



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



sid.inpe.br/mtc-m21d/2021/09.17.05.34-TDI

**CLASSIFICAÇÃO BASEADA EM GEOBIA E SÉRIES
TEMPORAIS DE IMAGENS SENTINEL-2 PARA
IDENTIFICAR PASTAGEM ARBUSTIVA E HERBÁCEA
NO CERRADO**

Marcos Antônio de Almeida Rodrigues

Dissertação de Mestrado do
Curso de Pós-Graduação em
Sensoriamento Remoto, orientada
pelos Drs. Thales Sehn Körting,
e Leila Maria Garcia Fonseca,
aprovada em 16 de agosto de 2021.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34T/45EL588>>

INPE
São José dos Campos
2021

PUBLICADO POR:

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

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):

Presidente:

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

Membros:

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)

BIBLIOTECA DIGITAL:

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

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

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

EDITORAÇÃO ELETRÔNICA:

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/2021/09.17.05.34-TDI

**CLASSIFICAÇÃO BASEADA EM GEOBIA E SÉRIES
TEMPORAIS DE IMAGENS SENTINEL-2 PARA
IDENTIFICAR PASTAGEM ARBUSTIVA E HERBÁCEA
NO CERRADO**

Marcos Antônio de Almeida Rodrigues

Dissertação de Mestrado do
Curso de Pós-Graduação em
Sensoriamento Remoto, orientada
pelos Drs. Thales Sehn Körting,
e Leila Maria Garcia Fonseca,
aprovada em 16 de agosto de 2021.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34T/45EL588>>

INPE
São José dos Campos
2021

Dados Internacionais de Catalogação na Publicação (CIP)

Rodrigues, Marcos Antônio de Almeida.

R618c Classificação baseada em GEOBIA e séries temporais de imagens Sentinel-2 para identificar pastagem arbustiva e herbácea no Cerrado / Marcos Antônio de Almeida Rodrigues. – São José dos Campos : INPE, 2021.

xxiv + 94 p. ; (sid.inpe.br/mtc-m21d/2021/09.17.05.34-TDI)

Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2021.

Orientadores : Drs. Thales Sehn Körting, e Leila Maria Garcia Fonseca.

1. Sensoriamento remoto. 2. Sentinel-2. 3. Cubo de dados.
4. OBIA. 5. Séries temporais. I.Título.

CDU 528.854(213.54)



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 DISSERTAÇÃO DE MARCOS ANTÔNIO DE ALMEIDA RODRIGUES**
BANCA Nº 215/2021, REG 819660/2019

No dia 16 de agosto de 2021, às 14h00min, 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 Mestre em Sensoriamento Remoto. O trabalho precisa da incorporação das correções sugeridas pela Banca Examinadora e revisão final pelo(s) orientador(es).

Título: “CLASSIFICAÇÃO BASEADA EM GEOBIA E SÉRIES TEMPORAIS DE IMAGENS SENTINEL-2 PARA IDENTIFICAR PASTAGEM ARBUSTIVA E HERBÁCEA NO CERRADO”

Membros da banca:

Dra. Ieda Del'Arco Sanches - Presidente - INPE

Dr. Thales Sehn Körting - Orientador - INPE

Dra. Leila Maria Garcia Fonseca - Orientadora - INPE

Dr. Hugo do Nascimento Bendini - Membro Interno - FIP Cerrado / FUNDEP / MCTIC / INPE

Dr. Cesare Di Girolamo Neto - Membro Externo - Instituto Tecnológico Vale/Desenvolvimento Sustentável



Documento assinado eletronicamente por **Hugo do nascimento bendini (E), Usuário Externo**, em 17/08/2021, às 11:10 (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 17/08/2021, às 11:26 (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 **Thales Sehn Korting, Pesquisador**, em 17/08/2021, às 14:58 (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 **Cesare di girolamo neto (E), Usuário Externo**, em 17/08/2021, às 18:43 (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 **Leila Maria Garcia Fonseca, Tecnologista**, em 18/08/2021, às 12:27 (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 **7990397** e o código CRC **7CDEC7F4**.



Referência: Processo nº 01340.005322/2021-90

SEI nº 7990397

*“O que se leva desta vida é a vida que se leva...” [e as relações que se
constrói].*

Barão de Itararé

*À Antônio, meu pai, à Maiza, minha mãe, à Izabela, minha irmã, e ao meu
cunhado, Gabriel, minha família, com quem eu posso contar em todos os
momentos.*

AGRADECIMENTOS

Agradeço a Deus que me deu vida, saúde e força para chegar até aqui.

Ao Thales, meu orientador, por ter depositado sua confiança em mim, me dando a oportunidade de cursar o mestrado no INPE e por sempre se mostrar disponível para conversas e reuniões, regadas com seu bom humor e alto astral.

À Leila, minha co-orientadora, por ter se disposto a participar da orientação, e pelas valiosas contribuições que deu ao trabalho, sempre muito pertinentes.

À CAPES, pelo apoio financeiro.

Ao Laércio, pelo apoio com a Análise por Componentes Principais e ao Egídio, com o Modelo Linear de Mistura Espectral.

À Laís, da biblioteca, que sempre foi muito atenciosa com todos os pedidos de impressão ainda durante a fase de disciplinas.

Aos professores e pesquisadores do INPE, que tanto contribuíram para a minha formação acadêmica com a transmissão do seu conhecimento.

Ao Raian, pela amizade e pela parceria do dia a dia na sala do prédio ASA.

Ao Hugo, por tanto ter contribuído com este trabalho com diversas conversas desde a sua concepção, e principalmente com as métricas fenológicas.

Ao Cesare, pelas várias discussões sobre as pastagens e o projeto TerraClass.

À Alana, à Mikha, à Daiane e ao Rennam por tanto apoio, conversas e troca de conhecimento.

Ao Bruno e à Tati, meus irmãos acadêmicos, pela companhia do dia a dia.

Ao Eduardo Estiliano, pelo companheirismo na disciplina de Princípios Físicos e pelo apoio psicológico no momento em que mais precisei. Sou muito grato!

A cada colega da turma de 2019 do mestrado em sensoriamento remoto do INPE, pela oportunidade do convívio e o compartilhamento de tantas experiências.

À Rita, à Cris, ao Nem, ao Anderson e à Ângela, pela hospitalidade com que me acolheram durante o meu período em São José dos Campos.

E à Stelinha (que companhia agradável!), pela amizade dentro e fora do INPE.

RESUMO

O território brasileiro é um dos mais diversos do mundo em termos de clima e meio ambiente, e isso se deve às suas dimensões continentais. Essa diversidade também é observada internamente aos biomas. A savana brasileira, conhecida como o bioma Cerrado, ocupa aproximadamente 20% do território nacional, o que corresponde a uma área de mais de 2 milhões de km², e tem a agricultura e a pastagem como os principais usos do solo em termos de área ocupada, com a pastagem ocupando a maior extensão territorial. Devido a isso, o país possui grande potencial agropecuário e de produção de carne, tanto para consumo interno quanto para exportação. Apesar disso, a expansão das áreas de pastagem no Cerrado ocorre por meio do avanço sobre as áreas de vegetação natural. O mapeamento das áreas de pastagem pode subsidiar o monitoramento do Cerrado para evitar novas conversões de vegetação natural em uso antrópico, auxiliando ainda no alcance de metas e acordos internacionais para a redução do desmatamento. Assim, este trabalho visa diferenciar pastagem arbustiva de pastagem herbácea em uma área de estudo na porção sul do Cerrado brasileiro, combinando diferentes dados gerados a partir de imagens dos satélites Sentinel-2. Para alcançar esse objetivo, foram utilizados cinco cubos de dados gerados a partir de séries temporais do ano de 2018, utilizando Análise por Componentes Principais (ACP), Índice de Vegetação Melhorado - *Enhanced Vegetation Index* (EVI), Modelo Linear de Mistura Espectral (MLME), Métricas Espectro-Temporais (MET) e Métricas Fenológicas (MF). Por meio da classificação baseada em objetos, os dados combinados foram classificados com o algoritmo *Random Forest*, utilizando atributos espectrais, texturais e espaciais. Com esta metodologia foi possível obter acurácia global superior a 90% e Índice Kappa correspondente a 0,8. Os resultados mostraram que o uso da combinação de dados para a classificação é uma técnica promissora para separar diferentes tipos de pastagem no Cerrado.

Palavras-chave: Sensoriamento remoto. Sentinel-2. Cubo de dados. OBIA. Séries temporais. Métricas temporais.

OBJECT-BASED CLASSIFICATION ASSISTED BY TIME SERIES OF SENTINEL-2 IMAGERY FOR IDENTIFICATION OF SHRUBBY AND HERBACEOUS PASTURES IN CERRADO

ABSTRACT

Brazil holds one of the most diverse territories in the world in terms of climate and environment, due to its continental extension. This diversity is also observed inside the biomes. Brazilian Savannah, also known as Cerrado biome, occupy nearly half of the territory, and its land cover is comprised mainly of crop lands and pasture areas, the latter performing the great majority. For this sake, the country has great potential of livestock and meat production, as for internal consuming as for exportation. Despite this, it occurs at the expensive of deforestation of savannah vegetation. Mapping accurately pasture areas may assist so in avoiding the conversion of new areas of savannah vegetation, supporting in the fulfillment of international dealings for deforestation reduction. Thus, this paper aims to differentiate shrubby pasture from herbaceous pasture in the southern region of Brazilian savannah, combining different data derived from Sentinel-2 A/B satellite imagery. To tackle this goal, we derived five data cubes from time series of 2018, using Principal Component Analysis (PCA), Enhanced Vegetation Index (EVI), Linear Spectral Mixture Model (LSMM), spectra-temporal metrics (STM) and phenological metrics (PM) data. With the object-based approach, we classified the data combinations by means of Random Forest algorithm, using spectral, textural and spatial attributes. Among the tests we realized, we achieved accuracy higher than 90% at the best performance, with Kappa Index corresponding to 0.8. Results showed that the data combination technique is promising to boost the separation of the different types of pasture.

Key words: Remote sensing. Sentinel-2. Data cube. OBIA. Time series. Temporal metrics.

LISTA DE FIGURAS

	<u>Pág.</u>
Figura 2.1 - Gráfico da evolução das áreas de floresta natural no Brasil de 1985 a 2017.	6
Figura 2.2 - Gráfico da evolução das áreas de pastagem no Brasil de 1985 a 2017.	6
Figura 2.3 - Gráfico da evolução da taxa de crescimento das áreas de pastagem plantada na região Norte em três períodos distintos.	7
Figura 2.4 - Gráfico da evolução das áreas de agricultura no Brasil de 1985 a 2017.	8
Figura 2.5 - Séries temporais de NDVI para um pixel de pastagem arbustiva e um pixel de pastagem herbácea na Amazônia.	14
Figura 2.6 - Representação esquemática da série temporal de um pixel.	16
Figura 2.7 - Fases do processo de Descoberta de Conhecimento em Bancos de Dados (KDD).	18
Figura 2.8 - Exemplos de árvores de decisão utilizadas pelo algoritmo Random Forest.	19
Figura 3.1 - Localização da área de estudo.	23
Figura 3.2 - Municípios do Cerrado (à esquerda) e tiles do Sentinel-2 (à direita) por área ocupada por pastagens.	24
Figura 4.1 - Representação dos cubos de dados gerados a partir das séries temporais das bandas espectrais selecionadas dos satélites Sentinel-2.	26
Figura 4.2 - Representação de todas as observações dos satélites Sentinel-2 A e B em 2018.	27
Figura 4.3 - Fluxograma da metodologia.	28
Figura 4.4 - Mapeamento do uso e cobertura do solo na área de estudo de acordo com a classificação do Mapbiomas.	29
Figura 4.5 - Máscara de pastagem gerada a partir do mapeamento de uso e cobertura do solo elaborado pelo projeto Mapbiomas.	30
Figura 4.6 - Etapas do processo de geração de máscara de nuvens.	31

