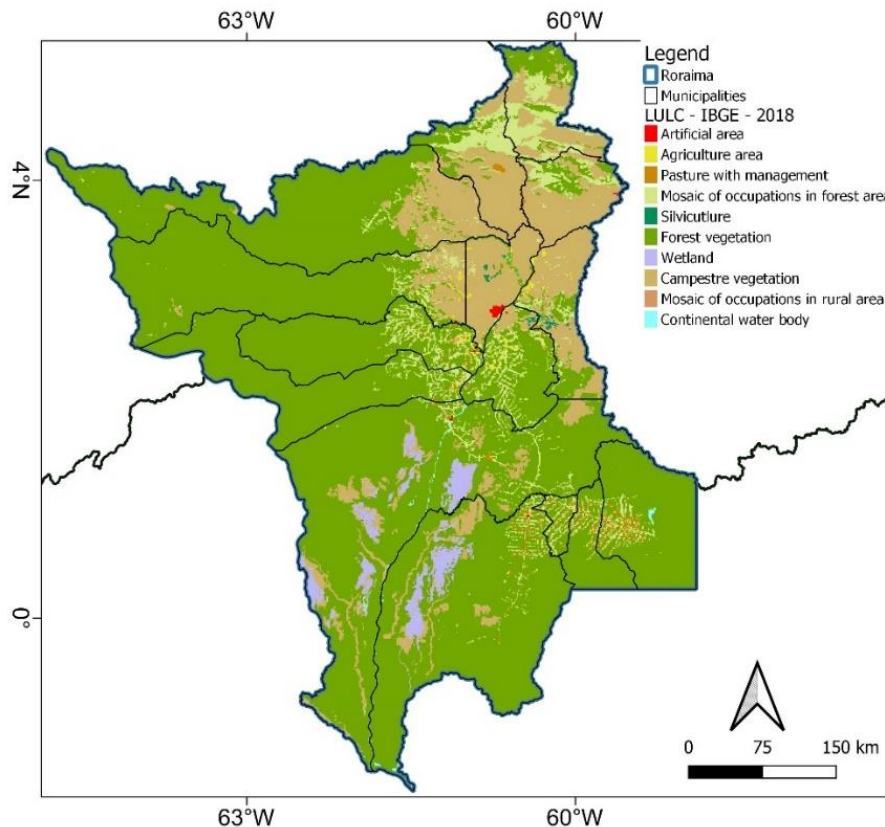
### 3.2 Existing LULC mapping programs

Regardless of the importance and recent expansion of agriculture in Roraima, there are few LULC maps available, often outdated or needing improvements. In Brazil, there are three LULC programs/projects at the state level: IBGE LULC (Figure 3.6), TerraClass (ALMEIDA et al., 2016) (Figure 3.7), and MapBiomias (SOUZA et al., 2020) (Figure 3.8). The last LULC map from IBGE (2020) is referent from 2018. They focused more on natural vegetation classes rather than providing a detailed land use classification. It was made by visual interpretation of optical images and used ancillary data, mainly crop statistics data obtained by the TerraClass (see Figure 3.7) and Deforestation



Monitoring in the Brazilian Legal Amazon (PRODES) (PRODES, 2018) projects. The product is available on the scale of 1:250,000, considering a grid of 1km<sup>2</sup>.

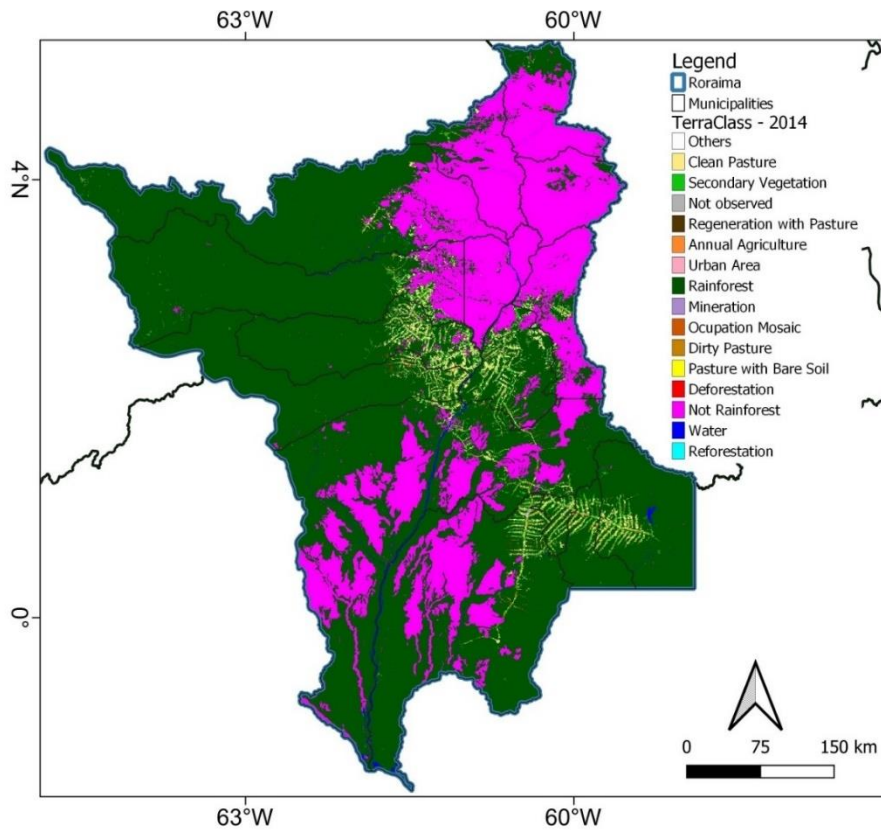
Figure 3.6 – Roraima’s land use and land cover (LULC) mosaic based on the IBGE initiative of 2018.



Source: Adapted from IBGE (2020).

TerraClass started as a complement of the PRODES program, to characterize the LULC classes in the deforested areas of the Amazon Rainforest (ALMEIDA et al., 2016), that is, to identify the land uses in previously deforested areas. More recently, TerraClass was expanded for the Cerrado biome (TERRACLASS CERRADO, 2018). Campinaranas and Savannas areas are not monitored by the TerraClass Amazon, since they are not rainforests. TerraClass Cerrado also does not monitor these classes in Roraima, because they are outside of the Cerrado biome. Campinarana and Savannas are not monitored by the TerraClass initiative yet (see Figure 3.7), but future versions will do it.

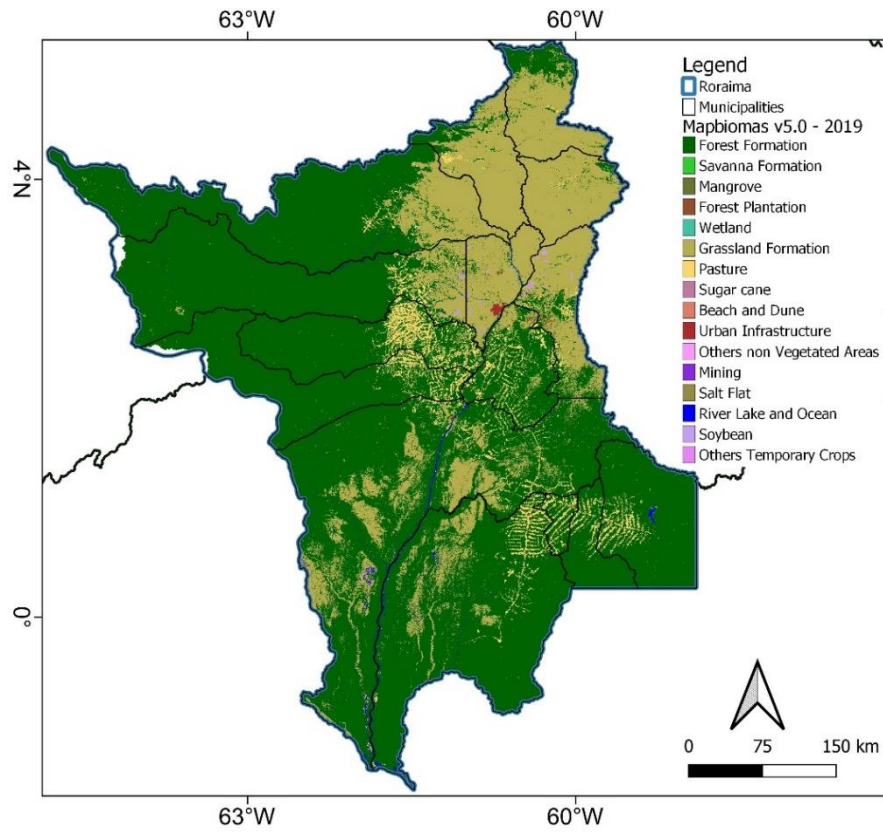
Figure 3.7 – Roraima LULC map, TerraClass Amazonia for 2014.



Source: Almeida et al. (2016).

MapBiomass is the most recent initiative of LULC mapping in Brazil (SOUZA et al., 2020). It provides annual LULC maps for the entire country, between 1985 and 2019 (v.5), with a spatial resolution of 30 meters. Although MapBiomass project maps the whole state of Roraima (Figure 3.8), it needs improvements to reach better results, for example, separation between savannas and campinaranas, classification of the Acacias formations, and improvement of the annual crop classes.

Figure 3.8 – MapBiomass v.5 LULC map for Roraima region for 2019.



Source: Author's production.

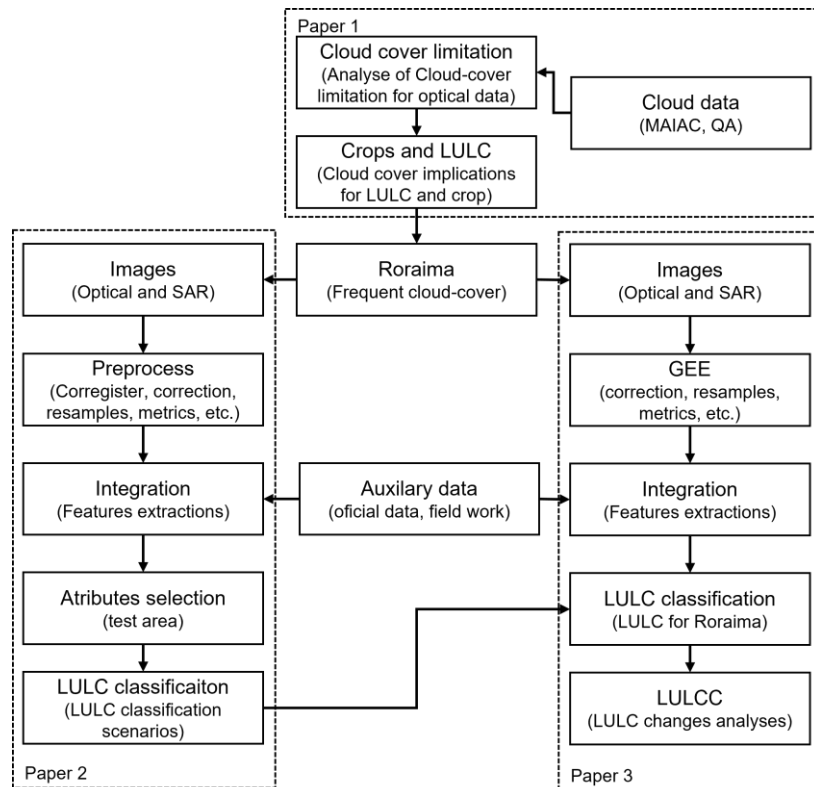
## **4. MATERIAL AND METHODS**

In this chapter, we present a general description of the material and methods used in this research. First, we describe about the thesis organization and show a general flowchart (4.1). We also describe the field work (4.2) conducted to collect data for the LULC classification process. In the last section of this chapter, we describe the Remote Sensing data (4.3) used in our analyses.

### **4. 1 Thesis organization**

This thesis is structured into a brief theoretical background, methods, three articles (Figure 4.1), general discussion, and general conclusion. First, we present our main study area (Roraima state), showing the LULC characteristics of the area and the limitations of using remote sensing data. In the same section, we presented a brief theoretical background to introduce the thesis topic. After, we describe our field campaign and the methodological steps. In our first paper, we present the study about the limitation of using optical data in South America and the implications for agriculture monitoring. The cloud cover assessment was useful to select the Roraima state as the study area for the analysis of the performance of optical-SAR-based LULC mapping. To define the methodology for LULC mapping in Roraima, we created scenarios and tested different classifiers, which are presented in Paper 2. In Paper 3, we show the probabilistic two-stage sampling approach to estimate the LULC, as well as the classification using roadside data. Next, we discuss the main findings of our research and highlight the main conclusions.

Figure 4.1 – Flowchart with the steps of data processing and papers division of this research.



Source: Author's production.

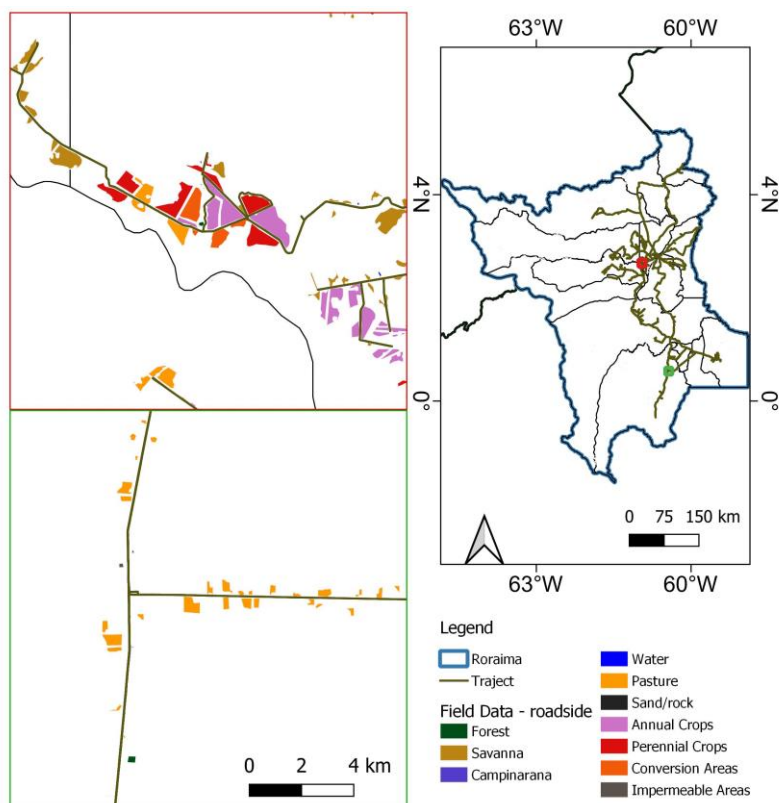
## 4.2 Field data

Accurate and representative field information is essential for LULC and LULCC classifications so that it is crucial to choose proper periods for field data collection. We performed fieldwork in Roraima in August-September, 2019, which correspond to the crop season and the end of the rainy season, allowing us to collect field data from all representative LULC classes, including rainfed crops.

We collected the LULC data (Figure 4.2) for classification training purposes, along the roadsides (WALDNER et al., 2019), using the Locus Map Pro applications. We collected a total of approximately 7000 points which was converted into polygons based on the QGIS software and Sentinel-2 MSI true color composites with 10 m spatial resolution. The polygons were drawn carefully to not include mixed pixels in the edges using Python routines and the QGIS

Earth Engine plug-in (GEE COMMUNITY, 2020) to access the MSI optical images from GEE to QGIS. In this step, we used images converted into reflectance at the top of the atmosphere (TOA), with 10 m (RGB true color) and 20 m (RGB Normalized Difference Vegetation Index -SWIR-RED) spatial resolutions. We used TOA data to draw the polygons because atmospherically corrected images became available only after December 2018 in the GEE platform. We incorporated the median images from each month to avoid cloud cover effects. Besides, we used Normalized Difference Vegetation Index (NDVI) images to better discriminate agriculture and pasture from natural fields. With this approach we certificate, looking for images from each month, that the classes identified in the field, indeed were on the images.

Figure 4.2 – Field data polygons are drawn after collecting roadside data in August and September of 2019 in Roraima.



Source: Author's production.

During the fieldwork, we obtained georeferenced photos using a Nikon Coolpix S9700 digital camera as well as some videos with a SJ7CAM (ActionCam) onboard. In Figure 4.2 and Figure 4.3 we show some details about our field work.

In total, we draw 5,300 polygons to training our models (Figure 4.2). These polygons were used to define the LULC classes for 2017 and 2018, using the visual interpretation of Sentinel-2 MSI images. It is important to mention that it was not possible to determine the crop species from the past years, thus we defined the class of this species as agriculture (other annuals) in general.







#### **4.2.1 LULC classes**

Based on previous literature and LULC programs (ALMEIDA et al., 2016; PAVANELLI et al., 2018; SOUZA et al., 2020) we divided into 39 LULC classes (Figure 3.3). After, we grouped these classes into three classification levels according to similarities. The more detailed level is the third one, with the 39 classes that we identified in the field. After we grouped some classes according to the similarities and got the second level, which is more similar to the LULC programs legends (see Section 3.2). We decided to identify Conversion and Sand/rock areas, even that classes are less representative in terms of area. Sand/rock has some confusion with annual crop in MapBiomass (Section 6) and identify areas that are being convert from natural formations to anthropic use is an important information for the local governments and researchers. Finally, we provide a macro group, the first level, which was divided only into Natural or Non-natural classes.

In Figure 4.3 we provide the details about our LULC classes and some examples/pictures from some classes.







Figure 4.3 – Panoramic field photos of land use and land cover classes in the Roraima State obtained during the field campaign in 2019.

		
<p><b>Forest:</b> Natural tree formations. This class has Forest, Buriti, and parkland cerrado.</p>	<p><b>Savanna:</b> include clean, dirty, and cerrado fields.</p>	<p><b>Campinarana:</b> Campinas and Campinarana together.</p>
		
<p><b>Water:</b> All the water formations, lakes, rivers, and dams.</p>	<p><b>Sand/rocks:</b> Sandbank, common in rivers, and outcrop.</p>	<p><b>Annual Crops:</b> Areas that are used to cultivate temporary crops. It was identified soybean, corn, beans, rice, follow lands, and other annuals crops.</p>

continue



Figure 4.3 – Conclusion.

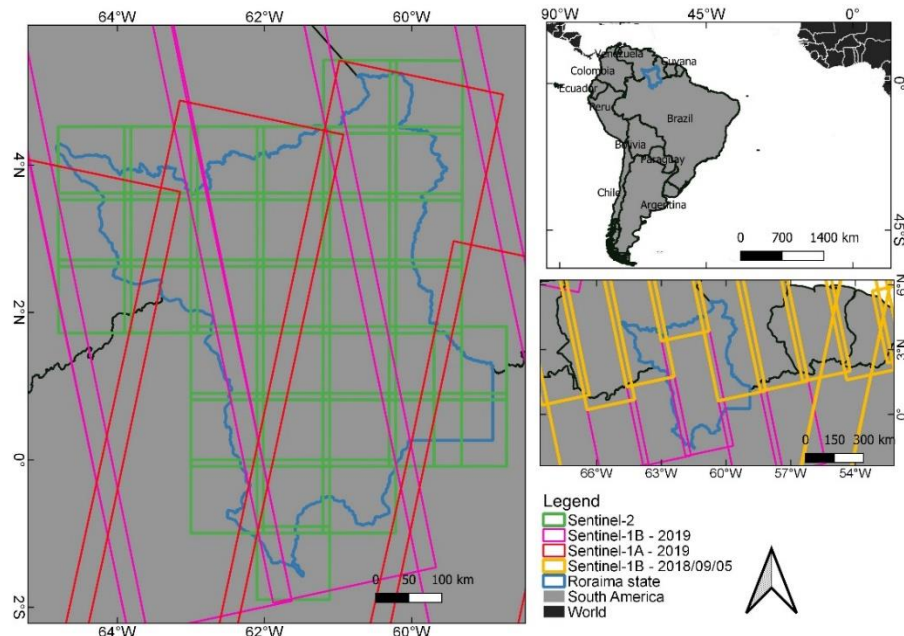
		
<p><b>Perennial Crops:</b> Perennial crops and silviculture. We identified acacia, papaya, cashew, dende, and mango.</p>	<p><b>Pasture:</b> Cultivated pasture.</p>	<p><b>Conversion:</b> Areas in conversion in the analyzed year. Areas that were being converted from natural formations to something else.</p>
		
<p><b>Impermeable:</b> Urban areas, roads, and other buildings.</p>		

Source: Author's production.

### 4.3 Remote sensing data

The orbit files for Sentinel-1 and Sentinel-2 satellites over the Roraima state are presented in Figure 4.4. To cover the entire state, it is necessary 47 MSI-Sentinel-2 MSI and 7 Sentinel-1 orbits.

Figure 4.4 – Sentinel-1 (considering 2018 and 2019) and Sentinel-2 (2019) orbits over the Roraima state.



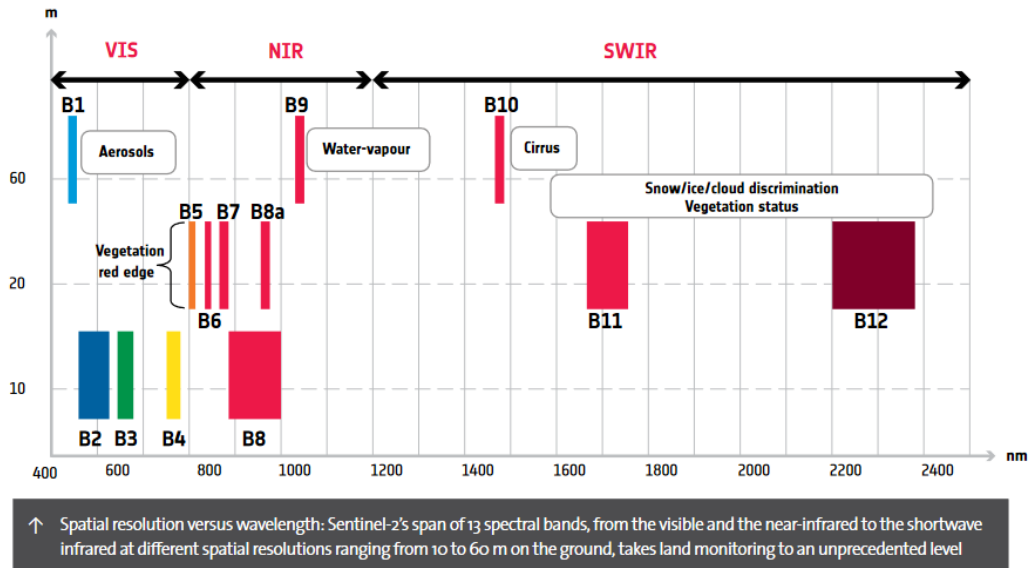
Source: Author's production.

#### 4.3.1 Sentinel-2

With both Sentinel-2A and Sentinel-2B satellites operating after June, 2017, the revisiting time over Roraima became five days. Onboard this constellation, there is the MSI sensor with 13 bands in the visible, near infrared (NIR) and shortwave infrared (SWIR) spectral bands (Figure 4.5 These sensors operate with 10 m of spatial resolution for bands in the visible (B2-Blue, B3-Green, and B4-Red) and in the NIR (B8-NIR) and 20 m spatial resolution for the red-edge bands (B5-Red Edge 1, B6-Red Edge 2, B7-Red Edge 3), NIR (B8A-

NIRA) and SWIR regions (B11-SWIR1, B12-SWIR2). Aerosols (B1), water-vapour (B9) and cirrus (B10) bands have spatial resolution of 60 m (ESA, 2015).

Figure 4.5 – Spectral and spatial resolutions of the Multispectral Instrument (MSI) sensor on board of the Sentinel-2 satellite.



Source: ESA (2015).

For vegetation studies, the bands, Red-edge 1, Red-edge 2, Red-edge 3, Near-Infrared (NIR), NIR-A, Shortwave Infrared 1 (SWIR 1), and Shortwave infrared 2 (SWIR 2) are the most used. Vegetation Indexes (VIs), as the NDVI (ROUSE et al., 1973), and the Land Surface Water Index (LSWI) (XIAO et al., 2002, 2004), brought important information to the vegetation studies. The NDVI is one of the most know VI and is associated with vegetative vigor. LSWI is sensitive to the water presence and dry matter and is used for flood mapping (DONG et al., 2013; TORBICK et al., 2016). NDVI (4.1) and LSWI (4.2) Equations are presented in the following:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad 4.1$$

$$LSWI = \frac{NIR - SWIR}{NIR + SWIR} \quad 4.2$$

where NIR is the near-infrared region (MSI/band 8 - 842 nm), Red (visible region, MSI/band 4 - 665 nm), and SWIR is the shortwave infrared region (MSI/band 11 - 1610 nm), are surface reflectance in each spectra.

#### 4.3.2 Sentinel-1

Sentinel-1 SAR has dense timeseries data with 12 days of temporal resolution for each satellite (Sentinel-1A and Sentinel-1B). Before September-2018, Sentinel-1B does not cover the entire Roraima state (Figure 4.4). Sentinel-1B started to cover the entire Roraima state after September 2018, after a demand to the ESA and configuring a temporal resolution of 4 and 8 days. Sentinel-1 most common polarization are VH (vertical transmitted and Horizontal received) and VV (vertical transmitted and vertical received). Moreover, IW mode, and Ground Range Detected (GRD) data types are the most utilized.

These data can be used to improve and/or develop new methods for mapping and monitoring LULC (TAMM et al., 2016). The ratio between the VH and VV polarizations (4.3), is also frequently used for LULC monitoring and/or classification. The ratio is less affected by environmental factors or image acquisition mode: thus it might have more stability than single polarizations VH or VV images (VELOSO et al., 2017; HARFENMEISTER; SPENGLER; WELTZIEN, 2019).

$$Ratio = \frac{VH}{VV} \quad 4.3$$

## **5. IMPACTS OF CLOUD COVER OVER SOUTH AMERICA OPTICAL REMOTE SENSING MONITORING – CROPLANDS AND RORAIMA CASE<sup>1</sup>**

Our first paper is about a cloud-cover study for the South America region. With this approach, it was possible to understand where and when Optical Remote sensing has more limitations due to the cloud-cover. The findings from this paper, together with agricultural expansion and data available from LULC programs data, helped us to choose the Roraima state, where a multisensor approach would be useful.

### **5.1 Introduction**

Population growth, biodiversity loss, and climate change are leading to an unparalleled challenge for global crop production (NELSON et al., 2009; FOLEY et al., 2011). Consequently, large-scale agricultural monitoring becomes a relevant research area for a better understanding of food production and its environmental impacts. South America (SA) is one of the world leaders in agricultural production, with approximately 8% of the worldwide croplands (1.5 million km<sup>2</sup>), especially Brazil and Argentina (ZHONG et al., 2017). Due to the large extension and dynamics of cropland areas, satellite remote sensing (RS) has proved to be a powerful tool for agricultural monitoring. Advances in Optical Remote Sensing (ORS) contribute to the development and improvement of operational systems focusing on cropland monitoring (ATZBERGER, 2013).

ORS data have been considered essential for agricultural monitoring, providing precise and in time information about crop development (WHITCRAFT; BECKER-RESHEF; JUSTICE, 2015b; VELOSO et al., 2017). However, tropical croplands have high crop dynamics, caused by climatic, socio-economic, and infrastructure factors (SANCHES et al., 2018a). Hence, the use of RS for crop monitoring is more complex and challenging in these areas (SANCHES et al., 2018b). Moreover, the acquisition of ORS cloud-free images is one of the biggest

---

<sup>1</sup> Most part of this chapter was published in the article in . Remote Sensing Applications: Society and Environment, 20, Prudente, V.H.R., Martins, V.S., Vieira, D.C., Silva, N.R. de F. e, Adami, M., Sanches, I.D, Limitations of cloud cover for optical remote sensing of agricultural areas across South America, Page Nos, Copyright Elsevier (2020) <https://doi.org/10.1016/j.rsase.2020.100414>

challenges for tropical regions (WHITCRAFT et al., 2015b). In SA, cloud cover is influenced by tropical climate features and atmospheric circulation, such as the South Atlantic Convergence Zone (SACZ) (CARVALHO; JONES; LIEBMANN, 2004; SILVA; KOUSKY, 2012), SA low-level jet east of the Andes (VERA et al., 2006), and seasonal migration of intertropical convergence zone (COOK, 2009). Following the rainfall season, frequent cloud cover is typically observed in the austral summer. This is because the convergence zone and intense convection oriented northwest to southeast in the subtropics near the southeastern coast of Brazil protrudes into the adjoining South Atlantic ocean as SACZ (ZHOU; LAU, 1998).

The SA rainy season helps to provide plant water demands to summer crops. However, this is the worst scenario to obtain cloud-free images (ASNER, 2001; MCNAIRN et al., 2009; WHITCRAFT et al., 2015b; EBERHARDT et al., 2016). Some studies demonstrated that cloud frequency could affect the potential of earth observations at regional (ASNER, 2001; SANO et al., 2007; SUGAWARA; RUDORFF; ADAMI, 2008; EBERHARDT et al., 2016) and global scales (WHITCRAFT et al., 2015b; WULDER et al., 2015). Specifically for cropland monitoring during the growing season, Whitcraft et al. (2015a; 2015a, 2015b) showed a global cloud cover pattern and reported that there are few opportunities for cloud-free ORS data during the vegetative cycle, especially at the end of the flowering season and at the beginning of plant senescence. However, these studies do not consider how crop types are distributed in cropland areas. This difficult analysis focused on the cloud cover impact on specific crop monitoring, mainly in tropical regions.

Roraima has potential to be a new Brazilian agriculture frontier. Agricultural exploration has been encouraged by combining the climatic, altitude, and water availability factors, as well as affordable land prices and government subsidies (agriculture and livestock) (CARVALHO; MUSTIN, 2017). Moreover, the crop calendar, with harvest during the off-season for the other Brazilian states (April-September), favors better prices and facilitates production chain logistics. Thus, it is necessary to monitor and regulates this activity; otherwise, it can affect the availability of natural resources (e.g. water, soil) (CARVALHO; MUSTIN,

2017). However, Roraima altitudes range between the Negro River to the Roraima Mountain (BARBOSA; BACELAR-LIMA, 2008), act as a natural barrier, blocking the moisture brought by the trade winds along the Intertropical Convergence Zone (ITCZ). The results are a precipitation gradient and a frequent cloud cover frequency in the state, which greatly limits the use of satellite optical images (BARBOSA et al., 2007).

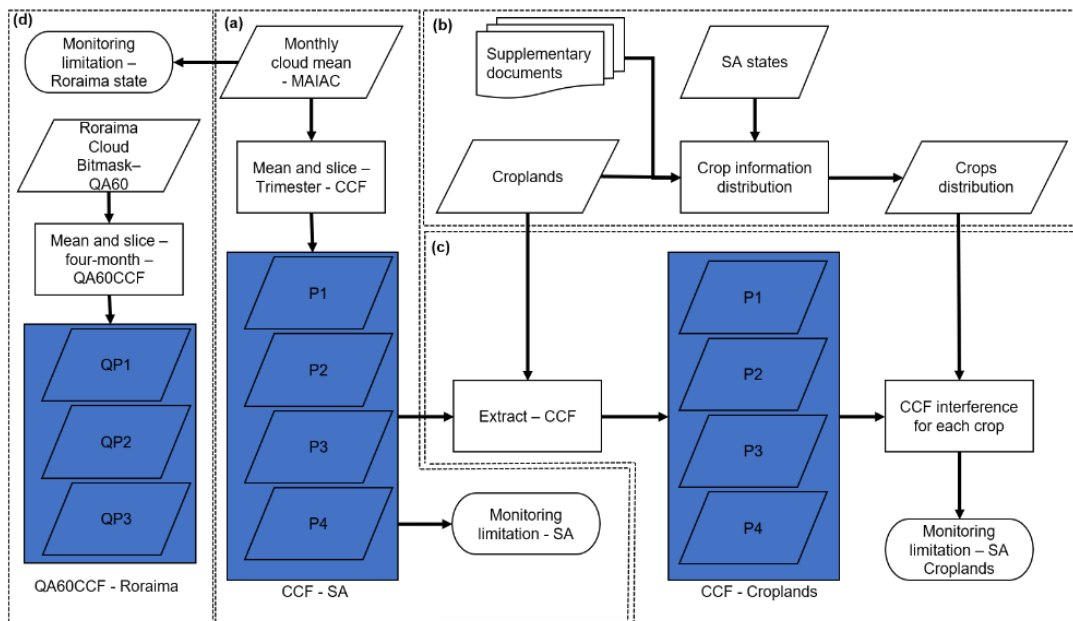
The understanding of the limitations imposed by cloud cover is essential for planning ORS applications. This cloud limitation can affect the LULC mapping process in some regions. Besides, few studies have focused on this subject for SA cropland. In this context, this study evaluates the cloud cover frequency of SA to identify the challenges of using ORS. Our goals were to identify regions where clouds limit the use of ORS in general and focused on its implications for SA croplands. For the proposed analysis, long-term monthly cloud cover was computed from the MODIS MCD19A2 product (2000 – 2015) over SA. The Sentinel-2 bitmask band QA60 was used to analyze the impact of cloudiness over the Roraima state, a region with frequent cloud cover. The discussion highlights the influence of cloud cover on the RS of the main crops in different seasons. This continental-scale analysis for agriculture is the first evaluation that combines crop type and seasonality with cloud cover variability. Also, this study was the base to analyze how clouds interfere in the use of ORS data in LULC approach in the Roraima state, a possible new agriculture frontier.

## **5.2 Material and methods**

Cloud cover is primary information for the assessment of the potential of optical remote sensing (ORS) for Earth observation applications. To assess the impact extent of cloudiness, the cloud cover information was extracted from the MODIS MCD19A2 for the entire South America (SA) and Sentinel-2 Quality band (QA60) for Roraima state. MODIS MCD19A2 was combined with the agricultural calendar of the main SA crops (e.g. maize, soybean, and cotton) and QA60 was analyzed for Roraima state. We divided the methodology into four sections (Figure 5.1). In Section 5.2.1 we described our study area, South America (SA) and Roraima. In Section 5.2.2 we showed the cloud cover products obtained from

the MODIS-MCD19A2 product and bitmask QA60. In Section 5.2.3 we presented the crop type and spatial and temporal cropland distribution over SA. In Section 5.2.5 we created the cloud cover frequency (CCF) of SA, using MODIS-MCD19A2, and analyzed it for the croplands. Finally, in Section 5.2.5 we create a CCF-QA for Roraima state and analyzed it for the entire state.

Figure 5.1 – Flowchart showing the four methodological steps: process to create the cloud cover frequency (CCF) for South America (SA).



(a), the process to understand the crop distribution over SA (b), cloud cover frequency over croplands, and the impact in the ORS monitoring (c), and detailed approach to understand the cloud cover impact for Roraima (d).

Source: Author's production.

### 5.2.1 Study area

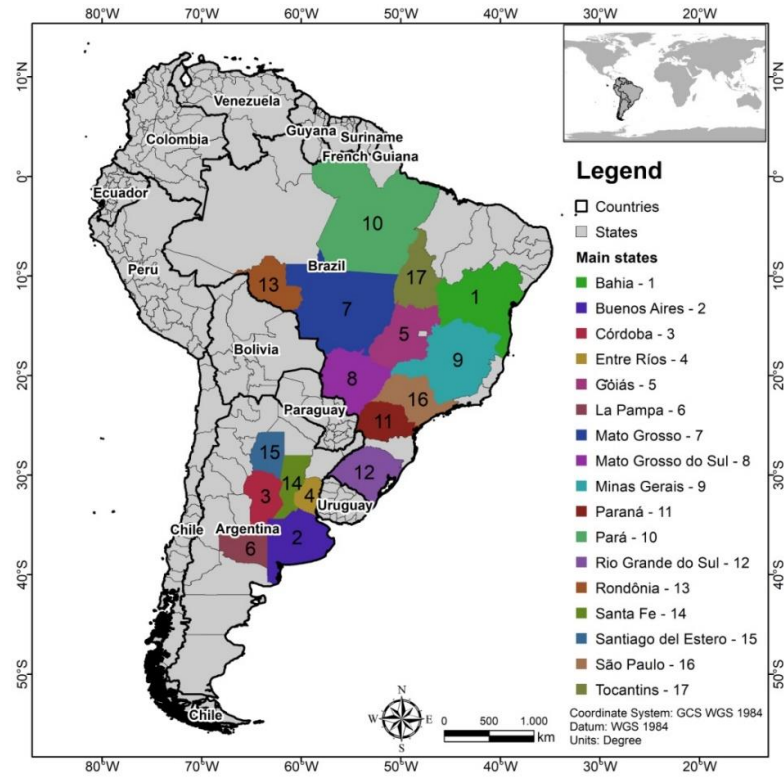
SA has a territorial extension of 17.8 million km<sup>2</sup> between latitudes 11° N and 55° S, and longitudes between 34° W and 82° W. The continent includes 12 countries, with 242 internal territorial divisions (states, provinces, departments, or regions) (Figure 5.2). Regarding altitude, SA can be divided into three parts. In the east, there are low altitude plateaus with the basins of the main Brazilian rivers. In the west, the Andes Mountain range presents high altitudes with peaks



achieving, approximately, 7,000 meters. In the middle region, there is the Central South American Depression, formed by plains with excerpts flooded as the Orinoco River in Venezuela, the Mamoré-Beni in Bolivia, and Paraguay river in Brazil, Paraguay, and Argentina (OLSON et al., 2001; GRIMM; PAL; GIORGI, 2007; ROSS, 2016; ADAMI et al., 2018). According to Koppen-Geiger classification, there is a large variability of climate patterns across SA, where the central northern region has tropical monsoon phenomena and the southern region shows humid subtropical climate (MECHOSO et al., 2004; VERA et al., 2006; PEEL; FINLAYSON; MCMAHON, 2007; SILVA; KOUSKY, 2012; VUILLE et al., 2012; ALVARES et al., 2013). Such differences allow the development of intensive agriculture, rendering the continent one of the largest agricultural producers in the world (FAOSTAT, 2018).

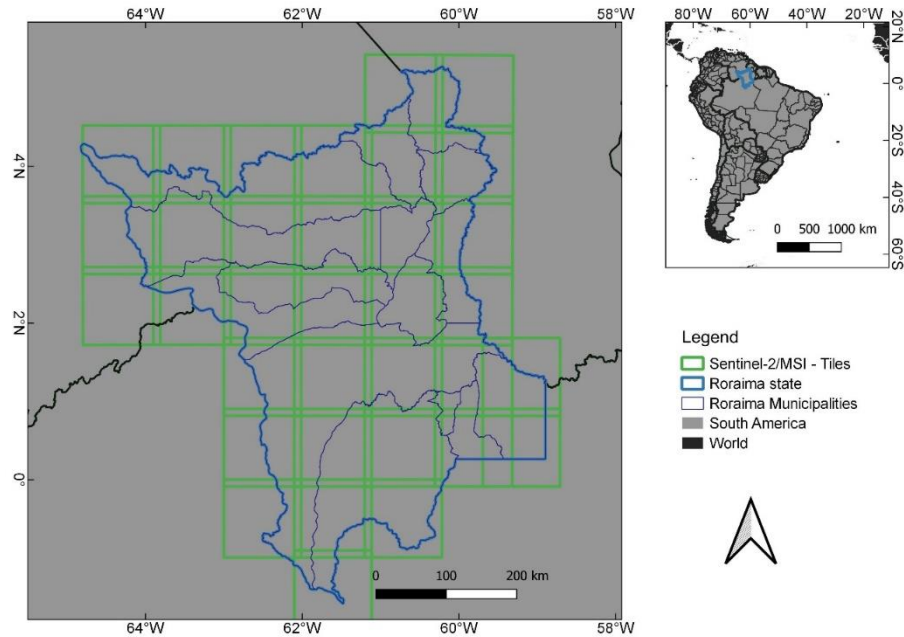
The Roraima state, located in the northern part of Brazil (Figure 5.3), is not one of the large Brazilian agriculture producers. However, this state shows large potential to crop expansion, mainly in the savanna's region. A detailed description of Roraima can be found in Chapter 3.

Figure 5.2 – Study area. The colored states represent states/provinces/departments with more than 1% of the South American cropland in 2015 (FAOSTAT, 2018).



Source: Author's production.

Figure 5.3 – Location of Roraima state locations in SA, including the Sentinel-2 MSI tiles over the state.



Source: Author's production.