Figura 4.7 - Cubo de dados gerado para a primeira principal componente de cada data da série temporal.....	33
Figura 4.8 - Cubo de EVI gerado para a série temporal.....	35
Figura 4.9 - Perfis espectrais dos endmembers escolhidos para as componentes do MLME gerado para a data de 15/07/2018.....	36
Figura 4.10 - Frações espectrais das componentes utilizadas no MLME e cubo de dados da imagem-fração sombra.....	37
Figura 4.11 - Esquema da geração das métricas espectro-temporais e cubo de dados resultante.....	39
Figura 4.12 - Imagens das métricas espectro-temporais utilizadas para a geração do cubo de dados.....	40
Figura 4.13 - Métricas fenológicas geradas pelo TIMESAT utilizando os ciclos sazonais da vegetação.....	41
Figura 4.14 - Exemplos dos ciclos sazonais de pastagem encontrados na área e estudo, sendo um mais elevado e de maior duração e outro menos elevado e de menor duração.....	43
Figura 4.15 - Imagens das métricas fenológicas extraídas do Timesat para o primeiro ciclo.....	44
Figura 4.16 - Esquema representando a extração de atributos.....	46
Figura 4.17 - Representação esquemática do processo de amostragem.....	51
Figura 4.18 - Detalhamento das amostras de pastagem arbustiva (em vermelho) e pastagem herbácea (em verde), e sua relação com a segmentação.....	52
Figura 4.19 - Classes de amostras conforme visão detalhada de campo: pastagem herbácea (A) e pastagem arbustiva (B).....	52
Figura 4.20 - Representação do esquema de agregação de dados utilizado nas classificações combinadas.....	54
Figura 5.1 - Percentual de cobertura de nuvens em cada data de observação da série temporal.....	55
Figura 5.2 - Quantidade de observações livres de nuvens pela quantidade de pixels observados.....	56

Figura 5.3 - Distribuição espacial da quantidade de pixels livres de nuvens observados.....	57
Figura 5.4 - Valores de acurácia global atingidos nas classificações individuais, com e sem seleção de atributos.....	60
Figura 5.5 - Valores de acurácia global atingidos nas classificações combinadas, com e sem seleção de atributos.....	61
Figura 5.6 - Classificação com a maior acurácia e localização dos quadrantes de ampliação.....	63
Figura 5.7 - Ampliações para detalhamento e comparação entre as classificações individuais, com e sem seleção de atributos.....	64
Figura 5.8 - Ampliações para detalhamento e comparação entre as classificações combinadas, com e sem seleção de atributos.....	65
Figura 5.9 - Gráficos boxplot representando a dispersão de cada classe de amostra por tipo de atributo e conjunto de dado.....	66
Figura 5.10 - Comparação entre a segmentação gerada com o algoritmo Multiresolution Segmentation (linhas vermelhas) e os superpixels gerados utilizando o algoritmo SLIC (linhas amarelas).....	70
Figura 5.11 - Relação entre os talhões das propriedades, o limite dos imóveis rurais e objetos provenientes da segmentação.....	70

LISTA DE TABELAS

	<u>Pág.</u>
Tabela 2.1 - Dinâmica do rebanho bovino (milhões de cabeças) e das áreas de pastagens (milhões de hectares) no Brasil e grandes regiões em 1975 e 2006.	7
Tabela 2.2 - Sistematização das características dos satélites Sentinel-2.	21
Tabela 4.1 - Métricas fenológicas extraídas do Timesat.	42
Tabela 4.2 - Atributos geométricos utilizados, por tipo de medida.	49
Tabela 5.1 - Quantidade total de atributos extraídos em valores absolutos e relativos, com e sem seleção de atributos, por tipo de conjunto de dados.....	58
Tabela 5.2 - Quantidade total de atributos selecionados, por tipo de atributo e conjunto de dados. Os valores mais elevados indicam maior importância para a classificação.	59

LISTA DE SIGLAS E ABREVIATURAS

ABIEC	Associação Brasileira das Indústrias Exportadoras de Carnes
ACP	Análise por Componentes Principais
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EVI	<i>Enhanced Vegetation Index</i>
FBDS	Fundação Brasileira para o Desenvolvimento Sustentável
GEOBIA	<i>Gographic-based Image Analysis</i>
GLCM	<i>Gray Level Co-occurrence Matrix</i>
HLS	<i>Harmonized Landsat Sentinel-2</i>
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
IBGE	Instituto Brasileiro de Geografia e Estatística
INPE	Instituto Nacional de Pesquisas Espaciais
LAPIG	Laboratório de Processamento de Imagens e Geoprocessamento
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
MET	Métricas Espectro-Temporais
MF	Métricas Fenológicas
MLME	Modelo Linear de Mistura Espectral
MMA	Ministério do Meio Ambiente
MODIS	<i>Moderate-Resolution Imaging Spectroradiometer</i>
NASA	<i>National Aeronautics and Space Administration</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
OBIA	<i>Object Based Image Analysis</i>
PIB	Produto Interno Bruto
PNMC	Política Nacional sobre Mudança do Clima
RF	<i>Random Forest</i>
RL	Reserva Legal
UFU	Universidade Federal de Uberlândia
UNFCCC	<i>The United Nations Framework Convention on Climate Change</i>

SUMÁRIO

	<u>Pág.</u>
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA.....	5
2.1. Dinâmica das pastagens.....	5
2.2. O projeto TerraClass	9
2.3. GEOBIA (Geographic Object-Based Image Analysis)	11
2.4. Cubos de dados.....	14
2.5. Séries temporais de imagens de satélite	15
2.6. Mineração de dados	17
2.7. Imagens ópticas de sensoriamento remoto	19
3. ÁREA DE ESTUDO.....	22
4. PROCEDIMENTOS METODOLÓGICOS	25
4.1. Complexidade do espaço de atributos e máscara de nuvens.....	28
4.2. Análise por componentes principais	31
4.3. Índices de espectrais	33
4.4. Modelo linear de mistura espectral	35
4.5. Métricas espectro-temporais.....	37
4.6. Métricas fenológicas	40
4.7. Segmentação de imagens	45
4.8. Extração de atributos	45
4.8.1. Atributos estatísticos.....	47
4.8.2. Atributos de textura.....	47
4.8.3. Atributos geométricos	48
4.9. Seleção de atributos	49
4.10. Processo de amostragem	50
4.11. Classificação e avaliação.....	52
5. RESULTADOS E DISCUSSÃO	55
5.1. Análise dos resultados.....	55
5.1.1. Cobertura de nuvens	55
5.1.2. Seleção de atributos	57

5.1.3. Assertividade dos modelos.....	59
5.2. Discussão dos resultados.....	62
5.2.1. Classificações e dispersão amostral.....	62
5.2.2. Acurácia global e conjuntos de dados	67
5.2.3. Aderência dos objetos às classes.....	68
6. CONCLUSÕES E PERSPECTIVAS.....	72
REFERÊNCIAS BIBLIOGRÁFICAS	73
APÊNDICE A - COBERTURA DE NUVENS NA SÉRIE TEMPORAL	84
APÊNDICE B - MATRIZES DE CONFUSÃO	88

1. INTRODUÇÃO

As pastagens formam o tipo de uso do solo mais comum no Brasil (ALENCAR et al., 2020). Junto com a agricultura, essas duas classes cobrem cerca de 30% do território brasileiro (SOUZA JUNIOR; AZEVEDO, 2017). Embora as áreas de pastagem ocorram em todos os biomas brasileiros, elas são mais evidentes no Cerrado, uma vez que este bioma apresenta a maior extensão de pastagens em relação aos demais biomas (AZEVEDO et al., 2018). Segundo dados oficiais do governo (INPE, 2020), o Cerrado perdeu cerca de 50% de sua vegetação original, contudo, existem estudos que indicam uma estabilização das taxas de desmatamento a partir de meados da última década (INPE; 2020; PARENTE et al., 2017). De acordo com Parente e Ferreira (2018), o desmatamento no Brasil é realizado principalmente para suprir as demandas da agropecuária, haja vista as áreas desmatadas serem convertidas em agricultura e pastagem (PARENTE et al., 2019).

Com a extensa área desmatada utilizada pela agropecuária, o Brasil possui o maior rebanho bovino do mundo (IBGE, 2017), sendo o maior exportador mundial de carne bovina (MAPA, 2018) e o segundo maior em termos de produção, ficando atrás apenas dos Estados Unidos (OECD/FAO, 2019). Em 2017, a agropecuária representou cerca de 22% do Produto Interno Bruto (PIB) brasileiro (ABIEC, 2018). Além disso, a pecuária foi responsável por 31% do PIB do agronegócio no mesmo ano (ABIEC, 2018).

Apesar da importância do agronegócio para a economia brasileira, este crescimento ocorreu às custas do desmatamento da vegetação natural, resultando em danosas consequências, como a emissão de gases de efeito estufa e a intensificação do aquecimento global (MONTEIRO et al., 2018). O Cerrado é importante dentro deste cenário devido à importância de sua reserva de biomassa e estoque de carbono (GOMES et al., 2020), além do fato de ser considerado um dos biomas mais biodiversos do mundo (SANO *et al.*, 2007), tanto em termos de flora quanto de fauna, e ser a fonte de recarga de um dos principais aquíferos brasileiros, o Guarani (CAMBRAIA NETO; RODRIGUES,

2021). Além disso, o Brasil assumiu compromissos internacionais para promoção da conservação do clima e do meio ambiente, como a Convenção-Quadro das Nações Unidas sobre a Mudança no Clima (UNFCCC), comprometendo-se a implementar ações para a redução de emissões de gases de efeito estufa. A assinatura deste compromisso resultou na criação da Política Nacional sobre Mudança do Clima - PNMC (MMA, 2009), que instituiu como meta a redução do desmatamento no Cerrado em 40% até 2020 (PPCERRADO, 2010). Isso mostra a importância da proposição e execução de medidas de redução do desmatamento.

Analisando as tendências do uso e cobertura da terra no Brasil entre 1985 e 2017, Parente *et al.* (2019) identificaram uma redução gradativa das áreas de vegetação natural e um crescimento contínuo no mesmo período das áreas de agricultura e pastagem. Essa tendência começou a mudar a partir de 2004, coincidindo com o início da implantação de planos elaborados para auxiliar na redução do desmatamento no Brasil (PPCDAM, 2004). O projeto PRODES (INPE, 2020) desempenha um papel fundamental dentro deste contexto, uma vez que os mapas de desmatamento produzidos anualmente são usados pelo governo brasileiro para monitorar a vegetação dos biomas Amazônia e Cerrado. Desde 2018, os dados PRODES estão sendo produzidos também para os outros biomas Caatinga, Mata Atlântica, Pampa e Pantanal (FUNCATE, 2021). O projeto DETER (DINIZ *et al.*, 2015) também tem importante atuação no controle do desmatamento, pois o DETER emite alertas diários de desmatamento, subsidiando as ações governamentais de fiscalização realizadas pelo IBAMA.

Portanto, os mapeamentos de uso e cobertura da terra são importantes ferramentas para auxiliar na implantação de políticas de preservação da vegetação dos biomas brasileiros (SANO *et al.*, 2010). Existem no Brasil algumas iniciativas para mapeamento do uso e cobertura do solo a nível nacional e regional, como os mapas produzidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2020), pela Fundação Brasileira para o Desenvolvimento Sustentável - FBDS (CHIAVARI *et al.*, 2020), pelo projeto Mapbiomas (ALENCAR *et al.*, 2020) e pelo projeto TerraClass (ALMEIDA *et al.*, 2016). O TerraClass é desenvolvido pelo Instituto Nacional de Pesquisas Espaciais

(INPE) em colaboração com a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), e visa mapear o uso e cobertura do solo nas áreas que já foram desmatadas nos biomas Amazônia e Cerrado. Os mapeamentos LULC (*Land Use and Land Cover*), geralmente, são baseados na classificação de imagens de uma única data ou de um mosaico temporal, utilizando interpretação visual, o que demanda um grande número de especialistas na área de sensoriamento remoto para produzir os mapas. Segundo Seyler et al. (2002), as pastagens no Cerrado são uma tipologia de uso do solo de difícil classificação baseada em séries temporais e na análise por pixel porque elas são compostas por uma variedade de tipos de vegetação como, por exemplo, gramíneas, arbustos e árvores. Apesar disso, métodos de classificação de pastagens baseados na análise de séries temporais e técnicas orientadas a objeto (GEOBIA) ainda são pouco explorados (BENDINI et al., 2020; GIROLAMO NETO et al., 2020). A vantagem de trabalhar com o GEOBIA, cuja unidade mínima de mapeamento são os segmentos de imagem, é que essa abordagem possibilita a utilização de outros tipos de atributos na classificação além dos espectrais, como os atributos de textura e geometria, ampliando as perspectivas de classificação das pastagens.

Coutinho et al. (2020) propuseram uma divisão da pastagem cultivada no Cerrado entre as classes de pasto arbustivo/arbóreo e pasto herbáceo. Estas classes são difíceis de serem discriminadas por meio de classificação automática de imagens de sensoriamento remoto, pois o aspecto que elas apresentam nas imagens é bastante similar (MULLER et al., 2015; RODRIGUES et al., 2020a). De acordo com Coutinho et al. (2020), o que diferencia a pastagem arbustiva da pastagem herbácea é a incidência dos indivíduos arbustivos e arbóreos. Enquanto a pastagem arbustiva é caracterizada pelo predomínio da vegetação arbustiva e arbórea, na pastagem herbácea há o predomínio da vegetação herbácea. De acordo com Neves et al. (2017), diferenciar as classes pasto arbustivo e pasto herbáceo no mapeamento de pastagem é fundamental para identificar pastos com diferença de produtividade (NEVES et al., 2017; RUFIN et al., 2015), o que enfatiza a importância da discriminação entre essas classes.

Devido à dificuldade em separar pasto arbustivo de pasto herbáceo, os mapeamentos anteriormente citados não separam os diferentes tipos de pastagem, identificando este tipo de uso em níveis mais gerais, ao passo que para coberturas como vegetação e agricultura, os mapeamentos geralmente apresentam maior detalhamento temático (BENDINI et al., 2019). Identificada esta lacuna, este trabalho tem como objetivo diferenciar pastagem arbustiva de pastagem herbácea em uma região no sul do Cerrado brasileiro, utilizando dados derivados de séries temporais de imagens de satélite. Para isso, foram utilizadas as imagens dos satélites Sentinel-2, adquiridas em 2018. A partir dessas imagens, foram gerados cubos de dados utilizando Análise por Componentes Principais (ACP), Índice de Vegetação Melhorado (*Enhanced Vegetation Index* - EVI), Modelo Linear de Mistura Espectral (MLME), Métricas Espectro-Temporais (MET) e Métricas Fenológicas (MF) no processo de classificação baseada em objeto, utilizando o algoritmo *Random Forest* - RF (BREIMAN, 2001).

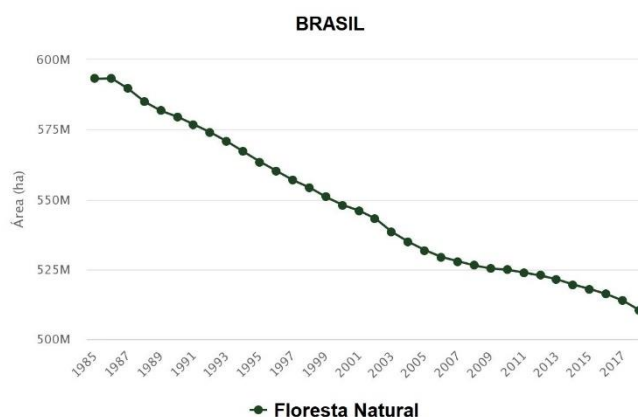
2. REVISÃO DE LITERATURA

2.1. Dinâmica das pastagens

O sistema brasileiro de produção agropecuária é tradicionalmente caracterizado pelo modelo extensivo, em que o gado é criado solto e se alimenta das forrageiras que crescem espontaneamente, e por isso tem um custo de produção geralmente baixo (DIAS FILHO, 2014). Em virtude disso, esse modo de produção necessita da disponibilização de grandes áreas para que se possa desenvolver a pastagem. Há outros países, contudo, considerados desenvolvidos, como os Estados Unidos, por exemplo, cuja produção agropecuária é pautada pelo sistema intensivo, em que o gado é criado em confinamento, dispensando a ocupação de grandes áreas. Como a produção nesse sistema necessita de maiores investimentos, os custos envolvidos geralmente são também mais elevados.

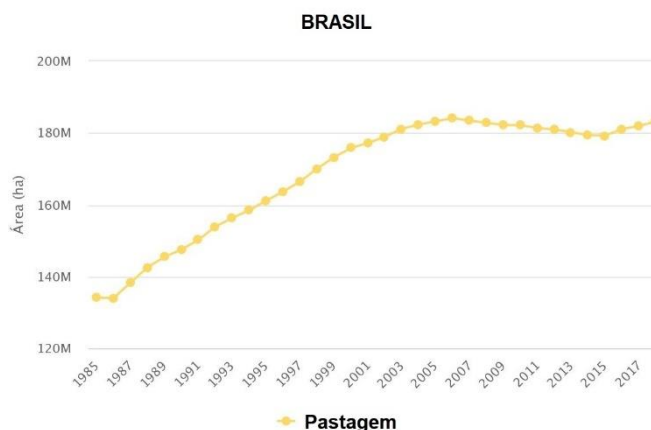
Diante disso, pode-se afirmar que as pastagens estão na vanguarda do desmatamento no Brasil, haja vista que, geralmente, o primeiro uso ao qual se destinam as áreas desmatadas é a pastagem (PPCERRADO, 2018). De acordo com os dados de uso e cobertura do solo do Mapbiomas (2020), entre 1985 e 2017, as áreas de floresta natural no Brasil apresentaram uma contínua tendência decrescente (Figura 2.1), enquanto as áreas de pastagem apresentaram tendência inversa (Figura 2.2). O Cerrado ocupa posição peculiar nessa discussão, uma vez que mais da metade de sua vegetação natural já foi convertida para uso antrópico (INPE, 2020) e é o bioma brasileiro com a maior área ocupada por pastagens (ALENCAR et al., 2020). Há estudos, contudo, que apontam para o início de uma possível mudança no sistema de produção agropecuária brasileiro (PARENTE; FERREIRA, 2018).

Figura 2.1 - Gráfico da evolução das áreas de floresta natural no Brasil de 1985 a 2017.



Fonte: Mapbiomas (2018).

Figura 2.2 - Gráfico da evolução das áreas de pastagem no Brasil de 1985 a 2017.



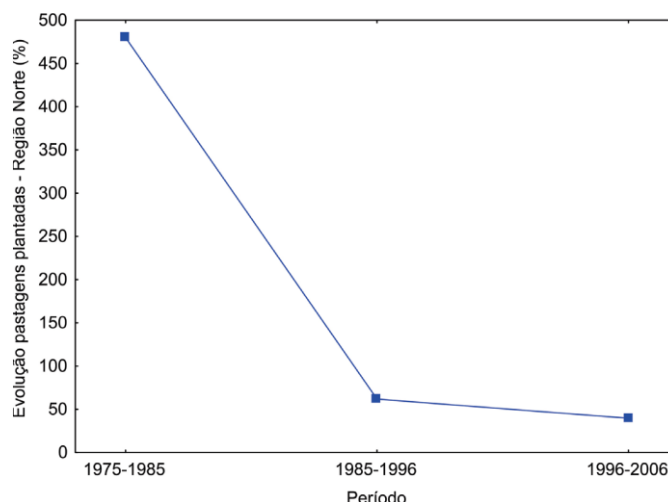
Fonte: Mapbiomas (2018).

Parente et al. (2019) identificaram uma desaceleração no ritmo de crescimento das áreas de pastagem no Brasil a partir de 2004, aproximadamente (Figura 2.2), que não foi acompanhada pela produção agropecuária. Enquanto as taxas de crescimento das pastagens diminuíram (

Figura 2.3), o Brasil se manteve como o maior exportador mundial de carne bovina (MAPA, 2018) e o segundo maior na produção (OECD/FAO, 2019). Somado a isso, foi observado um crescimento no número de cabeças de gado

(DIAS FILHO, 2014), que também não acompanhou a desaceleração no ritmo de crescimento das áreas de pastagem (Tabela 2.1).

Figura 2.3 - Gráfico da evolução da taxa de crescimento das áreas de pastagem plantada na região Norte em três períodos distintos.



Fonte: Adaptado de Dias Filho (2014).

Tabela 2.1 - Dinâmica do rebanho bovino (milhões de cabeças) e das áreas de pastagens (milhões de hectares) no Brasil e grandes regiões em 1975 e 2006.

Brasil e grandes regiões	1975		2006		Variação 1975/2006 (%)	
	Cabeças	ha	Cabeças	ha	Rebanho	Pastagem
Norte	2,11	5,28	41,06	32,63	1.845,9	517,9
Nordeste	18,29	30,62	27,88	32,65	52,4	6,6
Sudeste	35,58	47,27	39,21	32,07	10,2	-32,2
Sul	21,67	21,16	27,2	18,14	25,5	-14,3
Centro-Oeste	24,89	61,31	70,53	56,84	183,4	-7,3
Brasil	102,53	165,65	205,88	172,33	100,8	4

Fonte: Adaptado de Dias Filho (2014).

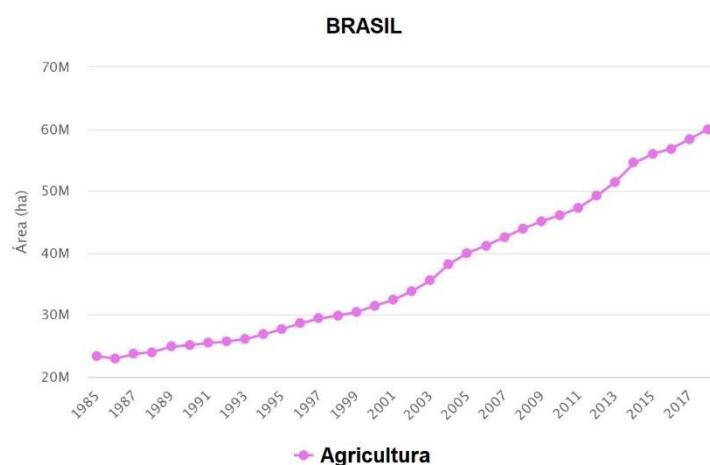
A essa desaceleração podem ser associados alguns fatores. A intensificação da regulamentação quanto ao desmatamento ocorreu aproximadamente no mesmo período que o início do declínio no ritmo de crescimento das áreas de pastagem,

sobretudo a partir da implementação de instrumentos legais como o PPCDAm (2004) e o PPCerrado (2010). O aumento no preço das terras nas regiões Sul, Sudeste e Centro-Oeste freou o incremento das áreas de pastagem nessas regiões e tornou mais atrativo o desenvolvimento das atividades agropecuárias nas regiões Norte e Nordeste, onde o preço da terra é menos elevado, e onde foi observado maior aumento no número de cabeças de gado (DIAS FILHO, 2014).

No mesmo período em que foi observado o declínio do crescimento das áreas de pastagem, as áreas de agricultura continuam em ascensão (

Figura 2.4), o que revela uma substituição das áreas de pastagem por áreas de cultivo agrícola. Isso reforça a afirmação de Parente et al. (2019), de que no Brasil as áreas de pastagem atualmente têm servido como reserva de terra para serem convertidas em outros tipos de uso. A agricultura de baixo carbono, que se baseia na sustentabilidade e pressupõe a manutenção dos sistemas florestais junto com os sistemas de cultivo agrícola, por meio da integração lavoura-pecuária, também possui papel fundamental nessa mudança observada. Todos esses fatores em conjunto apontam para a intensificação no sistema de produção agropecuária brasileiro.

Figura 2.4 - Gráfico da evolução das áreas de agricultura no Brasil de 1985 a 2017.