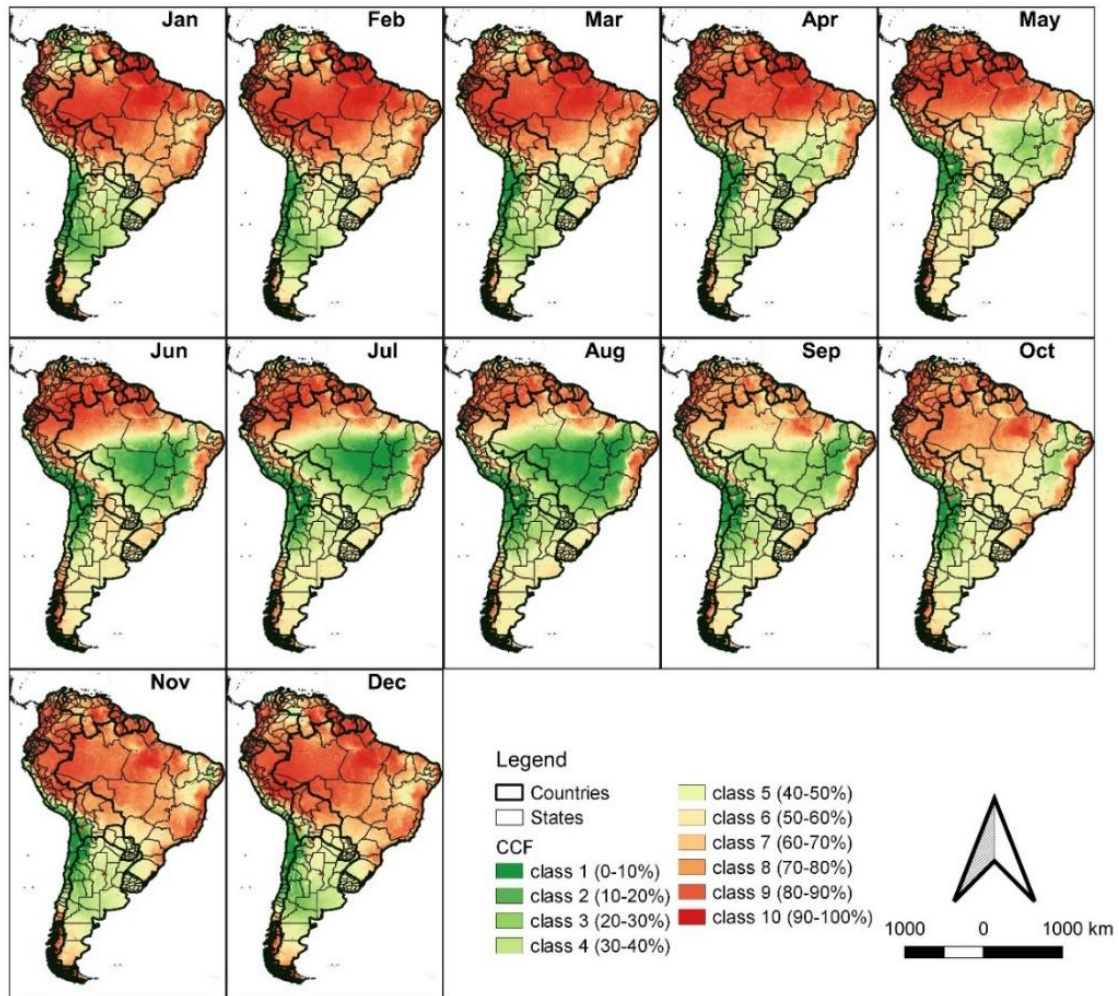
## 5.2.2 Cloud cover

MODIS sensors were launched aboard two NASA satellites: Terra, in December 1999 (10:30 LT descending node), and Aqua, in May 2002 (13:30 LT ascending node) (JUSTICE et al., 1998, 2002). MODIS provides near-daily global observations in 36 channels (0.415 – 14.235  $\mu\text{m}$ ), wide swaths (2230 km), and three spatial resolutions (250, 500, and 1,000 m). The suite of spectral bands has been used for the retrieval of several biophysical parameters and atmospheric properties. Recently, the MCD19A2 product derived from the Multi-Angle Implementation for Atmospheric Correction (MAIAC) algorithm has been released as part of the MODIS collection 6 product. It combined MODIS Terra and Aqua products containing atmospheric parameters at 1 km resolution, such as aerosol optical depth, columnar water vapor, and cloud cover. The MAIAC cloud mask implements a dynamic detection algorithm with a series of spectral thresholds, including thermal and cirrus bands. The time series and block-pixel analysis in MAIAC improve cloud screening. Several results have emphasized

the accuracy of the MAIAC cloud mask, especially in the Amazon region (HILKER et al., 2012; MARTINS et al., 2018a). A detailed description of the MAIAC cloud mask is found in Lyapustin et al. (2012, 2018). In this study, we used a daily MCD19A2 product, between 2000 and 2015, to obtain cloud cover across SA. The product was acquired in Level-1 and downloaded from the Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC) (<https://ladsweb.modaps.eosdis.nasa.gov>) and delivered as Hierarchical Data Format-Earth Observing System (HDF-EOS) file. The cloud mask was derived from the Quality Assurance (QA) layer of the MODIS Terra dataset for each tile (1200 x 1200 km) (see tiles in Lyapustin et al., 2018). Based on the daily files, we computed the monthly cloud cover fraction using 15-year MODIS products (Figure 5.4).

Sentinel-2 MSI has a bitmask quality band (QA60) that enables the identification of cloudy and cloud-free pixels with 60 m spatial resolution. QA band was used to analyze the cloud cover limitation over Roraima state during 2019. Bit 10 is for opaque clouds (0: no opaque clouds, 1: presence of opaque clouds), and Bit 11 is for cirrus clouds (0: No cirrus clouds are present, 1: cirrus clouds is present) (ESA, 2021).

Figure 5.4 – Cloud cover frequency considering MAIAC and MODIS.



Source: Author's production.

### 5.2.3 Cropland information

Cropland distribution and crop types across SA are shown in Figure 5.5. We obtained this information from the Worldwide Croplands project (ZHONG et al., 2017), in which global agricultural areas were mapped between 2000 and 2015 using Landsat 30-m data (THENKABAIL et al., 2012; TELUGUNTLA et al., 2015; ZHONG et al., 2017). In the Worldwide Croplands project, croplands are represented by cultivated and fallow areas. Cultivated areas are the lands with plants harvested for food, feed, and fiber, including both seasonal crops (e.g.,

wheat, rice, maize, soybeans, cotton) and continuous plantations (e.g., coffee, tea, rubber, cocoa, palm oil). Fallow-lands are defined as uncultivated lands during a season or a year but they remain farmlands and are equipped for cultivation, including plantations (e.g., orchards, vineyards, coffee, tea, rubber) (ZHONG et al., 2017; TELUGUNTLA et al., 2018). In the present study, we resampled the cropland product with the nearest method to the spatial resolution of 1km<sup>2</sup>, to match the spatial resolution of the cloud cover MCD19A2 product). In SA, Brazil presents approximately 42.1% of the croplands, followed by Argentina (25.3%), Uruguay (7.7%), Paraguay (6.1%), and Venezuela (5.0%) (ZHONG et al., 2017). Among the 242 SA states, six Argentinian states (Buenos Aires, Cordoba, Entre Rios, La Pampa, Santa Fe, and Santiago del Estero) and eleven Brazilian states (Mato Grosso, Mato Grosso do Sul, Minas Gerais, Goias, Sao Paulo, Rio Grande do Sul, Bahia, and Parana) stand out for having large, cultivated areas. Buenos Aires (27 Mha) in Argentina and Mato Grosso (23 Mha) in Brazil has the largest cropland extensions in SA (ZHONG et al., 2017).

According to the Food and Agriculture Organization (FAO), in 2015, the main crops cultivated in SA were soybeans, maize, sugarcane, wheat, rice, beans, coffee, cassava, sorghum, sunflower, cotton, barley, cocoa, and potatoes (FAOSTAT, 2018). These crops are differently distributed along the continent, as illustrated in Figure 5.5. South American cropland calendars, when existent, lack subnational growing season characterization and are not spatially explicit. Besides, they might be unreliable, due to unclear, out-of-date, and poorly documented data sources (WHITCRAFT; BECKER-RESHEF; JUSTICE, 2015b). Therefore, we used supplementary documents from different source levels (continent, national, and state) and created a summary of the agricultural calendar for the main crops cultivated in SA (Figure 5.6) (EMATER, 1983; SILVA NETO et al., 2001; OTSUBO; LORENZI, 2004; CARDOZO, 2009; FIALHO; ANDRADE; VIEIRA, 2013; INIA, 2014; MESQUITA, 2016; SILVA et al., 2017; SODRÉ, 2017; CONAB, 2017; OLIVEIRA NETO, 2017; AMIS, 2019; INTA, 2019). Based on the crop phenological plant cycle, sowing, and harvesting dates, we divided the year into four different intervals (periods).

**Period 1 – P1 (September-November):** represents the spring season in the southern hemisphere when there is the occurrence of soybean, maize, rice, and beans seeding; and also, the main wheat harvest and cotton seeding.

**Period 2 – P2 (December-February):** represents the summer season in the southern hemisphere when there is the occurrence of soybean, maize, rice, and beans harvest in the main croplands; and cotton seeding for some regions;

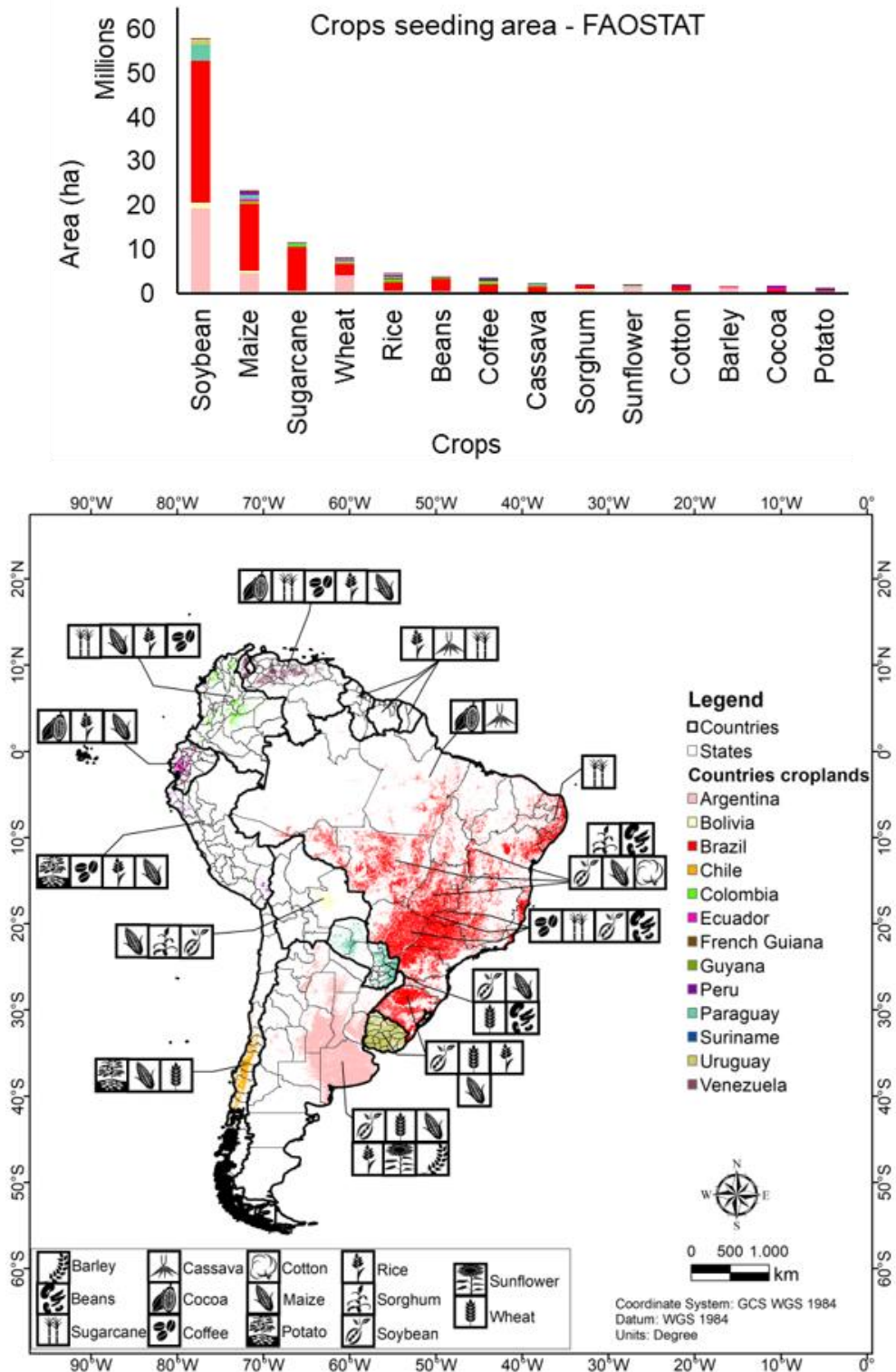
**Period 3 – P3 (March-May):** represents the autumn season in the southern hemisphere. It corresponds to the end of soybean and maize harvest over most South American croplands, and it is the period when the wheat sowing is concentrated;

**Period 4 – P4 (June - August):** represents the winter season in the southern hemisphere. It is when the wheat senescence and the beginning of harvest occur. For some regions, it corresponds to the beginning of the maize seeding.

For the analysis of perennial (cocoa and coffee) and semi-perennial (sugarcane and cassava) crops, we studied the four periods together. We did likewise for potato, an annual cycle crop that is planted and harvested year-round. For these crops, harvest happens in different quarters. Moreover, it is important to mention that the crop calendar in Roraima has harvest occurring during the off-season of the other Brazilian states (April-September). Thus, for Roraima, the periods are different (see Section 5.2.5 for details).



Figure 5.5 – Crop types distribution based on USGS and FAOSTAT data. Symbols: adapted from GEOGLAM.

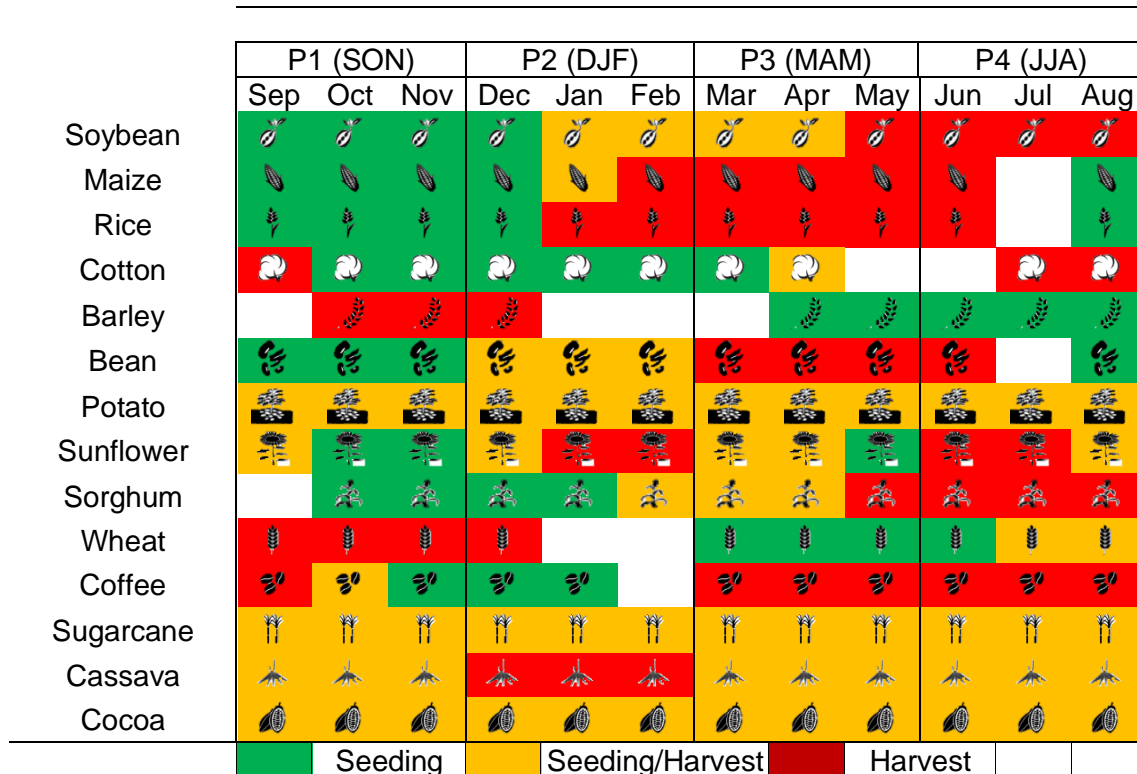


USGS: United States Geological Survey

Source: Author's production.



Figure 5.6 – Crop Calendar for the main crops cultivated in SA. Symbols: adapted from GEOGLAM.



Source: Author's production.

#### 5.2.4 MODIS cloud cover frequency

Using the monthly cloud cover over the entire SA territory (MCD19A2 product), we calculated the cloud cover mean (CCM) for each period (P1, P2, P3, and P4 – Section 5.2.3). Afterwards we sliced all CCM in percentiles (class 1: 0-10%; class 2: 10-20%; class 3: 20-30%; class 4: 30-40%; class 5: 40-50%; class 6: 50-60%; class 7: 60-70%; class 8: 70-80%; class 9: 80-90%; and class 10: 90-100%), creating a cloud cover frequency (CCF) over SA. The lower the CCF in one area, the bigger the chances of obtaining cloud-free ORS data for the area. To better understand the cloud cover pattern, we first analyzed the CCF over the entire SA Figure 5.2, and afterward, we focused on the cropland CCF areas Figure 5.5. To evaluate the most appropriate periods for monitoring each crop using ORS, we analyzed the CCF data together with the crop calendar data

(Figure 5.6) and the spatial distribution of crop areas (Figure 5.5). Therefore, for this step, we used the cropland areas to extract the CCF classes for each period. As Brazil and Argentina have the largest cropland extensions (67%) in SA, we analyzed the croplands CCF for each one of the 17 main agricultural state/province producers of these countries (Figure 5.2). Therefore, for these countries, we sliced the CCF into three interference classes:

- Low: CCF between 0% and 30%. In this class the CCF may have a low impact when using the ORS data to monitor the croplands;
- Medium: CCF between 30% and 70%. In this class, the cropland monitoring with ORS data could be more impacted, but its use may not be entirely restrictive;
- High: CCF between 70% and 100%. This is the worst class for crop monitoring with ORS data. There will possibly be just a few observations and monitoring with optical data may be restricted.

#### **5.2.5 QA60 cloud cover frequency**

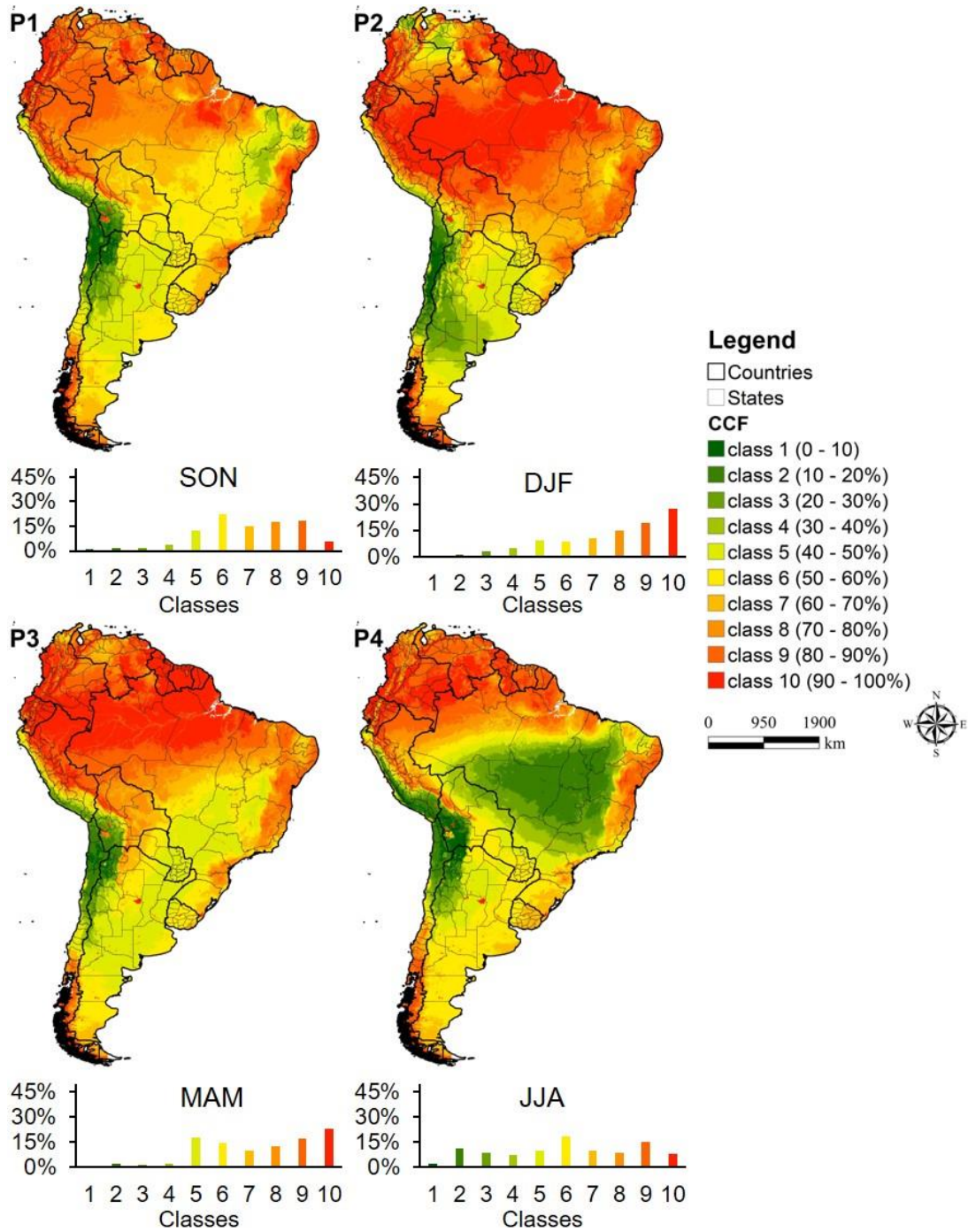
Considering the Roraima state (Figure 5.3), we calculate two cloud cover frequency products. First, we used the CCF from MODIS (Section 5.2.4) for a general approach. After, we created the QA60 Cloud Cover Frequency (QA60CCF) for 2019 and 2017. Using the bitmask QA60, inside the Google Engine platform (GORELICK et al., 2017), QA60CCF was calculated three different times. The first period (QP1) represents the months from January to April. The second period (QP2) represents the months from May to August. QP2 represents the annual crop season in Roraima state (for more information see Chapter 3). Finally, QP3 represents the months from September to December. Like CCM, QA60CCF was sliced in percentiles (class 1: 0-10%; class 2: 10-20%; class 3: 20-30%; class 4: 30-40%; class 5: 40-50%; class 6: 50-60%; class 7: 60-70%; class 8: 70-80%; class 9: 80-90%; and class 10: 90-100%).

## 5.3 Results

### 5.3.1 Cloud cover frequency – South America

CCF across SA depends on geographic location and periods (Figure 5.7). In P1, 73.5% of SA extension has CCF between class 5 and 9. For the austral summer (DJF), approximately 47% of SA territory presents CCF class 9 or 10, which highly limits the use of ORS data. In P3, there are fewer areas with CCF class 9 or 10 and more areas (32% of SA) with CCF between classes 5 and 6. The austral winter (JJA), compared to the other periods, is the one with the smallest area (23% of SA) with CCF above class 9, and with the largest area (29% of SA) with CCF under class 4. Considering the entire year, frequent cloud cover is observed on the eastern Brazilian coast, northern (as Roraima state), and extreme southwestern coast of SA. That means these regions have less chance of acquiring cloud-free ORS data. Moreover, the SA central region in P4 is one of the most favorable periods for the use of ORS data.

Figure 5.7 – CCF classes over the entire SA for four different periods.



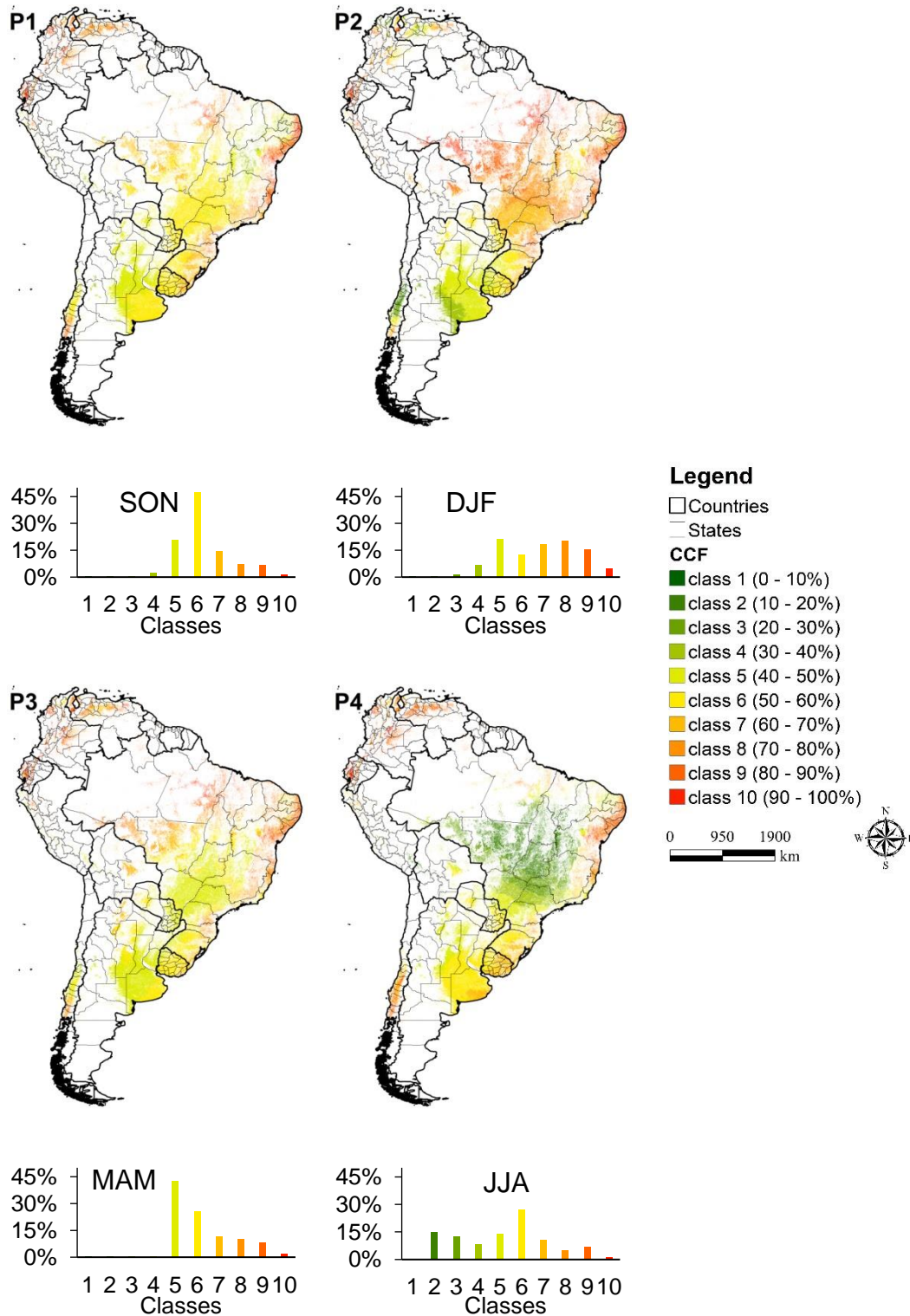
Source: Author's production.

### 5.3.2 Cloud cover frequency – South America croplands

Spatial distribution and CCF class concentration over SA croplands, for each period are presented in Figure 5.8. Details about cropland CCF over some of the main agricultural regions are shown in Figure 5.9 and Figure 5.10. In **P1**, almost 68% of the South American croplands have CCF classes 5 and 6. Argentinian croplands (Figure 5.9d) are concentrated (59%) in class 5, while approximately 52% of Brazilian (Figure 5.9a-c) and 60% of Bolivian (Figure 5.9c) croplands are in class 6. Uruguay and Paraguay have more than 75% of their croplands (Figure 5.9c) in class 6, and Chile has 40% of their croplands in the same CCF class 6 (Figure 5.9d). Venezuela and Colombia have more than 90% and 78%, respectively, of their croplands between CCF classes 8 and 9 (Figure 5.10). For **P2**, South American croplands are distributed among five CCF classes. The higher concentrations correspond to class 5 (21%) and 8 (20%). In this period, most Argentinian croplands (67%) correspond to class 5. Brazilian croplands are concentrated in classes 7 (27%) and 8 (32%), and Bolivian croplands in class 8 (36%) and 9 (44%). CCF remains between classes 4 and 6 in more than 90% of the Uruguayan croplands. Meanwhile, 73% of Paraguayan croplands have class 6. Chile is the only one with more croplands (54%) under CCF class 3. Venezuelan croplands are concentrated in class 5 (21%) and 8 (20%). Colombian croplands are concentrated in class 6 (24%) and 7 (25%).

In **P3**, Argentinian (71%), Brazilian (36%), Paraguayan (69%), and Chilean (41%) croplands are concentrated in CCF class 5, representing 42% of SA croplands. On the other hand, Bolivia has more croplands in class 7 (70%), Uruguay in class 6 (77%), and Venezuela between classes 7 (32%) and 8 (37%). For **P4**, SA croplands are concentrated between classes 4 and 6 (41% of croplands). However, in this period there is a higher cropland concentration (27%) between classes 1 and 3. Argentina and Paraguay have more croplands in class 7, with 67% and 62%, respectively. Brazilian croplands are concentrated in class 2 (25%) and 3 (21%). Croplands from Bolivia, Chile, Uruguay, and Venezuela are concentrated in classes 5 (45%), 8 (46%), 7 (58%) and 9 (69%), respectively.

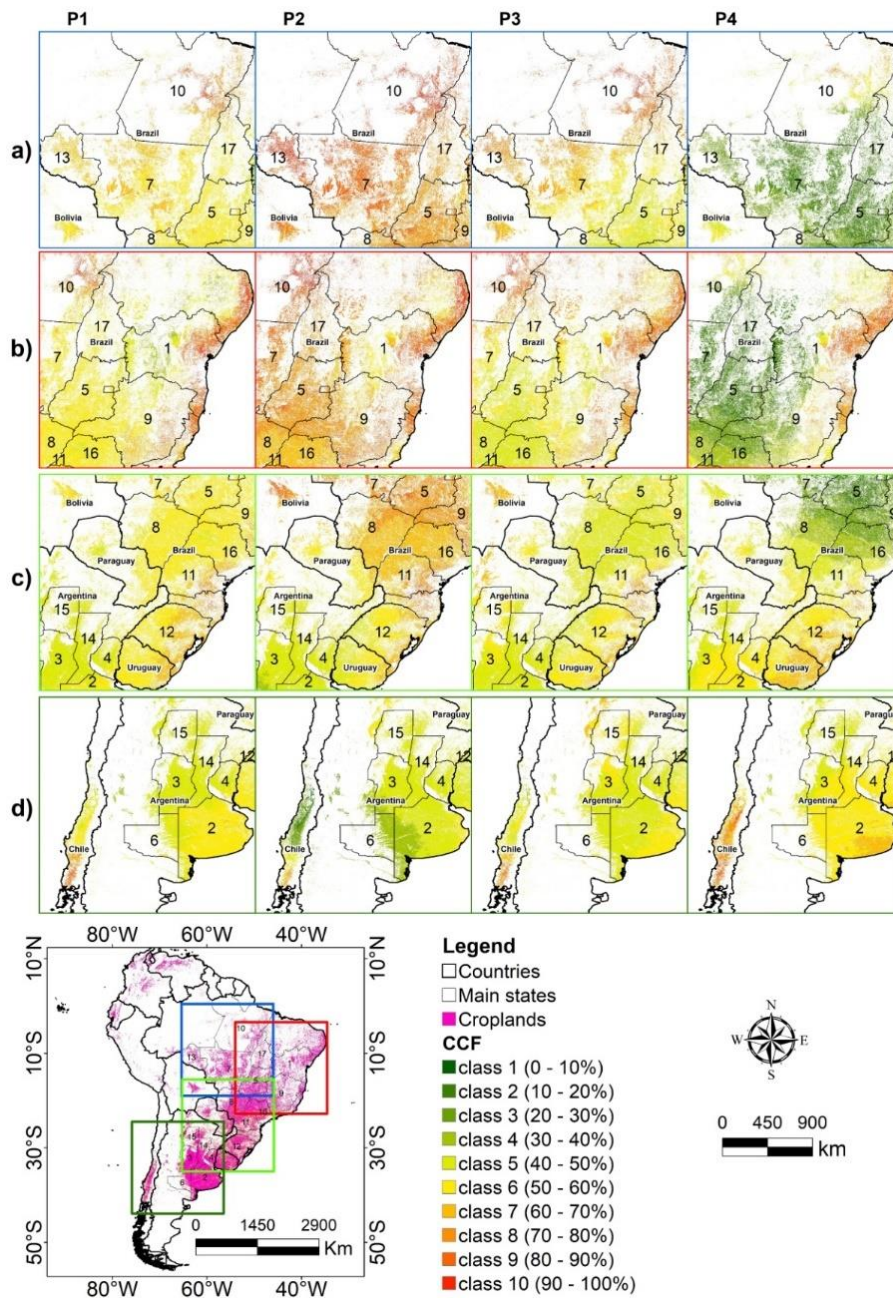
Figure 5.8 – Spatial and temporal CCF for croplands in each period.



Source: Author's production.



Figure 5.9 – Detail about the spatial and temporal CCF distribution over the Brazilian states croplands.



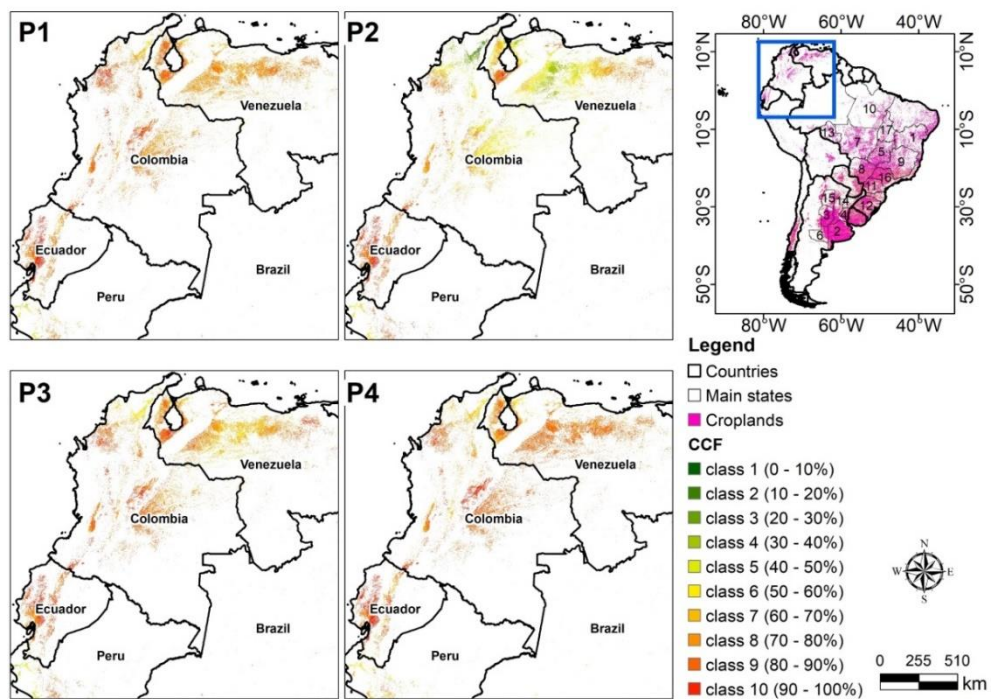
**a)** (13) Rondônia, (7) Mato Grosso, (10) Para, **b)** (1) Bahia, (9) Minas Gerais, (5) Goias, (17) Tocantins, **c)** (8) Mato Grosso do Sul, (16) São Paulo, (11) Parana and (12) Rio Grande do Sul; and Argentina states croplands of **d)** (2) Buenos Aires, (6) La Pampa, (3) Cordoba, (15) Santiago del Estero, (14) Santa Fe, and (4) Entre Rios.

Source: Author's production.

Summer crops comprise P1 and P2 in almost all SA croplands. During P1, CCF is between 40% and 60% for almost all croplands, not an ideal situation but

cloud-free images can still be obtained for some regions. In P2, the CCF occurs from class 5 to class 9 in the croplands. Winter crops comprise P3 and P4 in almost all SA croplands. For P3, CCF concentration over croplands is between classes 5 and 6, with a higher percentage in class 5. In P4, CCF decreases in the central region of SA, corresponding mainly to croplands located in the central-west parts of Brazil and of Bolivia.

Figure 5.10 – Detail about the spatial and temporal CCF distribution over Ecuador, Colombia, and Venezuela croplands.



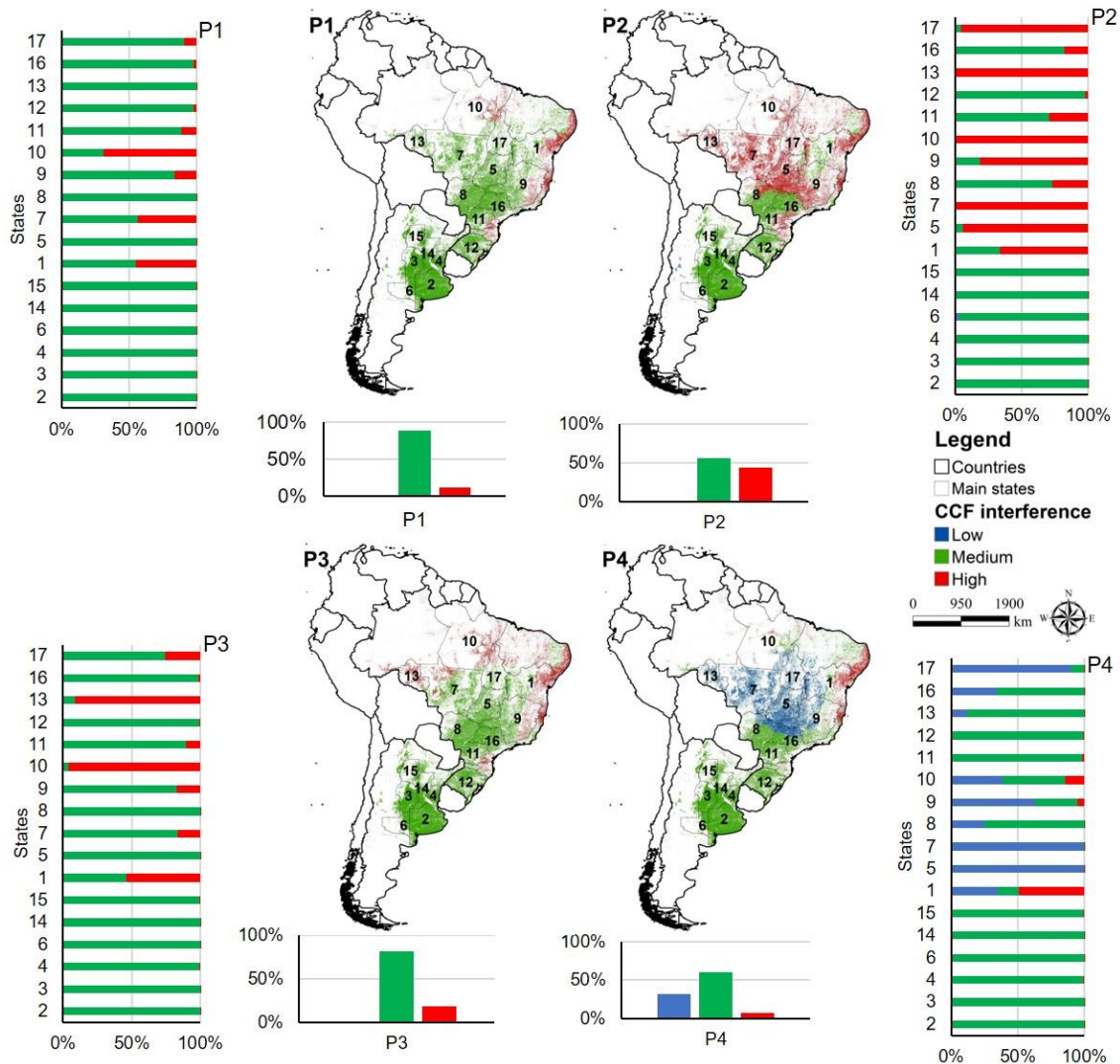
Source: Author's production.

We used CCF to create a CCF interference (Figure 5.11) to apply ORS data to cropland monitoring for each one of the 17 main Argentinian and Brazilian provinces/states. In **P1**, the Brazilian and Argentinian states have a predominance of CCF classes 5, 6, and 7, configuring medium CCF interference to cropland monitoring with ORS data. For the Brazilian states, medium CCF interference in class 6 is predominant in the Tocantins (62%), Goias (94%), Mato Grosso do Sul (94%), Parana (72%), and Sao Paulo (78%) croplands. Bahia state



has predominantly medium CCF interference in the western part, class 5 (24%), and high CCF interference in the eastern part, class 9 (30%). Mato Grosso state presents croplands in classes 6 (56%) and 7 (44%), configuring as medium CCF interference. Croplands in Rondonia states have the predominance of class 7 (82%), medium CCF interference. For Para croplands, we found that CCF interference is medium, class 7 (20%), for the southern region and high, for the rest of the state, classes 8 (34%), and 9 (28%). In Minas Gerais and Rio Grande do Sul croplands, the predominant classes are 6 (51% and 48%, respectively) and 7 (20% and 50%, respectively), defined as medium cloud cover interference. Moreover, for Argentinian provinces, we defined CCF interference as medium for all provinces, where Córdoba, Entre Rios, La Pampa, Santa Fe, and Santiago del Estero have most croplands classified as CCF class 5, with 96%, 75%, 84% and 97% of croplands, respectively. Only in Buenos Aires are croplands predominant (82%) in class 6. In **P2**, medium (CCF classes 5, 7) and high CCF interference (class 8) are predominant in Brazilian states and Argentinian provinces croplands. At the Brazilian state level, we found that a high CCF interference (class 8), predominates in Goias (87%), Tocantins (67%), and Minas Gerais (61%) croplands. Bahia has high CCF interference and its croplands are divided between CCF classes 8 (29%) and 9 (33%). For Mato Grosso croplands class 9 (68%) are predominant and for Para and Rondonia class 10 (61% and 70%, respectively), both with high CCF interference for the use of ORS data. We found medium CCF interference predominantly in class 7 for Mato Grosso do Sul (70%), Parana (66%), and Sao Paulo (81%) states, and in classes 6 (51%) and 7 (28%) for Rio Grande do Sul croplands. Similar to P1, we found medium CCF interference for all Argentina provinces, with CCF class 5. In Buenos Aires (65%), Cordoba (86%), Entre Rios (99%), Santa Fe (99%) and Santiago del Estero (79%) croplands class 5 predominates, while, in La Pampa croplands, class 4 is predominant (96%). Therefore, there are limitations to using ORS data for crop monitoring in these areas, mainly for P2, in the Brazilian central-west croplands.

Figure 5.11 – CCF interference for cropland monitoring with ORS data for each period, for Argentina provinces.



(2) Buenos Aires, (3) Cordoba,(4) Entre Rios (6) La Pampa, , (14) Santa Fe and (15) Santiago del Estero, and Brazilian states of (1) Bahia, (5) Goias, (7) Mato Grosso, (8) Mato Grosso do Sul, (9) Minas Gerais, (10) Para, (11) Parana, (12) Rio Grande do Sul,(13) Rondonia, (16) Sao Paulo and(17) Tocantins

Source: Author's production.

In **P3**, there is an overall predominance of CCF class 5 among Brazilian and Argentinian states croplands, with medium CCF interference. Goias, Mato Grosso do Sul, Parana, and Sao Paulo have most of their croplands with CCF class 5, with 72%, 93%, 59%, and 84% respectively. Croplands inside Bahia state are predominant in classes 8 (25%) and 9 (29%), and Para and Rondonia have a predominance of classes 9 (54%) and 8 (67%), configuring high CCF

interference. On the other hand, we found medium interference for Minas Gerais, classes 5 (37%) and 6 (33%), Mato Grosso, classes 6 (34%) and 7 (39%), Tocantins, class 6 (41%), and Rio Grande do Sul, class 6 (59%). Buenos Aires, Cordoba, Entre Rios, La Pampa, and Santa Fe Argentinian provinces have croplands with a predominance of class 5, with 70%, 70%, 95%, 99%, and 98% respectively. Only Santiago del Estero shows most croplands (85%) in class 6. Hence, we configured the Argentinian provinces as medium CCF interference. In **P4**, the majority of Brazilian states have low CCF interference. Goias (76%) and Mato Grosso (72%) croplands have class 2 of CCF. Similarly, in Minas Gerais and Tocantins class 3 predominates, with 59% and 89% respectively. For croplands in western Bahia, we found low CCF interference, class 2 (27%), however, for the eastern part, we found high CCF interference, class 9 (31%). Para croplands are divided among several classes, mainly class 2 (20%) and 3 (18%). Class 4 predominates in Rondonia croplands (71%), defined as high CCF interference. For part of Sao Paulo, class 3 (34%), Mato Grosso do Sul, class 3 (42%), we found low CCF interference. However, for Parana, class 5 (63%), and other parts of Sao Paulo, class 4 (52%), and Mato Grosso do Sul, class 5 (33%), we found medium CCF interference. Rio Grande do Sul is more similar to Argentina, with a cropland concentration in classes 6 (56%) and 7 (44%), characterizing a medium CCF interference. In Argentinian states, Buenos Aires (80%), Entre Rios (98%), and Santa Fe (99%) have the most croplands in class 6. Santiago del Estero has croplands predominantly (99%) in class 5. Cordoba has its croplands divided between classes 5 (42%) and 6 (58%). Thus, it is possible to see that in almost all Brazilian croplands the limitation to use ORS data for crop monitoring is lower than that of previous periods. Argentinian croplands remain in a similar condition.

In almost all Brazilian states, where soybean and maize are grown (considering only the first-crop – summer crop), the monitoring using ORS data is hindered by the medium to high CCF interference. During the seeding period of these crops in P1, concentrated at the end of September and October, a medium CCF interference is predominant in the Brazilian and Argentinian croplands. For the maximum vegetative development and harvest periods,

encompassing P2, which goes until March (beginning of P3), we found high CCF interference in the Brazilian central -western states. This cloud pattern decreases the chance of obtaining cloud-free ORS data during summer (P2), corresponding to the maximum development of these crops. However, Brazilian state Rio Grande do Sul and Argentinian provinces have a medium cloud interference for monitoring crops, owing to the fact that cloud frequency values are almost constant between classes 5 and 6. It is important to emphasize that even with a high CCF interference some “windows” might occur when crops may be identified (i.e., mapped).

Wheat is cultivated mainly in P2, it is sown between March and June (P2 and beginning of P3) and harvested from July to December (P3 and beginning of P4). Rio Grande do Sul and Parana Brazilian states and Argentinian provinces are the greatest wheat producers in SA. These regions have medium CCF interference; therefore, ORS data might be used to provide wheat monitoring in part of the cycle. Rice seeding season occurs between September to December (P1 and the beginning of P2), and its maximum development is near P2. We found medium CCF interference for the biggest rice producers, Argentinian provinces and Rio Grande do Sul state. Bean production is concentrated in Minas Gerais and Paraná states in Brazil. When cultivated during the first-crop, its maximum development occurs in November and December (end of P1 and beginning of P2). We found medium CCF interference during P1, and high CCF interference during P2, for the bean cropland locations in both states. Thus, in general, there is medium to high CCF interference for the monitoring of beans with ORS data.

The use of RS data to monitor crops with perennial and/or semi-perennial cycles is more feasible because these crops remain in the field for a longer period, increasing the probability of obtaining cloud-free ORS data during plant development. The fact that these crops remain in the field for a long time allows the acquisition of cloud-free images at different times. However, in some South American croplands, there is a high CCF interference throughout the year, affecting even the perennial and/or semi-perennial crop monitoring. In general, we found medium CCF interference in the monitoring of sugarcane (semi-

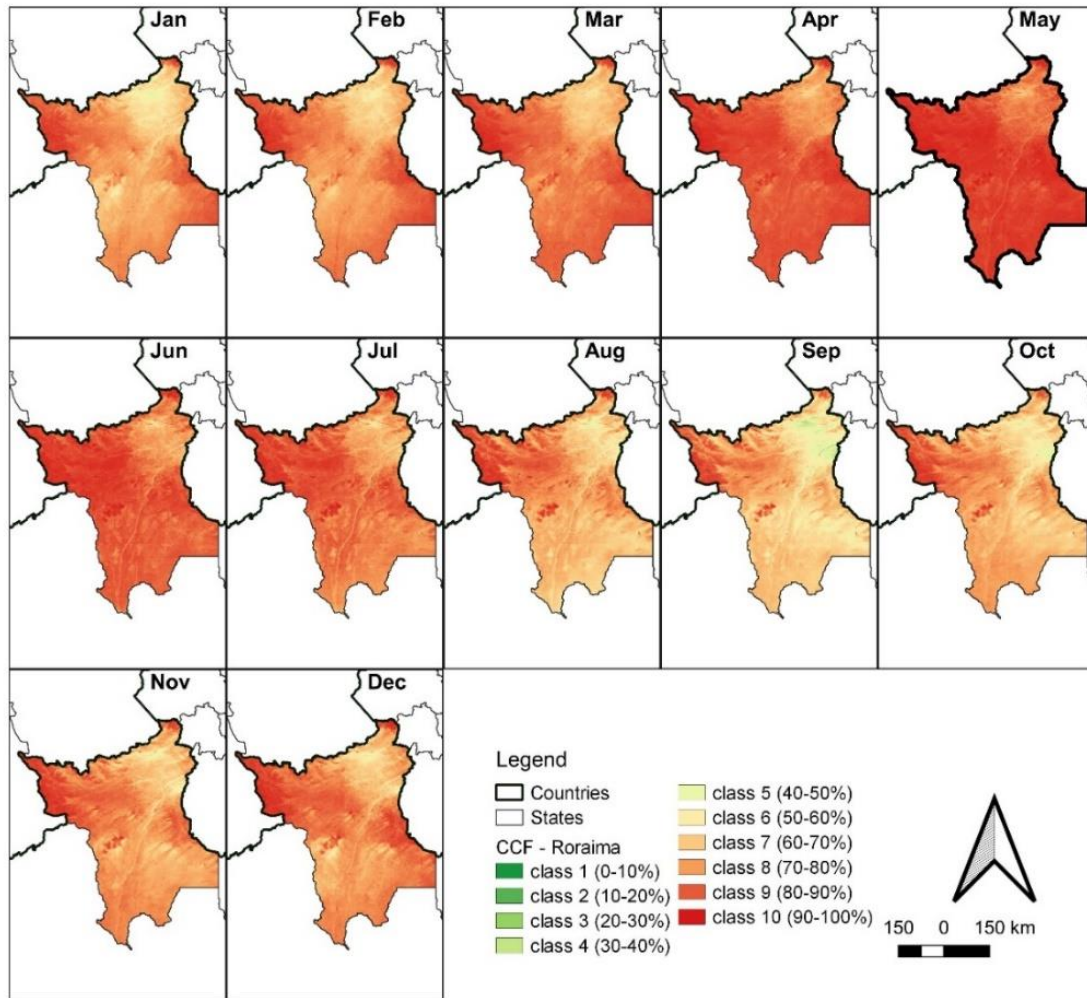
perennial) and coffee (perennial) crops in Brazil using ORS data. However, for specific regions, such as the eastern part of Bahia state, the persistent cloud frequency throughout the year, caused probably due to sea breeze (SILVA; KOUSKY, 2012), resulted in a high CCF interference when using ORS data.

### **5.3.3 Cloud cover frequency – Roraima case**

SA northern region, near Equator line, has a frequent CCF. Roraima state (Figure 5.12), located in this region, show monthly CCF of over 70% for the entire year. From April to July, inside the QP2, this limitation is higher. During this period, it is almost impossible to obtain cloud-free data. On the other hand, the months of September and October have the lower CCF, which means more probability to obtain ORS cloud-free images. Regions with high elevation (Figure 5.3), as the Roraima Mountain on the northern part, have the higher CCF.

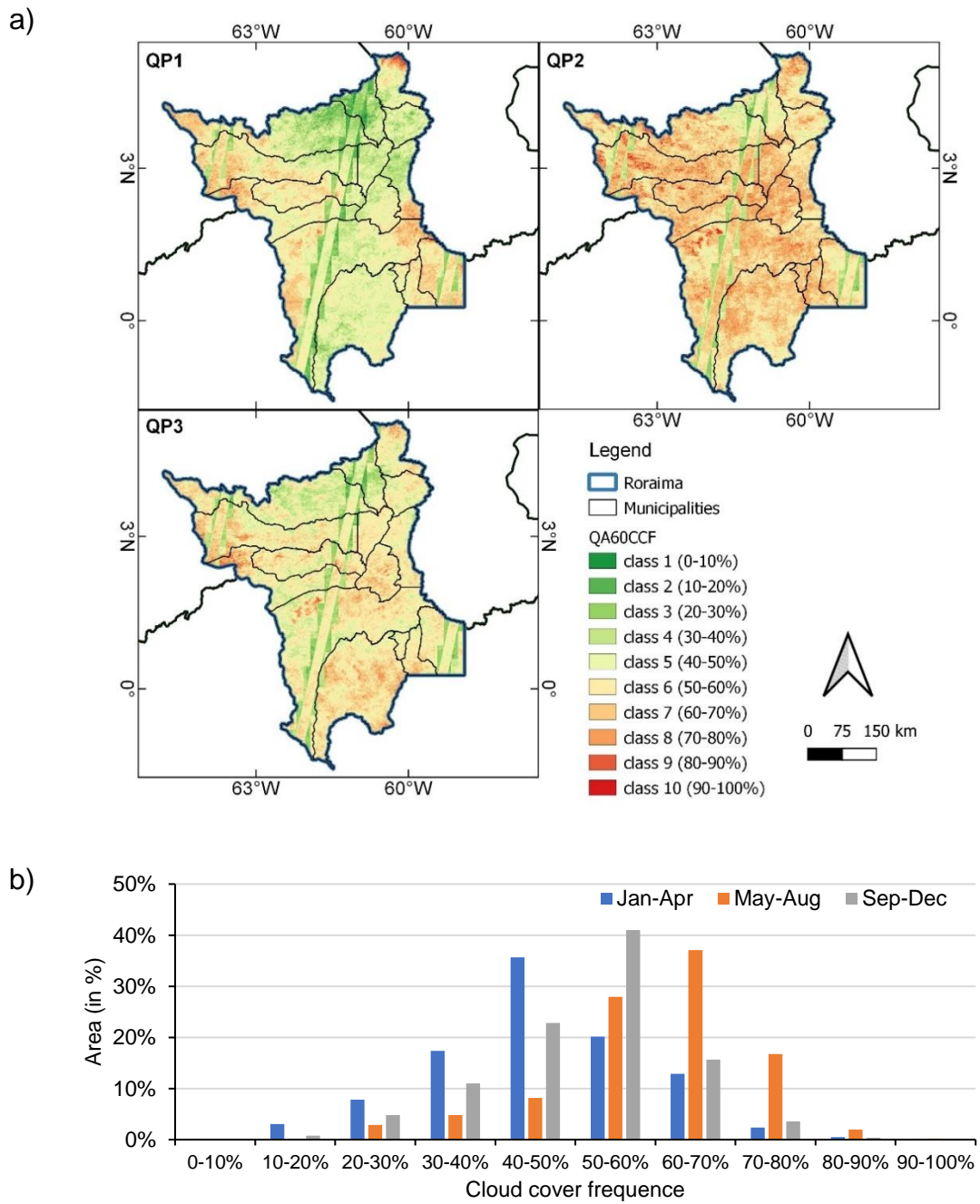
According to the QA60CCF (Figure 5.13), for the QP2 there are a few cloud-free observations for the Sentinel-2 MSI data, with 56% of Roraima territory with more than 60% of QA60CCF. QP1, on the other hand, has de lower QA60CCF, showing a better scenario of using ORS data with more than 64% of Roraima with less than 50% of QA60CCF. During the QP3, 63% of Roraima territory presented QA60CCF between 40% to 60%. For 2017 (Figure 5.15), QP1 presented the worst scenario, with more than 60% of Roraima with QA60CCF higher than 70%.

Figure 5.12 – Cloud cover frequency considering MAIAC and MODIS for Roraima state.



Source: Author's production.

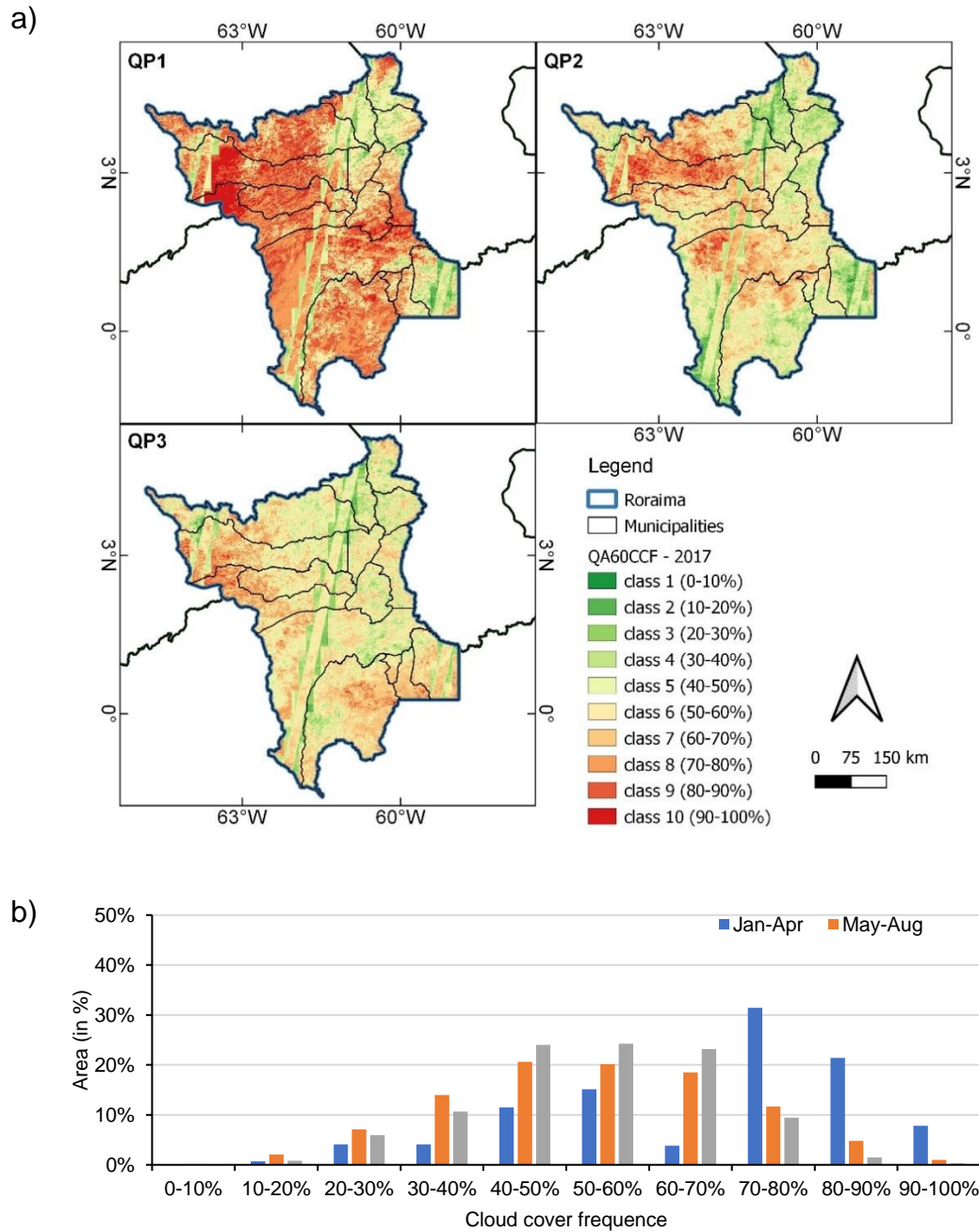
Figure 5.13 – QA60 Cloud cover frequency considering Sentinel-2 MSI optical data over three different periods of 2019 (QP1 – Jan-Apr, QP2 – May-Aug, QP3 – Sep-Dec).



Source: Author's production.



Figure 5.14 – QA60 Cloud cover frequency considering Sentinel-2 MSI optical data over three different periods of 2017 (QP1 – Jan-Apr, QP2 – May-Aug, QP3 – Sep-Dec).



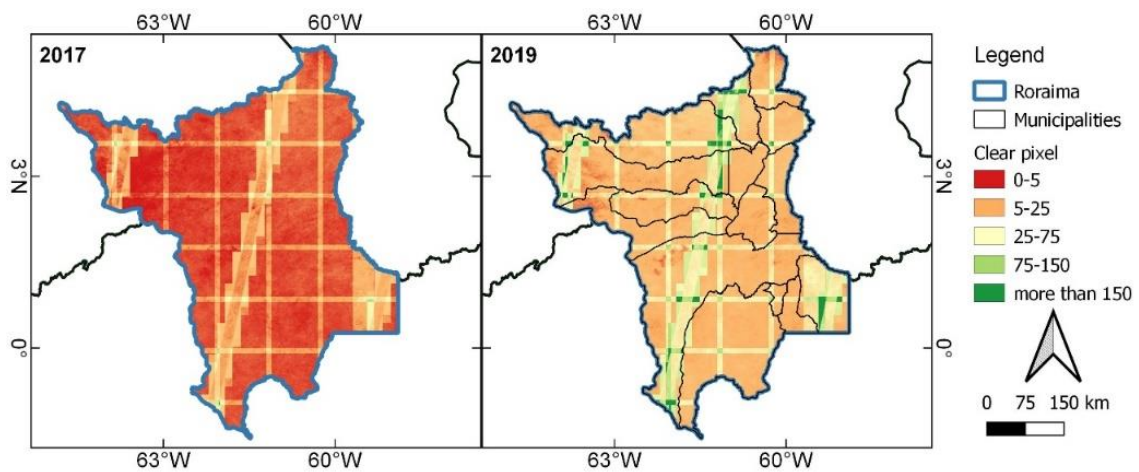
Source: Author's production.

The year 2017 has less free-cloud observation than 2019 (Figure 5.16) because MSI/Sentinel-2B was launched in June 2017, restricting the free-cloud



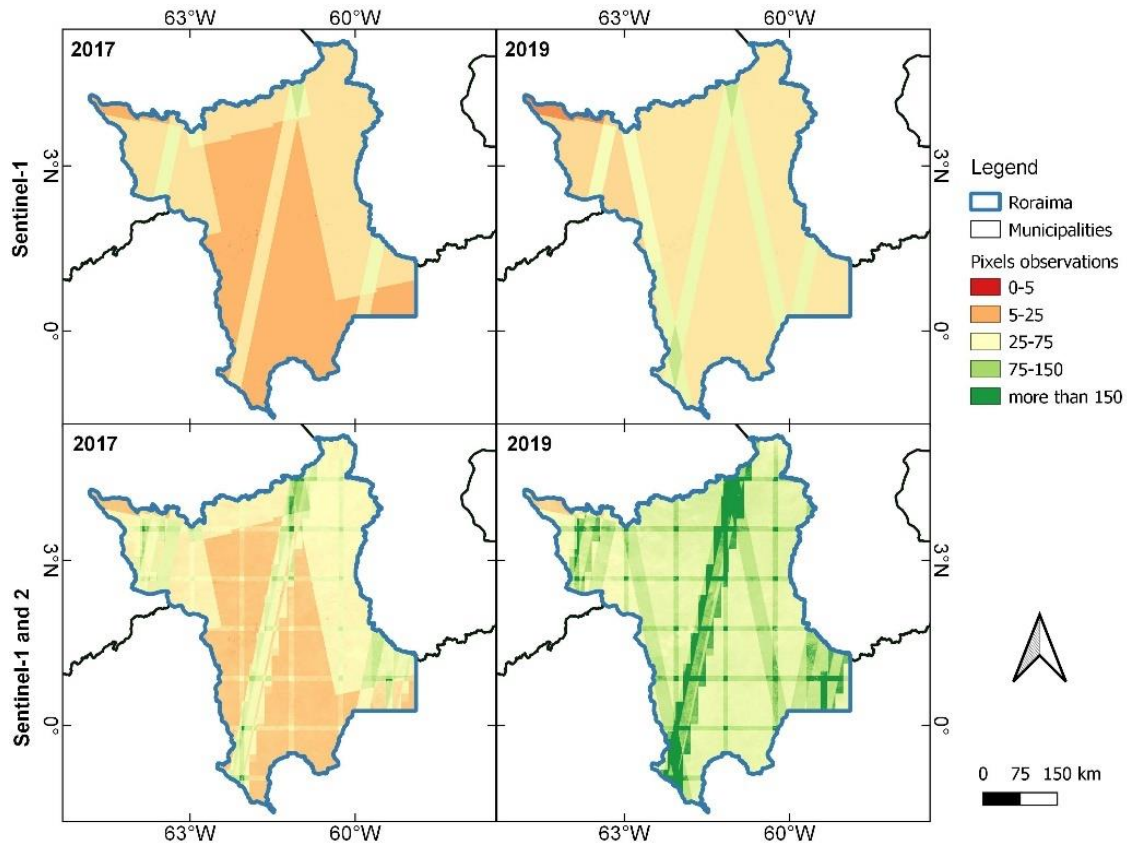
observation mainly in QP1 for 2017 (Figure 5.15). When combining Sentinel-1 and Sentinel-2 MSI it is possible to improve the probability to obtain useful data for 2017 (Figure 5.17). For 2019, the combination of Sentinel-1 and Sentinel-2 (Figure 5.17), produces a dense time series.

Figure 5.15 – Sentinel-2/MSI optical cloud-free cover observation data for 2017 (left map) and 2019 (right map).



Source: Author's production.

Figure 5.16 – Sentinel-2/MSI optical cloud-free (top images) and Sentinel-1 (bottom images) data observations over 2017 (left images) and 2019(right images).



Source: Author's production.

## 5.4 Discussion

Our results have shown that the use of ORS data for earth observation is a challenge in some parts of SA, due to frequent cloud cover. We found a frequent CCF over the Amazon and Brazilian coastal regions, making it very difficult to obtain ORS cloud-free data (Figure 5.7). This limitation is aggravated in the Equatorial zones (between 15° N and 14 °S), where there is a higher CCF when compared to that of regions with latitudes between 15°N and 40° N. This is due to the South Atlantic Convergence Zone – SACZ and Intertropical Convergence Zone – ITCZ (WHITCRAFT et al., 2015b), which reflect in the distribution of clouds and precipitation throughout the year (GRIMM; BARROS; DOYLE, 2000; GRIMM, 2003). In this sense, Asner (2001), Wulder et al. (2015), and Martins et al. (2018a) demonstrated this difficulty in obtaining cloudless Landsat data

available for tropical regions. We found that it is challenging to use optical sensors even in the months with lower cloud frequency, P4, which represents medium CCF interference (CCF near 40-50%) in most of SA, as we showed in Figure 5.11. Similar to our results, Sano et al. (2007) observed that it is less likely to obtain cloud-free Landsat images over Brazilian tropical Savannas (e.g. Brazilian central part), from October to March, corresponding to our P1 and P2. In contrast, it is more likely to obtain cloud-free Landsat images in July and August, our P4.

Agricultural monitoring, temporally dynamic targets, via ORS is challenging. This is because most croplands are cultivated in the rainy season, with an annual crop vegetation cycle and a higher cloud frequency. Whitcraft et al. (2015b) discussed the challenges for worldwide cropland monitoring. According to these authors, for annual crops, the major limitations are also from the beginning to the middle of the crop cycles. To overcome it and provide an 8-day image with 95% of the cloud-free sky of the croplands, Whitcraft et al., (2015a) defined that the global revisit frequency required is 1.7 – 2.6 days. The spatial distribution of South American croplands is heterogeneous. Brazil, Argentina, Uruguay, Paraguay, and Venezuela, concentrated 86.2% of the cropland areas (ZHONG et al., 2017). Moreover, the annual crops in these countries are cultivated in different periods. In general, to provide an 8-day revisit with more than 70% of cloud-free observation in South American croplands, Whitcraft et al. (2015b) mentioned that a 2-day temporal resolution is required.

In our results we noted that the low CCF interference occurs at the end of crop cultivation, end of P2, or out of the crop cycle, P3, in the Brazilian central-western cropland regions. In this region, during the P2 the high CCF interference is predominant. Therefore, is unlikely to have proper soybean and corn monitoring with ORS data for this region. Sano et al. (2008) also mentioned that due to less cloud-free ORS data between October and March (corresponding to ours P1 and P2), it is more difficult to provide an appropriate summer crop monitoring in the Brazilian central-western regions (Mato Grosso, Mato Grosso do Sul, Goias, Minas Gerais, Tocantins and west of Bahia states).

According to our study, Argentinian and southern Brazilian croplands have medium CCF interference during the entire year and represent a possible limitation for the annual crop monitoring system using ORS sensors. For the summer crops, we found that Parana and Sao Paulo as higher CCF interference than the Rio Grande do Sul and Argentina. Consequently, use ORS data to crop monitoring (mainly soybean, corn, and sugarcane) is more challenge in these states. However, for the winter crops (mainly wheat) Argentinian and southern Brazilian croplands have similar CCF patterns. In this context, is likely the winter crop monitoring in these areas. Eberhardt et al. (2016) described the difficulties for agricultural monitoring with ORS data in some Brazilian states (Parana, Rio Grande do Sul, and Sao Paulo). Likewise, these authors found limitations due to CCF interference for crop monitoring, such as soybean, and maize, during the quarter DJF, our P2, which corresponds to the main crop growing period. In contrast, these authors mentioned that the monitoring of winter crops, our P3 and P4, in the Parana state are more likely to be successful when compared to that of the summer crops in these same areas. Moreover, Eberhardt et al. (2016) reported that in Sao Paulo and Parana states, cloud coverage increased from September to March (our P1 and P2), corresponding to the spring and summer crops. Besides, for these authors, Rio Grande do Sul state had a stable and high cloud cover throughout the year. These patterns are like those found in our research. Sugawara et al. (2008) mentioned that it was even impossible to map the soybean crop in the whole Parana state between the years 2000 and 2007. This occurred even when using Landsat-5 Thematic Mapper (TM) and Landsat-7 Enhanced Thematic Mapper Plus (ETM+) data, with an 8-day temporal resolution. The authors found fewer cloud-free images from December to February, the main quarter for summer crops. King et al. (2017) using a stratified approach, found limitations to the estimated soybean in Argentina in 2013/2014. However, these authors cited that cloud cover remains a limiting factor in tracking crop phenology, and even mapping, rainfed crop types with time-series ORS data. They missed the soybean greenup and senescence monitoring in some areas, due to the cloud cover presence in Landsat 7 and 8 data, January 22nd, and February 15th, and after March 10th.

Sugarcane cropland monitoring in the Brazilian coastal region using only ORS data is rather complex. These areas have high CCF interference and are unlikely to obtain cloud-free data. However, for the areas with sugarcane in Sao Paulo and northern Parana, the CCF interference is medium and has lower CCF for the P3 and P4. According to Aguiar et al. (2011), sugarcane crop cultivation in Sao Paulo state tends to reach its maximum vegetative development in April, when the harvest season starts and continues until December. Rudorff et al. (2010) mentioned that the images from January to April better provide the spatial distribution of sugarcane in Sao Paulo state. According to these authors, images from January and February (our P2) have more probability to have clouds than images from March to April (our P3), as observed in our study for Sao Paulo croplands. Aguiar et al. (2011) mentioned that images over the entire harvest period (April to February) are necessary to define the harvest type. These authors analyzed the sugarcane harvest type for Sao Paulo state between the years 2006 and 2011, using ORS data from satellite Landsat-5 TM. They found that TM images often presented clouds that precluded image classification, even when most of the sugarcane harvesting is performed during the dry season with lower cloud frequency and is relatively favorable for acquiring cloud-free ORS data.

Considering the CCF interference over each crop in each state, we created a table to summarize the feasibility of ORS data use. It is important to highlight that this summary is to provide accurate crop monitoring and does not reflect the crop mapping likelihood. We defined three situations:

- Very likely: low CCF interference during almost the entire vegetative crop cycle;
- Likely: medium CCF interference in most of the vegetative crop cycle;
- Unlikely: high CCF interference in most of the vegetative crop cycle.

Table 5.1. Summary of ORS data use feasibility: very likely (+), likely ( $\pm$ ), and unlikely (-) considering the main crops in the main producer states of Brazil and Argentina: Regions, where the crop is inexpressive, are blank.

		Soybean	Maize	Wheat	Rice	Beans	Sugarcane	Coffee
Brazil	MT	-	-					
	MS	$\pm$	$\pm$					
	MG	$\pm$	$\pm$			$\pm$	+	+
	GO	-	-				+	
	SP	-	-				+	+
	RS	$\pm$	$\pm$	$\pm$	$\pm$			
	BA <sub>1</sub>	-	-				-	
	PR	-	-	$\pm$		$\pm$	+	
	PA	-	-					
	RO	-	-					+
	TO	-	-					
Argentina	BA <sub>2</sub>	$\pm$	$\pm$	$\pm$	$\pm$			
	CO	$\pm$	$\pm$	$\pm$	$\pm$			
	SF	$\pm$	$\pm$	$\pm$	$\pm$			
	ER	$\pm$	$\pm$	$\pm$	$\pm$			
	LP	$\pm$	$\pm$	$\pm$	$\pm$			
	SE	$\pm$	$\pm$	$\pm$	$\pm$			

Legend					
+	Very likely	+ -	Likely	-	Unlikely

MT: Mato Grosso; MS: Mato Grosso do Sul; MG: Minas Gerais; GO: Goias; SP: Sao Paulo; RS: Rio Grande do Sul; BA<sub>1</sub>: Bahia; PR: Parana; PA: Para; RO: Rondonia; TO: Tocantins; BA<sub>2</sub>: Buenos Aires; CO: Córdoba; SF: Santa Fe; ER: Entre Rios; LP: La Pampa; SE: Santiago del Estero.

Source: Author's production.

Roraima state has more frequent cloud-cover during the QP2 period, impacting the use of Sentinel-2-MSI for crop season. Pavanelli et al. (2018) mentioned that due to the frequent cloud cover, mainly during the crop seasons, it is almost impossible to obtain clear sky observations with optical data in this region. The study of Martins et al. (2018a) also illustrated how difficult is to have cloud-free data in this Amazon region. Thus, for the LULC mapping process, ORS data from QP1 and QP3 could be not enough, because the Annual crops will be not well representing. Besides, the Roraima QA60CCF could be characterized as medium interference for ORS optical data use during the entire 2019. Also, it is important to highlight that during 2019 there are Sentinel-2A and 2B, but before June 2017, data from 2B was not available, resulting in a temporal resolution of 10 days and consequently less data available. That increase the proportion of cloudy pixel and restricting more the ORS cloud-free data for 2017.

One possible way to overcome the cloud limitation in areas with medium cloud cover interferences is to combine different optical sensors, creating a hypothetical constellation (WHITCRAFT et al., 2015a; WULDER et al., 2015), to improve ORS temporal resolution during agricultural seasons. However, in regions with high cloud frequency during the phenological crop cycle, it is difficult to create an agricultural monitoring system based entirely on optical sensor data. Even with a better temporal resolution, there could be gaps without ORS data. Thus, one alternative for these areas is to use microwave Synthetic Aperture Radar (SAR) (WHITCRAFT et al., 2015a) which is less affected by cloud cover. SAR data could be used alone or combined with ORS. For Roraima state, the use of SAR data, combining with an optical sensor, help to improve the time-series data that can be used in the Land Use and Land Cover (LULC) classification.

## **5.5 Conclusion**

We evaluated the implications of CCF for monitoring agriculture in SA and LULC in Roraima state using ORS. SA presents limitations for the use of continuous ORS data due to high cloud frequency. This is more evident in equatorial zones (Amazon), as Roraima example, and Brazilian coastal regions,

where the frequent cloud cover limits the use of ORS data even for LULC mapping. Agricultural monitoring with ORS data, during the rainy season, QP2 for Roraima, and P1 and P2 (September-February) for others Brazilian states, becomes difficult in practically the entire territory. In P4 (June-August), we observed a greater probability of obtaining cloud-free ORS data from the middle to the end of the crop cycle, in the central part of SA. Soybean and maize are the crops with the highest CCF interference for ORS monitoring. Wheat, rice, and beans have in general medium CCF interference for the use of ORS data, thus making it possible to monitor part of the crop cycle. For perennial and semi-perennial crops, like sugarcane and coffee, that remain in the field throughout the whole year, there is a better chance of obtaining some cloud-free images for cropland monitoring. However, for the croplands on the Brazilian coast, the monitoring of sugarcane and coffee crops with ORS data has high CCF interference. In Roraima, the frequent cloud cover limits the use of ORS, even to monitor small parts of the crops cycle. The few cloud-free observations in Roraima are a big challenge to provide the LULC mapping or monitoring. The development of sensors with a better temporal resolution, the use of microwave sensors, and the combination of optical and microwave sensors are presented as methods to overcome the limitations of data availability for the LULC monitoring in Roraima state and the monitoring of the agricultural croplands highlighted in this study.



## **6. MULTISENSOR APPROACH TO LAND USE AND LAND COVER MAPPING IN NORTHERN AMAZON, BRAZIL<sup>2</sup>**

This chapter described how we choose the approach to provide the LULC maps. It was used in a small area to test the methodology and verify the importance of the use of SAR data.

### **6.1 Introduction**

Nowadays, combining food production with minimizing the environmental impact is one of the biggest challenges to be faced by the agriculture sector, government, and researchers (SEARCHINGER et al., 2015; ESTES et al., 2016). Part of this food demand is met by agriculture expansion, resulting in severe environmental impacts (LAMBIN; GEIST; LEPERS, 2003), and contributing to global climatic changes (DENG; ZHAO; YAN, 2013). In this sense, Land Use and Land Cover (LULC) mapping is fundamental to land use management and to understand the environmental effects at local, regional, and global scales (ADAMI et al., 2018; PAVANELLI et al., 2018). Thus, Remote Sensing (RS) technology is widely utilized for synoptic and continuous LULC monitoring, allowing identification of the LULCC (WULDER et al., 2015; VELOSO et al., 2017).

Traditionally, ORS data is used to map and characterize LULC, but it suffers limitations due to cloud cover (ASNER, 2001; WULDER et al., 2015; MARTINS et al., 2018a). High cloud frequency is a persistent difficulty found in the Amazon region, which highly compromises LULC mapping based on ORS data (MARTINS et al., 2018a; PRUDENTE et al., 2020a), especially in agricultural areas whose growth occurs during the rainy season (PRUDENTE et al., 2020a). Cloud cover, along with landscapes fragmentations and transitions among vegetation types (LU et al., 2012; LAURIN et al., 2013), and quick LULCC (WHITCRAFT et al., 2015b; EBERHARDT et al., 2016) make the use of ORS data challenging to map LULC in Amazon regions (SANO et al., 2007;

---

<sup>2</sup> A paper based on this chapter is currently under review at the ISPRS Journal of Photogrammetry and Remote Sensing journal

WHITCRAFT et al., 2015b; EBERHARDT et al., 2016; ZHANG; SKAKUN; PRUDENTE, 2020).

In this sense, one option to overcome the cloud cover limitation is the use of Synthetic Aperture Radar (SAR) (WHITCRAFT et al., 2016). SAR sensors work in almost all-weather conditions, independently of sunlight illumination, been less influenced by cloud cover frequency than optical data (LIU et al., 2013; MOREIRA et al., 2013; WULDER et al., 2015). The Sentinel-1 SAR, which is free, has dense time-series data, can be used to improve and/or develop new methods for mapping and monitoring LULC (TAMM et al., 2016). However, since SAR data interpretation is complex, its use is not widespread as ORS data (OLDONI et al., 2020). Thus, approaches that integrated both optical and SAR data have been explored in LULC studies, allowing to take advantage of each sensor (INGLADA et al., 2016; CLERICI; VALBUENA CALDERÓN; POSADA, 2017; TORBICK et al., 2017a; REICHE et al., 2018; VAN TRICHT et al., 2018).

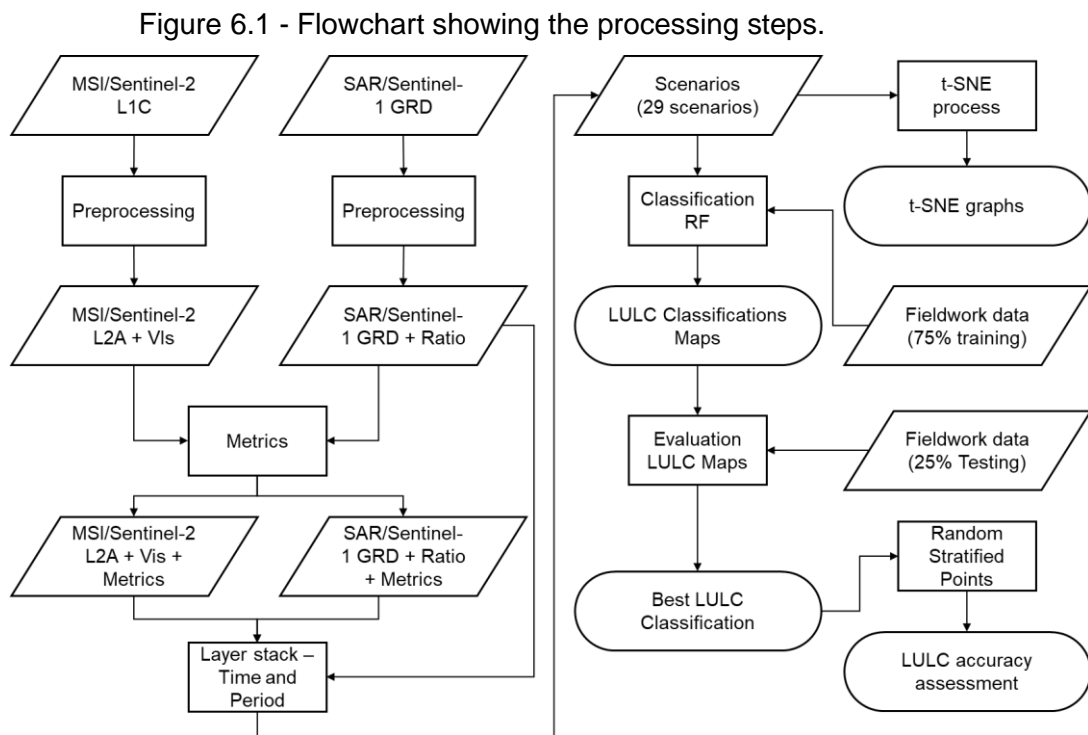
On SAR-Optical approach, Random Forest (RF) (RODRIGUEZ-GALIANO et al., 2012; INGLADA et al., 2016; TORBICK et al., 2016, 2017a, 2017b; CLERICI; VALBUENA CALDERÓN; POSADA, 2017; ZHOU et al., 2017; PAVANELLI et al., 2018) and Multilayer Perceptron (MLP) (SKAKUN et al., 2016; CAMARGO et al., 2019) are two of the most commonly non-parametric classifiers used to provide LULC classification (LU et al., 2012; CLERICI; VALBUENA CALDERÓN; POSADA, 2017). The RF classifier is highlighted due to the robustness and capability to hold a high number of variables (JHONNERIE et al., 2015; DINIZ; GAMA; ADAMI, 2020) and high data dimensionality (TORBICK et al., 2017b). Meanwhile, the MLP is a feed-forward artificial neural network (NN) trained by the backpropagation method, designed to map a set of input vectors to a set of output vectors (SKAKUN et al., 2016; CAMARGO et al., 2019).

Even though LULC map information is highly important for the management of tropical areas, there is a lack of information for some regions in the Brazilian Amazon. Savannas and Campinarana areas are not considered in programs designed for forest monitoring despite their ecological importance (CARVALHO; MUSTIN, 2017). In this context, this study aims to classify the

LULC in a tropical area in Roraima state, in the Brazilian Amazon region, providing a robust methodology to fill this gap of information. We have investigated the benefits of combining SAR with optical data in the LULC mapping process. In our study, we used RF and the MLP classifiers and Sentinel-1 SAR and Sentinel-2 MSI (Multispectral Instrument) optical images acquired in different periods along the year 2019.