Fonte: Adaptado de Mapbiomas (2018).

A identificação dos tipos de pastagem possui papel importante na desestimulação do desmatamento. Conforme observado por Dias Filho (2014), o avanço das áreas de pastagem sobre as florestas é motivado pela improdutividade do pasto, principalmente na região de fronteira agrícola, onde o custo da terra, como dito, é mais baixo. Assim, ao se exaurir a capacidade do pasto, o que pode ocorrer pelo próprio pisoteio animal, busca-se novas áreas onde se possa continuar o desenvolvimento da atividade. No Cerrado, essa situação se agrava ainda mais devido aos fatores climáticos, pois o sistema irregular de chuvas torna mais lenta a regeneração natural do pasto.

Segundo Dias Filho (2014), a exposição do solo e o surgimento de espécies invasoras de porte arbustivo ou arbóreo no pasto são sinais de degradação da pastagem. De acordo com Coutinho et al. (2020), o pasto arbustivo é caracterizado pela predominância de espécies arbustivas e arbóreas que se desenvolvem entremeadas a espécies forrageiras, ao passo que o pasto herbáceo estaria associado à predominância de gramíneas. Considerando, portanto, estas definições, é possível estabelecer uma conexão entre o pasto arbustivo e o pasto degradado, e entre o pasto limpo e o pasto não degradado.

Com isso, pode-se afirmar que as pastagens plantadas são importantes não apenas para a recuperação das áreas degradadas no pasto, podendo ser vistas também como meio indireto de evitar o desmatamento. Portanto, a criação de políticas públicas que visem a criação de metodologias para identificação e mapeamento das áreas degradadas, caracterizadas pelo pasto arbustivo, pode ser vista como uma forma de incentivo à redução do desmatamento.

2.2. O projeto TerraClass

Em 2007 foi criado o Programa de Monitoramento Ambiental dos Biomas Brasileiros (PMABB), tendo como objetivo o mapeamento e monitoramento do desmatamento, a avaliação do seu avanço, e a qualificação da cobertura vegetal e uso do solo. Um dos desdobramentos deste programa foi a criação do Projeto TerraClass, que se destina ao mapeamento do uso do solo nas áreas indicadas pelo PRODES como desmatamento (ALMEIDA et al., 2016). Este projeto foi implementado pela primeira vez em 2010, tendo como foco inicialmente a

Amazônia Legal, e sendo posteriormente ampliado para abranger também o Cerrado.

O TerraClass Cerrado 2013 (MMA, 2015a), foi então executado a partir de uma parceria firmada entre diversas instituições públicas brasileiras, como o Instituto Nacional de Pesquisas Espaciais (INPE), a Universidade Federal de Uberlândia (UFU), a Universidade Federal de Goiás (UFG), através do Laboratório de Processamento de Imagens e Geoprocessamento (LAPIG), a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) e o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA). A Portaria 365 do Ministério do Meio Ambiente (MMA, 2015b), lançou estratégias para continuar a execução do projeto e expandir para os demais biomas brasileiros, e com isso, uma nova versão do TerraClass Cerrado foi lançada em 2021 considerando o ano-base de 2018.

A classificação de imagens é uma função de aprendizado que visa o agrupamento de dados em classes previamente definidas (WEISS; KILIOKOWSKI, 1991; HAND, 1981). No caso do TerraClass, como se tratam de biomas distintos e considerando que cada um possui suas especificidades, é natural que as classes do TerraClass Amazônia e do TerraClass Cerrado não fossem as mesmas. Apesar disso, e por se tratar de um tipo de uso do solo presente em todo o território brasileiro, a classe de pastagem está presente em ambos os mapeamentos. Vale salientar, contudo, a existência de uma variação. Enquanto no projeto executado para a Amazônia a classe de pastagem é subdividida entre pasto limpo, pasto sujo e pasto com solo exposto, no executado para o Cerrado, a classe de pastagem não é subdividida.

A legenda do mapa do TerraClass Cerrado 2013 descreve a classe “pastagem” da seguinte forma:

Áreas formadas, prioritariamente, por plantio de forrageiras exóticas perenes. O solo é coberto por vegetação de gramíneas e/ou leguminosas, com altura variando de alguns decímetros a alguns metros. A principal atividade desenvolvida sobre elas é a pecuária, visando a produção de animais de grande porte com objetivos econômicos (DPI/OBT, 2013).

As classes de pasto limpo e pasto sujo do TerraClass Amazônia 2008 (Coutinho et al., 2013) são descritas da seguinte forma:

Pasto Limpo (Pastagem Cultivada Herbácea): Áreas de pastagem em processo produtivo, com predomínio de vegetação herbácea. Refere-se às áreas recém-implantadas ou com baixa infestação por invasoras herbáceas e arbustivas. Ausência de indivíduos arbóreos.

Pasto Sujo (Pastagem Cultivada Arbustiva): Áreas de pastagem em processo produtivo, com predomínio da vegetação herbácea e com a presença de vegetação arbustiva esparsa, frequentemente presença de indivíduos arbóreos. Envolve diferentes estágios de degradação, com a presença significativa de invasoras arbustivas e presença de alguns indivíduos arbóreos.

Como pode ser observado, a diferença básica entre essas duas classes está na incidência de espécies invasoras de porte arbustivo ou arbóreo. De acordo com Neves (2017, p. 7), “o aumento da quantidade de vegetação invasora em pastagens diminui a produtividade devido ao menor acesso às forrageiras e à maior competição por nutrientes”. Isso mostra que a identificação das classes de pastagem arbustiva e pastagem herbácea no Cerrado também importante.

2.3. GEOBIA (Geographic Object-Based Image Analysis)

A classificação automática de imagens de satélite é feita utilizando basicamente dois tipos de abordagem, um deles baseado na análise de pixels, e o outro baseado na análise de objetos (BLASCHKE, 2010). Existe uma miríade de trabalhos que foram feitos utilizando a classificação baseada em pixels, o que mostra que é uma abordagem consolidada, todavia, apresenta como desvantagem o fato de não considerar o contexto espacial existente entre os pixels adjacentes. Além disso, os mapas resultantes de classificações baseadas em pixels geralmente apresentam aspecto ruidoso (CAMPOS et al., 2013). Como os objetos englobam vários pixels em um mesmo polígono, considerando a relação de vizinhança entre eles, este tipo de abordagem apresenta a possibilidade de sanar as inconsistências inerentes à abordagem de pixels.

Apesar disso, existem meios pelos quais é possível diminuir o efeito ruidoso (*salt-and-pepper noise*) apresentado pelas classificações baseadas em pixel, como a filtragem espacial, por exemplo. Este efeito se manifesta através dos pixels espacialmente inconsistentes e que guardam pouca ou nenhuma relação com os pixels vizinhos. A fim de tentar melhorar a consistência espacial, pode ainda ocorrer a refiltragem de um produto que já passou por essa operação (VALE; VACOVAIS; ZANIN, 2014), fazendo com que a característica espectral, que levou um pixel a receber determinada classificação, seja desconsiderada. Além de aumentar a carga de trabalho no pós-processamento e ainda assim não ser uma solução efetiva, passa-se a trabalhar com um produto que deixou de ser fiel à identidade espectral originária do pixel. Dentro deste contexto, a abordagem de objetos se apresenta como um novo paradigma que se propõe a gerar um produto espacialmente consistente, uma vez que são disponibilizadas ferramentas para lidar com as relações de vizinhança existentes entre os pixels, trazendo o contexto espacial para a análise (BLASCHKE et al., 2014).

Os objetos, que constituem a base dessa abordagem, também podem ser chamados de polígonos. Há na literatura uma discussão a respeito da nomenclatura recebida por essa nova forma de abordagem. Existem aqueles que a chamam de OBIA (*Object Based Image Analysis*) e também aqueles que defendem chamá-la GEOBIA (*Geographic - ou Geospatial - Object-Based Image Analysis*). Apesar dessa discussão ontológica, o importante é saber que ambas as nomenclaturas se referem à mesma forma de abordagem, e que tanto uma quanto outra se prestam a analisar as imagens de satélite focando não mais em cada pixel individualmente, mas sim no conjunto de pixels que guardam entre si relações intrínsecas de similaridade tanto espectral quanto espacial (HAY; CASTILLA, 2008).

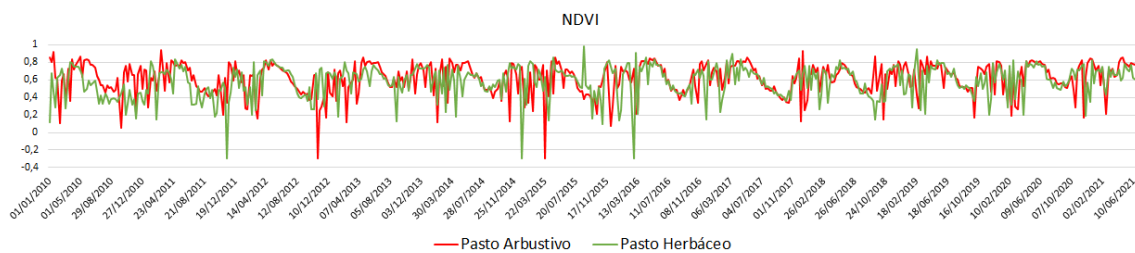
Essa forma de abordagem traz diversas vantagens em relação à baseada no pixel, a começar pela consistência espacial, já que considera as relações de vizinhança entre os pixels, com a consequente eliminação do ruído, evitando assim a existência de pixels isolados fora de contexto (CHEN et al., 2020). Uma outra grande vantagem é que ela possibilita utilizar na classificação outros tipos de atributos além dos espectrais, os quais também podem corroborar para a

distinção entre os alvos. Exemplos são os atributos geométricos. Neste aspecto, vale salientar que essa abordagem depende de uma etapa prévia, a qual se denomina “segmentação” (Seção 4.7), etapa essa em que a imagem é particionada em diversos “segmentos”, que são os próprios objetos. Idealmente, estes objetos contêm pixels similares, e se distinguem dos objetos adjacentes. É nesta etapa que o contexto espacial e a relação de vizinhança entre os pixels são considerados (CASTRO et al., 2019).

Vale ainda dizer que as diferentes tipologias de uso às quais as diferentes formas de cobertura da terra são submetidas, condicionam a ocupação do solo de diferentes maneiras, descrevendo distintos padrões espaciais. São justamente essas características que são exploradas para diferenciar os alvos na análise dos valores de área, perímetro, entre outros atributos geométricos dos polígonos, quando se considera a abordagem baseada em objetos (LI et al., 2015).

Para a classificação de pastagens o GEOBIA se mostra bastante promissor devido às características deste tipo de cobertura do solo. Como a vegetação gramínea está presente tanto na pastagem arbustiva quanto na pastagem herbácea, a análise por pixels dificulta a separação destes tipos de classes. A título de exemplo, a Figura 2.5 apresenta séries temporais de pixels do sensor MODIS, extraídas das classes de pasto arbustivo e pasto herbáceo, conforme o mapa de uso e cobertura da terra da Amazônia proveniente do projeto TerraClass 2008. Conforme pode ser observado, a abordagem por pixels gera amostras semelhantes para cada uma das classes. Como essas classes se caracterizam pela densidade e pelo porte dos indivíduos que as compõem, para diferenciá-las é, então, necessário utilizar uma abordagem que considere as relações espaciais entre os pixels, e a essa demanda o GEOBIA tem a capacidade de atender.

Figura 2.5 - Séries temporais de NDVI para um pixel de pastagem arbustiva e um pixel de pastagem herbácea na Amazônia.



As séries temporais foram geradas considerando os pixels do sensor MODIS dos satélites Terra e Aqua, na resolução espacial de 250 metros.

Fonte: Embrapa (2021).

2.4. Cubos de dados

Com o constante crescimento da quantidade de dados surgiu a necessidade da criação de sistemas que fossem capazes de organizar e sistematizar esses dados, de modo que fosse possível recuperá-los sem maiores dificuldades (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). A Análise de Dados Prontos, termo mais conhecido em inglês como *Analysis-Ready Data* (ARD), consiste em uma forma de organização de dados em que o esforço do usuário no processamento deve ser o mínimo possível, otimizando tempo e recursos na análise dos dados (SIQUEIRA et al., 2019).