## 6.2 Material and methods

This section is divided into four sub-sections. In Section 6.2.1 we describe the study area and the fieldwork to collect ground truth data. In Section 6.2.2, we describe remote sensing data and pre-processing steps. In Section 6.2.3 is described an experimental setup with multiple scenarios tested. Finally, in Sections 6.2.4 and 6.2.5, we focus on multi-dimension feature visualization using the t-Distributed Stochastic Neighbor Embedding (t-SNE) approach, and LULC classification, respectively. Figure 6.1 shows an overall flowchart describing various steps.

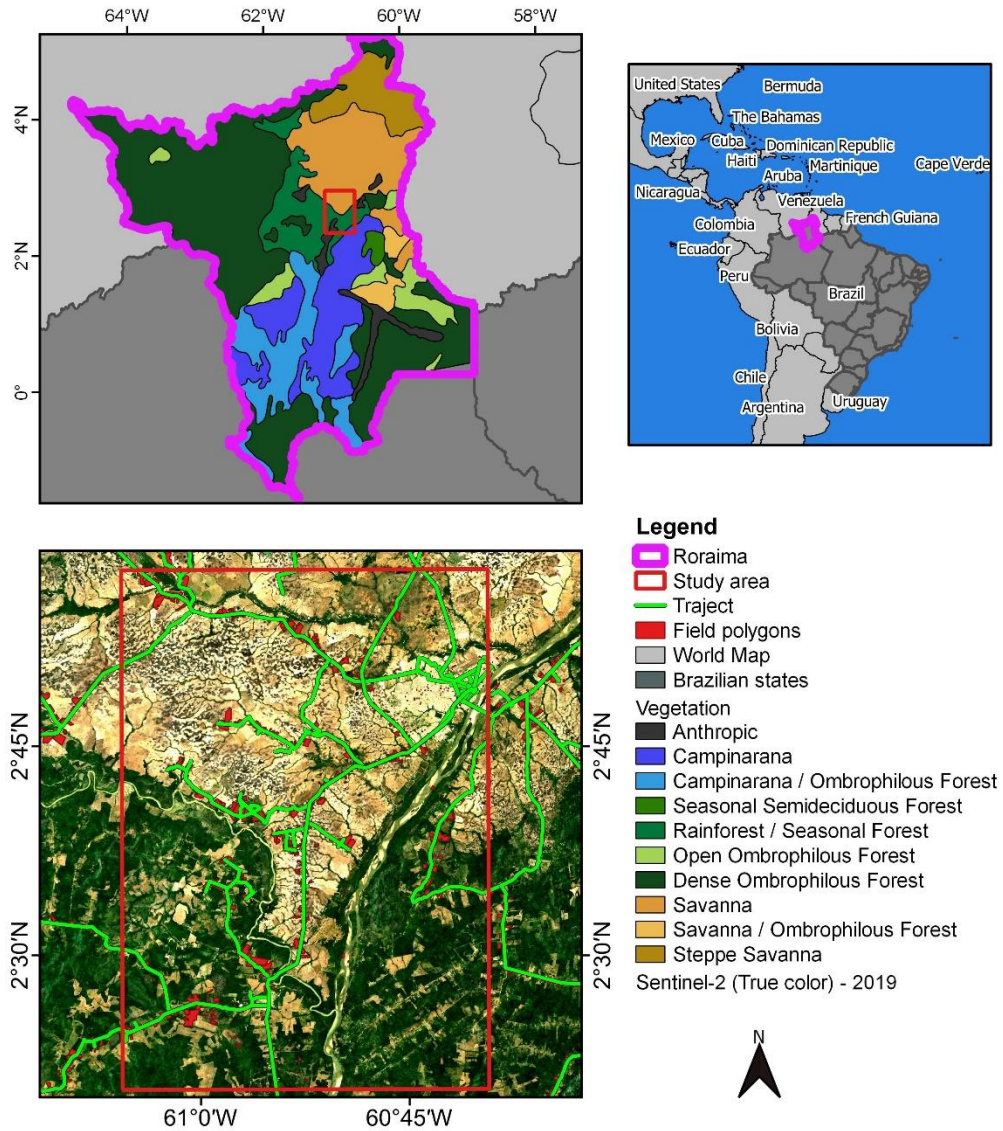


Source: Author's production.

### 6.2.1 Study area

To test our approach, we selected a test area located near the ecologic tension between Savannas and Forest (see Chapter 3 for a full description of Roraima). This area is among the coordinates 61° 5' 44.9" W, 2° 20' 20.5" N; 60° 39' 17.43" W, 2° 57' 43.14" N, with 3,381.7 km<sup>2</sup> (Figure 6.1). We selected this study area due to the LULC heterogeneity, ecologic tension, and presence of agricultural and pasture lands. To provide accurate information about the LULC classes, we used data from our fieldwork (see more details in 4.1). In total, we used 719 polygons (Figure 6.2), representing 10 LULC classes: forest, savanna, Campinarana, water, sand/rock, annual crops, perennial crops, pasture, conversion, impermeable (see Figure 6.3 to see pictures and description for each class).

Figure 6.2 – Roraima study area, with natural formations (upper-left) and fieldwork detail (down-left).



Source: Adapted of Barcelar-Lima (2008) and IBGE (2012).

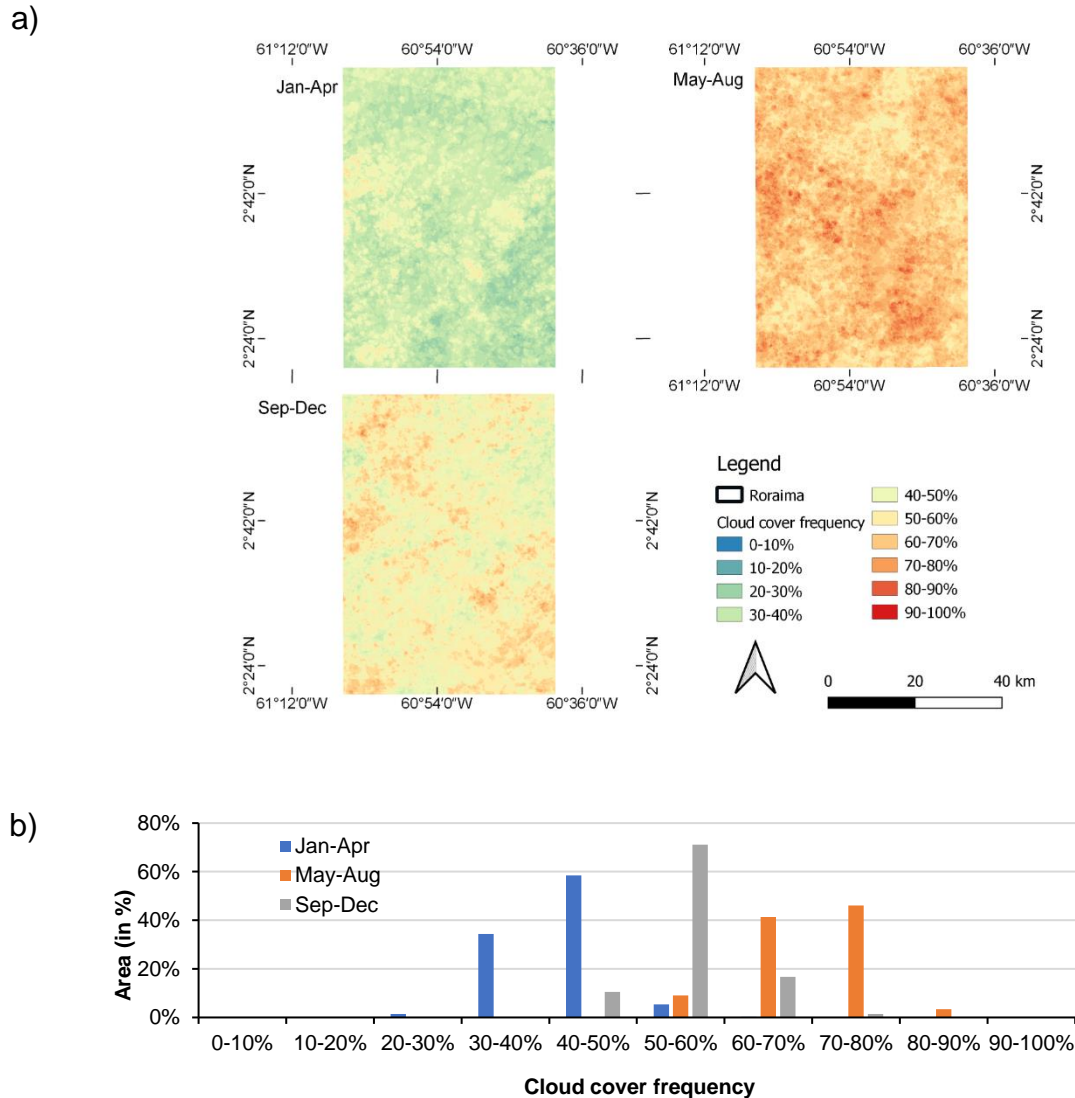
### 6.2.2 Remote sensing data

We used images from Sentinel-2 MSI optical and Sentinel-1 SAR sensors, obtained from the Copernicus Open Access Hub (also known as the Sentinels Scientific Data Hub) (<https://scihub.copernicus.eu/>), with an open-source toolbox in a python routine, named SentinelSat (<https://github.com/sentinelsat/sentinelsat>) (HU; GHAMISI; ZHU, 2018).

Due to the frequent cloud cover, mainly during the crop seasons, it is almost impossible to obtain cloud-free optical data in this region (PAVANELLI et al., 2018; PRUDENTE et al., 2020a). We considered images with less than 70% of cloud cover, totaling 23 Sentinel-2 images (Figure 6.3). The optical data was download at Level-1C, without atmospheric correction. We used the bands Blue, Green, Red, Red-edge 1, Red-edge 2, Red-edge 3, Near-Infrared (NIR), NIR-A, Short Infrared 1 (SWIR 1), and Short infrared 2 (SWIR 2) (ESA, 2020). In the optical preprocessing, we provided images atmospheric correction, clouds mask and coregister. The atmospheric correction was processed with the Sen2Cor algorithm, using sen2cor v2.8 (available in the SNAP software). In parallel to the atmospheric correction, it proceeded with the cloud mask using the Fmask v4.0 software (QIU; ZHU; HE, 2019).

After the atmospheric correction and cloud cover mask, we resampled the data from 20 meters to 10 meters. With all the data at 10 meters of spatial resolution, we co-registered images (SKAKUN et al., 2017), using the image from 2019 April 1<sup>st</sup>, which presented a less percentage of cloud cover, as our reference. We also calculated two VIs, NDVI (ROUSE et al., 1973), and LSWI (XIAO et al., 2002, 2004). The NDVI is one of the most know VI and is associated with vegetative vigor. LSWI is sensitive to the water presence and is used to flood mapping (DONG et al., 2013; TORBICK et al., 2016). NDVI (Equation 4.1) and LSWI (Equation 4.2) equations are presented in Section 4.3.

Figure 6.3 – Cloud cover frequency map (a) and graph (b) over the study area for MSI data, using quality data from google earth engine considering three periods of 2019 (January to April, May to August, and September to December).



Source: Author's production.

For SAR data, we used Sentinel 1A and 1B (TORRES et al., 2012) data, acquired at VH and VV polarization in the IW and GRD image acquisition mode. Sentinel-1 has a temporal resolution of 12 days for each satellite, and 4 and 8 days considering both satellites. In total, we used 59 Sentinel-1 images. The preprocessing was performed using the Sentinel Application Platform-SNAP applications and python routines. For defined the preprocessing, we ran several

tests and adapted the methods proposed by Dey et al., (2020). Our preprocessing included: apply orbit file; thermal noise removal without re-introduction; calibration: gamma enough; multi look with ground square pixel; speckle filter: Refined Lee; terrain correction: bilinear resample method, 10 meters of spatial resolution; and convert the results to decibels (dB). We also calculated the ratio between the polarizations VH and VV (Equation 4.3). The ratio is less affected by environmental factors or acquisition systems, thus might has more stability than single polarizations VH or VV (VELOSO et al., 2017; HARFENMEISTER; SPENGLER; WELTZIEN, 2019).

We calculated average, median, mean, standard deviation, variance, range, and percentiles (25%, and 75%) metrics for optical (bands and VIs) and SAR (polarization and ratio) data. For this step, we used the panda library in the python routine. Due the different angles from Sentinel-1A and Sentinel-1B interfere in the backscattering values, we proceed with metrics separately.

To process the data, we used two different machines. It was used a precision 5280 tower computer with processor Intel (R) Xeon(R) W-2145 CPU @3.70GHz, 64 GB internal memory, Windows 10 Enterprise – 64 Bits, 5 TB of Hard Drive (MegaRAID 9460-16i). Also, it was used a laptop Dell intel i7 8<sup>th</sup>, 32GB RAM, 6GB NVIDIA, 1 TB SSD, 1TB HDD in this process.

### **6.2.3 Scenarios**

To verify the SAR data influence on the LULC classification process, we split our data into six different dataset combinations, during five different time intervals (Table 6.1). As datasets, we used only optical data (D1 and D2), optical plus SAR data (D3 and D4), and only SAR data (D5 and D6). For optical data, we used metrics instead of single optical images to minimize cloud influence. We performed the integration between SAR and optical data at feature levels (INGLADA et al., 2016; SKAKUN et al., 2016; GÓMEZ, 2017; PAVANELLI et al., 2018; ZHANG et al., 2018b). We assumed that there is no consistent shift between the SAR and the optical images.



We divided the year in three different periods: Period 1 – P1 (January to April), Period 2 – P2 (May to August), Period 3 – P3 (September to December), Period 4 – P4 (data for the whole year), and Period 5 – P5 (the periods P1, P2, and P3 combined). P1 and P3 represents the dry periods. P2 represents the rain season when Annual crops are presents in Roraima. We combine P1, P2, and P3 into the P5 to represent the seasonal variation from each period. P4, represent the data for the whole year (January to December), without consider the variation inside each period.

Combining the different periods with the datasets, we created 29 different scenarios (Table 6.1). For D6, with Sentinel-1 data without metrics (average, median, mean, standard deviation, variance, range, and percentiles (25%, and 75%)), the P5 and P4 are the same.

Table 6.1. Details about the 29 scenarios formed with different datasets and periods.

Datasets	Descriptions	Periods
Dataset 1 (D1)	Metrics for bands optical data	P1, P2, P3, P4, P5
Dataset 2 (D2)	Metrics for optical bands and VI	P1, P2, P3, P4, P5
Dataset 3 (D3)	Metrics for optical bands and VI plus polarization and ratio for SAR data	P1, P2, P3, P4, P5
Dataset 4 (D4)	Metrics for optical bands and VI, plus polarizations, ratio, and metrics for SAR data	P1, P2, P3, P4, P5
Dataset 5 (D5)	Polarizations, ratio, and metrics for SAR data	P1, P2, P3, P4, P5
Dataset 6 (D6)	SAR polarization and ratio	P1, P2, P3, P4

P1: period 1 (January to April), P2: period 2 (May to August), P3: period 3 (September to December), P4: period 4 (the whole year), P5: period 5 (the periods P1, P2, and P3 combined).

Source: Author's production.

#### 6.2.4 t-Distributed Stochastic Neighbor Embedding (t-SNE)

We used the t-SNE technique to reduce our high dimensionality data and to help understand the possibility of class separation before the classification

process). We used t-SNE from the scikit-learn library (PEDREGOSA et al., 2011), with interactions equal to 300, perplexity equal to 30.

### **6.2.5 LULC classification scenarios**

We used two machine learning classifiers, Random Forest (RF), and Multi-Layer Perception (MLP), using a python routine with the scikit-learn library (PEDREGOSA et al., 2011). After testing different parameters (PRUDENTE et al., 2020b), we used numbers of trees equal to 30 for the RF classifier, and layers size equal to 50, alpha equal to 0.01, and learning rate values of 0.005 for the MLP classifier. The field data (polygons) were randomly separated into 75% for training and 25% for the validation process. For the validation and to find the best LULC classification, we analyzed the confusion matrix, the overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) (OLOFSSON et al., 2014).

An independent validation was used for the best classification result (scenario and classifier), we used Stratified Random Points (SRP). The goal of SRP is to have a practical design that satisfies the accuracy assessment objectives and most of the desirable design criteria. SRP affords the option to increase the sample size in classes that occupy a small proportion of the area, to reduce the standard errors of the class-specific accuracy estimates for these rare classes (OLOFSSON et al., 2014). We randomly sampled 1,158 points, stratified according in Forest: 262, Savannas: 272, Campinarana: 75, Water: 75, Pasture: 99, Sand/rocks: 75, Annual Crops: 75, Perennial Crops: 75, Conversion: 75, and Impermeable: 75. This guarantees that the sample size in a small class will be large enough to represent that class. This approach allows us to estimate the errors in terms of the area along with uncertainties.

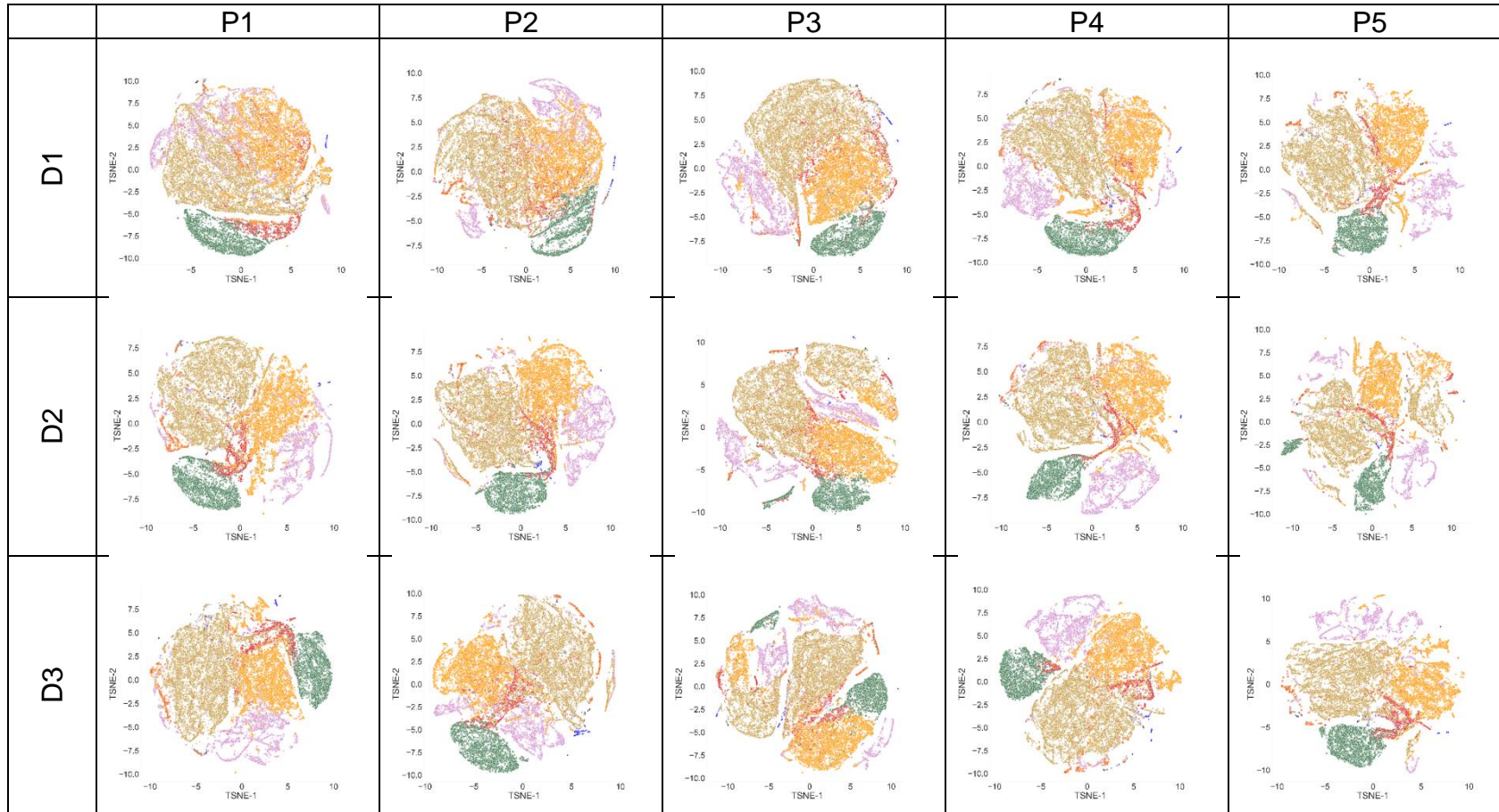
### **6.3 Results**

In this section, we present the results of inputs feature visualization using the t-SNE algorithm (Section 6.3.1). In Section 6.3.2, we analyze classification results for the 29 different scenarios and independent validation of the final map.

### 6.3.1 Visualization of input features using t-SNE

The t-SNE graphs show the differences between the classes for each database in each period (Figure 6.4). Optical data, visualize give better separation of the classes Forest, Savanna, and Pasture. For SAR, we could not verify the same separation as in the optical data. Classes with similar geometrical structures are closer or even mixed. We found that Forest and Perennial Crops are mixed, meanwhile, Savannas and Pasture represent two other groups, where has mix with Annual Crops, Conversion areas. Considering the periods and datasets, we found that the t-SNE technique showed better results (Figure 6.4) for P4 and P5, with data from the entire year, and when combining data from optical and SAR sensors (D3 and D4). Scenario D3P5 (Figure 6.5) has the best visual results, with better separation to Forest, Savanna, Pasture, and Annual Crops classes. Perennial Crops class has confusion with Forest, Savanna, and Pasture classes. Besides, Pasture has overlapped with Savannas and Annual Crops classes, and Campinarana was confused with Pasture. Impermeable class is concentrated on the left-middle ( $x=-8$ ;  $y=-4$ ) of the graph (Figure 6.5). Sand/rock has two small groups, one with Impermeable class and the other in the middle-bottom ( $x=-1$ ;  $y=-4$ ) of the graph.

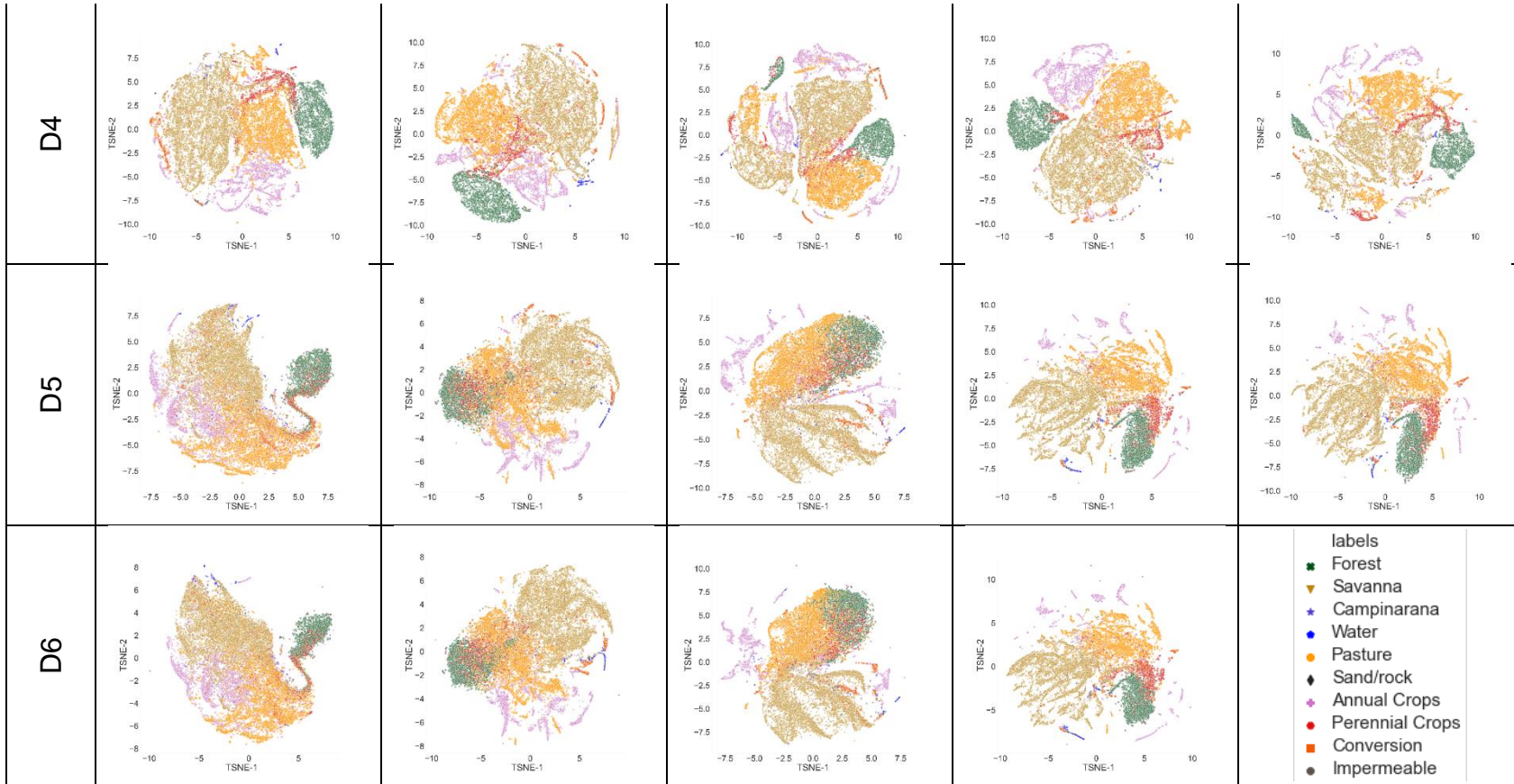
Figure 6.4 – t-SNE results for each one of the 29 scenarios.



Continue

2

Figure 6.4- Conclusion.

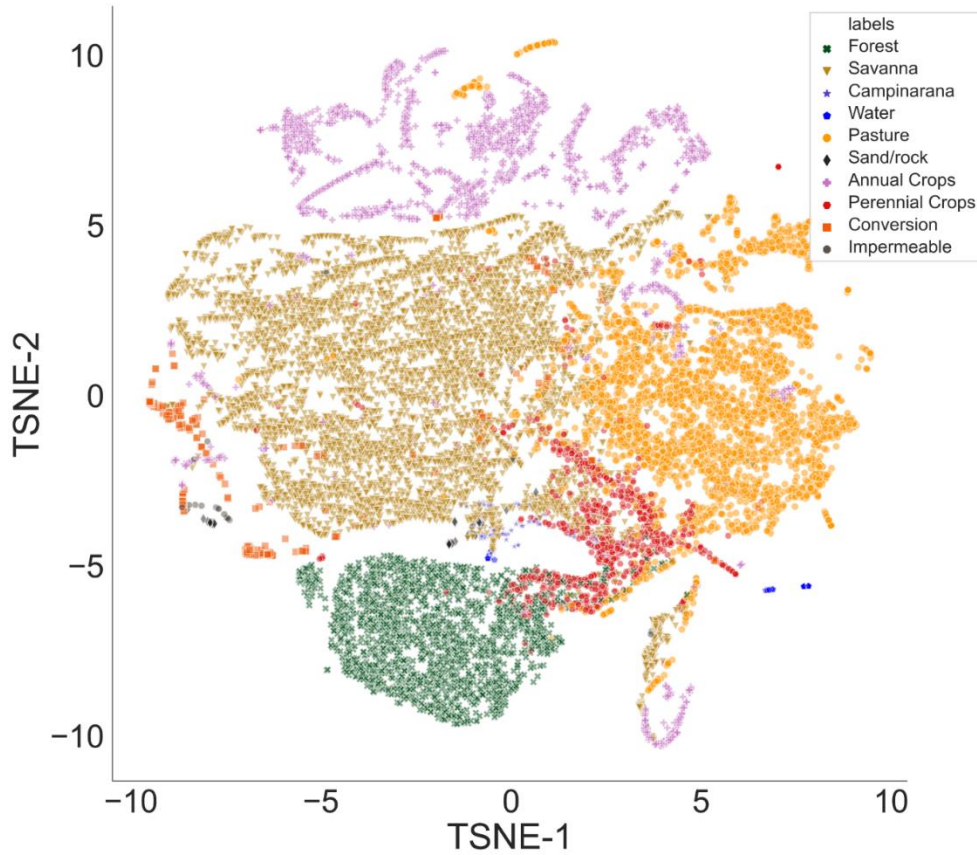


3

Source: Author's production.

4

Figure 6.5 – t-SNE graph for the scenario with the best separation (D3P5).



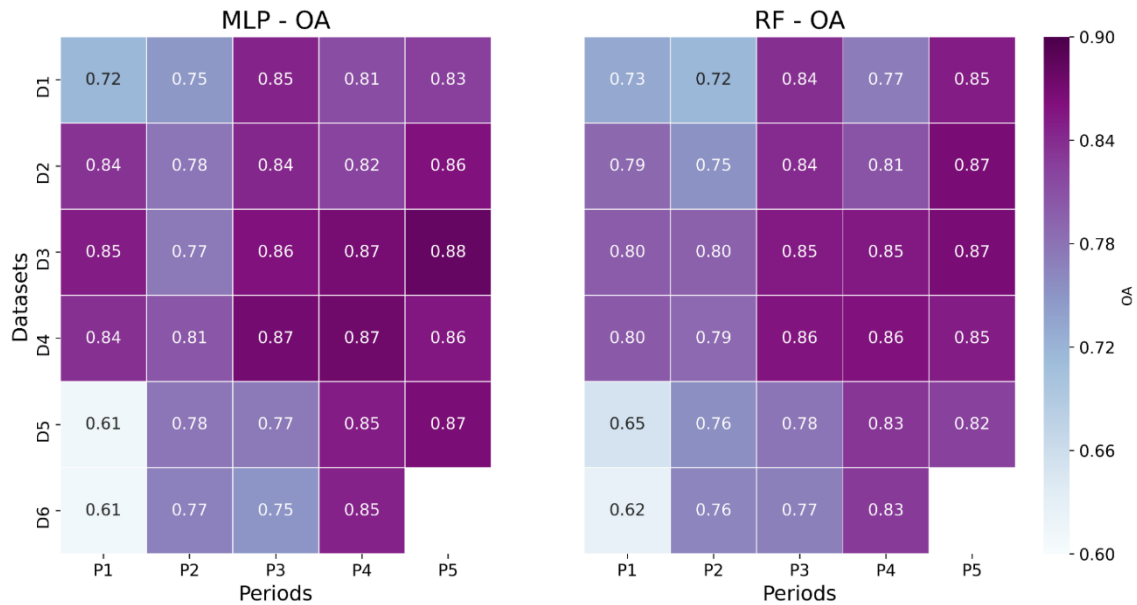
Source: Author's production.

### 6.3.2 Classification

The Overall Accuracy (OA) of all scenarios for MLP and RF classifiers is shown in Figure 6.6. In general, we verified that the MLP classifier shown slightly higher OA than RF. Moreover, when using data from a single sensor, we found better results using optical data (D1 and D2) than SAR data (D5 and D6) for P1 and P3. For the periods analyzed, when we used data for the whole year (P4 and P5) we had higher accuracies than when using only data for P1 and P2. Considering the different periods of the year, the P2 has the lower and P3 has the higher OA values for optical data (D1 and D2). Meanwhile, for SAR datasets (D5 and D6) the higher value is for P2 and the lower for the P1. Finally, considering the multisensory SAR-Optical approach (D3 and D4), we found that the OA values increased in almost all periods. Overall, the best result was achieved for the dataset D3, using data for the whole year P5, using MLP.



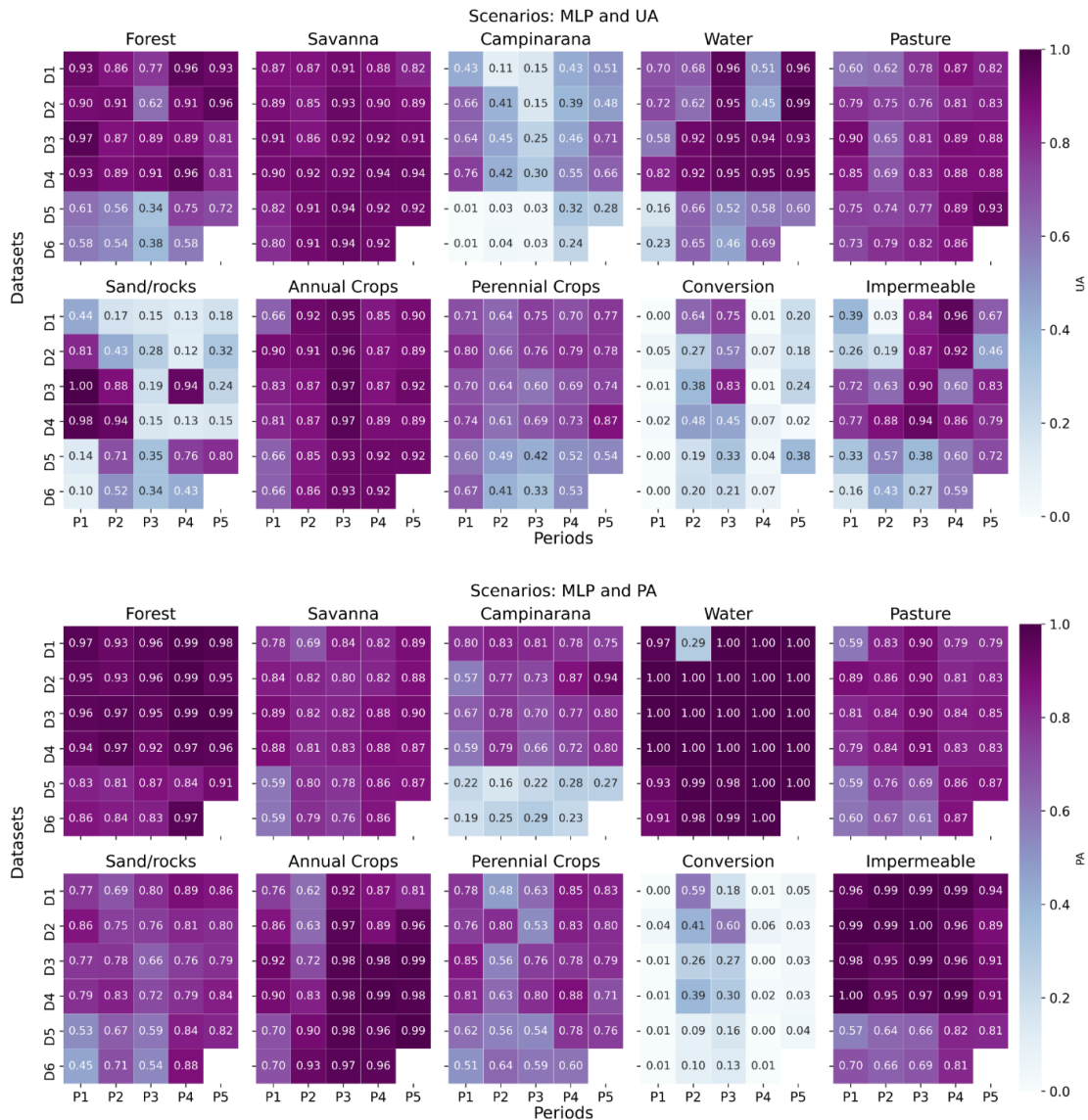
Figure 6.6 – Overall Accuracy (OA) of Random Forest (RF) and Multi-layer Perceptron (MLP) classifiers.



Source: Author's production.

In Figure 6.7 and Figure 6.8, we show the Users (UA) and Producers (PA) Accuracies for MLP and RF classifiers, respectively. Considering that the MLP approach has better overall accuracies, we showed the UA and the PA of this analysis in detail. SAR dataset shows higher UA and PA for Savanna, Pasture, and Annual Crops, but has lower UA and PA for Campinarana and Conversion classes. In general, optical data has better results than SAR. However, the UA for Sand/rocks and Campinarana, and the PA for Conversion classes remain low. For the datasets with multisensor data (D3 and D4) from the entire year (P4 and P5), we verified an increase in the UA and PA values. We also highlighted the better UA and PA values for P3 compared to P2 and P1.

Figure 6.7 – Users (UA) and Producers (PA) accuracies for each class in each scenario considering the MLP classifier.



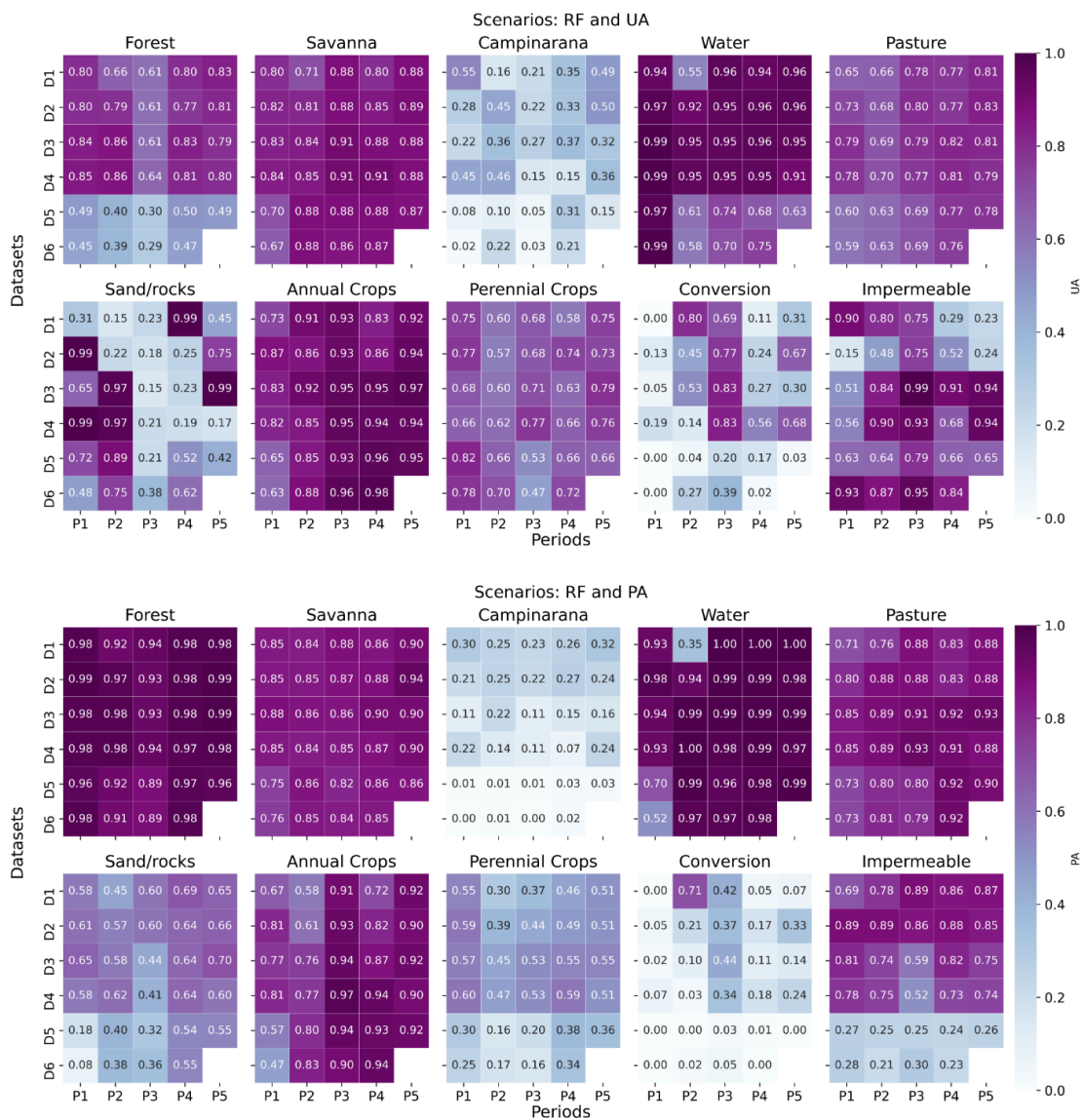
Source: Author's production.

MLP and RF maps are shown in Figure 6.9 and Figure 6.10, respectively. In general, maps based only on SAR data shown more salt-pepper effects, meanwhile maps based only on optical data presented cloud interference. The problems were reduced when using the multisensor approach. Our best classification (Figure 6.11) was obtained using optical metrics and SAR data (D3) combining data from the different periods (P5). The predominant class is



Savannas (119,936ha ~ 34.17%), followed by Forest (115,155ha ~ 32.81%) and Pasture (43,309ha ~ 12.34%). Perennial Crops (17,859ha ~ 5.09%), Water (17,658ha ~ 5.03%), and Annual crops (14,215ha ~ 4.05%) are the following classes in terms of area. The classes with less area are Sand/rocks (1,700ha ~ 0.48%), Conversion (2,347ha ~ 0.67%), Impermeable (8,724ha ~ 2.49%), and Campinarana (10,115ha ~ 2.88%).