Dentro do contexto do sensoriamento remoto, o ARD se aplica ao processamento das imagens de satélite abarcando todas as suas etapas, desde a obtenção dos dados de radiância até a correção atmosférica das imagens e sua disponibilização em reflectância de superfície (GIULIANI et al., 2017). O termo cubo de dados se refere ao conjunto de imagens de sensoriamento remoto, organizadas no tempo e que compartilham a mesma posição geográfica (SIMOES et al., 2021; POTAPOV et al., 2020). Segundo Appel e Pebesma (2019), um cubo de dados possui três dimensões, X, Y e T, as quais correspondem, respectivamente, à latitude, longitude e ao tempo (GREGORY, 2003).

Um dos benefícios que o cubo de dados apresenta é a integração de produtos de diferentes sensores e plataformas, permitindo a criação de séries temporais

densas. Contudo, muito comumente imagens de satélites diferentes irão apresentar características incompatíveis, podendo-se citar como exemplo a resolução espacial, o ajuste espacial e a geometria de iluminação e observação. Técnicas de reamostragem, registro e harmonização devem, portanto, ser empregadas para tornar os dados compatíveis (SKAKUN et al., 2018). Um exemplo de cubo de dados harmonizado é o HLS (*Harmonized Landsat Sentinel-2*), desenvolvido pela NASA para a América do Norte e algumas áreas piloto pelo mundo (CLAVERIE et al., 2018). O cubo de dados australiano (LEWIS et al., 2017) foi o pioneiro e acabou desencadeando o surgimento de outros como o suíço (GIULIANI et al., 2017), o armênio (ASMARYAN et al., 2019), o catalão (MASO et al., 2019) e o regional africano (KILLOUGH, 2019). O projeto brasileiro de cubo de dados é conhecido como *Brazil Data Cube* (BDC) e foi desenvolvido pelo INPE, integrando dados do Landsat, MODIS, Sentinel e CBERS (FERREIRA et al., 2020).

2.5. Séries temporais de imagens de satélite

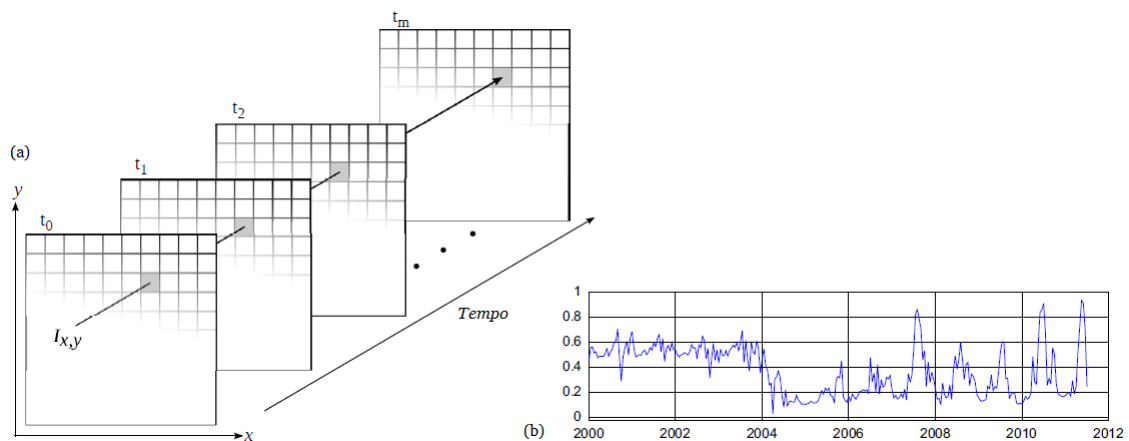
Os estudos de mudança da cobertura da terra, bem como os mapeamentos de uso do solo têm sido potencializados com a adesão das séries temporais aos métodos de análise. Essa potencialização vem justamente da incorporação do fator tempo como uma das variáveis a serem consideradas. Levando em conta que cada tipo de uso da terra possui uma dinâmica temporal própria, este seria um aspecto fundamental para diferenciá-lo dos demais. A análise do uso e cobertura do solo por meio de séries temporais permite, portanto, realizar essas diferenciações, pois disponibiliza ferramentas que permitem explorar o comportamento espectral dos alvos ao longo do tempo (PLETSCH et al., 2019; ASSIS, 2017).

Existem, contudo, duas formas distintas de se abordar a questão do tempo nas análises de mudança de cobertura da terra. A primeira leva em consideração uma data inicial e uma final, de modo que as observações entre essas duas datas não são consideradas (ASSIS et al., 2017). Esse tipo de análise é amplamente utilizado e permite a constatação de mudanças, além da mensuração da magnitude em que as mesmas ocorreram (ASSIS et al., 2017).

Entretanto, não possibilita a identificação de padrões e tendências, já que neste método de análise existem janelas temporais para as quais não se tem informação a respeito dos alvos (MÜLLER et al., 2015).

A outra maneira de realizar essa abordagem é por meio de séries temporais densas e regulares, isto é, que possuem uma certa frequência de observações ao longo de um determinado período. Como nesta abordagem as séries apresentam regularidade, através delas é possível realizar análise de padrões, tendências, sazonalidade e ciclos, além de identificar anomalias (COSTA; FONSECA; KÖRTING, 2015). Neste trabalho serão abordadas estas séries, que serão doravante chamadas de “séries temporais”. Dado um cubo de dados, uma série temporal será determinada pelos valores apresentados por um pixel considerando todas as imagens que formam este cubo (MAUS et al., 2016). A Figura 2.6 ilustra um vetor de pixels ao longo de várias matrizes de imagem, e a série temporal resultante.

Figura 2.6 - Representação esquemática da série temporal de um pixel.



Fonte: Adaptado de Maus *et al.* (2016) e Korting (2013).

Vale ainda dizer que dependendo do período do ano e da região em que as observações forem realizadas, a imagem, no todo ou em parte, pode ter seus pixels contaminados pela presença de nuvens. Essa é a causa mais comum da ocorrência dos ruídos verificados nas séries temporais, pois as nuvens

apresentam valor de reflectância elevado, fazendo com que os pixels contaminados se comportem como *outliers* em relação aos valores que vinham apresentando quando o alvo de fato estava sendo imageado.

Existem meios que permitem reduzir o impacto destes ruídos, que são os filtros, os quais possuem métodos distintos de suavização, podendo ser citados o filtro da média, o filtro de Gauss e o filtro Savitzky-Golay (SAVITZKY; GOLAY, 1964). Todavia, tais filtros devem ser usados com parcimônia para evitar uma suavização excessiva, de modo que a reflectância originária dos alvos seja descaracterizada (MATOSAK et al., 2019).

Há casos em que pode ser mais indicada a aplicação de máscara de nuvens nas imagens contaminadas, ou mesmo o descarte das mesmas. Nestes casos, inevitavelmente as séries se tornarão irregulares e algumas características temporais serão perdidas. Entretanto, com uma série densa, isto é, com muitas observações ao longo do período, este impacto poderá ser minimizado.

2.6. Mineração de dados

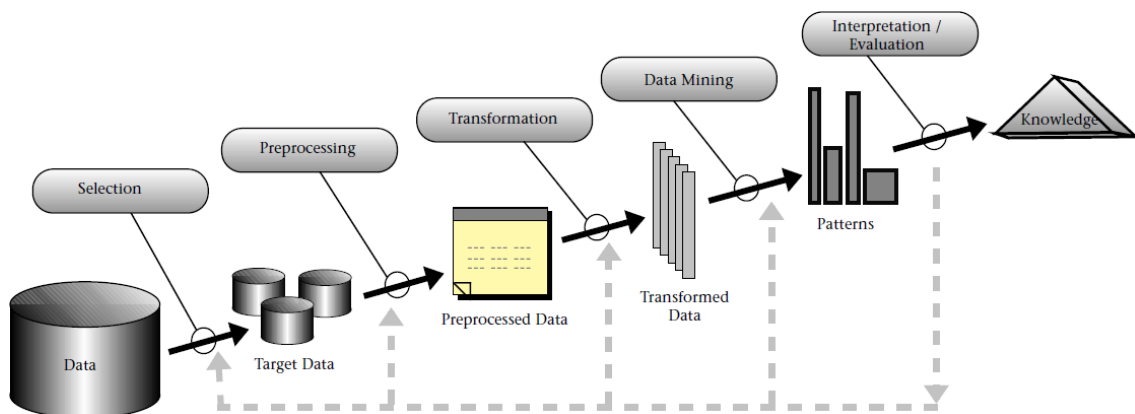
Com o constante aumento do volume, variedade e velocidade na geração de dados, os três V's do *Big Data* (LANEY, 2001), nas mais variadas áreas, incluindo o sensoriamento remoto, sua análise exclusivamente manual se tornou impraticável, passando a ser cada vez mais automatizada. A mineração de dados, dentro deste contexto, surge como uma ferramenta capaz de dar o suporte necessário para que grandes volumes de dados possam ser sistematizados sem grande consumo de tempo, garantindo eficiência e acurácia às análises.

O termo “mineração de dados” é bastante confundido com o processo de Descoberta de Conhecimento em Bancos de dados (*Knowledge Discovery in Databases* - KDD), quando na verdade, consiste em apenas uma de suas etapas. O KDD pode ser entendido como um processo interativo e iterativo que visa a coleta, a análise, o ordenamento e a sistematização de dados, a fim de transformá-los em informações úteis, que façam sentido para quem está interpretando (FAYYAD et al., 1996a). Dentro deste processo, a mineração de dados é a etapa na qual são utilizados algoritmos de *Machine Learning*

(Aprendizado de Máquina) para analisar os dados velozmente, identificando a repetição de padrões, os quais serão posteriormente interpretados, produzindo conhecimento. Por ser a etapa mais corriqueira do KDD, o processo inteiro acaba levando o nome da mineração de dados, sendo que as demais etapas são tão importantes quanto esta. O processo é interativo pois o analista está sempre presente, dando significado ao que está sendo produzido, e iterativo pois é possível avançar ou recuar nas etapas quantas vezes for necessário. A

Figura 2.7, a seguir, ilustra todas as etapas do processo de KDD.

Figura 2.7 - Fases do processo de Descoberta de Conhecimento em Bancos de Dados (KDD).



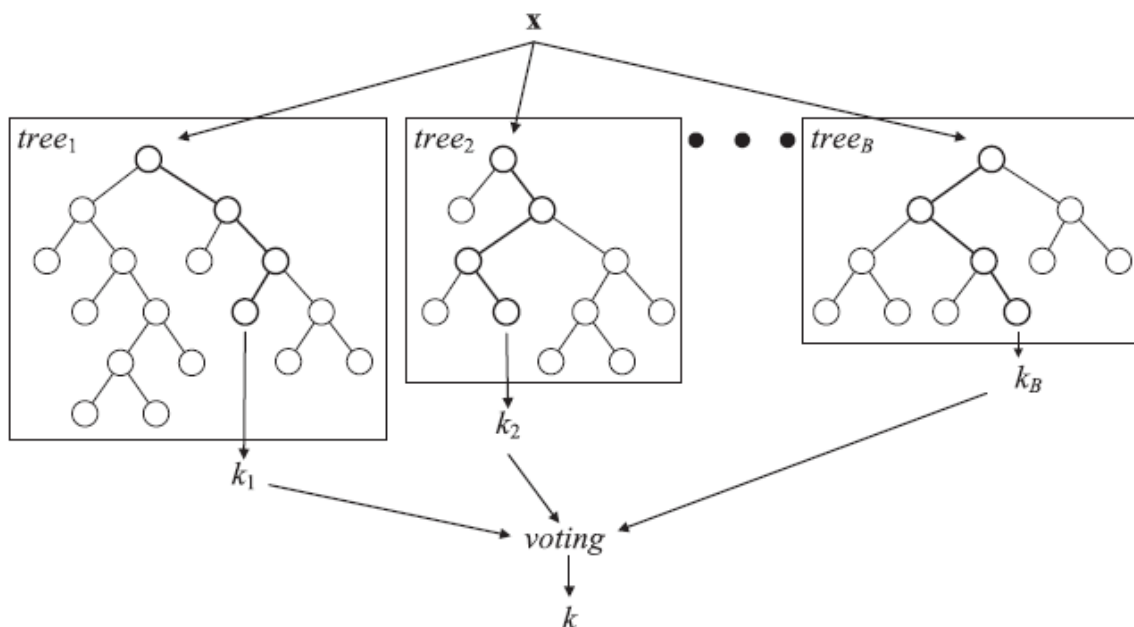
Fonte: Fayyad et al. (1996a).

O KDD possui aplicação em diversas áreas, e uma delas é a classificação de imagens de satélite, que é feita a partir de um conjunto de amostras. Dentre os algoritmos que realizam a mineração de dados, estão aqueles baseados em árvores de decisão, dos quais o *Random Forest* (BREIMAN, 2001), traduzido como Florestas Aleatórias, é um dos mais utilizados. Este algoritmo, ao invés de usar apenas uma árvore, utiliza várias, até centenas, cada uma possuindo um voto para a rotular um pixel ou um objeto, sendo atribuída a classe que tiver sido mais votada.

O *overfitting* se trata de um ajuste excessivo dos dados e diminui a capacidade preditiva dos algoritmos (XU et al., 2014). De acordo com Belgiu e Dragut (2016)

o *Random Forest* é pouco sensível ao sobreajuste, sendo, portanto, um algoritmo eficaz para classificações de imagem (GUAN et al., 2013). Após a escolha do número de árvores, ao fim da classificação o algoritmo apresenta um indicador, o *Out-of-bag*, que é utilizado para avaliar os erros da predição. Para diminuí-los, e conseqüentemente aumentar a qualidade da classificação, é possível realizar a “poda” das árvores, que consiste na retirada dos nós que apresentaram os maiores erros. A Figura 2.8, a seguir, ilustra o funcionamento do *Random Forest*.

Figura 2.8 - Exemplos de árvores de decisão utilizadas pelo algoritmo *Random Forest*.



Fonte: Verikas et al. (2011).

2.7. Imagens ópticas de sensoriamento remoto

Neste trabalho estão sendo utilizados os dados dos satélites Sentinel-2 A e B, sensor MSI (*MultiSpectral Instrument*). As imagens Sentinel-2 são disponibilizadas pela ESA (*European Space Agency*) em dois níveis distintos de processamento: o Nível-1C e o Nível-2A (ESA, 2015). Em ambos os níveis, as imagens são ortorretificadas e georreferenciadas ao datum WGS84, sistema UTM de coordenadas. O primeiro nível corresponde a imagens de reflectância

no topo da atmosfera - TOA (*Top-Of-Atmosphere*), isto é, sem correção atmosférica. Já o segundo corresponde a imagens de reflectância de superfície - BOA (*Bottom-Of-Atmosphere*), ou seja, já corrigidas atmosféricamente.

Cada um dos satélites Sentinel-2 (A e B), possui resolução temporal de 10 dias, com tempo de revisita intercalado um com o outro. Enquanto um dos satélites observa uma órbita, a mesma só será revisitada por este satélite após 10 dias, ao passo que o segundo satélite visita esta órbita na metade deste período, mantendo também um intervalo 10 dias. Contudo, vale salientar que isso é válido apenas para a região intertropical, pois o tempo de revisita diminui com o aumento da latitude (LI ; ROY, 2017). Assim, uma área no Equador é observada por um dos satélites Sentinel-2 a cada 5 dias e com isso, cada órbita é imageada em um ano 73 vezes. As características dos satélites Sentinel-2 foram sistematizadas e são apresentadas na Tabela 2.2.

Tabela 2.2 - Sistematização das características dos satélites Sentinel-2.

Banda	Região	Resolução Espacial			Resolução Radiométrica	Resolução Temporal	SWATH	Resolução Espectral (nm)	Largura da Faixa Espectral (nm)	Comprimento de Onda Central
		10 m	20 m	60 m						
1	Aerossol	-	-	x	12 bits	10 dias	290 km	433 - 453	20	443
2	Azul	x	x	x	12 bits	10 dias	290 km	458 - 523	65	490
3	Verde	x	x	x	12 bits	10 dias	290 km	543 - 578	35	560
4	Vermelho	x	x	x	12 bits	10 dias	290 km	650 - 680	30	665
5	Red Edge 1	-	x	x	12 bits	10 dias	290 km	698 - 713	15	705
6	Red Edge 2	-	x	x	12 bits	10 dias	290 km	733 - 748	15	740
7	Red Edge 3	-	x	x	12 bits	10 dias	290 km	773 - 793	20	783
8A	NIRn	-	x	x	12 bits	10 dias	290 km	855 - 875	20	865
8	NIR	x			12 bits	10 dias	290 km	785 - 900	115	842
9	Vapor D'água	-	-	x	12 bits	10 dias	290 km	935 - 955	20	945
10	Cirrus	-	-	-	12 bits	10 dias	290 km	1360 - 1390	30	1375
11	SWIR 1	-	x	x	12 bits	10 dias	290 km	1565 - 1655	90	1610
12	SWIR 2	-	x	x	12 bits	10 dias	290 km	2100 - 2280	180	2190
Scene Classification		-	x	x	-	-	-	-	-	-
Aerossol		x	x	x	-	-	-	-	-	-
True Color Image		x	x	x	-	-	-	-	-	-
Water Vapor		x	x	x	-	-	-	-	-	-
Quality Band		-	x	x	-	-	-	-	-	-

*A banda 10, utilizada para detecção de nuvens Cirrus só é disponibilizada no Nível-1C de processamento.

**Os satélites Sentinel-2 não possuem banda pancromática.

Fonte: European Space Agency (ESA)

Fonte: Produção do autor.

3. ÁREA DE ESTUDO

O Cerrado é um dos biomas mais atingidos pelo desmatamento, tendo cerca de 50% de sua vegetação original sido convertida para diferentes usos do solo (INPE, 2020). Embora os dados PRODES (INPE, 2020; ASSIS et al., 2019) apontem uma estabilização das áreas desmatadas nos últimos anos, este bioma ainda está sob a ameaça do desmatamento, sofrendo pressão principalmente das áreas de pastagem, que, segundo dados do PPCerrado (2018), é a principal responsável pela substituição da vegetação nativa. O Cerrado é o bioma brasileiro com a maior área ocupada por pastagens, totalizando cerca de 23% de sua área total (ALENCAR et al., 2020). O Código Florestal brasileiro prevê a manutenção da vegetação nativa do Cerrado em 20% da área de uma dada propriedade rural, o que desfavorece a preservação da vegetação (MMA, 2012). Em comparação com a Amazônia, por exemplo, este percentual é de 80% (MMA, 2012).

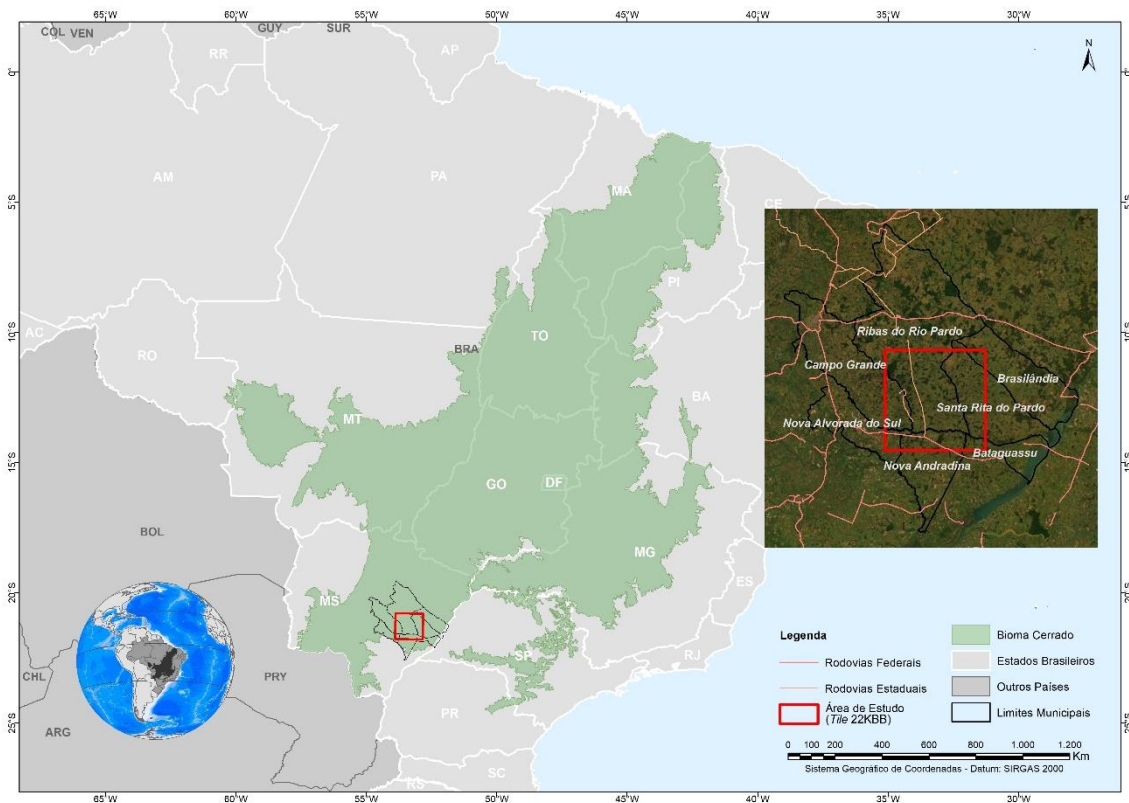
Diante da importância que tem o Cerrado no cenário nacional, a área de estudo proposta para este trabalho corresponde a um *tile* de 100km x 100km da articulação MGRS (*Military Grid Reference System*) que se localiza no Cerrado, cobrindo sete municípios do estado do Mato Grosso do Sul (MS), os quais podem ser observados na

Figura 3.1. Esse *tile* foi escolhido por abranger os municípios que possuem as maiores extensões territoriais ocupadas por pastagem no Cerrado dentro do estado do MS (Figura 3.2), segundo dados do Mapbiomas (AZEVEDO et al., 2018), sendo um local propício para a aplicação da metodologia ora proposta.

Os municípios de Ribas do Rio Pardo e Campo Grande ocupam a primeira e a terceira posição, respectivamente, entre os municípios com as maiores áreas de pastagem no Cerrado (SOUZA JUNIOR, 2019), sendo que o primeiro detém as maiores taxas de desmatamento entre todos os municípios que compõem o Cerrado (PPCERRADO, 2018). Juntos, os municípios representam mais de 50% da área do *tile*, sendo o restante representado pelos outros cinco municípios.