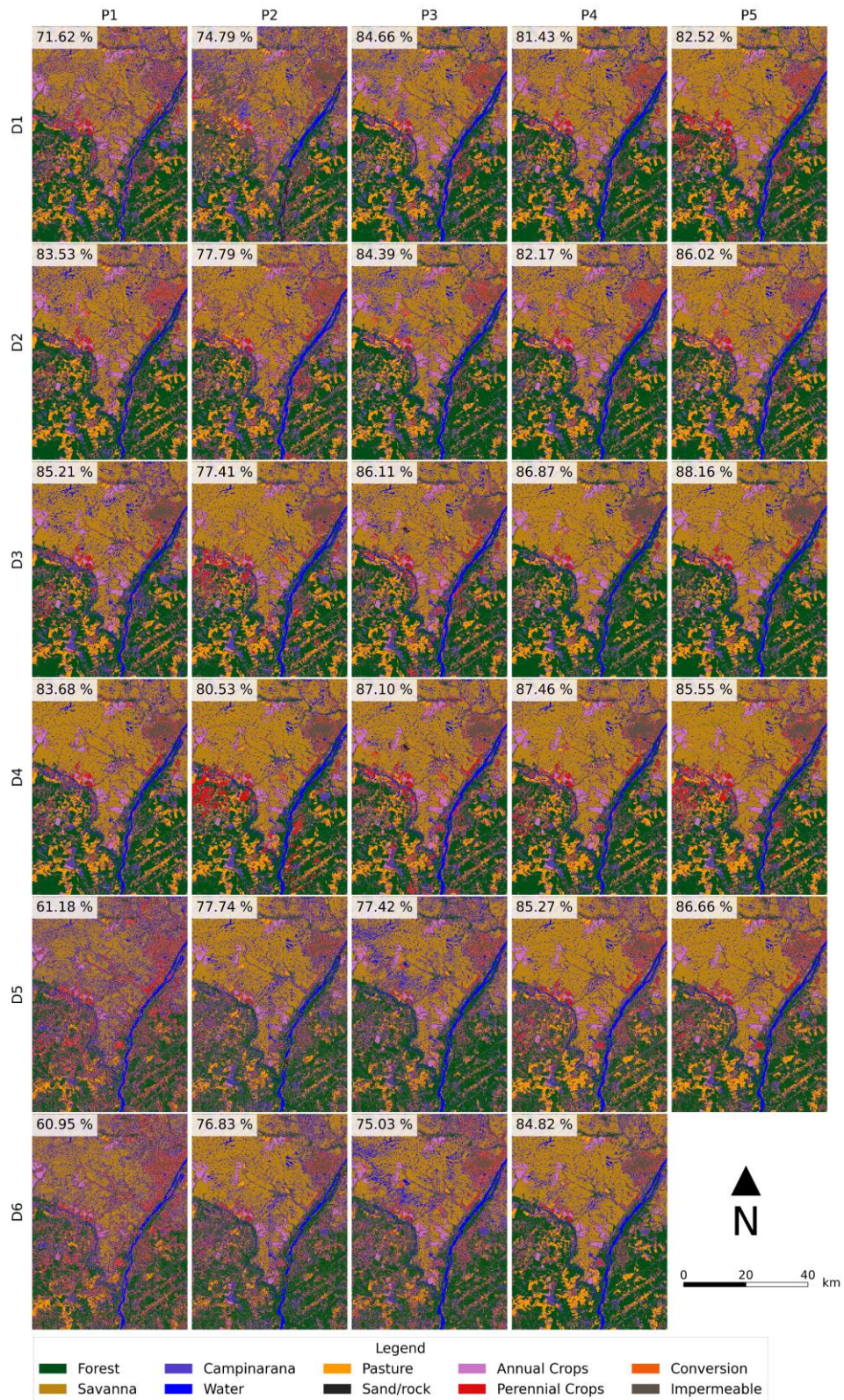
Figure 6.8 – Users (UA) and Producers (PA) accuracies for each class in each scenario considering the RF classifier.



Source: Author's production.

To further evaluate the LULC classification accuracy, we performed the accuracy assessment using an independent set of points. The OA accuracy, using SRP was  $OA = 86.41 \pm 1.74\%$ , reaching almost the same value from the field points ( $OA = 88.15\%$ ). To compare the accuracies for each class, we provide the confusion matrixes (Figure 6.12 and Figure 6.13) along with estimated UA and PA (Figure 6.14). For Forest and Savannas classes, we observed an improvement in the UA (81% to  $95.10 \pm 1.3\%$ , and 91% to  $96.30 \pm 1.1\%$ , respectively) and a decrease in the PA (99% to  $89.3\% \pm 1.3\%$ , 90% to  $86.8 \pm 1.4\%$ , respectively) when using SRP. For Annual Crops and Impermeable classes, similar UA value (92% to  $93.30 \pm 2.9\%$ , and 83% to  $86 \pm 4\%$ , respectively) and decreased PA (99% to  $90.2 \pm 4.7\%$ , and 91% to  $75 \pm 5.9\%$ , respectively) are observed. For the Water class, similar values from UA (93% to  $90.7 \pm 3.4\%$ ) and PA (100% to  $100 \pm 0\%$ ). For Campinarana, UA = 71% and PA = 80% from field data, decrease to UA = 36% ( $\pm 5.6\%$ ) and PA = 42.8% ( $\pm 7.3\%$ ) with SRP. Considering the Pasture class, UA decreased 88% (field data) to  $73.7 \pm 4.4\%$  (SRP), however, PA stayed with similar values (85% to  $82.1 \pm 2.9\%$ ) using SRP. For the Sand/rock class, UA value increased by 24%, with field data, to  $77.3 \pm 4.9\%$  with SRP, and PA values increase by 79% with field data to  $84.8 \pm 12.9\%$  with SRP, but with the higher error value among the classes. UA values decreased from 74% to  $21.3 \pm 4.8\%$  and PA keeps similar values (79% to  $80.2 \pm 10.5\%$ ), considering the Perennial Crop class. For the Conversion class, UA increase by 24% to  $33.3 \pm 5.5\%$ , and PA had the higher increase values (3% to  $89.6 \pm 4.6\%$ ).

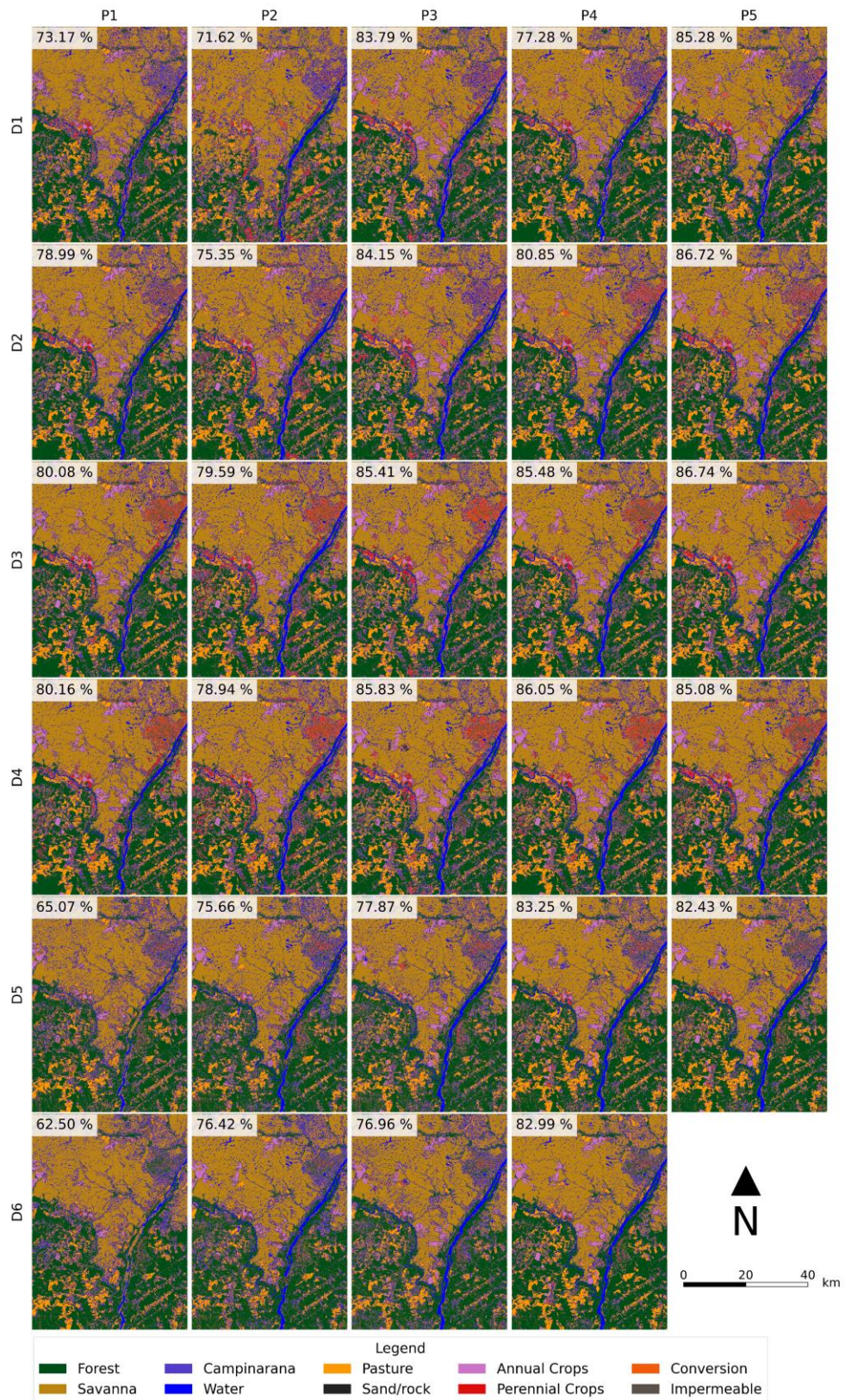
Figure 6.9 – Classification results over the 30 scenarios using MLP classifiers for 2019.



Source: Author's production.



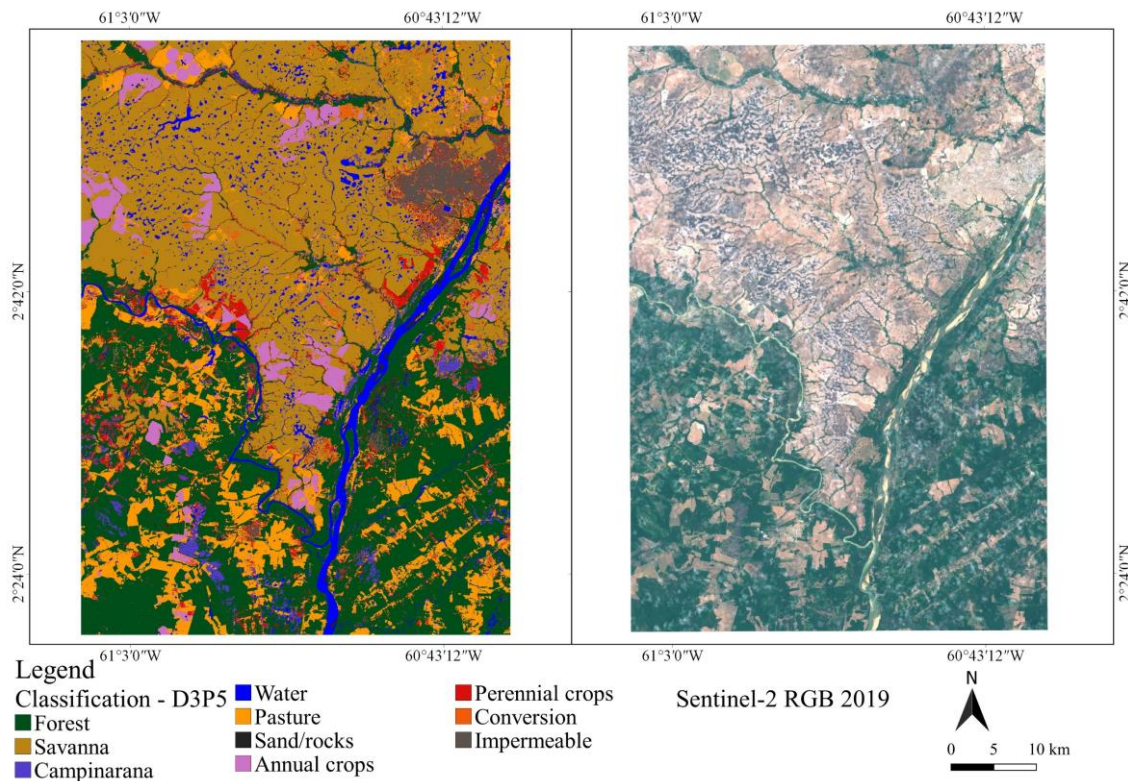
Figure 6.10 – Classification results over the 30 scenarios using RF classifier.



Source: Author's production.



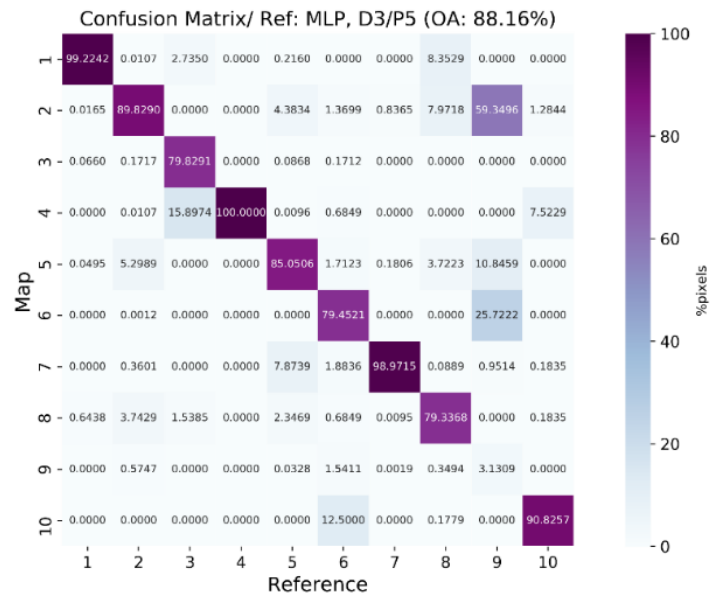
Figure 6.11 – Best LULC classification scenario (D3P5) map, using the MLP classifier (left), and a 2019 Sentinel-2 RGB True color image, using mean of the band values over the year (right).



Source: Author's production.

We compared our results with maps from the MapBiomas initiative (SOUZA et al., 2020). MapBiomas provides annual LULC maps for the entire Brazilian territory between 1985 and 2019, with a spatial resolution of 30 meters. MapBiomas is the only updated LULC program that maps the whole state. For our study area, MapBiomas LULC is classified as Forest Formation (116,875ha ~ 32.95%), Forest Plantation (40ha ~ 0.01%), Grassland Formation (166,925ha ~ 47.06%), Pasture (43,086ha ~ 12.15%), Urban Infrastructure (10,570ha ~ 2.98%), River, Lake, and Ocean (7,336ha ~ 2.07%), Soybean (7,522ha ~ 2.12%) and Others Temporary Crops (2,388ha ~ 0.67%) (Figure 6.15). We grouped Soybeans and Others Temporary Crops as Temporary Crops (9,910ha ~ 2.79%).

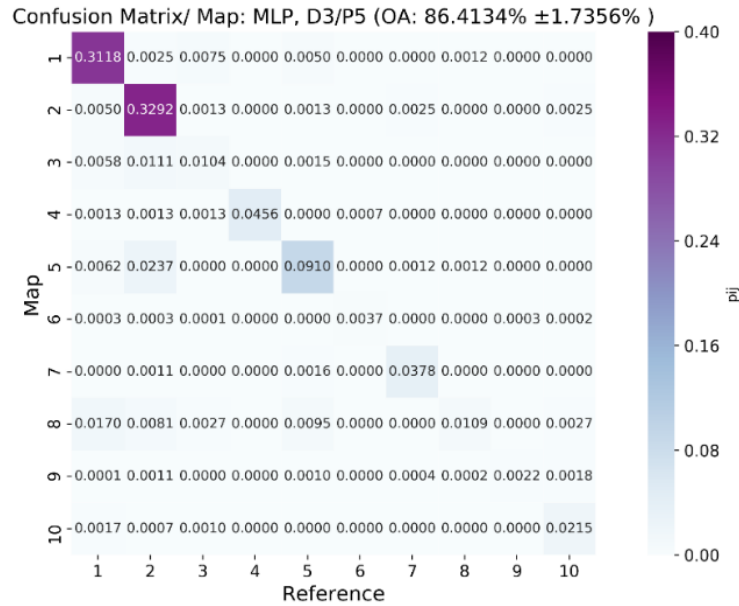
Figure 6.12 – Confusion matrix in terms of validation pixels from polygons derived through field campaign for the best LULC classification scenario (D3P5 – MLP), using field data.



Values are in reference percentage. 1: Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

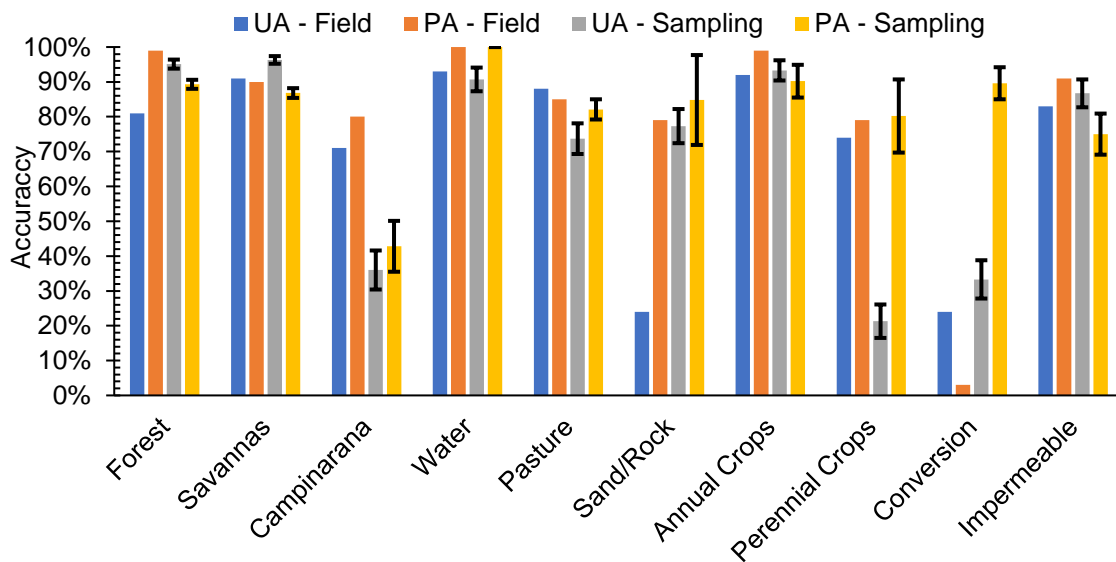
Figure 6.13 – Confusion matrix in terms of area percentage (OLOFSSON et al., 2014) for the best LULC classification scenario (D3P5 – MLP), using Stratified Random Points.



1: Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas

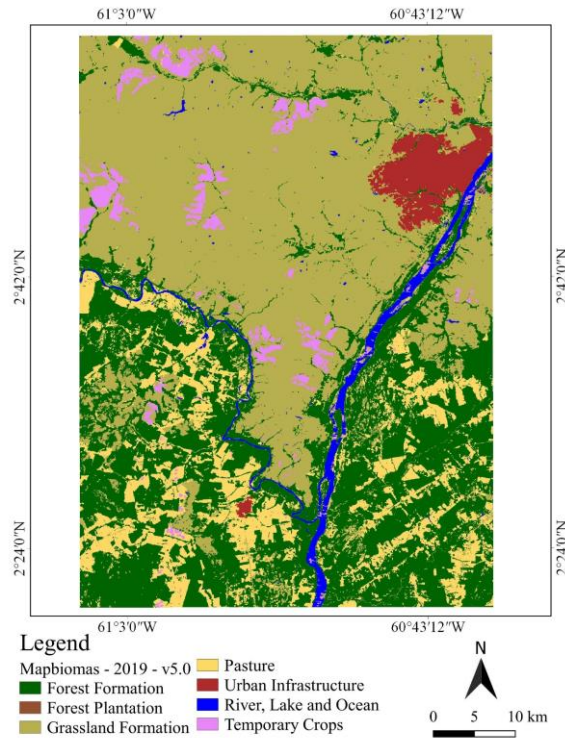
Source: Author's production.

Figure 6.14 – Comparison among accuracies from field data and sampling approach.



Source: Author's production.

Figure 6.15 – Land Use and Land Cover mapping provided by the MapBiomias v5.0 – 2019.

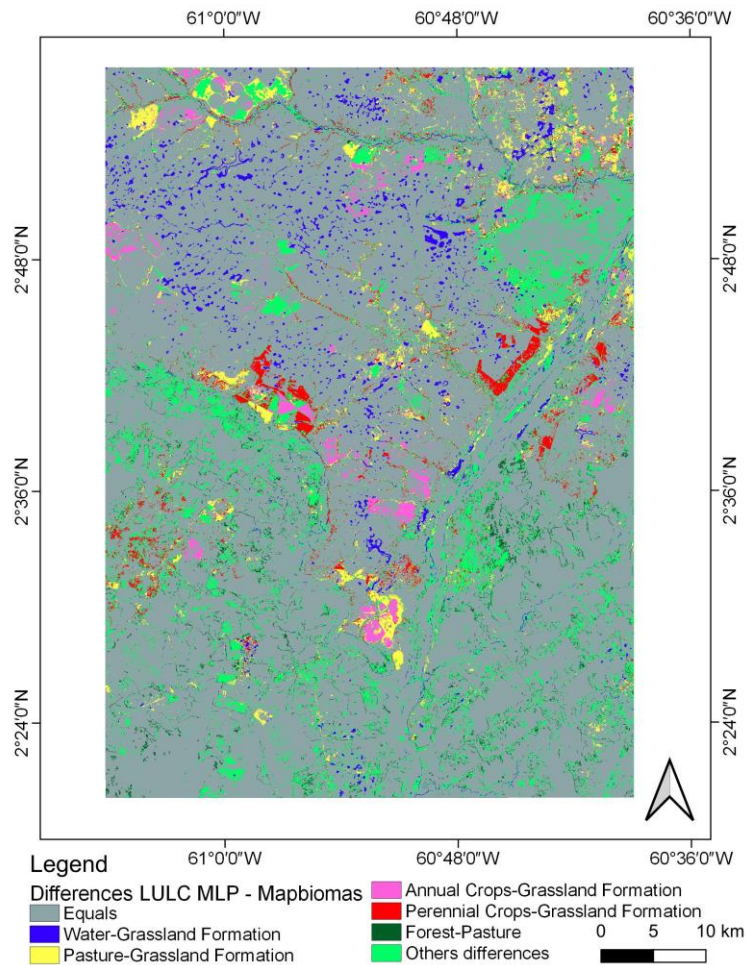


Source: Souza et al. (2020).

The difference map between our approach and the MapBiomias initiative is shown in Figure 6.16. The equal area, considering Grassland formation equal to Savannas and Campinarana, is 277,083ha and the differences are 73,937ha. We can highlight that part of Grassland formation in MapBiomias LULC program is classified as Pasture (10,678ha), Perennial crops (8,690ha), Water (8,468ha), Annual crops (5,091ha), and Forest (4,719ha) classes in our approach. Besides, part of our Forest (5,813ha) classifications is shown as Pasture in MapBiomias.



Figure 6.16 – Map of the differences between our LULC mapping and the LULC provide for MapBiomas v5.0 – 2019 (SOUZA et al., 2020).



Source: Author's production.

## 6.4 Discussion

The Roraima state in Brazil is prone to periods of frequent cloud cover during the April-September rainy season, which severely limited the exploitation of optical satellite imagery. Thus, it is extremely difficult to accurately map Land Use and Land Cover (LULC) with optical data only (PRUDENTE et al., 2020a). Moreover, there is no accurate LULC map for the whole state. In this context, SAR data can be an alternative to improve the LULC classification. Therefore, the optical-SAR multisensor approach is a viable way out to provide LULC classification in regions with frequent cloud cover. To validate this, we used two

different classifiers algorithms in 29 scenarios, with optical, SAR, and SAR-optical datasets, in five different periods of the year 2019.

According to the t-SNE (t-Distributed Stochastic Neighbor Embedding), SAR-optical data (D3), combining the three different periods (P5) provided the best visual separation among the classes. Moreover, optical in general showed better visual separation than SAR datasets. In this sense, our study shows that t-SNE provides a good way to visualize high dimensionality RS data than classic technique, as scatter plots (PAVANELLI et al., 2018) and boxplots (TORBICK et al., 2016, 2017a). However, only a few studies have used this technique in the RS field (DEY et al., 2020; MARTINS et al., 2020), and we did not find any research that applied the t-SNE in a multisensor dataset.

RF is commonly used in a multisensor SAR-Optical approach to map LULC (TORBICK et al., 2016, 2017a; CLERICI; VALBUENA CALDERÓN; POSADA, 2017; ZHOU et al., 2017; PAVANELLI et al., 2018). Using Sentinel-1 SAR data for LULC mapping over a test area in the Brazilian Amazon forest, Diniz et al. (2020) had a better result using RF than Support Vector Machine. We found slightly higher OA values (1% or 2% in general) using MLP classifiers compared to RF. Besides, LULC classifications from MLP were visually better, with less salt-pepper effect when compared with RF classifier.

Considering the datasets analyzed, classification based exclusively on SAR data had lower accuracy than when using only optical data. We identified that aggregating VI metrics (D2) improves the results than using only MSI band metrics (D1). Also, the use of metrics for SAR data (D6) did not improve the OA when compared with only SAR data (D5). For SAR-optical data-based classifications, we found better overall accuracies than when analyzing the SAR and optical data separately. Besides, the LULC maps were visually better with the multisensor approach. Clerici et al. (2017) also found better results using optical (MSI/Sentinel-2) and SAR (Sentinel-1) data to map LULC in Colombia. They found that SVM (Support Vector Machine) had better accuracy than the RF classifier, achieving OA=88.75%. However, compared to our study, their study area was more homogeneous, with six LULC classes. Zhou et al. (2017), using

data from Landsat-8 OLI and Sentinel-1 SAR provided in-season winter wheat classification in China, achieving the best results with the multisensor approach. Torbick et al. (2016) integrated data from Landsat-8 OLI, Sentinel-1 SAR, and PALSAR-2 images to mapping forest plantations in Myanmar and Indonesia. They obtained the best accuracy when integrating the data from the three sensors. In the following work, Torbick et al. (2017a) used data from the same sensors to classify/update the LULC in Myanmar, and also found the best accuracy using data from optical and SAR data combined.

Considering the periods analyzed, the classifications performed to P3 had better results than P1 and P2. That could be because during P2 we had less optical data available, resulting in a poor classification. Also, the fieldwork was realized at the end of P2 and the beginning of P3, helping to better identify some classes in the field. Pavanelli et al. (2018) and Lu et al. (2011), who worked with LULC classification in the Amazon regions, described the difficulty to get cloud-free data during the rainy season. However, they used only one optical and one SAR image from the dry season. We achieved better accuracy (OA = 88.16%) with the data from the P5, aggregating the data from P1 (January to April), P2 (May to August), and P3 (September to December). That is because with this approach it was considered the seasonality existent in the classes. Pavanelli et al. (2018) mentioned that using only data from the dry season is not enough to discriminate all LULC classes. In our study, the Sand/rocks class was better identified in the dry period (P1), because the sandbanks appear inside the Rio Branco River in this period. The Water class, in general, had a better performance for P3, due to the end of the rainy season in P2, formed the small lakes in the lavrados regions, favoring the identification in the P3. During the P3, the Conversion class also had better discrimination, mainly due to fieldwork in this period. We expected that because conversion is a dynamic class, occurring during part of the year, thus some of these areas were Savannas or Pasture in the past, in a different period of the year. It is important highlight that this conversion class covers only the change from natural to anthropic class, what was assured by field work and temporal image interpretation. Besides, during the

Conversion process occur the soil turning, creating confusion with the Sand/rocks class.

Due to the frequent cloud cover in the P2, in this period, the Annual Crops class is better discriminated with SAR than optical data. Besides, Pasture is present during this time, beginning a source of confusion between these classes. Forest, Savanna, Pasture, and Annual Crops classes were better discriminated when we used data from P4 and P5. That happens because it is used data for the entire year, registering the vegetation variability and mitigating the cloud cover frequency interference (PRUDENTE et al., 2020a). Perennial Crops is a heterogeneous class, formed by Acacias, Carswell, Eucalyptus, etc. Therefore, this class is confused with Forests and Savanna classes, due to the similar structure.

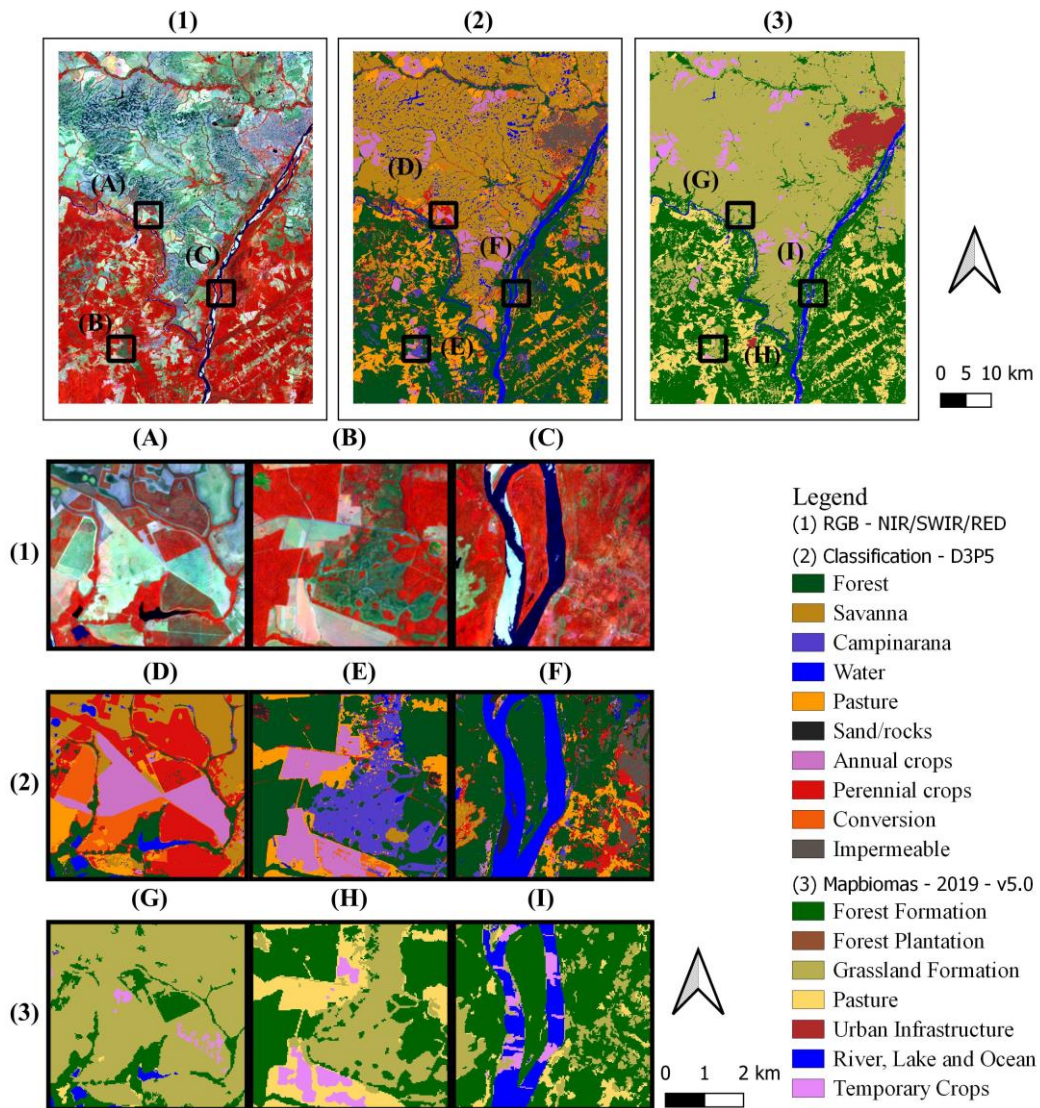
Overall, our best LULC classification accuracy was obtained using the MLP classifier applied to the data from optical and SAR data combined (D3), considering the three periods together (P5). We used the benefits of optical and SAR data associate with the seasonality from different periods. We assessed the accuracy of this classification by performing Stratified Random Points (SRP). The approach with field data and SRP were similar in respect of overall accuracy, however, UA and PA had different values for Campinarana, Pasture, and Conversion classes. These classes with small areas were better represented with SRP. Waldner et al. (2019), mapping the croplands, describe that the overall accuracy is similar using roadside and random sampling points. However, they explained that SRP better represents the classes than roadside sampling.

We compared our best result with the MapBiomass maps initiative (SOUZA et al., 2020). MapBiomass provide an annual LULC map, using a Google Earth Engine (GEE) platform (GORELICK et al., 2017), and Landsat-8 OLI, Landsat-7 Enhanced Thematic Mapper Plus (ETM+), and Landsat-5 Thematic Mapper (TM) Landsat optical images at 30 meters of spatial resolution. They do a hierarchical classification scheme, using a RF classifier and combined different classification methods (date range, input data, etc.) (SOUZA et al., 2020). It is classified each

LULC class separately and joins them, generating three different LULC map levels.

For our study area, our map has more detailed classes than MapBiomas Level 3 (Figure 6.17). Grassland Formations in the MapBiomas class represents our Savannas and Campinarana classes. Also, we identified areas with Acacia and Cashew (Perennial Crops) that were not present inside the Forest Plantation class in MapBiomas LULC map. Acacia is a common Forest plantation in Roraima's Lavrados, and now it has been converted into agricultural areas. Also, we have the Conversion class, and this is an important class to identify where and when is occurring the changes in the Roraima state. Sand/rocks and Conversion classes do not exist. Moreover, Sandbanks inside of the Rio Branco River are identified as Temporary Crops in the MapBiomas map. Additionally, our results showed the Lakes formations in the Savannas regions. Therefore, our results provided a better overview of the natural resources, conversion areas, and land use, classifying all in a single step. This information is important to monitor and regulate anthropic activity; otherwise, it can affect the availability of natural resources (e.g., water, soil) (CARVALHO; MUSTIN, 2017) and carbon stocks (BARNI et al., 2016).

Figure 6.17 – Comparison between our result and the MapBiomass v.5 LULC map.



Source: Author's production.

However, it is important to highlight MapBiomass is a nationwide initiative, which annual LULC maps that are improving each version. In this sense, we believe that our results could be helpful to improve the next MapBiomass LULC version, mainly to better discriminate Annual Crops, Perennial Crops, and the Lakes formations. Besides, we did not compare our results with other wide LULC programs as TerraClass Amazon (ALMEIDA et al., 2016) and IBGE (2017). do not cover the Savannas and Campinarana regions (BARNI et al., 2020b), it is not

observed in the TerraClass and is classified as Grassland and Wetlands in IBGE, and are not updated, 2014 for both maps.

Pavanelli et al. (2018) provided LULC classification in a small area in this region. They considered 17 LULC classes for 2015 and used one Landsat-8 OLI optical image and one ALOS/PALSAR-2 (Phased Array L-band Synthetic Aperture Radar-2) SAR image. They found OA of 83% using RF classifier, slightly lower value than ours. However, our results with MLP show a higher OA and a better visual map with 10 meters of resolution. Also, in our study we used free data, that was not fully available in the area for 2015. Lu et al. (2011), using a Maximum Likelihood classifier, found a better LULC map accuracy (OA = 81.13%) using only optical data than SAR data or integrating SAR and Optical data. However, they studied a different tropical region in Brazil Amazon and used only one image per sensor, ALOS PALSAR, RADARSAT-2, and Landsat-5 TM, during the dry season. Moreover, these authors utilized fusion approaches to integrating the data (principal component analysis, normalized multiplication, high-pass filter resolution-merging, and Wavelet) and considered six classes (Forest, Pasture, Water, Wetland, Urban and Succession Vegetation).

## **6.5 Conclusion**

Our approach represents an advance for heterogenous LULC mapping in tropical regions with Earth remote sensing data. Optical and SAR multisensor data had better OA than only optical or SAR data. Besides, SAR data is an important source of data, mainly during the rainy season (P2), when cloud cover frequency limits the availability of the optical imagery. The use of different periods allowed us to identify when could be a concentrated effort to map specific classes. SAR-optical data for the P5, combining P1 (January to April), P2 (May to August), and P3 (September to December), show better LULC maps results. Moreover, the MLP classifier performed higher OA than RF and better-smoothed maps. Minor and similar classes are difficult to estimate, even with SAR and optical data. For validation, data from fieldwork can hide patterns that will be better represented with random sample points. As suggestion for future works, red-edge vegetation

indexes and coherence could be tested. For future processes, this approach will be applied to the whole state in different years to analyze the LULC changes.



## **7. GOOGLE EARTH ENGINE IN A TWO-STAGE SAMPLING MULTISENSOR APPROACH TO ESTIMATE THE LAND USE AND LAND COVER CLASSES OVER RORAIMA STATE**

### **7.1 Introduction**

Update and continuous Land Use and Land Cover (LULC) information help to understand and measure LULCC environmental effects at local, regional, and global scales (ADAMI et al., 2018; PAVANELLI et al., 2018). Anthropogenic expansion, like agriculture areas, if not well planned and can generate severe environmental impacts (LAMBIN; GEIST; LEPERS, 2003), and contributes to the global climatic changes (DENG; ZHAO; YAN, 2013).

ORS data are traditionally used to identify and characterize LULC (CHAVES; PICOLI; SANCHES, 2020). However, high cloud cover frequency (ASNER, 2001; WULDER et al., 2015; MARTINS et al., 2018a), along with quick land-use changes (WHITCRAFT et al., 2015b; EBERHARDT et al., 2016), hampers the use of ORS (see more in Chapter 5). On top of that, tropical landscapes fragmentations and transitions among vegetation types (LU et al., 2012; LAURIN et al., 2013), make the scenario even worst for LULC mapping based on optical sensors data (SANO et al., 2007; SUGAWARA; RUDORFF; ADAMI, 2008; WHITCRAFT et al., 2015b; EBERHARDT et al., 2016). As result, such products for tropical areas, as South America, are not available or need improvements (SONG et al., 2021a).

Microwave data collection, as provided by Synthetic Aperture Radar (SAR) sensors, has a significant potential to be utilized for mapping and monitoring LULC (STEELE-DUNNE et al., 2017). SAR are active sensors and can provide useful data under almost all-weather conditions, independently of light conditions, been less affected by atmospheric and cloud conditions (LIU et al., 2013). Approaches that integrated ORS and SAR data, denominate multisensor approach, allow us to use the well know ORS knowledge with the SAR climate limitations alternative solution. However, the quantity of data generated demands

high computation power and large space for storage. A way to overcome this is by using cloud processing platforms (GORELICK et al., 2017; PALAZZO et al., 2018; ZHANG et al., 2018a). To decrease data volume, it is possible to use different metrics to explore the temporal variation from the remote sensing data. This static transformation of an image time-series is called multi-temporal metrics (SONG et al., 2017), which can be very useful for LULC characterization (BECKER et al., 2021). Moreover, sampling approaches can be helpful to mitigate cloud cover (EBERHARDT et al., 2016; KING et al., 2017b).

Combining cloud processing (GORELICK et al., 2017) with metrics to explore the temporal variation from remote sensing data (SONG et al., 2017), and sample-based methods (SONG et al., 2017) can be a viable way to provide LULC in less time than traditional LULC classification approaches. In this sense, this research has as objective to characterize the LULC for Roraima state in three years (2017, 2018, and 2019). For that, we propose two different LULC classification methods. First, we provide a more classical design, training, and classifying the entire state at once, using field roadside data. Second, we used a sampling approach to estimate and classify the LULC.

## **7.2 Material and methods**

To show the procedure adopted to map LULC for Roraima, we divided this section into four parts. i) An overview of the study area. ii) The data used. iii) The roadside classification and the survey approach. iv) The LULC changes.

### **7.2.1 Study area**

Our study area is Roraima, located in the north of Brazil (see more details in Chapter 3).

### **7.2.2 Data**

We used images from the satellites Sentinel-2 MSI optical sensor, and Sentinel-1 SAR microwave sensor. In total, it was used 47 tiles from the MSI sensor and seven orbits from Sentinel-1 SAR. Sentinel-2 has a temporal

resolution of 5 days over Roraima, considering the sensors onboard of Sentinel-2A and Sentinel-2B. We used data without atmospheric correction (L1C) because data with atmospheric correction (L2A) are only available after December 2018 in the GEE platform. Also, we used NDVI (Normalized Difference Vegetation Index (Equation 4.1), proposed by Rouse et al. (1973), and LSWI (Land Surface Water Index) (XIAO et al., 2002, 2004) (Equation 4.2). For SAR data, the temporal resolution is 12 days for each satellite and 4 and 8 days when combining Sentinel-1A and Sentinel-1B. We used the IWD swath, GR data, and VV and VH polarizations. The ratio between the polarizations VH and VV (Equation 4.3) was used because it might have more stability than single polarizations VH or VV, being less affected by environmental factors or acquisition systems, as reported in the literature (VELOSO et al., 2017; HARFENMEISTER; SPENGLER; WELTZIEN, 2019).

A field mission was conducted to collect road-side data and validation points (see Section 4.1 for details), during August and September 2019. With the QGIS software and Sentinel-2 images, we draw the polygons and labels the classes for 2017 and 2018. In total, we got 5,3 thousand polygons, distributed in 10 different classes (forest, savanna, Campinarana, water, sand/rock, annual crops, perennial crops, pasture, conversion, impermeable).

### **7.2.3 LULC classification**

GEE platform (GORELICK et al., 2017) was used for our two LULC classification approaches. We adapted the methodology used in our test are (Chapter 6). Although MLP had slightly better results, we use the classifier RF since MLP is not available in the GEE platform. RF is commonly used in a multisensor SAR-Optical approach to map LULC (TORBICK et al., 2016, 2017a; CLERICI; VALBUENA CALDERÓN; POSADA, 2017; ZHOU et al., 2017; PAVANELLI et al., 2018). The classification results were generated with 30 meters instead of 10 meters (Chapter 6) of pixel size. Inside GEE, we filtered all Sentinel-2 and Sentinel-1 images available from Roraima. Also, we used all images with less than 50% of cloud cover. To remove the clouds and shadows,

we filter using the Bits and Sentinel-2 cloud probabilistic (GOOGLE, [s.d.]), with a threshold of 65%.

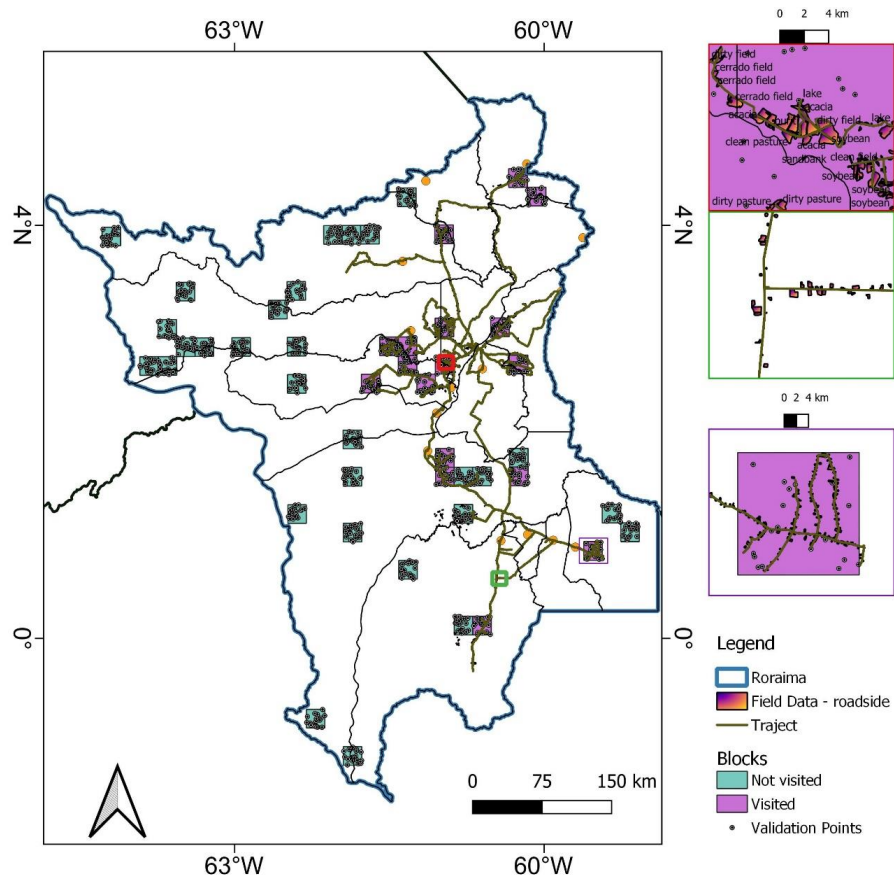
To summarize the images collections, we used standard deviation, minimum, maximum, and median metrics for Sentinel-1 and Sentinel-2 data. For Sentinel-1, we used the three different periods; P1: January to April; P2: May to August; and P3: September to December, for each year (2017, 2018, and 2019). P2 represent the rain season and P1 and P3 represent the dry season in Roraima. Due to the low number of cloud-free pixels (see Section 5.3.3), mainly for 2017, we adopted to use Sentinel-2 data for the entire year and not only of each period. For Sentinel-1, the metrics were calculated separately for each satellite (Sentinel-1A and Sentinel-1B). As post-classification, we used a mode filter (kernel in circle and radius equal to 1) in a post-classification step to smooth the results and avoid isolate misclassified pixels. Finally, the results were exported in GeoTiff format.

In our first approach to classifying LULC, we split our roadside data samples (Figure 7.1) into 75% for training and 25% for validation, for each year. We used a number of trees equal to 100. The metrics from both sensors were combined in a stack and then the RF models were trained using this stack.

For our second approach, to determine the Roraima LULC maps for 2017, 2018, and 2019, we adapted a two-stage sampling design used for soybean area estimate (KING et al., 2017a; SONG et al., 2017). For that, we provided blocks and samples, in two stages. Different from previous studies in the literature, we did not use a different stratum. In the first stage, blocks of 20 km x 20 km were randomly sampled. According to King et al. (2017) and Song et al. (2017), this block size allows a second-stage sample to be completed in one day. Roraima state was divided into a regular grid with 437 blocks. After, we randomly select 20 blocks to visit in situ and analyzed the possibility of accessing these areas in the field. Because Roraima state has few roads, and several times it was not possible to reach the blocks selected. In these cases, more blocks were randomly sampled. In total 48 blocks were previously selected and among them, 17 blocks were visited during one field campaign.

In the second sampling stage, we used a random sampling of 20 points in each block, to quantify the LULC. In total, 940 points were randomly sampled. To make the field work easier and faster, we delineated the field polygon around each point to visit in situ. and (see Section 4.1) we drawn routes connecting the points and the blocks. A total of 198 LULC valid points were collected. When it was not possible to identify the LULC class (impossible to reach the point or see the area with binoculars), we classified the LULC class using satellite data (NDVI time series from MSI and MODIS and visual interpretation of RGB image compositions).

Figure 7.1 – Details about the field data mission using a two-stage sampling approach.



Source: Author's production.

In our study, we used one stratum to sample the blocks. To estimate the area for each class ( $j$ ), we calculated the class area in each block  $i$  ( $Y_{i,j}$ ). Thus,

$Y_{i,j}$  is given by the relationship between the area of block  $i$  ( $A_{j,i}$ ) and the proportion of the samples in the class ( $j$ ) among the samples inside block  $i$  (7.1). The total area, for each class ( $\widehat{Y}_j$ ), is given by the sum of the areas  $Y_{i,j}$  of the class  $j$  in each block  $i$ , where  $n_j$  is the sample and  $N_j$  is the population size for the class  $j$  (7.2). Finally, it was calculated the variance from each class using 7.3.

$$\widehat{Y}_{j,i} = A_{j,i} \widehat{P}_{j,i} \quad 7.1$$

$$\widehat{Y}_j = \frac{N_j}{n_n} \sum_{i=1}^{n_j} \widehat{Y}_{j,i} \quad 7.2$$

$$V(\widehat{Y}_j) = N_j^2 \left(1 - \left(\frac{n_j}{N_j}\right)\right) \left(\frac{s_j^2}{n_j}\right) \quad 7.3$$

$s_j^2$  represent the sample variance for each class  $j$ . The Standard Error (SE) is the root square of the variance  $V(\widehat{Y}_j)$ .

After, we classified the LULC inside each block using the same methodology mentioned for the roadside classification. This step is useful to correct the areas estimations of the first step. For last, LULC classification was performed at once to the entire state. Accuracy verification was performed in these two last steps using the validated points. It was used the Overall Accuracy (OA), Users (UA), and Producers (PA) Accuracies, derived from the confusion matrix. These metrics were computed in python routine, using the sklearn package.

#### 7.2.4 Determination of LULC changes

LULCC classifications avoid carrying on misclassification over each LULC process, generating better results than analysis of the changes between LULC classifications (ZALLES et al., 2019). Thus, we classified the LULCC between 2017 and 2019. To classify such changes, we follow the same procedure that was used for the roadside classification in the previous section. The only

difference was that we used Sentinel-1 and Sentinel-2 data from 2017 to 2019, and not for different periods.

### 7.3 Results

We present this section divided into three parts. i) Results from the roadside approach. ii) Results for the approach using the two-stage sampling process. iii) Results of the LULCC from 2017 to 2019.

#### 7.3.1 Roadside LULC classification

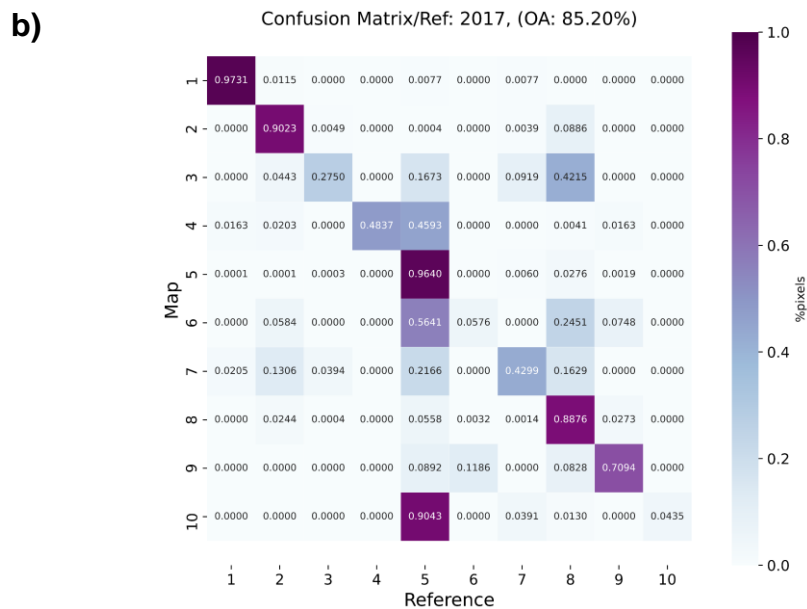
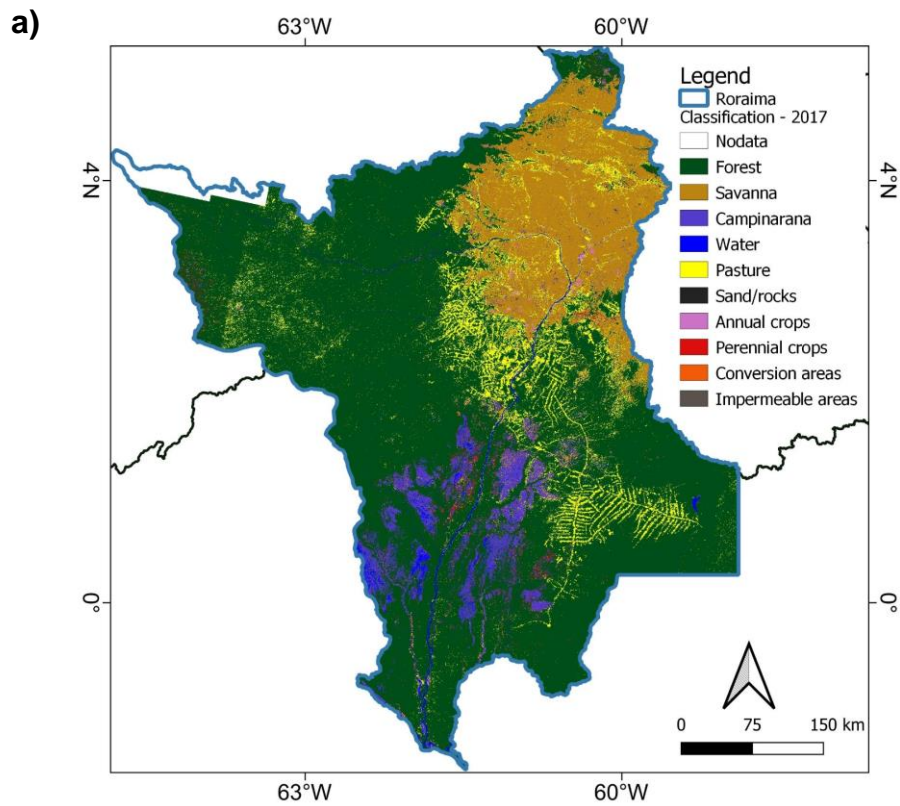
We used part of our roadside data to train the RF classifier. Forest (Table 7.1) is the predominant class, meanwhile, Pasture is the main anthropic class. The OA of the classifications were 85.20%, 89.26%, and 89.94%, for 2017 (Figure 7.2), 2018 (Figure 7.3), and 2019 (Figure 7.4), respectively. As Sentinel-1A does not cover the entire state, there are a spot without data in the northwestern part. Using the validation dataset, the Users (UA) and Producers (PA) Accuracies were computed for each year (Figure 7.5).

Table 7.1. Area (in km<sup>2</sup>) for each Roraima LULC class, for 2017, 2018, and 2019, using the roadside data.

Id	Class	2017		2018		2019	
		Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%	Area (km <sup>2</sup> )	%
1	Forest	156163	70.0	157590	70.6	154030	69.0
2	Savannas	32539	14.6	33026	14.8	33121	14.8
3	Campinaranas	10800	4.8	14185	6.4	17207	7.7
4	Water	3094	1.4	2981	1.3	1860	0.8
5	Pasture	17844	8.0	12662	5.7	15826	7.1
6	Sand/rocks	62	0.0	98	0.0	1	0.0
7	Annual Crops	390	0.2	703	0.3	556	0.2
8	Perennial Crops	2125	1.0	1801	0.8	440	0.2
9	Conversion	47.4894	0.0	53	0.0	42	0.0
10	Impermeable	0.5985	0.0	10	0.0	26	0.0

Source: Author's production.

Figure 7.2 – LULC classification map (a) and confusion matrix (b) for 2017 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.

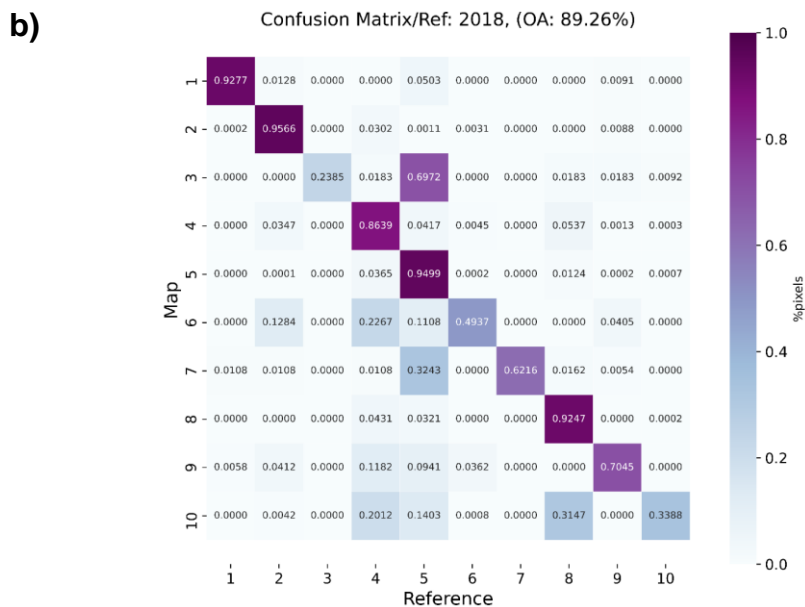
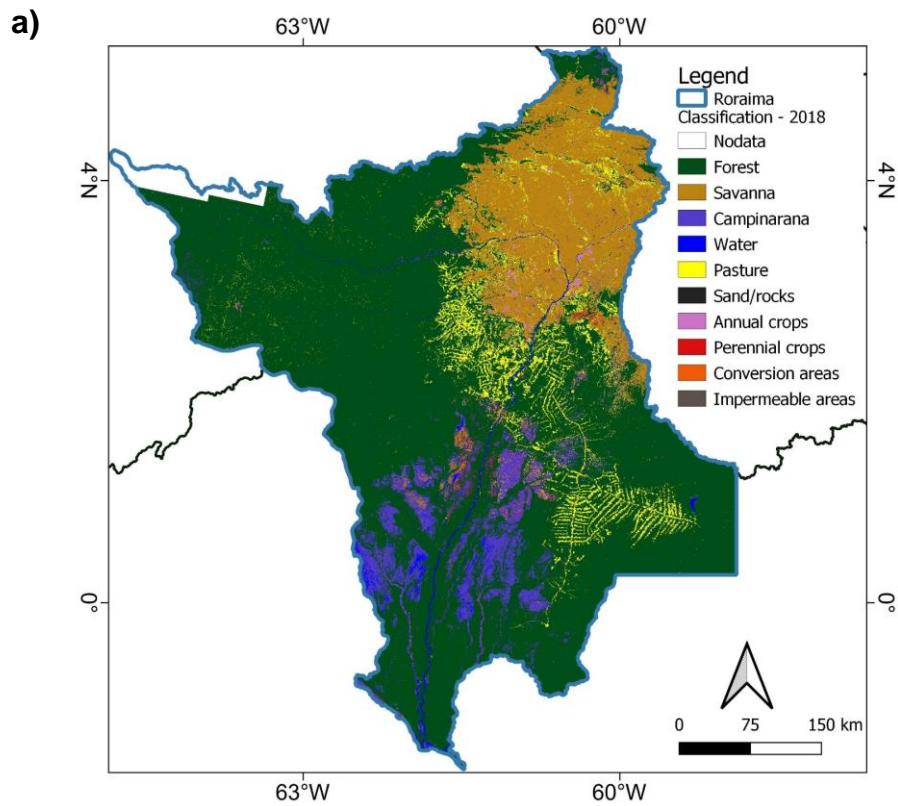


1: Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.



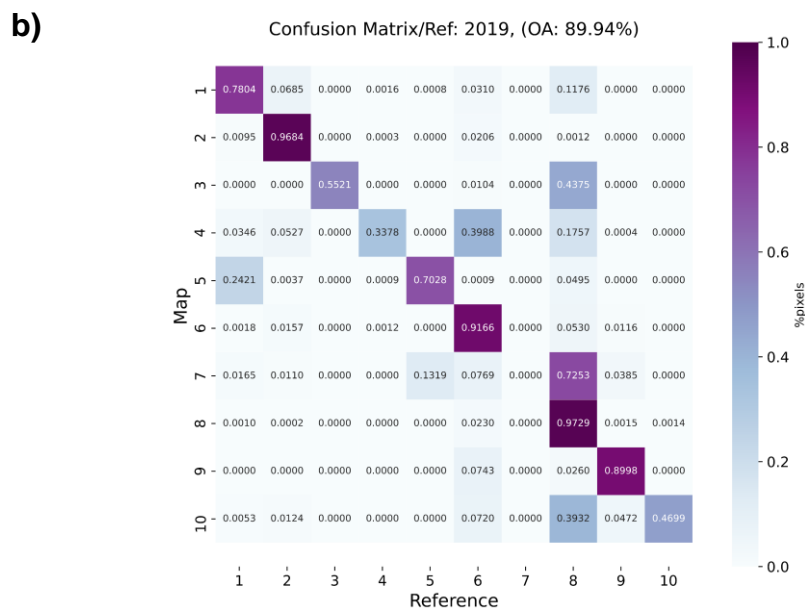
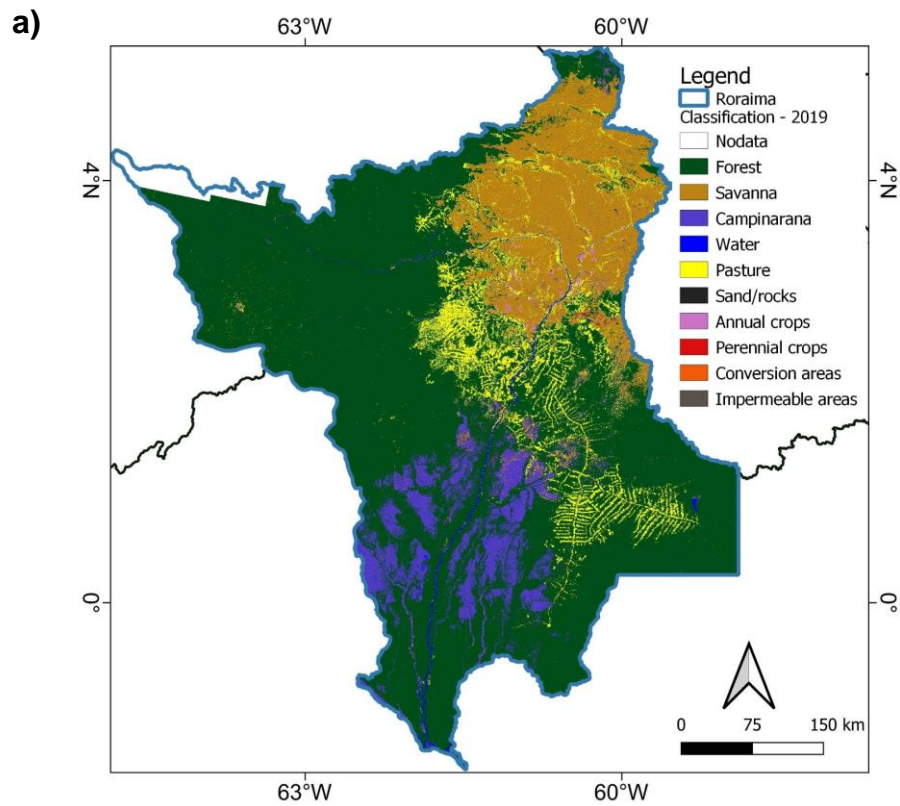
Figure 7.3 – LULC classification map (a) and confusion matrix (b) for 2018 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

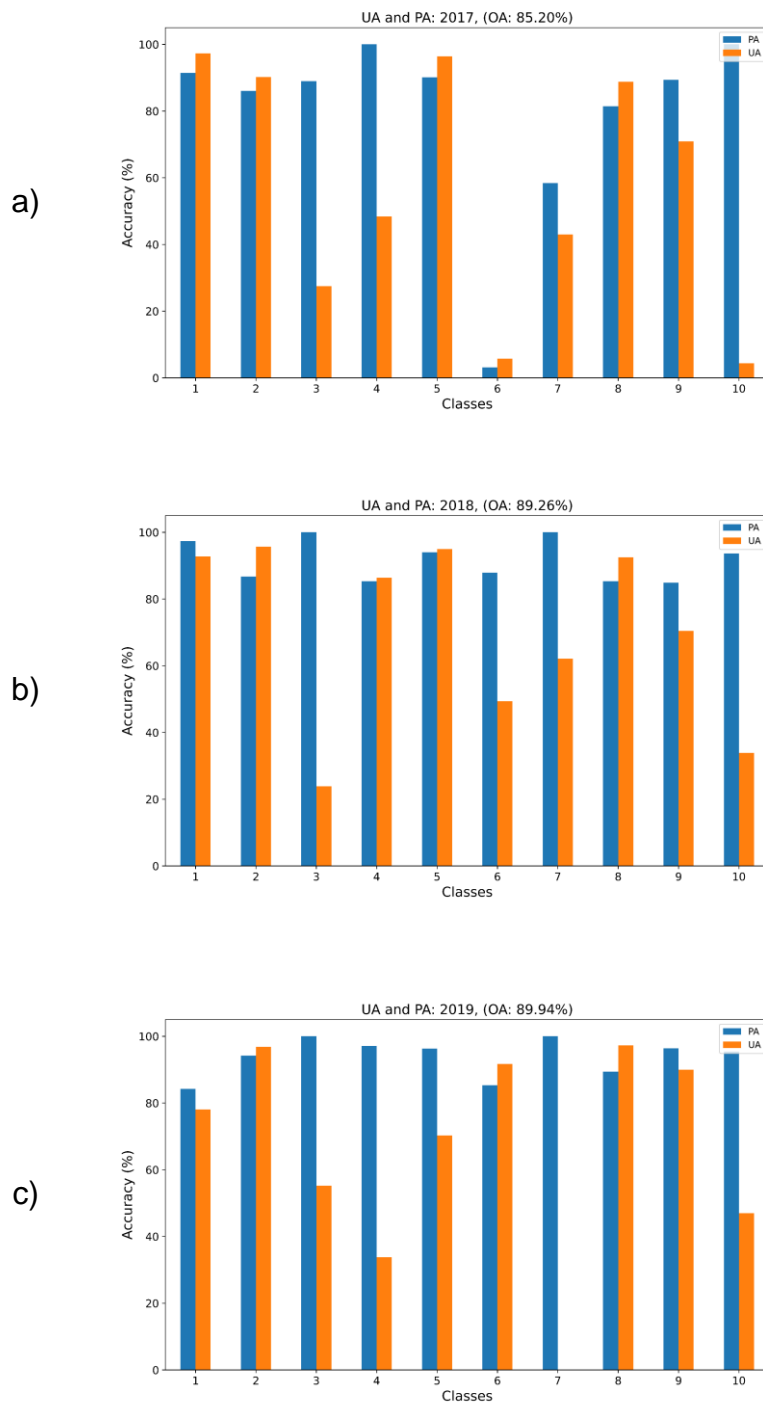
Figure 7.4 – LULC classification map (a) and confusion matrix (b) for 2019 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

Figure 7.5 – Users (UA) and Producers (PA) Accuracies for 2017 (a), 2018 (b), and 2019 (c) LULC classification, using testing and multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

### 7.3.2 Sample-based LULC classification

In the first stage of the two-stage sampling approach, it was possible to estimate the area of the LULC classes found on the blocks (Table 7.2). The predominant class in Roraima is Forest, followed by Savannas and Campinarana. Annual Crops represent near 1% of Roraima territory in 2019, meanwhile, Pasture represents around 6%. Sand/rocks does not contain any validate points and is not present in this first estimation.

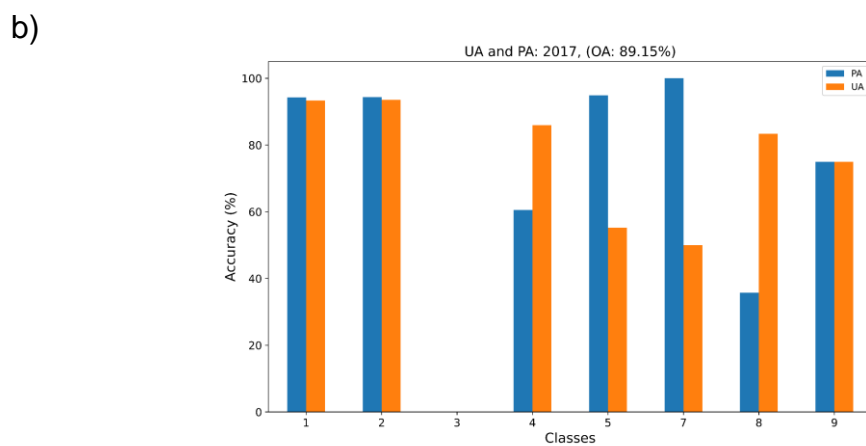
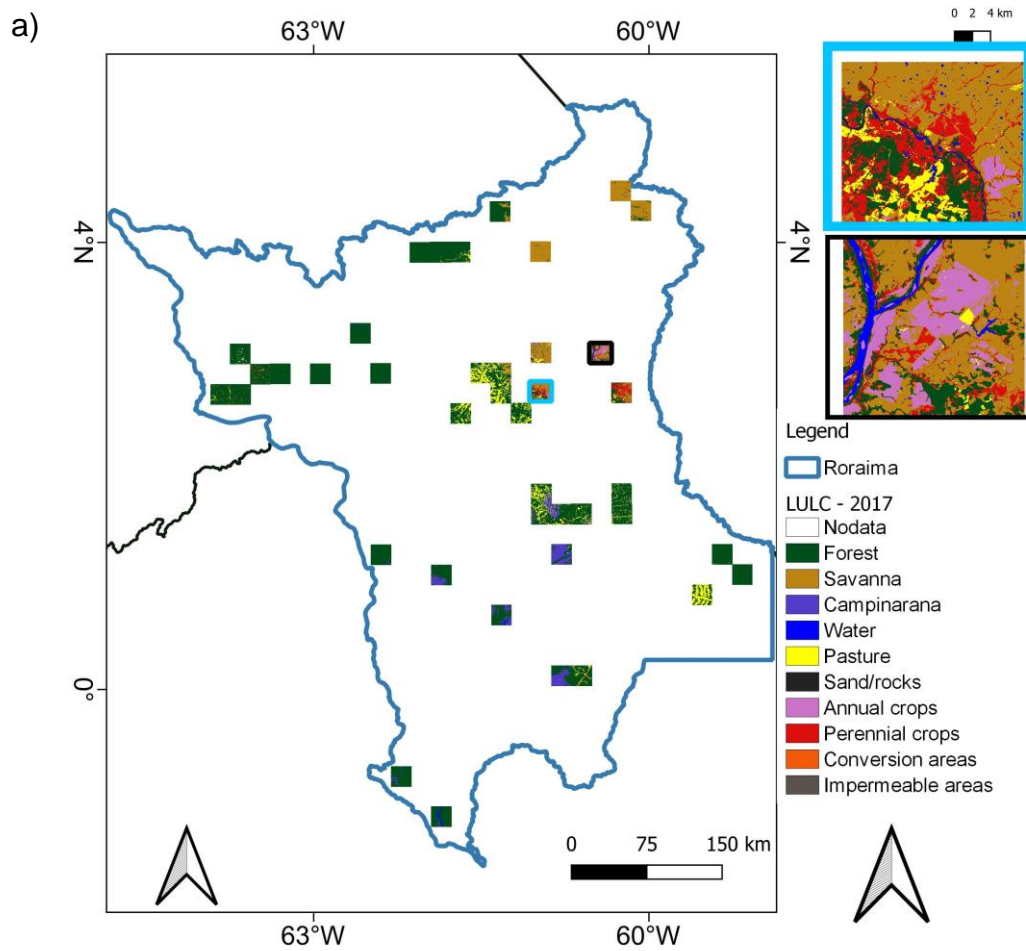
Table 7.2. Estimate area (in km<sup>2</sup>) for each LULC class for 2017 and 2019 considering the points from the block approach. SE: Standard Error.

Id	Class	2017			2018			2019		
		Area	SE	%	Area	SE	%	Area	SE	%
1	Forest	128497	7546	73.5	128125	7585	73.3	127195	7682	72.8
2	Savannas	19526	6516	11.2	19526	6516	11.2	19526	6516	11.2
3	Campinaranas	12459	3929	7.1	12459	3929	7.1	12459	3929	7.1
4	Water	744	340	0.4	744	340	0.4	744	340	0.4
5	Pasture	10600	2960	6.1	10971	3060	6.3	11901	3288	6.8
7	Annual Crops	1488	881	0.9	1860	1121	1.1	1860	1121	1.1
8	Perennial Crops	1116	892	0.6	1116	892	0.6	1116	892	0.6
9	Conversion	372	351	0.2	0	0	0.0	0	0	0.0

Source: Author's production.

From the 47 sampled blocks, six have only one LULC class inside. Thus, we used 41 blocks for the LULC classification process each year. After, using the validation points we generated the accuracies from these maps for each block. The OA values were 89.15% for 2017, 91.34% for 2018, and 90.24% for 2019. Some classes that were in our training data, were not in the validation set. Using the area for each class, in each block, we corrected the LULC estimate (Table 7.3).

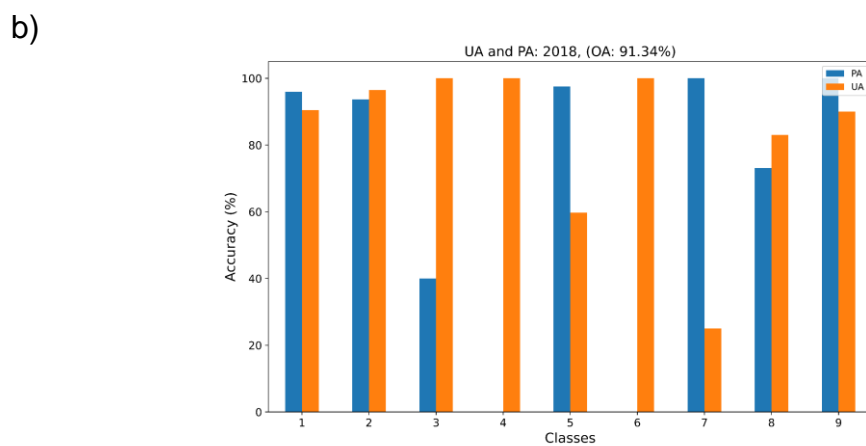
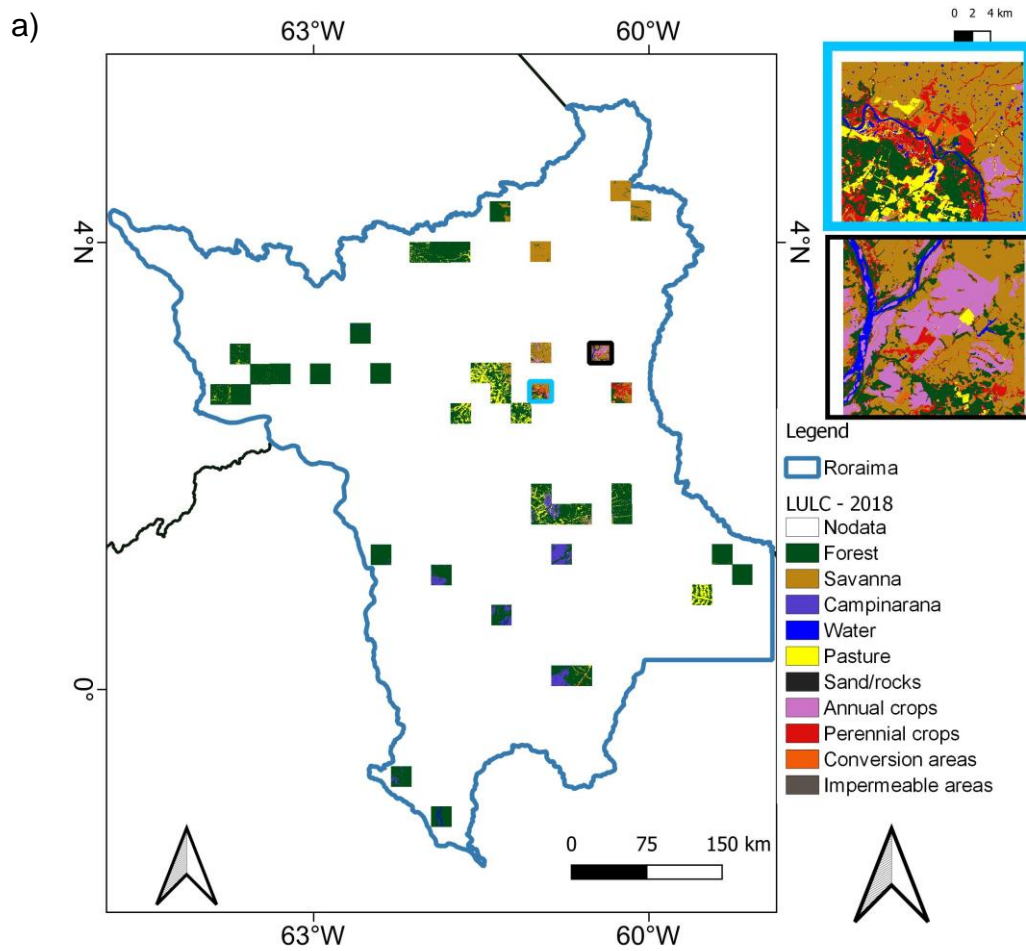
Figure 7.6 – LULC classification (a) and Users (UA) and Producers (PA) Accuracies (b) for the area of the blocks, for 2017, based on multisensor Sentinel-1 and Sentinel-2 data processed in GEE platform using RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

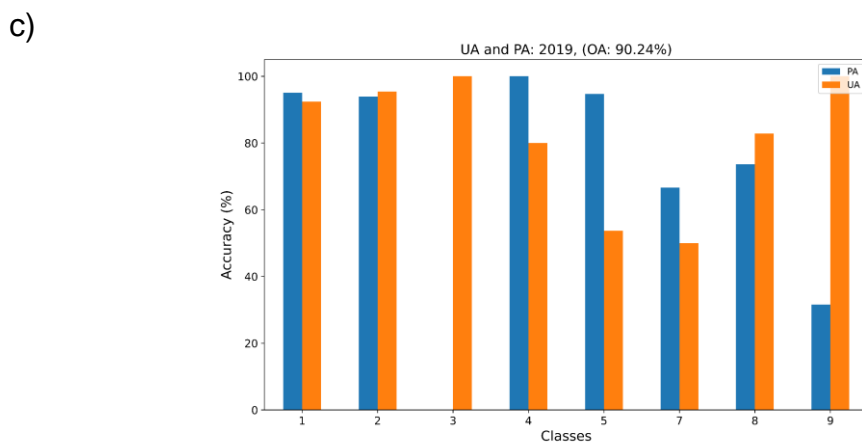
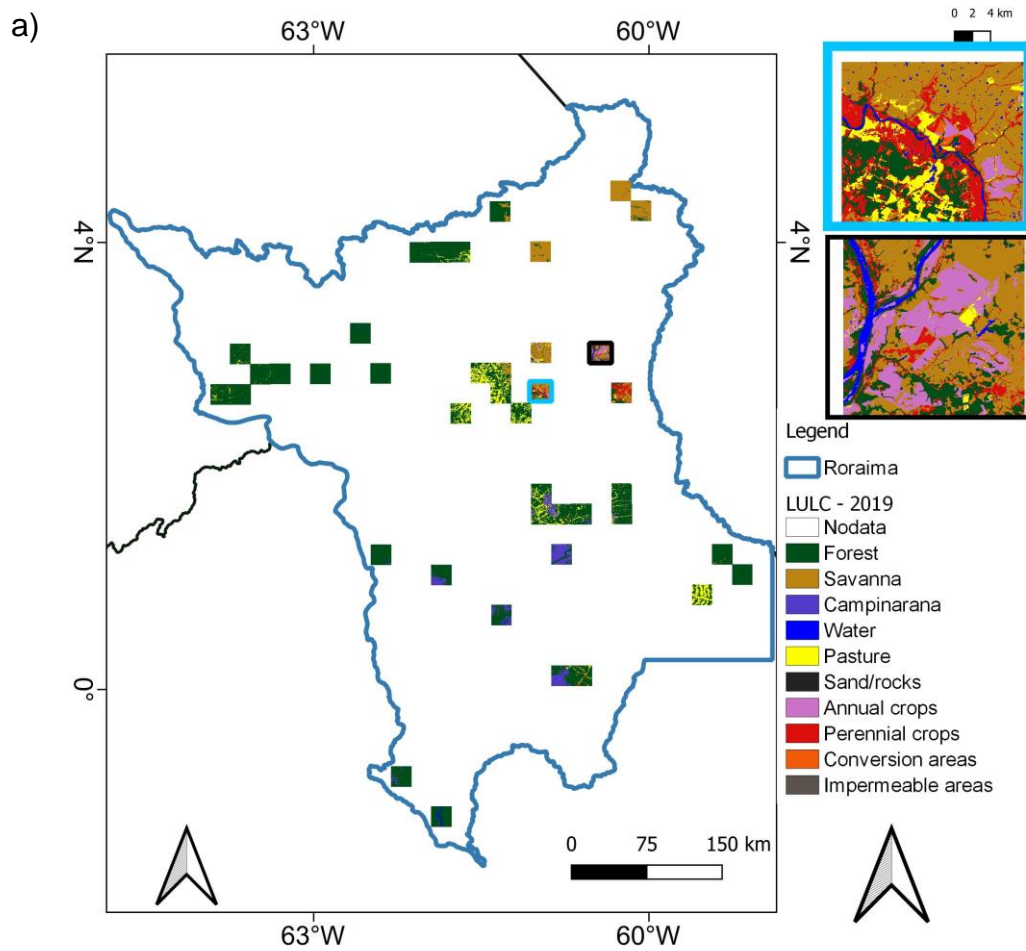
Figure 7.7 – LULC classification (a) and Users (UA) and Producers (PA) Accuracies (b) result from blocks to 2018 with multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

Figure 7.8 – LULC classification (a) and Users (UA) and Producers (PA) Accuracies (b) result from blocks to 2019 with multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

Table 7.3 – Estimate area (in km<sup>2</sup>) for each LULC class for 2017, 2018, and 2019, after the classification of each block. SE: Standard Error.

Id	Class	2017			2018			2019		
		Area	SE	%	Area	SE	%	Area	SE	%
1	Forest	127104	7606	72.7	127848	7424	73.1	126841	7577	72.6
2	Savannas	19010	6402	10.9	18767	6287	10.7	18572	6238	10.6
3	Campinaranas	9014	3234	5.2	8986	3213	5.1	9032	3286	5.2
4	Water	1098	368	0.6	1134	370	0.6	1067	364	0.6
5	Pasture	13869	3381	7.9	13588	3198	7.8	14515	3395	8.3
6	Sand/rocks	61	37	0.0	32	17	0.0	42	17	0.0
7	Annual Crops	1167	750	0.7	1564	944	0.9	1373	893	0.8
8	Perennial Crops	3055	1718	1.7	2587	1525	1.5	3013	1697	1.7
9	Conversion	424	249	0.2	295	164	0.2	344	110	0.2

Source: Author's production.

The last step consisted in the LULC classification, at once, for Roraima state. The LULC maps and confusion matrixes are present in Figure 7.9 to Figure 7.11. Figure 7.12 The PA and UA values of each one of these maps can be observed in Figure 7.12. The OA was 88.15%, 89.13%, and 89.12% for 2017, 2018, and 2019, respectively. Areas of each class in each classification are presented in Table 7.4.