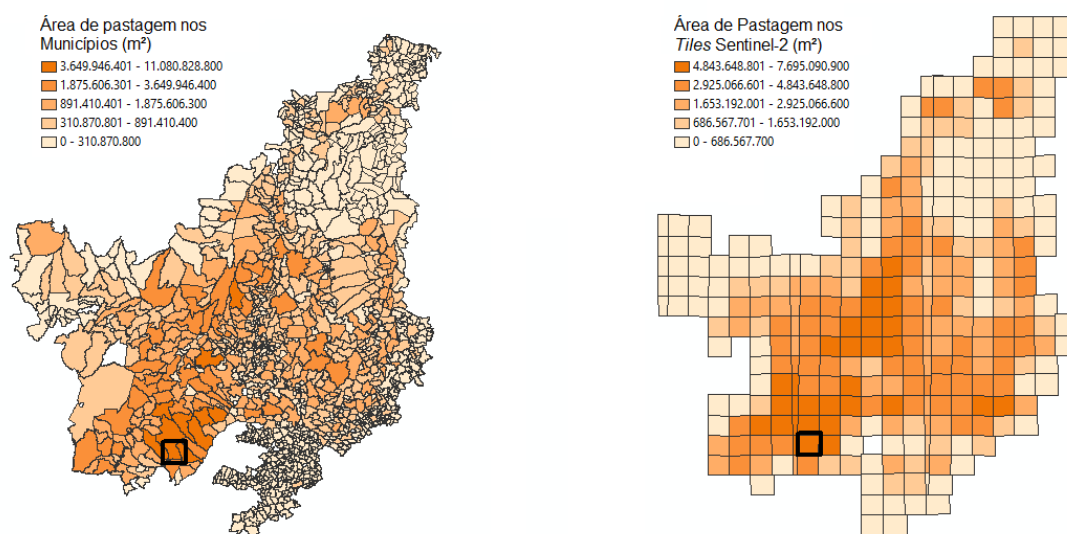
Diante disso, pode-se concluir que o *tile* 22KBB é representativo das pastagens do Cerrado no estado do MS, o que justifica sua escolha como área de estudo. Segundo dados da Pesquisa Pecuária Municipal (PPM, 2019), os municípios que compõem a área de estudo contribuem com 16,1% da produção pecuária estadual em termos de cabeça de gado. Ainda segundo esta pesquisa, com 20,9 milhões de cabeças, o MS é o quarto maior estado brasileiro em efetivo de rebanhos bovinos, e o município de Ribas do Rio Pardo é o terceiro maior do país, com 1,1 milhões de cabeças. De acordo com dados de 2017 da SEMAGRO (2020), a agropecuária é um importante setor para a economia do MS, uma vez que representa 17,6% do PIB estadual (SEMAGRO, 2020).

Figura 3.1 - Localização da área de estudo.



Fonte: Produção do autor.

Figura 3.2 - Municípios do Cerrado (à esquerda) e *tiles* do Sentinel-2 (à direita) por área ocupada por pastagens.



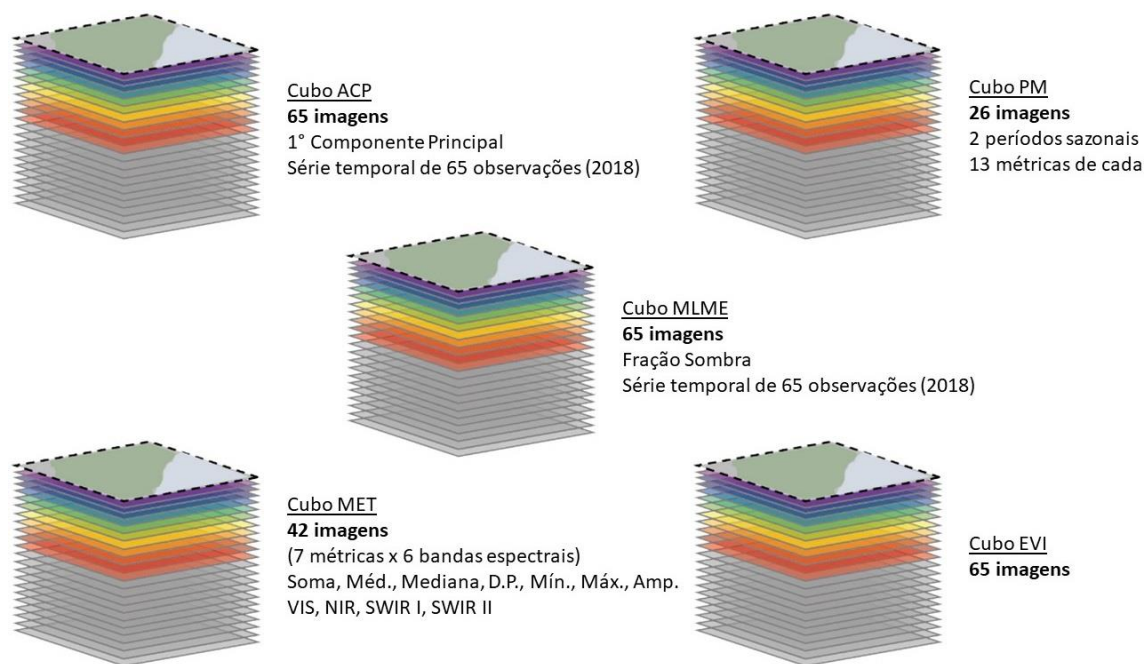
O cálculo da área de pastagem tanto nos municípios quanto nos tiles foi feito utilizando o mapa de uso e cobertura do solo elaborado pelo projeto Mapbiomas para o ano de 2018.

Fonte: Produção do autor.

4. PROCEDIMENTOS METODOLÓGICOS

Neste trabalho, imagens do Sentinel-2, adquiridas em 2018, foram utilizadas para a composição de séries temporais. No repositório de imagens da ESA, somente imagens de reflectância de superfície mais recentes estão disponíveis (a partir de 2020) *on the fly* (para outros períodos esta disponibilizada sob demanda). As demais imagens são disponíveis apenas em reflectância no topo da atmosfera. Deste modo, as imagens no Nível-1C do ano de 2018 foram selecionadas e a correção atmosférica, utilizando o Sen2Cor (ESA, 2018), disponível no *software* SNAP versão 7.0, foi realizada. As bandas utilizadas foram as seguintes: 2 (Azul), 3 (Verde), 4 (Vermelho), 8A (Near Infra-Red narrow - NIRn), 11 (SWIR 1) e 12 (SWIR 2), na resolução de 20 m. Utilizando essas imagens, foram geradas séries temporais para cada banda, e a partir dessas séries foram gerados novos produtos, constituindo diferentes cubos de dados (Figura 4.1). A partir das séries temporais das bandas, foram extraídas as Componentes Principais de cada data, o *Enhanced Vegetation Index* - EVI, o Modelo Linear de Mistura Espectral, as Métricas Espectro-Temporais e as Métricas Fenológicas. Em 2018 houve 70 passagens dos satélites Sentinel-2 sobre a área de estudo, entretanto, cinco dessas observações estiveram completamente contaminadas por nuvens, sendo necessário descartá-las da análise. Assim, passou-se a trabalhar com séries temporais formadas por 65 observações, conforme ilustra a Figura 4.2. A metodologia utilizada para desenvolver este trabalho é apresentada no fluxograma da Figura 4.3, e suas etapas são descritas na sequência.

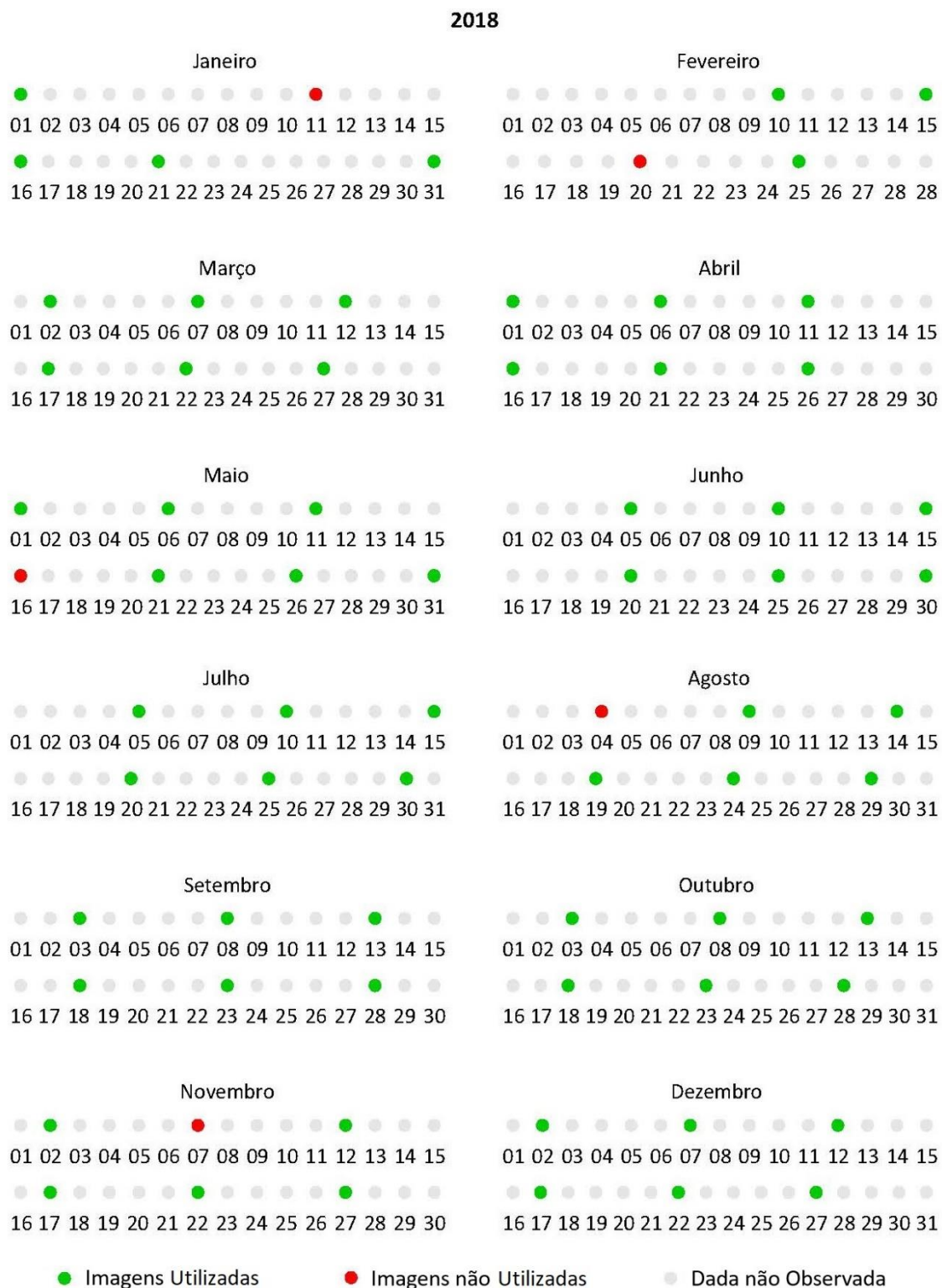
Figura 4.1 - Representação dos cubos de dados gerados a partir das séries temporais das bandas espectrais selecionadas dos satélites Sentinel-2.



ACP: Análise por Componentes Principais; EVI: *Enhanced Vegetation Index* (Índice de Vegetação Melhorado); MLME: Modelo Linear de Mistura Espectral; MET: Métricas Espectro-Temporais; MF: Métricas Fenológicas.

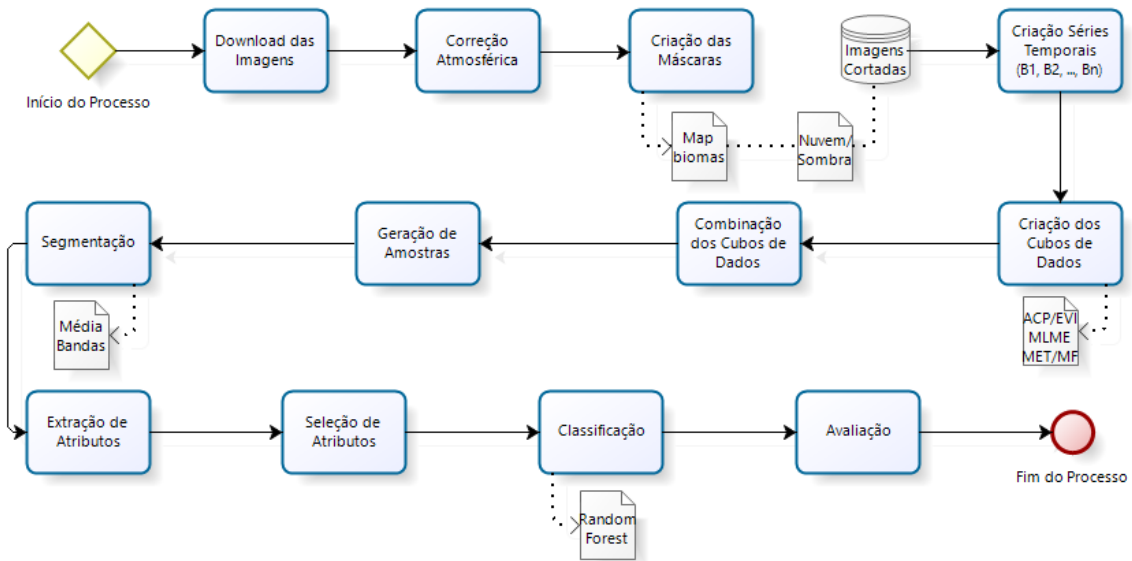
Fonte: Produção do autor.