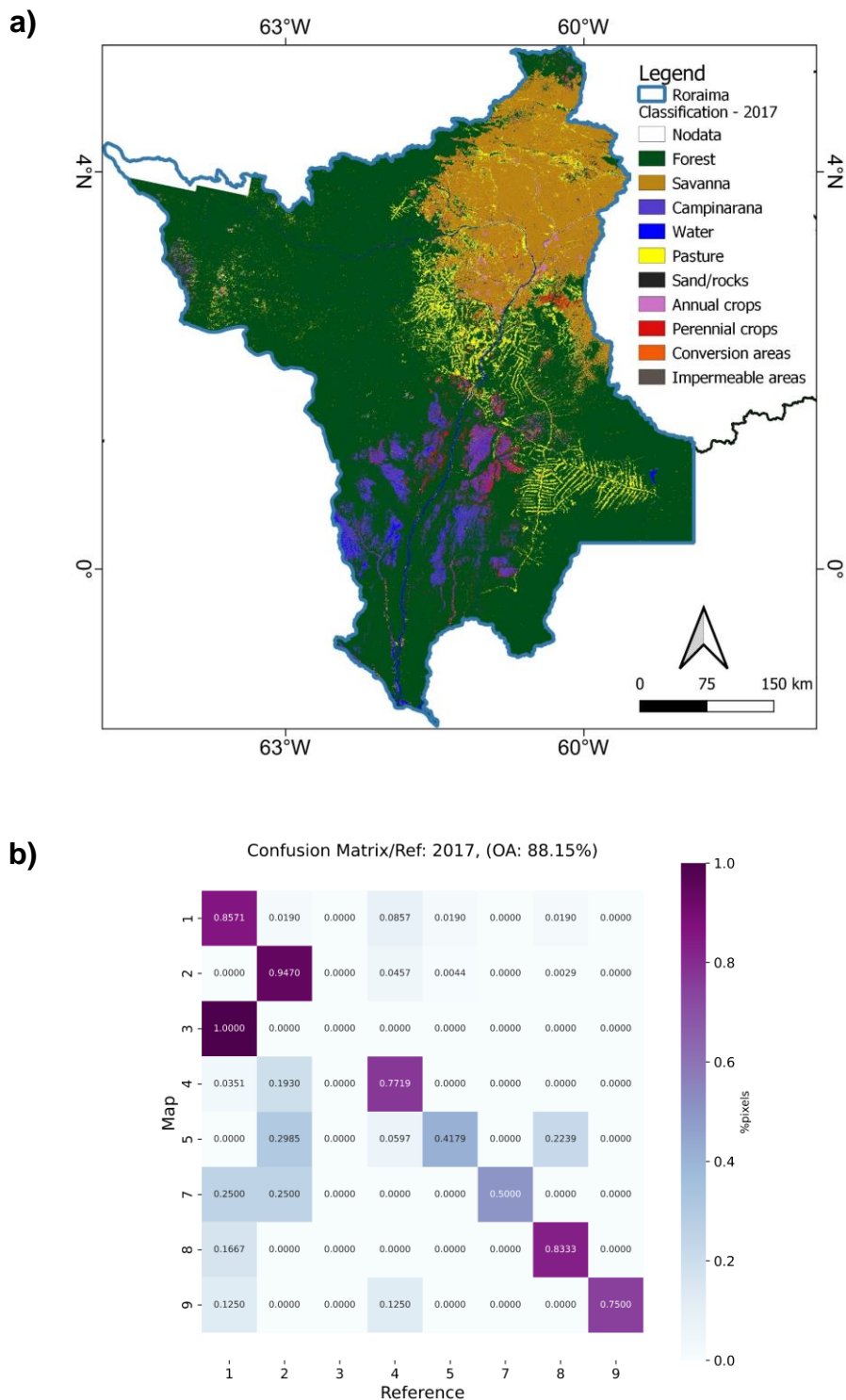
Table 7.4. Area (in km<sup>2</sup>) for each LULC class for 2017, 2018, and 2019 after the classification in each block.

Id	Class	2017		2018		2019	
		Area	%	Area	%	Area	%
1	Forest	159457	71.5	158801	71.2	160729	72.0
2	Savannas	32447	14.5	31317	14.0	31547	14.1
3	Campinaranas	9993	4.5	10767	4.8	9525	4.3
4	Water	2613	1.2	2751	1.2	2609	1.2
5	Pasture	13035	5.8	15094	6.8	13548	6.1
6	Sand/rocks	109	0.0	165	0.1	127	0.1
7	Annual Crops	641	0.3	621	0.3	671	0.3
8	Perennial Crops	4767	2.1	3560	1.6	4153	1.9
9	Conversion	2	0.0	32	0.0	200	0.1

Source: Author's production.



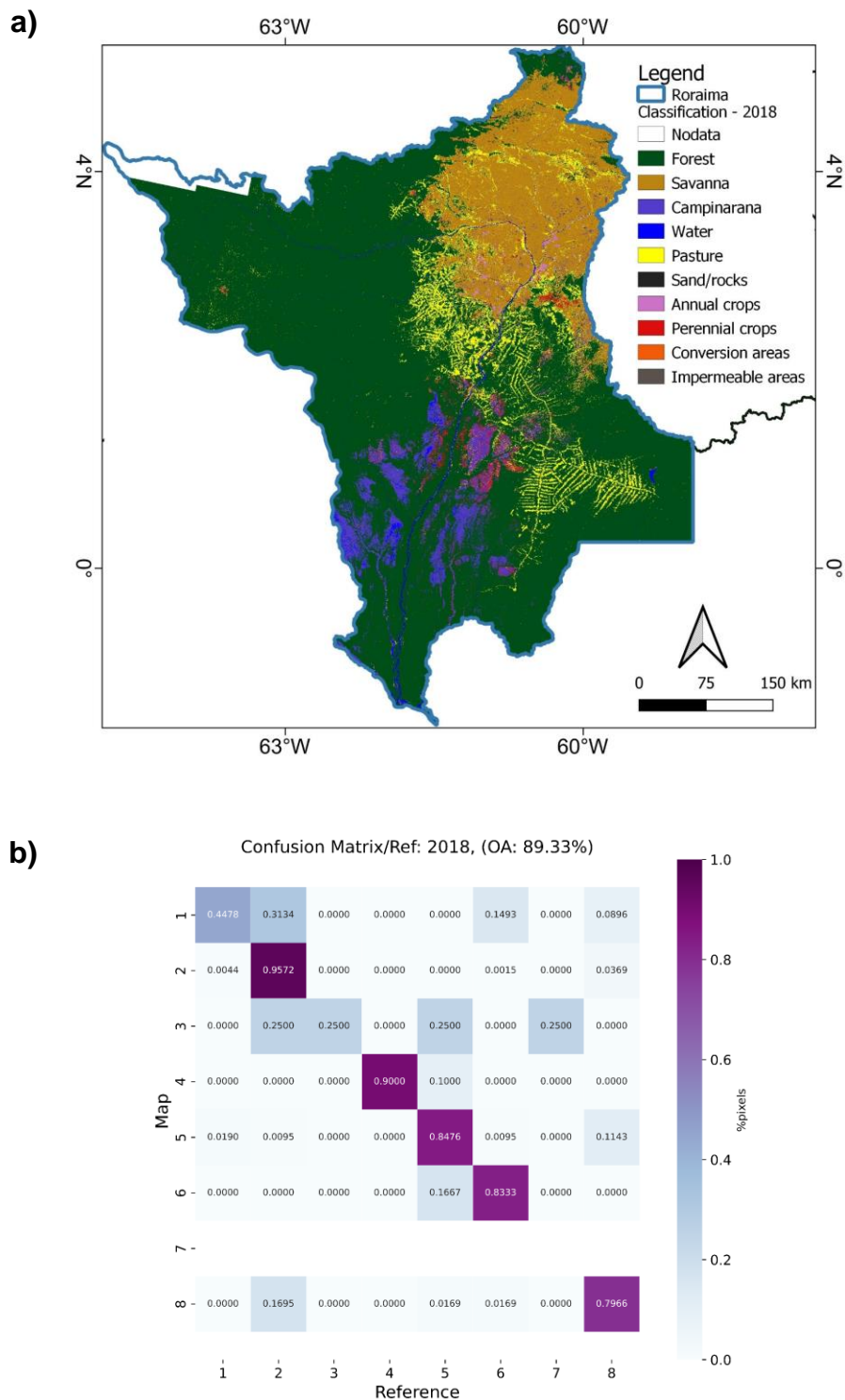
Figure 7.9 – Sample approach LULC classification map (a) and confusion matrix (b) for 2017 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

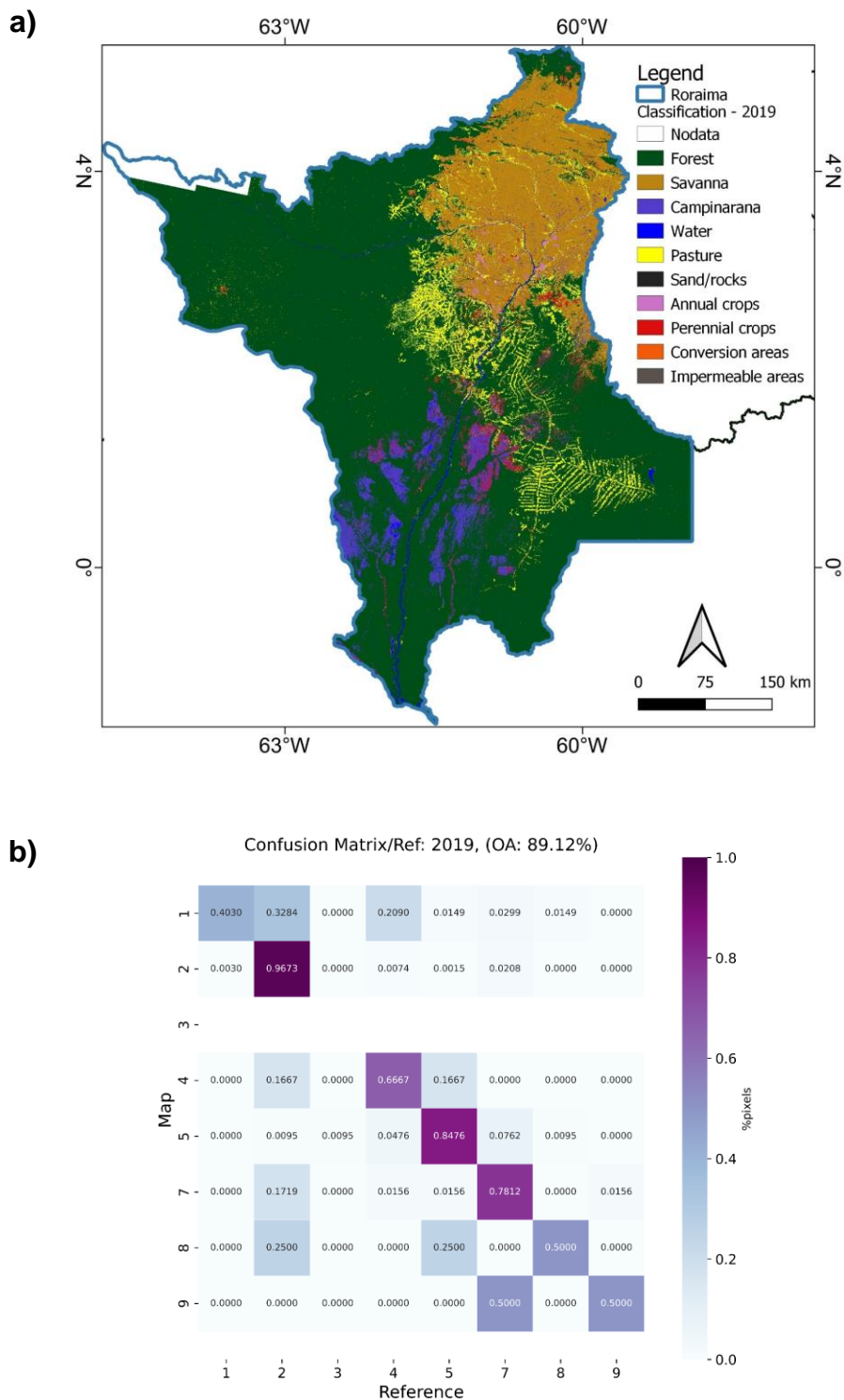
Figure 7.10 – Sample approach LULC classification map (a) and confusion matrix (b) for 2018 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

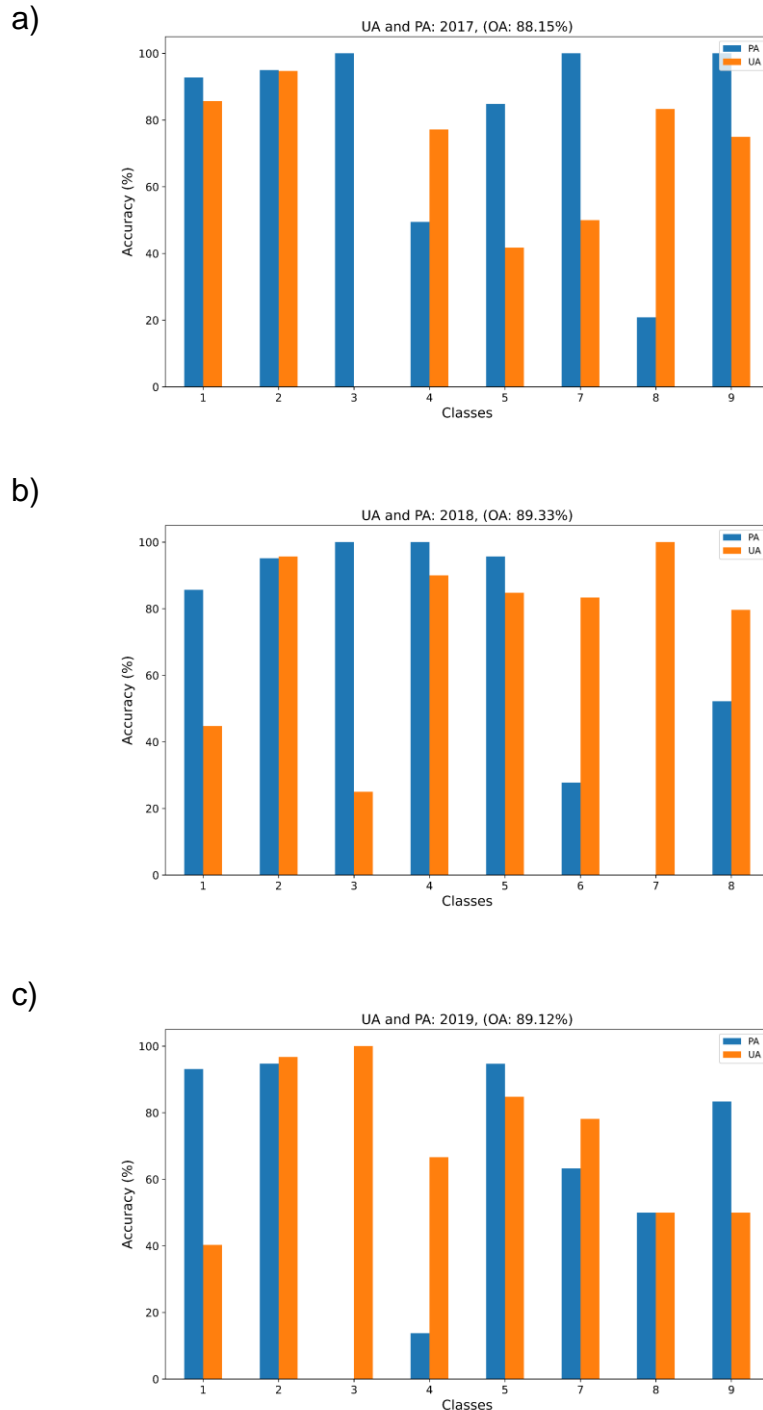
Figure 7.11 – Sample approach LULC classification map (a) and confusion matrix (b) for 2019 using Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

Source: Author's production.

Figure 7.12 – Users (UA) and Producers (PA) Accuracies for 2017 (a), 2018 (b), and 2019 (c) considering the blocks to the LULC classification, using testing and multisensor Sentinel-1 and Sentinel-2 data in GEE platform with RF classifier.



Forest, 2: Savannas, 3: Campinarana, 4: Water, 5: Pasture, 6: Sand/rocks, 7: Annual Crops, 8: Perennial Crops, 9: Conversion areas, 10: Impermeable areas.

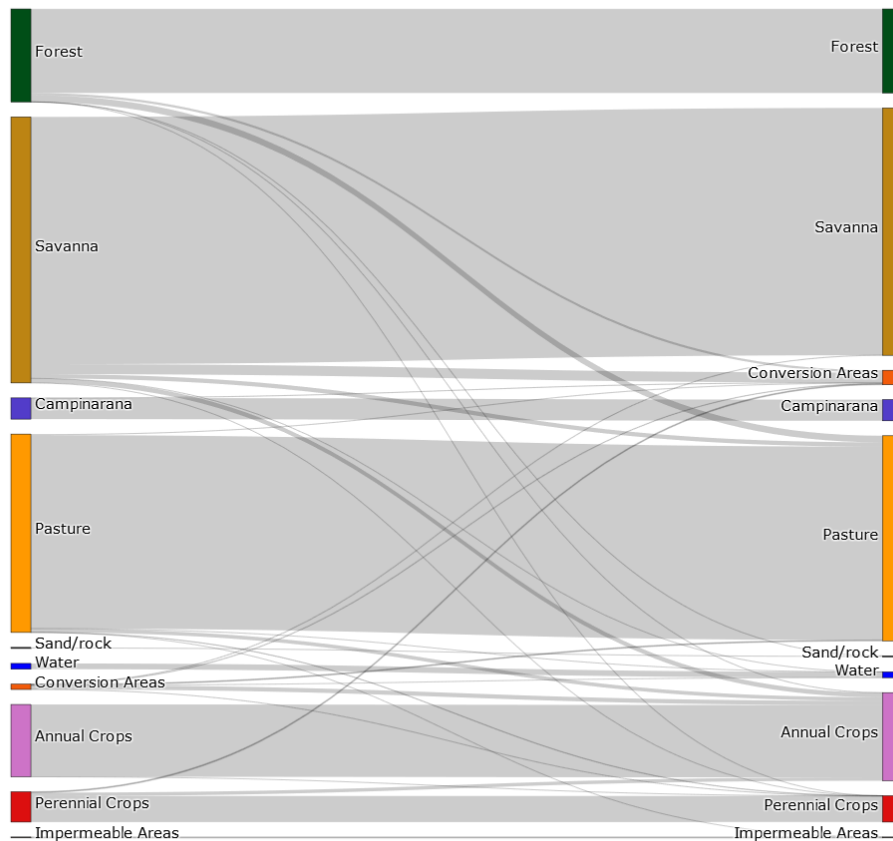
Source: Author's production.

### 7.3.3 LULC changes

To better understanding the LULC changes, we used the Sankey diagram (SCHMIDT, 2008). Analyzing our roadside dataset, it is possible to identify that Savannas and Forest are the natural predominant classes (Figure 7.13). On the other hand, Pasture, Annual and Perennial Crops are the non-natural predominant classes. Besides, with the field data we can see some of the changes as Forest to conversion areas or Pasture, Savannas to Pasture or Annual Crops, and Perennial Crops to Annual Crops.

Figure 7.13 – Sankey graph to represent the LULCC based on the roadside field data.

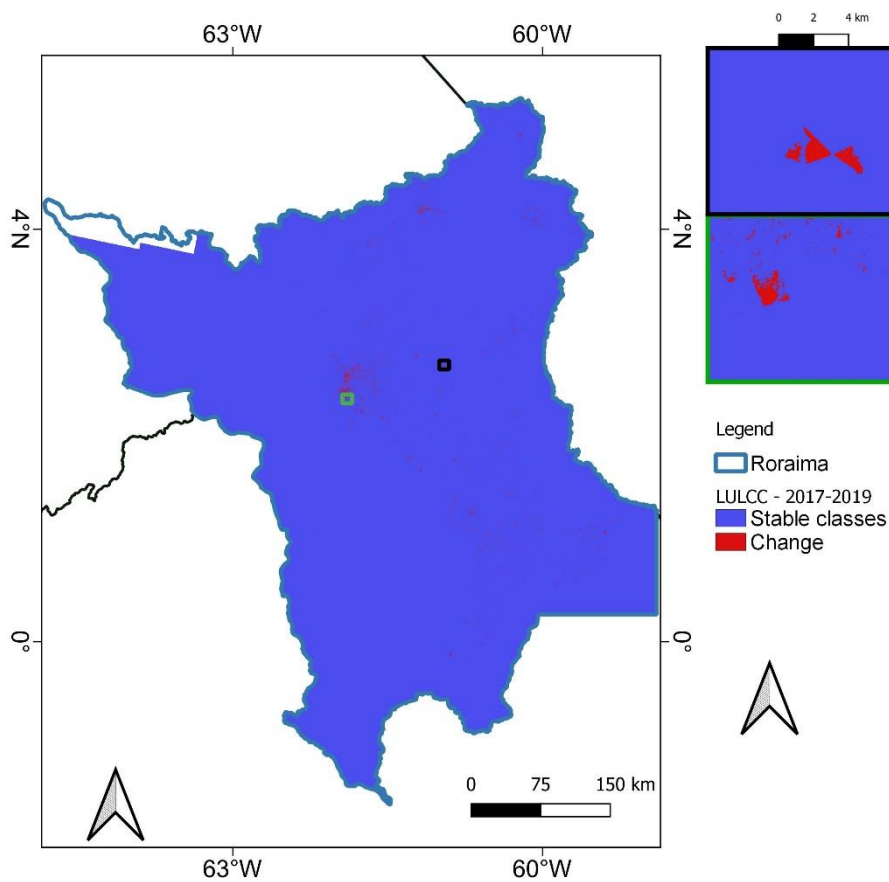
Land Use and Land Cover Changes - 2017 to 2019 - Roraima



Source: Author's production.

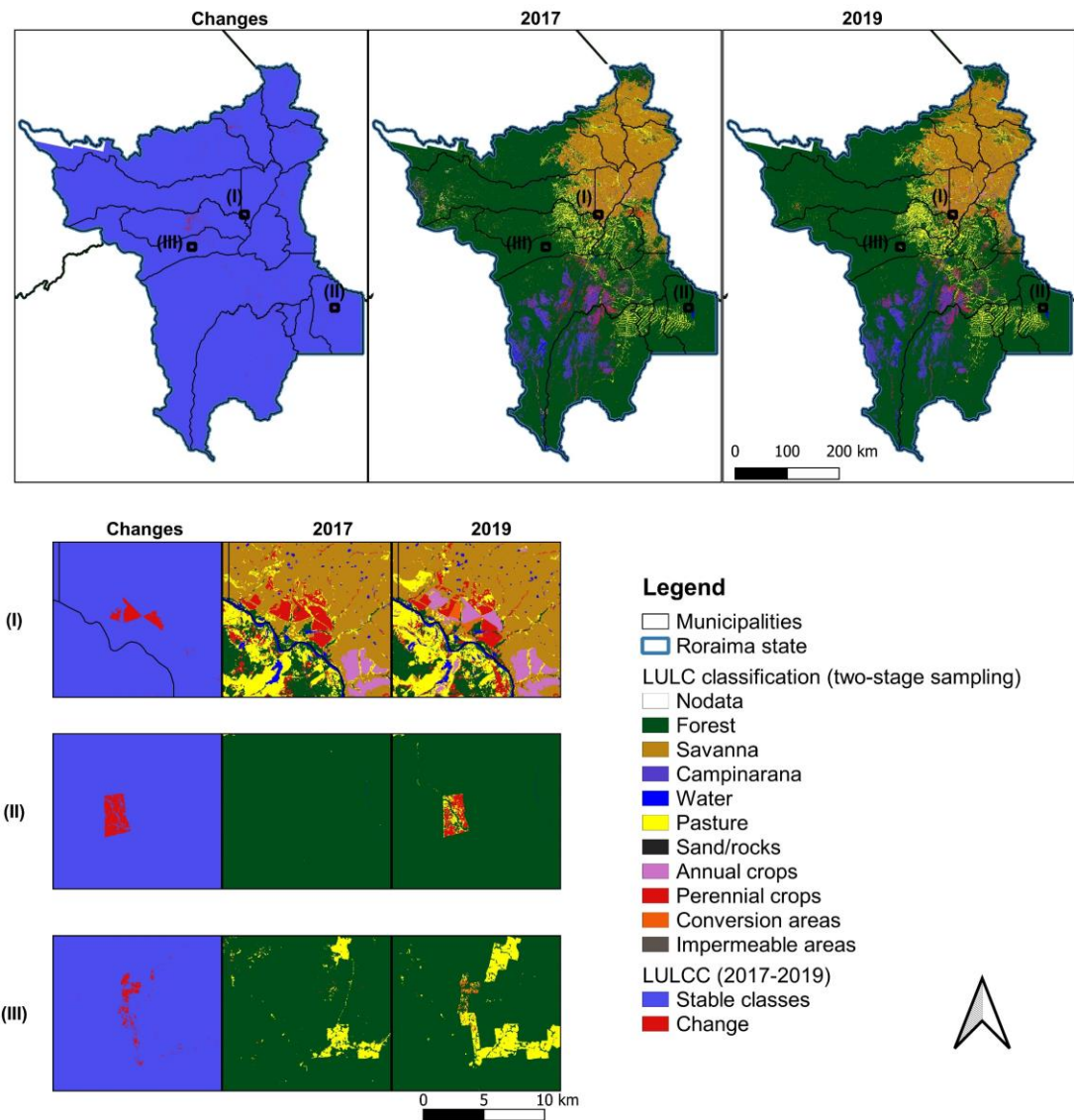
The LULCC classification based on the roadside data, for the whole state, using the GEE platform, is presented in Figure 7.14. In general, the stable class predominates (99.7%) over the change class. The changes detected are mainly from Forest to Pasture, Savannas to Pasture or Agriculture, and Perennial Crops (manly Acacias) to Annual Crops (Figure 7.15).

Figure 7.14 – LULCC classification from 2017 to 2019, using the roadside data. Change details on the right.



Source: Author's production.

Figure 7.15 – Details of LULCC classifications and the two-stage sampling LULC classification from 2017 to 2019.



Source: Author's production.

## 7.4 Discussion

Accurate and update LULC information is challenging in large tropical regions, as Roraima states. The use of remote sensing data is recommended for this task, but frequent cloud cover and huge data processing are some of the limitations. In this research, we proposed two approaches to try to minimize such issues. The use of SAR data combined with optical data, and the use of a two-stage sampling-based approach to estimate the LULC classes which have the

benefit to be quicker and less computational-consuming. To guarantee good training samples to the classification process, the samples from our study came from fieldwork and from visual interpretation of time series of optical images, by a specialist.

For the roadside LULC classification, our results showed that the OA is near 90% for 2018 and 2019. Lower OA was found for 2017, the year that had the fewer quantity of cloud-free data available (Figure 5.15). Although the use of roadside data for LULC classification has been a known approach (WALDNER et al., 2019), the process takes a considerable time, generating many polygons and demanding high computational power to train the classifiers. So, the time between the field campaign and the firsts LULC results will be longer than the two-stage sampling approach. With a stratified method, it is possible to provide the first results quicker than with the roadside approach. The first results are later corrected with the LULC classification of each block. Different from the roadside approach, with the two-stage sampling it is possible to estimate the proportion of the LULC areas in different levels, even before the classification to the entire area.

Forest and Savannas are the predominant classes in Roraima. The Campinarana located in the southern part of the state show some misclassification with Savannas and Perennial crops. One reason could be because there are the Campinas and Campinaranas inside our Campinarana class. Campinas is formed by grassland and small shrubs, similar to the Savannas. Moreover, Campinaranas is formed by small trees, that could be like some Perennial species, as Acacias.

The pasture class is the predominant non-natural class in Roraima. It is more present in the southern part, in the middle of the Forest class. Extensive cows' pasture is the activity more present in Roraima. In Perennial Crops, Acacias is the predominant species, present mainly in the lavrados region. In recent years has been replaced for Annual crops (soybean and corn). Dende, Orange, and Banana are common Perennial Crops species located more in the south. However, Orange and Bananas fields are small, and classify them with the



proposed methodology was challenging. Dende, used for palm oil, is produced on a more industrial scale, with bigger areas than orange groves, which facilitates the identification.

For Annual Crops, comparing (Table 7.5) the areas estimated using our approaches with the Municipal Agricultural Production (PAM) data from IBGE (2018a), the two-stage sampling approach has more similarity with PAM data for 2018 and 2019. The roadside approach result has the same pattern as PAM, although with lower values for 2017, and higher values for 2018. We do not compare our results with the Perennial Crops from PAM, because IBGE does not consider Acacia species as Perennial Crops.

Table 7.5 Area (in km<sup>2</sup>) from Municipal Agricultural Production (PAM) data from IBGE and our roadside and two-stage sampling-based classifications for the Annual Crops class.

<b>Approach</b>	<b>2017</b>	<b>2018</b>	<b>2019</b>
PAM	476	655	635
Roadside	390	703	556
Two-stage sampling	641	621	671

Source: Author's production.

The major changes that we observed were from Forest to Pasture and from Perennial Crops to Annual crops. Also, during the training process, it was possible to identify small deforestation areas near the rivers in the northwestern part of Roraima. Many of these areas are destined for mining activities. However, due to the mining process, rafts, and dredges, the identification using remote sensing data is difficult (LOBO et al., 2018).

Understanding where and when is LULCC is happening has an important role in food security and environmental protection. However, this type of information is not available for Roraima yet. In this sense, our study brings some insights to provide LULC classifications for this state. Our results showed a few LULCC from 2017 to 2019 in Roraima. Three years is a short period for LULCC in Roraima. Besides, it is important to highlight that the RS data used in this paper started later in 2017 (considering Sentinel-2A and 2B). Moreover, Sentinel-1 represents a new era, with SAR data for free of charge and better temporal and

spatial resolution. If the data were available, in early 2000, for example, the LULCC detection would be bigger. Our results indicate the usefulness of the methodology proposed to map LULC, which can help improve existing LULC mapping initiatives, as well as be helpful to identify and monitoring future LULCC.

## **7.5 Conclusion**

Our research explored remote sensing data from optical and SAR sensors (Sentinel-2 and Sentinel-1, respectively), and two methodologies, to map the LULC for Roraima state in 2017, 2018, and 2019. The multisensor data was necessary because of the high cloud cover of this region. Field samples collected during a field campaign, alongside the roads, and on selected blocks, were essential for the analysis conducted.

Based on our results, the two-stage sampling point has the potential to provide LULC information in fewer steps with less computational demand. However, as the blocks and points are randomly sampled, some classes of small areas might not be well represented. On other hand, the roadside approach has the advantage that it captures small classes, but it needs higher computational power, as well as more time to be run, to pre-process the field data. Overall, by applying the combination of SAR and optical data it was possible to obtain reasonable LULC maps.

## 8. OVERALL DISCUSSION

LULC information about Brazil has an important role for economics and environmental purposes. However, there is a lack of updates and accurate information for the entire territory. Without such knowledge about the spatial LULC distribution is not possible to monitor and understand the and use dynamics and impacts. Initiatives as TerraClass (ALMEIDA et al., 2016) (Figure 3.7) and MapBiomas (SOUZA et al., 2020) (Figure 3.8) are trying to change this scenario. However, for the Roraima state, the first one does not cover the entire state yet, and the second one needs some improvements. In this context, our approach shows some promising results that can help such initiatives to improve their methodology for mapping the Roraima state, as well other regions.

One of the reasons for the shortage of information about LULC is the frequent cloud cover over tropical regions. Our first study showed that the use of Optical Remote Sensing (ORS) data for earth observation is a challenge for most parts of South America (SA). This limitation is aggravated in the Equatorial zones (between 15° N and 14 °S), where there is a higher Cloud-Cover Frequency (CCF) when compared to that of regions with latitudes between 15° N and 40° N. We found that it is challenging to use optical sensors even in the months with lower cloud frequency (June to August), which represents medium CCF interference (CCF near 40-50%) in most of SA, as we showed in Figure 5.11.

Roraima state, located in a tropical region near the Equator line, has frequent cloud cover, which makes it difficult to sense this area with an optical sensor even if is the Sentinel-2A and 2B with a temporal resolution of 4 to 8 days. The period more affected is between April and August, the crop season (QP2). Corroborating with our study, Pavanelli et al. (2018) mentioned that due to the frequent cloud cover, mainly during the crop seasons, it is almost impossible to obtain clear sky observations with optical data in this region. Thus, for LULC mapping, ORS data from QP1 (January to April) and QP3 (September to December) are not enough, because the Annual crops class will be not well represented.

Associating the lack of accurate LULC information with the cloud cover limitation, and the recent agriculture expansion, turn Roraima LULC classification a challenging task. In this context, the use of microwave data, as Sentinel-1 SAR, is an alternative to improve the LULC mapping. Although SAR data is less affected by atmospheric than ORS data, the interpretation of these types of data is more complex and less widespread. Therefore, the key is to combine optical and –SAR, in a multisensor approach to take benefit of the complementary character of each type of sensor.

To test the optical-SAR approach, we choose a region in Roraima with ecological tension between Forest and Savannas, with different classes, as Annual and Perennial Crops, Pasture, Water, Campinarana, etc. We used two different classifiers algorithms in 29 scenarios, with optical, SAR, and SAR-optical datasets, in five different periods of the year 2019. Multi-Layer Perceptron (MLP) classifier had slightly higher Overall Accuracy (OA) values (1% or 2% in general) than Random Forest (RF). Besides, LULC classifications from MLP were visually better, with less salt-pepper effect when compared with RF classifier. Our best LULC classification was obtained using the MLP classifier applied to the dataset (D3P5) with optical and SAR data combined (D3), considering the three periods together (P5). These results showed the benefits of optical and SAR data associate with the seasonality from different periods.

Our best classification results have more detailed classes than MapBiomias (SOUZA et al., 2020) (Figure 6.17). We separate the Grassland Formations class from MapBiomias in Savannas and Campinarana classes. Besides, in our approach, it was possible to identify the areas with Acacia and Cashew (Perennial Crops), that were not well represented in the Forest Plantation class in the MapBiomias LULC map. Acacia is a common Forest plantation in Roraima Lavrados and is one of the sources of conversion to the annual crops' areas. In general, our results provided a better overview of the LULC in Roraima. Thus, our methodology could be helpful to better discriminate Annual Crops, Perennial Crops, and the Lakes formations in the next MapBiomias LULC version.

We classified the LULC of the whole Roraima state using the Google Earth Engine (GEE) platform (GORELICK et al., 2017). Although MLP had better results, we used RF because the GEE platform does not have MLP. With the two-stage sampling approach, it was possible to estimate the LULC areas from the validated points, blocks classifications, and the entire estate classification. Comparing with the classification of the entire state at once, the two-stage sampling provide the firsts information with less effort and gave some idea about the changes among the studied years. In the three years analyzed (2017 to 2019), the LULCC areas were mostly located in Savannas and Acacias, which were converted to croplands, and in Forest, areas turned into Pasture. Our results showed that the multisensor approach, combined with the two-stages sampling, could be a useful tool to monitor the LULC and provide information about the variations in LULC classes.

## 9. OVERALL CONCLUSION

Our approach represents an advance for heterogeneous LULC mapping in tropical regions using spaceborne remote sensing data. This area presents limitations for the use of continuous Optical Remote Sensing (ORS) data due to high cloud frequency. This is more evident in equatorial zones (Amazon), as are the case of Roraima State, and Brazilian coastal regions. Agricultural monitoring with ORS data, during P1 and P2 (September-February) for most of the country and QP2 (May-August) for Roraima, becomes difficult in practically the entire territory. Our approach with ORS and SAR multisensor data had better Overall Accuracy (OA) than comparing with only optical or SAR data. Besides, SAR data is an important source of data, mainly during the rainy season (P2), when cloud cover frequency limits the availability of the optical imagery. The use of different periods allowed us to identify when could be a concentrated effort to map specific classes. SAR-optical data for the P5, combining P1 (January to April), P2 (May to August), and P3 (September to December), show better LULC maps results. Moreover, the MLP classifier performed higher OA than RF and better-smoothed maps. Minor and similar classes are difficult to estimate, even with SAR and optical data. To expand the approach to the entire state, we use RF inside the Google Earth Engine (GEE). Moreover, we test a two-stage sampling approach to estimate the LULC area. A two-stage sampling point has the potential to provide LULC information in a few steps with less computational demands. However, as the blocks and points are random samplings, some classes could not be well represented. On the other hand, the roadside approach needs higher computational power, as well as more time, to pre-process the field data. The roadside approach has the advantage that it captures small classes, not depending on where the class is. Both methods need accurate input and Remote Sensing data to process the LULC classification. Thus, multisensor with SAR and optical data is a way to work around the cloud limitation, a common situation of tropical areas.

## REFERENCES

ADAMI, M.; BERNARDES, S.; ARAI, E.; FREITAS, R. M.; SHIMABUKURO, Y. E.; ESPÍRITO-SANTO, F. D. B.; RUDORFF, B. F. T.; ANDERSON, L. O. Seasonality of vegetation types of South America depicted by moderate resolution imaging spectroradiometer (MODIS) time series. **International Journal of Applied Earth Observation and Geoinformation**, v. 69, p. 148–163, 2018.

AGRICULTURAL MARKET INFORMATION SYSTEM (AMIS). **WHEAT**: planting and harvesting calendar. Rome, 2019. Available from: <http://www.amis-outlook.org/amis-about/calendars/>.

AGUIAR, D. A.; RUDORFF, B. F. T.; SILVA, W. F.; ADAMI, M.; MELLO, M. P. Remote sensing images in support of environmental protocol: monitoring the sugarcane harvest in São Paulo State, Brazil. **Remote Sensing**, v. 3, n. 12, p. 2682–2703, 2011.

ALMEIDA, C. A.; COUTINHO, A. C.; ESQUERDO, J. C. D. M.; ADAMI, M.; VENTURIERI, A.; DINIZ, C. G.; DESSAY, N.; DURIEUX, L.; GOMES, A. R. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. **Acta Amazonica**, v. 46, n. 3, p. 291–302, 2016.

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.

ASNER, G. P. Cloud cover in Landsat observations of the Brazilian Amazon. **International Journal of Remote Sensing**, v. 22, n. 18, p. 3855–3862, 2001.

ATZBERGER, C. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. **Remote Sensing**, v. 5, n. 2, p. 949–981, 2013.

BARBOSA, R. I.; BACELAR-LIMA, C. G. Notas sobre a diversidade de plantas e fitofisionomias em Roraima através do banco de dados do herbário INPA. **Amazônia: Ciência & Desenvolvimento**, v. 4, n. 7, p. 131–154, 2008.

BARBOSA, R. I.; CAMPOS, C.; PINTO, F.; FEARNSSIDE, P. M. The “Lavrados” of Roraima: biodiversity and conservation of Brazil's Amazonian Savannas. **Functional Ecosystems and Communities**, v. 1, n. 1, p. 29–41, 2007.

BARBOSA, R. I.; KEIZER, E.; PINTO, F. Ecosistemas terrestres de Roraima: área e modelagem espacial da biomassa. In: BARBOSA, R. I. (Ed.). **Roraima: homem, ambiente e ecologia**. 2. ed. Boa Vista: Femact, 2010. p. 347–368

BARNI, P. E.; BARBOSA, R. I.; XAUD, H. A. M.; XAUD, M. R.; FEARNSIDE, P. M. Precipitation in northern Amazonia: spatial distribution in Roraima, Brazil. **Sociedade & Natureza**, v. 32, p. 439–456, 2020a.

BARNI, P. E.; MANZI, A. O.; CONDÉ, T. M.; BARBOSA, R. I.; FEARNSIDE, P. M. Spatial distribution of forest biomass in Brazil's state of Roraima, northern Amazonia. **Forest Ecology and Management**, v. 377, p. 170–181, 2016.

BARNI, P. E.; RORAIMA, U. E.; MANZI, A. O.; FEARNSIDE, P. Simulated deforestation versus satellite data in Roraima, Northern Amazonia, Brazil. **Sustainability in Debate**, v. 11, p. 78–94, 2020b.

BECKER-RESHEF, I.; JUSTICE, C.; SULLIVAN, M.; VERMOTE, E.; TUCKER, C.; ANYAMBA, A.; SMALL, J.; PAK, E.; MASUOKA, E.; SCHMALTZ, J.; HANSEN, M.; PITTMAN, K.; BIRKETT, C.; WILLIAMS, D.; REYNOLDS, C.; DOORN, B. Monitoring global croplands with coarse resolution Earth observations: the Global Agriculture Monitoring (GLAM) Project. **Remote Sensing**, v. 2, n. 6, p. 1589–1609, 2010.

BECKER, W. R.; LÓ, T. B.; JOHANN, J. A.; MERCANTE, E. Statistical features for land use and land cover classification in Google Earth Engine. **Remote Sensing Applications: Society and Environment**, v. 21, e 100459, 2021.

CAMARGO, F. F.; SANO, E. E.; ALMEIDA, C. M.; MURA, J. C.; ALMEIDA, T. A Comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical Savanna using ALOS-2/PALSAR-2 Polarimetric Images. **Remote Sensing**, v. 11, n. 13, p. 1600, 2019.

CARDOZO, E. **Calendario de siembra**. Asunción: Ministerio de Agricultura y Ganadería, 2009. Available from: [http://www.mag.gov.py/dgp/Calendario de siembra MAG.pdf](http://www.mag.gov.py/dgp/Calendario_de_siembra_MAG.pdf).

CARVALHO, W. D.; MUSTIN, K. The highly threatened and little known Amazonian savannahs. **Nature Ecology & Evolution**, v. 1, n. 4, e0100, 2017.

CARVALHO, L. M. V.; JONES, C.; LIEBMANN, B. The South Atlantic Convergence Zone: intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. **Journal of Climate**, v. 17, n. 1, p. 88–108, 2004.

CHATZIANTONIOU, A.; PSOMIADIS, E.; PETROPOULOS, G. Co-orbital Sentinel 1 and 2 for LULC mapping with emphasis on wetlands in a



mediterranean setting based on machine learning. **Remote Sensing**, v. 9, n. 12, p. 1259, 2017.

CHAVES, M. E. D.; PICOLI, M. C. A.; SANCHES, I. D. Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: a systematic review. **Remote Sensing**, v. 12, n. 18, p. 3062, 2020.

CLERICI, N.; VALBUENA CALDERÓN, C. A.; POSADA, J. M. Fusion of Sentinel-1A and Sentinel-2A data for land cover mapping: a case study in the lower Magdalena region, Colombia. **Journal of Maps**, v. 13, n. 2, p. 718–726, 2017.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **Calendário de plantio e colheita de grãos no Brasil 2017**. Brasília: CONAB, 2017. Available from: <https://www.conab.gov.br/institucional/publicacoes/outras-publicacoes/item/7694-calendario-agricola-plantio-e-colheita>.

COOK, K. H. South American climate variability and change: remote and regional forcing processes. In: VIMEAUX, F.; SYLVESTRE, F.; KHODRI, M. (Ed.). **Past climate variability in South America and surrounding regions: developments in paleoenvironmental research**. 14. ed. Dordrecht: Springer, 2009. p. 193–212.

DENG, X.; ZHAO, C.; YAN, H. Systematic modeling of impacts of land use and land cover changes on regional climate: a review. **Advances in Meteorology**, v. 2013, p. 1–11, 2013.

DEY, S.; MANDAL, D.; ROBERTSON, L. D.; BANERJEE, B.; KUMAR, V.; MCNAIRN, H.; BHATTACHARYA, A.; RAO, Y. S. In-season crop classification using elements of the Kennaugh matrix derived from polarimetric RADARSAT-2 SAR data. **International Journal of Applied Earth Observation and Geoinformation**, v. 88, e102059, 2020.

DINIZ, J. M. F. S.; GAMA, F. F.; ADAMI, M. Evaluation of polarimetry and interferometry of sentinel-1A SAR data for land use and land cover of the Brazilian Amazon Region. **Geocarto International**, p. 1–19, 2020.

DONG, J.; XIAO, X.; CHEN, B.; TORBICK, N.; JIN, C.; ZHANG, G.; BIRADAR, C. Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. **Remote Sensing of Environment**, v. 134, p. 392–402, 2013.

DRUSCH, M.; DEL BELLO, U.; CARLIER, S.; COLIN, O.; FERNANDEZ, V.; GASCON, F.; HOERSCH, B.; ISOLA, C.; LABERINTI, P.; MARTIMORT, P.; MEYGRET, A.; SPOTO, F.; SY, O.; MARCHESE, F.; BARGELLINI, P. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. **Remote Sensing of Environment**, v. 120, p. 25–36, 2012.

EBERHARDT, I. D. R.; SCHULTZ, B.; RIZZI, R.; SANCHES, I. D.; FORMAGGIO, A. R.; ATZBERGER, C.; MELLO, M. P.; IMMITZER, M.; TRABAQUINI, K.; FOSCHIERA, W.; LUIZ, A. J. B. Cloud cover assessment for operational crop monitoring systems in tropical areas. **Remote Sensing**, v. 8, n. 3, p. 1–14, 2016.

EMPRESA DE ASSISTÊNCIA TÉCNICA E EXTENSÃO RURAL (EMATER). **Sistemas de produção para cana-de-açúcar**. [S.l.]: EMATER, 1983.

ESTES, L. D.; SEARCHINGER, T.; SPIEGEL, M.; TIAN, D.; SICHINGA, S.; MWALE, M.; KEHOE, L.; KUEMMERLE, T.; BERVEN, A.; CHANEY, N.; SHEFFIELD, J.; WOOD, E. F.; CAYLOR, K. K. Reconciling agriculture, carbon and biodiversity in a savannah transformation frontier. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 371, n. 1703, 2016.

EUROPEAN SPACE AGENCY (ESA). Colour vision for Copernicus: the story of Sentinel-2. **ESA Bulletin**, v. 155, p. 88, 2015.

EUROPEAN SPACE AGENCY (ESA). **Sentinel-1 operations**. Available from: [https://www.esa.int/Enabling\\_Support/Operations/Sentinel-1\\_operations](https://www.esa.int/Enabling_Support/Operations/Sentinel-1_operations). Access on: 30 June 2021.

EUROPEAN SPACE AGENCY (ESA). **MultiSpectral Instrument (MSI) overview**. Available from: <https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>. Access on: 16 Sept. 2020.

EUROPEAN SPACE AGENCY (ESA). **Cloud masks: Sentinel-2**. Available from: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-1c/cloud-masks>. Access on: 15 May 2021.

FIALHO, J. F.; ANDRADE, R. F. R.; VIEIRA, E. A. **Mandioca no Cerrado: questões práticas**. 2. ed. Brasília: Embrapa, 2013. 88 p. ISBN 9788570352057.

FOLEY, J. A.; RAMANKUTTY, N.; BRAUMAN, K. A.; CASSIDY, E. S.; GERBER, J. S.; JOHNSTON, M.; MUELLER, N. D.; O'CONNELL, C.; RAY, D. K.; WEST, P. C.; BALZER, C.; BENNETT, E. M.; CARPENTER, S. R.; HILL, J.; MONFREDA, C.; POLASKY, S.; ROCKSTRÖM, J.; SHEEHAN, J.; SIEBERT, S.; TILMAN, D.; ZAKS, D. P. M. Solutions for a cultivated planet. **Nature**, v. 478, n. 7369, p. 337–342, 2011.

FOOD AND AGRICULTURE ORGANIZATION (FAO). **FAO statistical database**. Available from: <http://www.fao.org/faostat/en/#data/QC>. Access on: 11 Nov. 2018.

FRITZ, S.; SEE, L.; BAYAS, J. C. L.; WALDNER, F.; JACQUES, D.; BECKER-RESHEF, I.; WHITCRAFT, A. K.; BARUTH, B.; BONIFACIO, R.;

CRUTCHFIELD, J.; REMBOLD, F.; ROJAS, O.; SCHUCKNECHT, A.; VAN DER VELDE, M.; VERDIN, J.; WU, B.; YAN, N.; YOU, L.; GILLIAMS, S.; MÜCHER, S.; TETRAULT, R.; MOORTHY, I.; MCCALLUM, I. A comparison of global agricultural monitoring systems and current gaps. **Agricultural Systems**, v. 168, p. 258–272, 2019.

FRITZ, S.; SEE, L.; MCCALLUM, I.; YOU, L.; BUN, A.; MOLTCHANOVA, E.; DUERAUER, M.; ALBRECHT, F.; SCHILL, C.; PERGER, C.; HAVLIK, P.; MOSNIER, A.; THORNTON, P.; WOOD-SICHTA, U.; HERRERO, M.; BECKER-RESHEF, I.; JUSTICE, C.; HANSEN, M.; GONG, P.; ABDEL AZIZ, S.; CIPRIANI, A.; CUMANI, R.; CECCHI, G.; CONCHEDDA, G.; FERREIRA, S.; GOMEZ, A.; HAFFANI, M.; KAYITAKIRE, F.; MALANDING, J.; MUELLER, R.; NEWBY, T.; NONGUIERMA, A.; OLUSEGUN, A.; ORTNER, S.; RAJAK, D. R.; ROCHA, J.; SCHEPASCHENKO, D.; SCHEPASCHENKO, M.; TEREKHOV, A.; TIANGWA, A.; VANCUTSEM, C.; VINTROU, E.; WENBIN, W.; VAN DER VELDE, M.; DUNWOODY, A.; KRAXNER, F.; OBERSTEINER, M. Mapping global cropland and field size. **Global Change Biology**, v. 21, n. 5, p. 1980–1992, 2015.

GEE COMMUNITY. **Google Earth Engine plugin for QGIS**. 2020. Available from: <https://github.com/gee-community/qgis-earthengine-plugin>.

GÓMEZ, M. G. C. **Joint use of Sentinel-1 and Sentinel-2 for land cover classification: a machine learning approach**. [S.I.]: Lund University, 2017.

GOOGLE. **Sentinel-2: cloud probability**. Available from: [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S2\\_CLOUD\\_PROBABILITY](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_CLOUD_PROBABILITY). Access on: 18 July 2021.

GORELICK, N.; HANCHER, M.; DIXON, M.; ILYUSHCHENKO, S.; THAU, D.; MOORE, R. Google Earth Engine: planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, v. 202, p. 18–27, 2017.

GRIMM, A. M. The El Niño impact on the Summer monsoon in Brazil: regional processes versus remote influences. **Journal of Climate**, v. 16, n. 2, p. 263–280, 2003.

GRIMM, A. M.; BARROS, V. R.; DOYLE, M. E. Climate variability in southern South America associated with El Niño and La Niña events. **Journal of Climate**, v. 13, n. 1, p. 35–58, 2000.

GRIMM, A. M.; PAL, J. S.; GIORGI, F. Connection between Spring conditions and peak Summer Monsoon rainfall in South America: role of soil moisture, surface temperature, and topography in eastern Brazil. **Journal of Climate**, v. 20, n. 24, p. 5929–5945, 2007.

HARFENMEISTER, K.; SPENGLER, D.; WELTZIEN, C. Analyzing temporal and spatial characteristics of crop parameters using Sentinel-1 backscatter data. **Remote Sensing**, v. 11, n. 13, p. 1569, 2019.

HILKER, T.; LYAPUSTIN, A. I.; TUCKER, C. J.; SELLERS, P. J.; HALL, F. G.; WANG, Y. Remote sensing of tropical ecosystems: Atmospheric correction and cloud masking matter. **Remote Sensing of Environment**, v. 127, p. 370–384, 2012.

HU, J.; GHAMISI, P.; ZHU, X. Feature extraction and selection of Sentinel-1 dual-pol data for global-scale local climate zone classification. **ISPRS International Journal of Geo-Information**, v. 7, n. 9, p. 379, 2018.

IENCO, D.; INTERDONATO, R.; GAETANO, R.; HO TONG MINH, D. Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 158, p. 11–22, 2019.

INGLADA, J.; VINCENT, A.; ARIAS, M.; MARAIS-SICRE, C. Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series. **Remote Sensing**, v. 8, n. 5, 2016.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Monitoramento da cobertura e uso da terra do Brasil**: 2016. Rio de Janeiro: IBGE, 2012. 271p. ISBN 978-85-240-4272-0.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Monitoramento da cobertura e uso da terra do Brasil**: 2000–2010–2012–2014: em grade territorial estatística. Rio de Janeiro: IBGE, 2017. 35 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Conheça cidades e estados do Brasil**. Available from: <https://cidades.ibge.gov.br/brasil/rr/panorama>. Access on: 19 July 2018a.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Pesquisa Agrícola Municipal – PAM**. Available from: <https://sidra.ibge.gov.br/tabela/1612>. Access on: 30 May 2018b.

INSTITUTO NACIONAL DE INVESTIGACIÓN AGROPECUARIA (INIA). **Meses en los cuales se realizan operativas de siembra y cosecha de cultivos y forrajeras en Uruguay, precipitaciones y temperaturas históricas mensual y estacional**. 2014. Available from: <http://www.ainfo.inia.uy/digital/bitstream/item/7115/1/Siembras.pdf>.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Câmeras imageadoras CBERS-3 e 4**. Available from: <http://www.cbbers.inpe.br/sobre/cameras/cbbers3-4.php>. Access on: 3 May 2018.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). PRODES. **Monitoramento da Floresta Amazônica brasileira por satélite**. Available from: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>. Access on: 20 July 2018.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Projeto TerraClass Cerrado**: mapeamento do uso e cobertura vegetal do Cerrado. Available from: <http://www.dpi.inpe.br/tccerrado/>. Access on: 20 July 2018.

INSTITUTO NACIONAL DE TECNOLOGIA AGROPECUARIA (INTA). **Cultivos por provincia**. 2019. Available from: <https://public.tableau.com/profile/gabi.ta1236#!/vizhome/Cultivosxprovincia/Cultivosporprovincias>.

JHONNERIE, R.; SIREGAR, V. P.; NABABAN, B.; PRASETYO, L. B.; WOUTHUYZEN, S. Random forest classification for mangrove land cover mapping using Landsat 5 TM and Alos Palsar imageries. **Procedia Environmental Sciences**, v. 24, p. 215–221, 2015.

JOINT EXPERIMENT FOR CROP ASSESSMENT AND MONITORING (JECAM). **Joint Experiment for Crop Assessment and Monitoring**. Available from: <http://www.jecam.org>. Access on: 31 July 2019.

JUSTICE, C.; TOWNSHEND, J. R.; VERMOTE, E.; MASUOKA, E.; WOLFE, R.; SALEOUS, N.; ROY, D.; MORISETTE, J. An overview of MODIS Land data processing and product status. **Remote Sensing of Environment**, v. 83, n. 1/2, p. 3–15, 2002.

JUSTICE, C. O.; VERMOTE, E.; TOWNSHEND, J. R. G.; DEFRIES, R.; ROY, D. P.; HALL, D. K.; SALOMONSON, V. V.; PRIVETTE, J. L.; RIGGS, G.; STRAHLER, A.; LUCHT, W.; MYNENI, R. B.; KNYAZIKHIN, Y.; RUNNING, S. W.; NEMANI, R. R.; ZHENGMING WAN; HUETE, A. R.; VAN LEEUWEN, W.; WOLFE, R. E.; GIGLIO, L.; MULLER, J.; LEWIS, P.; BARNESLEY, M. J. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. **IEEE Transactions on Geoscience and Remote Sensing**, v. 36, n. 4, p. 1228–1249, 1998.

KHANDELWAL, R. **T-distributed Stochastic Neighbor Embedding(t-SNE)**. Available from: <https://towardsdatascience.com/t-distributed-stochastic-neighbor-embedding-t-sne-bb60ff109561>. Access on: 10 June 2021.

KING, L. A.; ADUSEI, B.; STEHMAN, S. V.; POTAPOV, P. V.; SONG, X.-P.; KRYLOV, A.; DI BELLA, C.; LOVELAND, T. R.; JOHNSON, D. M.; HANSEN, M. C. A multi-resolution approach to national-scale cultivated area estimation of soybean. **Remote Sensing of Environment**, v. 195, p. 13–29, 2017.

KRYLOV, A.; STEININGER, M. K.; HANSEN, M. C.; POTAPOV, P. V.; STEHMAN, S. V.; GOST, A.; NOEL, J.; TALERO RAMIREZ, Y.; TYUKAVINA, A.; DI BELLA, C. M.; ELLIS, E. A.; ELLIS, P. Contrasting tree-cover loss and subsequent land cover in two neotropical forest regions: sample-based assessment of the Mexican Yucatán and Argentine Chaco. **Journal of Land Use Science**, p. 1–16, 6 fev. 2019

LAMBIN, E. F.; GEIST, H. J.; LEPERS, E. Dynamics of land-use and land-cover change in tropical regions. **Annual Review of Environment and Resources**, v. 28, n. 1, p. 205–241, 2003.

LAURIN, G. V.; LIESENBERG, V.; CHEN, Q.; GUERRIERO, L.; DEL FRATE, F.; BARTOLINI, A.; COOMES, D.; WILEBORE, B.; LINDSELL, J.; VALENTINI, R. Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa. **International Journal of Applied Earth Observation and Geoinformation**, v. 21, n. 1, p. 7–16, 2013.

LIU, C.; SHANG, J.; VACHON, P. W.; MCNAIRN, H. Multiyear crop monitoring using polarimetric RADARSAT-2 data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 51, n. 4, p. 2227–2240, 2013.

LOBO, F. L.; SOUZA-FILHO, P. W. M.; NOVO, E. M. L. M.; CARLOS, F. M.; BARBOSA, C. C. F. Mapping mining areas in the Brazilian Amazon using MSI/Sentinel-2 imagery (2017). **Remote Sensing**, v. 10, n. 8, p. 1178, 2018.

LU, D.; BATISTELLA, M.; LI, G.; MORAN, E.; HETRICK, S.; FREITAS, C. C.; DUTRA, L. V.; SANT'ANNA, S. J. S. Land use/cover classification in the Brazilian Amazon using satellite images. **Pesquisa Agropecuária Brasileira**, v. 47, n. 9, p. 1185–1208, 2012.

LU, D.; LI, G.; MORAN, E.; DUTRA, L.; BATISTELLA, M. A comparison of multisensor integration methods for land cover classification in the Brazilian Amazon. **GIScience & Remote Sensing**, v. 48, n. 3, p. 345–370, 2011.

LYAPUSTIN, A.; WANG, Y.; KORKIN, S.; HUANG, D. MODIS Collection 6 MAIAC algorithm. **Atmospheric Measurement Techniques**, v. 11, n. 10, p. 5741–5765, 2018.

LYAPUSTIN, A.; WANG, Y.; LASZLO, I.; KORKIN, S. Improved cloud and snow screening in MAIAC aerosol retrievals using spectral and spatial analysis. **Atmospheric Measurement Techniques**, v. 5, n. 4, p. 843–850, 2012.

MAATEN, L. VAN DER. Accelerating t-SNE using tree-based algorithms. **Journal of Machine Learning Research**, v. 15, p. 1–21, 2014.

MARKHAM, B.; STOREY, J.; MORFITT, R. Landsat-8 sensor characterization and calibration. **Remote Sensing**, v. 7, n. 3, p. 2279–2282, 2015.

MARTINS, V. S.; KALEITA, A. L.; GELDER, B. K.; NAGEL, G. W.; MACIEL, D. A. Deep neural network for complex open-water wetland mapping using high-resolution WorldView-3 and airborne LiDAR data. **International Journal of Applied Earth Observation and Geoinformation**, v. 93, e102215, 2020.

MARTINS, V. S.; NOVO, E. M. L. M.; LYAPUSTIN, A.; ARAGÃO, L. E. O. C.; FREITAS, S. R.; BARBOSA, C. C. F. Seasonal and interannual assessment of cloud cover and atmospheric constituents across the Amazon (2000–2015): insights for remote sensing and climate analysis. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 145, 2018a.

MARTINS, V. S.; SOARES, J. V.; NOVO, E. M. L. M.; BARBOSA, C. C. F.; PINTO, C. T.; ARCANJO, J. S.; KALEITA, A. Continental-scale surface reflectance product from CBERS-4 MUX data: assessment of atmospheric correction method using coincident Landsat observations. **Remote Sensing of Environment**, v. 218, p. 55–68, 2018b.

MCNAIRN, H.; BRISCO, B. The application of C-band polarimetric SAR for agriculture: a review. **Canadian Journal of Remote Sensing**, v. 30, n. 3, p. 525–542, 2004.

MCNAIRN, H.; CHAMPAGNE, C.; SHANG, J.; HOLMSTROM, D.; REICHERT, G. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 64, n. 5, p. 434–449, 2009.

MECHOSO, C. R.; ROBERTSON, A. W.; ROPELEWSKI, C. F.; GRIMM, A. M. The American Monsoon systems. In: INTERNATIONAL WORKSHOP ON MONSOONS, 3., 2004. **Proceedings...** 2004. p. 2–6.

MELO, V. F.; SCHAEFER, C. E. G. R.; VALE JUNIOR, J. F.; UCHÔA, S. C. P. Aspectos pedológicos e de manejo dos solos em Roraima. In: BARBOSA, R. I.; MELO, V. F. (Ed.). **Roraima: homem, ambiente e ecologia**. 2.ed. Boa Vista: Femact, 2010. p. 391–407.

MESQUITA, C. M. **Manual do café: implantação de cafezais**. Belo Horizonte: EMATER, 2016.

MOREIRA, A.; PRATS, P.; YOUNIS, M.; KRIEGER, G.; HAJNSEK, I.; PAPATHANASSIOU, K. A tutorial on Synthetic Aperture Radar. **IEEE Geoscience and Remote Sensing Magazine**, p. 1–43, 2013.

NELSON, G. C.; ROSEGRAN, M. W.; KOO, J.; ROBERTSON, R.; SULSER, T.; ZHU, T.; RINGLER, C.; MSANGI, S.; PALAZZO, A.; BATKA, M.; MAGALHAES, M.; VALMONTE-SANTOS, R.; EWING, M.; LEE, D. **Climate change: impact on agriculture and costs of adaptation**. Washington, 2009. Available from: <http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/130648>.



OLDONI, L. V.; PRUDENTE, V. H. R.; DINIZ, J. M. F. S.; WIEDERKEHR, N. C.; SANCHES, I. D.; GAMA, F. F. Polarimetric SAR data from Sentinel-1a applied to early crop classification. **ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 43, p. 1039–1046, 2020.

OLIVEIRA NETO, A. A. **A cultura do trigo**: análise dos custos de produção e da rentabilidade nos anos-safra 2008 a 2017. [S.l.]: Conab, 2017.

OLOFSSON, P.; FOODY, G. M.; HEROLD, M.; STEHMAN, S. V.; WOODCOCK, C. E.; WULDER, M. A. Good practices for estimating area and assessing accuracy of land change. **Remote Sensing of Environment**, v. 148, p. 42–57, 2014.

OLSON, D. M.; DINERSTEIN, E.; WIKRAMANAYAKE, E. D.; BURGESS, N. D.; POWELL, G. V. N.; UNDERWOOD, E. C.; D'AMICO, J. A.; ITOUA, I.; STRAND, H. E.; MORRISON, J. C.; LOUCKS, C. J.; ALLNUTT, T. F.; RICKETTS, T. H.; KURA, Y.; LAMOREUX, J. F.; WETTENGEL, W. W.; HEDAO, P.; KASSEM, K. R. Terrestrial ecoregions of the World: a new map of life on Earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. **BioScience**, v. 51, n. 11, p. 933–938, 2001.

ORYNBAIKYZY, A.; GESSNER, U.; CONRAD, C. Crop type classification using a combination of optical and radar remote sensing data: a review. **International Journal of Remote Sensing**, v. 40, n. 17, p. 6553–6595, 2019.

OTSUBO, A. A.; LORENZI, J. O. **Cultivo da mandioca na região centro sul do Brasil**. Cruz das Almas: Embrapa, 2004. Available from: [https://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Mandioca/mandioca\\_centrosul/colheita.htm](https://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Mandioca/mandioca_centrosul/colheita.htm).

PALAZZO, F.; ŠMEJKALOVÁ, T.; CASTRO-GOMEZ, M.; RÉMONDIÈRE, S.; SCARDA, B.; BONNEVAL, B.; GILLES, C.; GUZZONATO, E.; MORA, B. RUS: A new expert service for Sentinel users. **Proceedings**, v. 2, n. 7, p. 369, 2018.

PAVANELLI, J. A. P.; SANTOS, J. R.; GALVÃO, L. S.; XAUD, M. R.; XAUD, H. A. M. PALSAR-2/ALOS-2 and OLI/Landsat-8 data integration for land use and land cover mapping In Northern brazilian Amazon. **Boletim de Ciências Geodésicas**, v. 24, n. 2, p. 250–269, 2018.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, É. Scikit-learn: machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.



PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. Updated world map of the Köppen-Geiger climate classification. **Hydrology and Earth System Sciences**, v. 11, n. 5, p. 1633–1644, 2007.

PRUDENTE, V. H. R.; MARTINS, V. S.; VIEIRA, D. C.; SILVA, N. R. F. E.; ADAMI, M.; SANCHES, I. D. Limitations of cloud cover for optical remote sensing of agricultural areas across South America. **Remote Sensing Applications: Society and Environment**, v. 20, e 100414, 2020a.

PRUDENTE, V. H. R.; SANCHES, I. D.; ADAMI, M.; SKAKUN, S.; OLDONI, L. V.; XAUD, H. A. M.; XAUD, M. R.; ZHANG, Y. SAR data for land use land cover classification in a tropical region with frequent cloud cover. In: IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2020b, Hawaii. **Proceedings...** Hawaii: IEEE, 2020.

QIU, S.; ZHU, Z.; HE, B. Fmask 4.0: improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. **Remote Sensing of Environment**, v. 231, e111205, 2019.

REICHE, J.; HAMUNYELA, E.; VERBESSELT, J.; HOEKMAN, D.; HEROLD, M. Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. **Remote Sensing of Environment**, v. 204, p. 147–161, 2018.

RHYS, H. I. **Machine Learning with R, the tidyverse, and mlr**. New York: Manning, 2020. ISBN 9781617296574.

RODRIGUEZ-GALIANO, V. F.; GHIMIRE, B.; ROGAN, J.; CHICA-OLMO, M.; RIGOL-SANCHEZ, J. P. An assessment of the effectiveness of a random forest classifier for land-cover classification. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 67, n. 1, p. 93–104, 2012.

ROSS, J. L. S. O relevo brasileiro no contexto da América do Sul. **Revista Brasileira de Geografia**, v. 61, n. 1, 2016.

ROUSE, J. W.; HASS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In: EARTH RESOURCES TECHNOLOGY SATELLITE (ERTS) SYMPOSIUM, 3., 1973. **Proceedings...** 1973. p. 309–317.

RUDORFF, B. F. T.; AGUIAR, D. A.; SILVA, W. F.; SUGAWARA, L. M.; ADAMI, M.; MOREIRA, M. A. Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data. **Remote Sensing**, v. 2, n. 4, p. 1057–1076, 2010.

SANCHES, I. D.; FEITOSA, R. Q.; ACHANCCARAY DIAZ, P. M.; DIAS SOARES, M.; BARRETO LUIZ, A. J.; SCHULTZ, B.; PINHEIRO MAURANO, L.

E. Campo Verde database: seeking to improve agricultural remote sensing of tropical areas. **IEEE Geoscience and Remote Sensing Letters**, v. 15, n. 3, p. 369–373, 2018a.

SANCHES, I. D.; FEITOSA, R. Q.; ACHANCCARAY, P.; MONTIBELLER, B.; LUIZ, A. J. B.; SOARES, M. D.; PRUDENTE, V. H. R.; VIEIRA, D. C.; MAURANO, L. E. P. LEM benchmark database for tropical agricultural remote sensing application. **ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 42, p. 387–392, 2018b.

SANO, E. E.; FERREIRA, L. G.; ASNER, G. P.; STEINKE, E. T. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. **International Journal of Remote Sensing**, v. 28, n. 12, p. 2739–2752, 2007.

SANTOS, H. G.; CARVALHO JÚNIOR, W.; DART, R. O.; ÁGLIO, M. L. D.; SOUSA, J. S.; PARES, J. G.; FONTANA, A.; MARTINS, A. L. S.; OLIVEIRA, A. P. **O novo mapa de solos do Brasil**: legenda atualizada escala 1:5.000.000. Rio de Janeiro: Embrapa Solos, 2011. Available from: <http://www.cnps.embrapa.br/solosbr/publicacao.htm>.

SCHMIDT, M. The Sankey diagram in energy and material flow management. **Journal of Industrial Ecology**, v. 12, n. 1, p. 82–94, 2008.

SEARCHINGER, T. D.; ESTES, L.; THORNTON, P. K.; BERINGER, T.; NOTENBAERT, A.; RUBENSTEIN, D.; HEIMLICH, R.; LICKER, R.; HERRERO, M. High carbon and biodiversity costs from converting Africa's wet savannahs to cropland. **Nature Climate Change**, v. 5, n. 5, p. 481–486, 2015.

SILVA, S. D. A. E.; NAVA, D. E.; MONTERO, C. R. S.; STURZA, V. S. Sistema de produção de cana-de-açúcar para agricultura familiar. In: WOLFF, L. F.; MEDEIROS, C. A. B. (Ed.). **Alternativas para a diversificação da agricultura familiar de base ecológica**. Pelotas: Embrapa Clima Temperado, 2017. p. 47–55.

SILVA, V. B. S.; KOUSKY, V. E. The South American monsoon system: climatology and variability. In: WANG, S. Y.; GILLIES, R. (Ed.). **Modern climatology**. [S.l.]: InTech, 2012. p. 123–152.

SILVA NETO, P. J.; MATOS, P. G. G.; MARTINS, A. C. S.; SILVA, A. P. **Sistema de produção de cacau para a Amazônia brasileira**. [S.l.]: CEPLAC, 2001.

SKAKUN, S.; KUSSUL, N.; SHELESTOV, A. Y.; LAVRENIUK, M.; KUSSUL, O. Efficiency assessment of multitemporal C-Band Radarsat-2 intensity and Landsat-8 surface reflectance satellite imagery for crop classification in Ukraine.

**IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 9, n. 8, p. 3712–3719, 2016.

SKAKUN, S.; ROGER, J.-C.; VERMOTE, E. F.; MASEK, J. G.; JUSTICE, C. O. Automatic sub-pixel co-registration of Landsat-8 Operational Land Imager and Sentinel-2A Multi-Spectral Instrument images using phase correlation and machine learning based mapping. **International Journal of Digital Earth**, v. 10, n. 12, p. 1253–1269, 2017.

SOARES, M. D.; DUTRA, L. V.; COSTA, G. A. O. P.; FEITOSA, R. Q.; NEGRI, R. G.; DIAZ, P. M. A. A meta-methodology for Improving land cover and land use classification with SAR imagery. **Remote Sensing**, v. 12, n. 6, p. 961, 2020.

SODRÉ, G. A. **Cultivo do cacauzeiro no Estado da Bahia**. Ilheus: MAPA; Ceplac; Cepec, 2017. 126 p. ISBN 9788599169070.

SONG, W.; WANG, L.; LIU, P.; CHOO, K.-K. R. Improved t-SNE based manifold dimensional reduction for remote sensing data processing. **Multimedia Tools and Applications**, v. 78, n. 4, p. 4311–4326, 2019.

SONG, X.-P.; HANSEN, M. C.; POTAPOV, P.; ADUSEI, B.; PICKERING, J.; ADAMI, M.; LIMA, A.; ZALLES, V.; STEHMAN, S. V.; DI BELLA, C. M.; CONDE, M. C.; COPATI, E. J.; FERNANDES, L. B.; HERNANDEZ-SERNA, A.; JANTZ, S. M.; PICKENS, A. H.; TURUBANOVA, S.; TYUKAVINA, A. Massive soybean expansion in South America since 2000 and implications for conservation. **Nature Sustainability**, 2021a.

SONG, X.-P.; HUANG, W.; HANSEN, M. C.; POTAPOV, P. An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. **Science of Remote Sensing**, v. 3, e100018, 2021b.

SONG, X. P.; POTAPOV, P. V.; KRYLOV, A.; KING, L. A.; DI BELLA, C. M.; HUDSON, A.; KHAN, A.; ADUSEI, B.; STEHMAN, S. V.; HANSEN, M. C. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. **Remote Sensing of Environment**, v. 190, p. 383–395, 2017.

SOUZA, C. M.; Z. SHIMBO, J.; ROSA, M. R.; PARENTE, L. L.; A. ALENCAR, A.; RUDORFF, B. F. T.; HASENACK, H.; MATSUMOTO, M.; G. FERREIRA, L.; SOUZA-FILHO, P. W. M.; OLIVEIRA, S. W.; ROCHA, W. F.; FONSECA, A. V.; MARQUES, C. B.; DINIZ, C. G.; COSTA, D.; MONTEIRO, D.; ROSA, E. R.; VÉLEZ-MARTIN, E.; WEBER, E. J.; LENTI, F. E. B.; PATERNOST, F. F.; PAREYN, F. G. C.; SIQUEIRA, J. V.; VIERA, J. L.; FERREIRA NETO, L. C.; SARAIVA, M. M.; SALES, M. H.; SALGADO, M. P. G.; VASCONCELOS, R.; GALANO, S.; MESQUITA, V. V.; AZEVEDO, T. Reconstructing three decades

of land use and land cover changes in brazilian biomes with Landsat archive and Earth Engine. **Remote Sensing**, v. 12, n. 17, e2735, 2020.

STEELE-DUNNE, S. C.; MCNAIRN, H.; MONSIVAIS-HUERTERO, A.; JUDGE, J.; LIU, P.-W.; PAPATHANASSIOU, K. Radar remote sensing of agricultural canopies: a review. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 5, p. 2249–2273, 2017.

STEINHAUSEN, M. J.; WAGNER, P. D.; NARASIMHAN, B.; WASKE, B. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. **International Journal of Applied Earth Observation and Geoinformation**, v. 73, p. 595–604, 2018.

SUGAWARA, L. M.; RUDORFF, B. F. T.; ADAMI, M. Viabilidade de uso de imagens do Landsat em mapeamento de área cultivada com soja no Estado do Paraná. **Pesquisa Agropecuária Brasileira**, v. 43, n. 12, p. 1777–1783, 2008.

TAMM, T.; ZALITE, K.; VOORMANSIK, K.; TALGRE, L. Relating Sentinel-1 interferometric coherence to mowing events on Grasslands. **Remote Sensing**, v. 8, n. 10, p. 802, 2016.

TAVARES, P.; BELTRÃO, N.; GUIMARÃES, U.; TEODORO, A. Integration of Sentinel-1 and Sentinel-2 for classification and LULC mapping in the urban area of Belém, eastern brazilian Amazon. **Sensors**, v. 19, n. 5, p. 1140, 2019.

TELUGUNTLA, P.; THENKABAIL, P. S.; OLIPHANT, A.; XIONG, J.; GUMMA, M. K.; CONGALTON, R. G.; YADAV, K.; HUETE, A. A 30-m Landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 144, p. 325–340, 2018.

TELUGUNTLA, P.; THENKABAIL, P. S.; XIONG, J.; CONGALTON, R.; TILTON, J.; SANKEY, T. T.; MASSEY, R. Global Cropland Area Database (GCAD) derived from remote sensing in support of food security in the twenty-first Century: current achievements and future possibilities. In: THENKABAIL, P. S. (Ed.). **Land resources: monitoring, modeling, and mapping: advances over last 50 years and a vision for the future**. 2.ed. [S.l.]: Taylor, 2015. p. 1–45.

THENKABAIL, P. S.; KNOX, J. W.; OZDOGAN, M.; GUMMA, M. K.; CONGALTON, R. G.; WU, Z.; MILESI, C.; FINKRAL, A.; MARSHALL, M.; MARIOTTO, I.; YOU, S.; GIRI, C.; NAGLER, P. Assessing future risks to agricultural productivity, water resources and food security: how can remote sensing help? **Photogrammetric Engineering and Remote Sensing**, v. 78, n. 8, p. 773–782, 2012.

TORBICK, N.; CHOWDHURY, D.; SALAS, W.; QI, J. Monitoring rice agriculture across Myanmar using time series Sentinel-1 assisted by Landsat-8 and PALSAR-2. **Remote Sensing**, v. 9, n. 2, p. 119, 2017a.

TORBICK, N.; LEDOUX, L.; SALAS, W.; ZHAO, M. Regional mapping of plantation extent using multisensor imagery. **Remote Sensing**, v. 8, n. 3, p. 236, 2016.

TORBICK, N.; SALAS, W.; CHOWDHURY, D.; INGRAHAM, P.; TRINH, M. Mapping rice greenhouse gas emissions in the Red River Delta, Vietnam. **Carbon Management**, v. 8, n. 1, p. 99–108, 2017b.

TORRES, R.; SNOEIJ, P.; GEUDTNER, D.; BIBBY, D.; DAVIDSON, M.; ATTEMA, E.; POTIN, P.; ROMMEN, B.; FLOURY, N.; BROWN, M.; TRAVER, I. N.; DEGHAJE, P.; DUESMANN, B.; ROSICH, B.; MIRANDA, N.; BRUNO, C.; L'ABBATE, M.; CROCI, R.; PIETROPAOLO, A.; HUCHLER, M.; ROSTAN, F. GMES Sentinel-1 mission. **Remote Sensing of Environment**, v. 120, p. 9–24, 2012.

VAN DER MAATEN, L.; HINTON, G. Visualizing data using t-SNE. **Journal of Machine Learning Research**, v. 9, p. 2579–2625, 2008.

VAN TRICHT, K.; GOBIN, A.; GILLIAMS, S.; PICCARD, I. Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: a case study for Belgium. **Remote Sensing**, v. 10, n. 10, p. 1642, 2018.

VELOSO, A.; MERMOZ, S.; BOUVET, A.; LE TOAN, T.; PLANELLS, M.; DEJOUX, J. F.; CESCHIA, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. **Remote Sensing of Environment**, v. 199, p. 415–426, 2017.

VERA, C.; BAEZ, J.; DOUGLAS, M.; EMMANUEL, C. B.; MARENGO, J.; MEITIN, J.; NICOLINI, M.; NOGUES-PAEGLE, J.; PAEGLE, J.; PENALBA, O.; SALIO, P.; SAULO, C.; SILVA DIAS, M. A.; DIAS, P. S.; ZIPSER, E. The South American low-level jet experiment. **Bulletin of the American Meteorological Society**, v. 87, n. 1, p. 63–78, 2006.

VIOLANTE, A. **An introduction to t-SNE with Python example**. Available from: <https://towardsdatascience.com/an-introduction-to-t-sne-with-python-example-5a3a293108d1>. Access on: 10 June 2021

VUILLE, M.; BURNS, S. J.; TAYLOR, B. L.; CRUZ, F. W.; BIRD, B. W.; ABBOTT, M. B.; KANNER, L. C.; CHENG, H.; NOVELLO, V. F. A review of the South American monsoon history as recorded in stable isotopic proxies over the past two millennia. **Climate of the Past**, v. 8, n. 4, p. 1309–1321, 2012.

WALDNER, F.; BELLEMANS, N.; HOCHMAN, Z.; NEWBY, T.; DE ABELLEYRA, D.; VERÓN, S. R.; BARTALEV, S.; LAVRENIUK, M.; KUSSUL, N.; MAIRE, G. LE; SIMOES, M.; SKAKUN, S.; DEFOURNY, P. Roadside collection of training data for cropland mapping is viable when environmental and management gradients are surveyed. **International Journal of Applied Earth Observation and Geoinformation**, v. 80, p. 82–93, 2019.

WHITCRAFT, A. K.; BECKER-RESHEF, I.; JUSTICE, C. A framework for defining spatially explicit Earth observation requirements for a Global Agricultural Monitoring Initiative (GEOGLAM). **Remote Sensing**, v. 7, n. 2, p. 1461–1481, 2015a.

WHITCRAFT, A. K.; BECKER-RESHEF, I.; JUSTICE, C. O. Agricultural growing season calendars derived from MODIS surface reflectance. **International Journal of Digital Earth**, v. 8, n. 3, p. 173–197, 2015b.

WHITCRAFT, A. K.; BECKER-RESHEF, I.; KILLOUGH, B.; JUSTICE, C. Meeting Earth observation requirements for global agricultural monitoring: an evaluation of the revisit capabilities of current and planned moderate resolution optical Earth observing missions. **Remote Sensing**, v. 7, n. 2, p. 1482–1503, 2015a.

WHITCRAFT, A. K.; MCNAIRN, H.; LEMOINE, G.; LETOAN, T.; SOBUE, S. **The power of Synthetic Aperture Radar for global agricultural monitoring**. Ceos, 2016. Available from: [http://ceos.org/document\\_management/Ad\\_Hoc\\_Teams/GEOGLAM/AHTGEOGLAM\\_Power-of-Synthetic-Aperture-Radar-for-Global-Agricultural-Monitoring\\_Apr2016.pdf](http://ceos.org/document_management/Ad_Hoc_Teams/GEOGLAM/AHTGEOGLAM_Power-of-Synthetic-Aperture-Radar-for-Global-Agricultural-Monitoring_Apr2016.pdf).

WHITCRAFT, A. K.; VERMOTE, E. F.; BECKER-RESHEF, I.; JUSTICE, C. O. Cloud cover throughout the agricultural growing season: impacts on passive optical earth observations. **Remote Sensing of Environment**, v. 156, p. 438–447, 2015b.

WULDER, M. A.; HILKER, T.; WHITE, J. C.; COOPS, N. C.; MASEK, J. G.; PFLUGMACHER, D.; CREVIER, Y. Virtual constellations for global terrestrial monitoring. **Remote Sensing of Environment**, v. 170, p. 62–76, 2015.

XIAO, X.; BOLES, S.; LIU, J.; ZHUANG, D.; LIU, M. Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. **Remote Sensing of Environment**, v. 82, n. 2–3, p. 335–348, 2002.

XIAO, X.; HOLLINGER, D.; ABER, J.; GOLTZ, M.; DAVIDSON, E. A.; ZHANG, Q.; MOORE, B. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. **Remote Sensing of Environment**, v. 89, n. 4, p. 519–534, 2004.

ZALLES, V.; HANSEN, M. C.; POTAPOV, P. V.; STEHMAN, S. V.; TYUKAVINA, A.; PICKENS, A.; SONG, X.; ADUSEI, B.; OKPA, C.; AGUILAR, R.; JOHN, N.; CHAVEZ, S. Near doubling of Brazil's intensive row crop area since 2000. **Proceedings of the National Academy of Sciences**, v. 116, n. 2, p. 428–435, 2019.

ZALLES, V.; HANSEN, M. C.; POTAPOV, P. V.; PARKER, D.; STEHMAN, S. V.; PICKENS, A. H.; PARENTE, L. L.; FERREIRA, L. G.; SONG, X.; HERNANDEZ-SERNA, A.; KOMMAREDDY, I. Rapid expansion of human impact on natural land in South America since 1985. **Science Advances**, v. 7, n. 14, e abg1620, 2021.

ZHANG, H. K.; ROY, D. P.; YAN, L.; LI, Z.; HUANG, H.; VERMOTE, E.; SKAKUN, S.; ROGER, J.-C. Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences. **Remote Sensing of Environment**, v. 215, p. 482–494, 2018a.

ZHANG, X.; WU, B.; PONCE-CAMPOS, G.; ZHANG, M.; CHANG, S.; TIAN, F. Mapping up-to-date paddy rice extent at 10 M resolution in China through the integration of optical and Synthetic Aperture Radar images. **Remote Sensing**, v. 10, n. 8, p. 1200, 2018b.

ZHANG, Y.; SKAKUN, S.; PRUDENTE, V. Detection of changes in impervious surface using Sentinel-2 imagery. In: IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2020, Hawaii. **Proceedings...** Hawaii, 2020.

ZHONG, Y.; GIRI, C.; THENKABAIL, P. S.; TELUGUNTLA, P.; CONGALTON, R. G.; YADAV, K.; OLIPHANT, A.; XIONG, J.; VIJAV, V.; POEHNELT, J. **Global Food Security-support Analysis Data (GFSAD) cropland extent 2015 South America 30 m V001**. NASA EOSDIS Land Processes DAAC. Sioux Falls: USGS EROS, 2017. Available from: <https://lpdaac.usgs.gov/products/gfsad30sacev001/>.

ZHOU, J.; LAU, K.-M. Does a monsoon climate exist over South America? **Journal of Climate**, v. 11, n. 5, p. 1020–1040, 1998.

ZHOU, T.; PAN, J.; ZHANG, P.; WEI, S.; HAN, T. Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region. **Sensors (Switzerland)**, v. 17, n. 6, p. 1–16, 2017.