Figura 4.2 - Representação de todas as observações dos satélites Sentinel-2 A e B em 2018.



Fonte: Produção do autor.

Figura 4.3 - Fluxograma da metodologia.

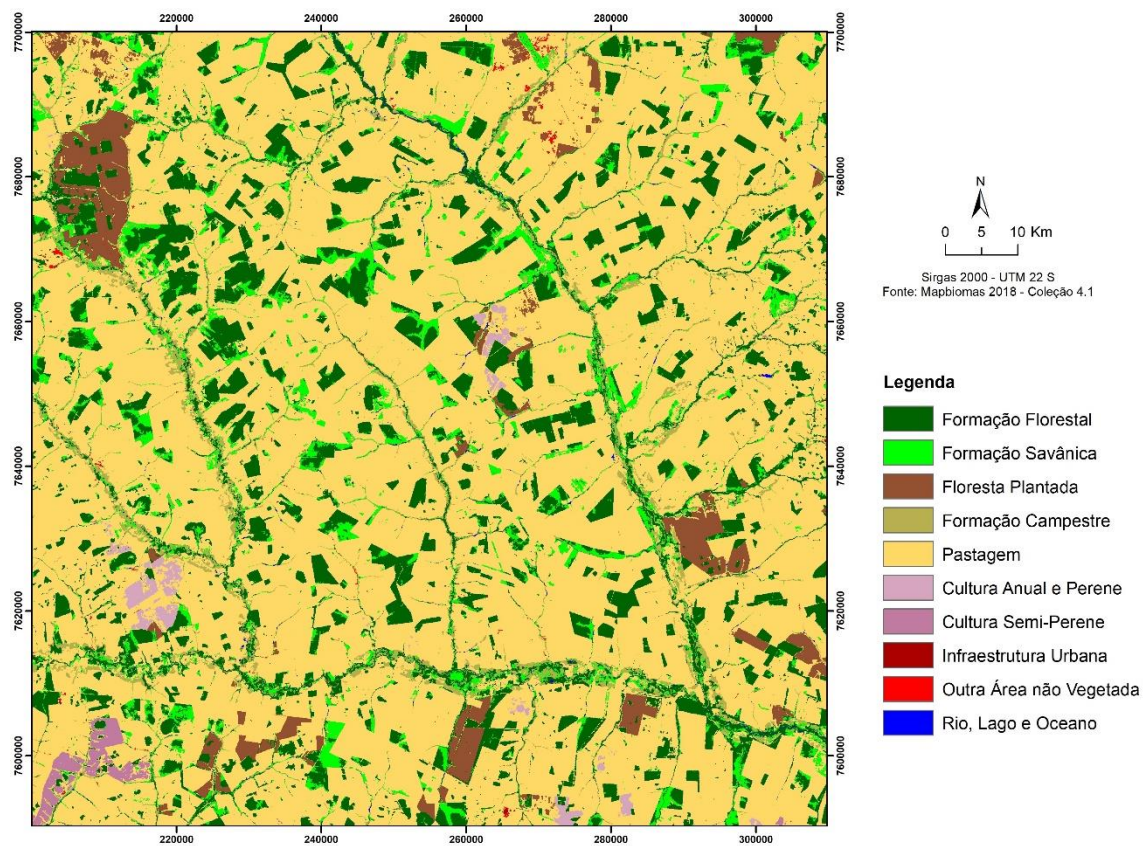


Fonte: Produção do autor.

4.1. Complexidade do espaço de atributos e máscara de nuvens

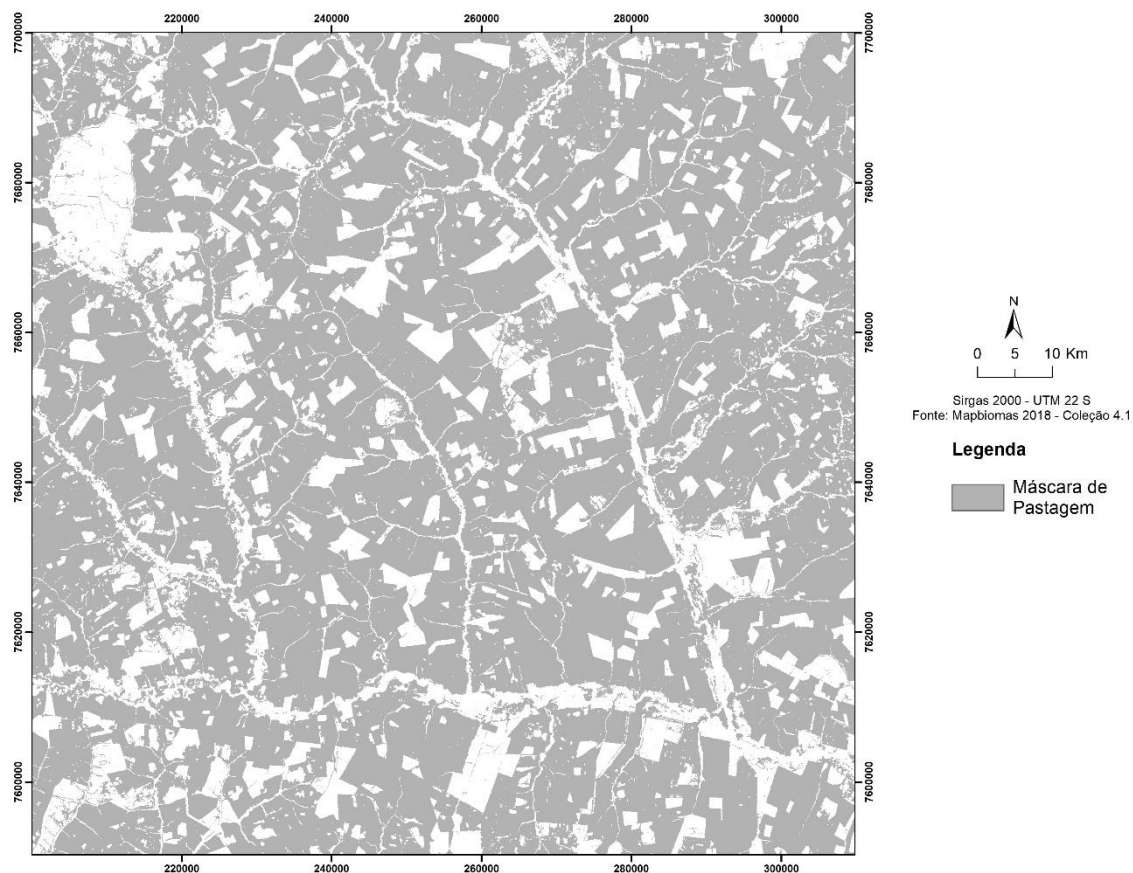
A área de estudo é composta por diversos tipos de cobertura do solo, além da pastagem. Como este trabalho tem foco em diferenciar tipos de pastagem, foi feita uma máscara para eliminar das imagens todos os pixels que representam alvos diferentes deste. Para isso, considerou-se a classificação executada pelo projeto Mapbiomas relativa ao ano de 2018. O Mapbiomas corresponde a um projeto que tem como objetivo o mapeamento do uso e cobertura da terra em todo o território brasileiro, utilizando mosaicos anuais de imagens do Landsat a partir de 1985. No ano considerado neste trabalho, o Mapbiomas identificou as classes de uso do solo conforme apresentado na Figura 4.4. A área correspondente à classe de pastagem foi, então, extraída, gerado uma máscara (Figura 4.5) que foi aplicada a todas as imagens, a fim de reduzir a complexidade do espaço de atributos.

Figura 4.4 - Mapeamento do uso e cobertura do solo na área de estudo de acordo com a classificação do Mapbiomas.



Fonte: Mapbiomas (2018).

Figura 4.5 - Máscara de pastagem gerada a partir do mapeamento de uso e cobertura do solo elaborado pelo projeto Mapbiomas.



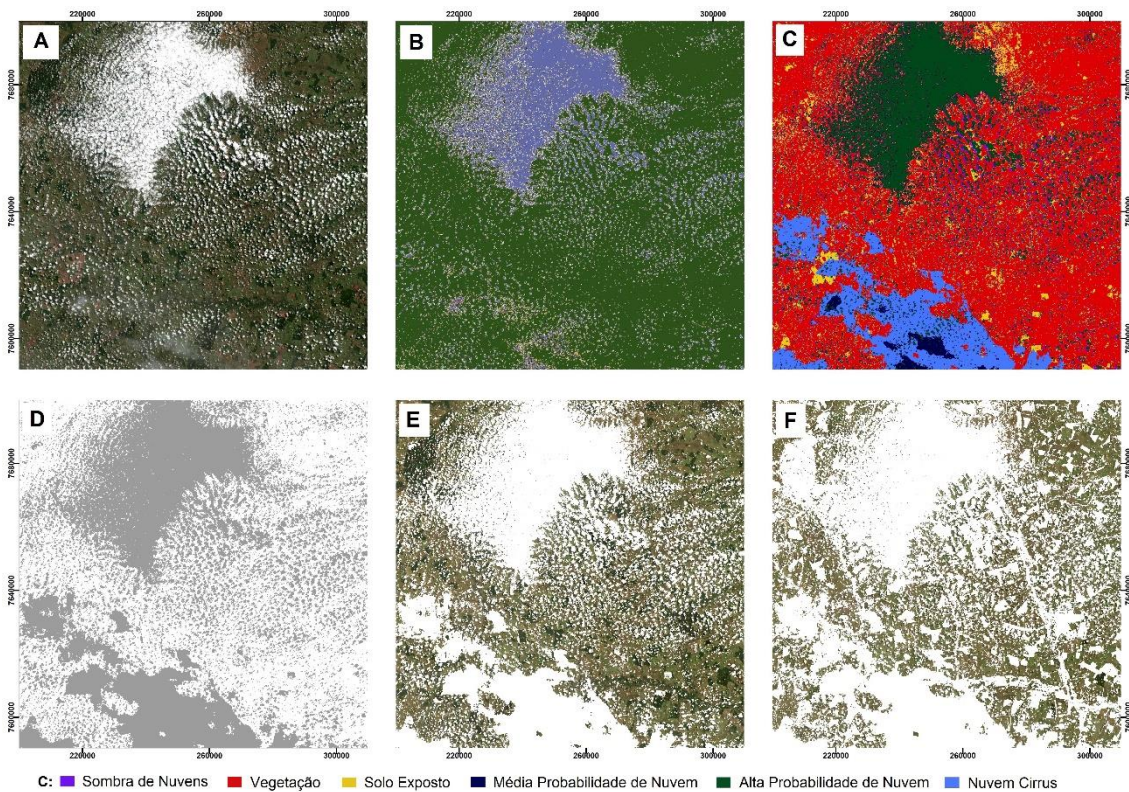
Fonte: Adaptado de Mapbiomas (2018).

Em trabalhos anteriores, Rodrigues et al. (2020a) e Rodrigues et al. (2020b) notaram que a cobertura de nuvens prejudica a análise de séries temporais gerando *outliers* e impedindo a observação dos alvos na superfície. Em virtude disso, neste trabalho optou-se por eliminar das imagens os pixels contaminados por nuvens ou sombra de nuvens. Para isso, foi utilizada a máscara de nuvens gerada com o Sen2Cor (ESA, 2018) para cada data da série temporal. Essa máscara é gerada por meio de um mapa de probabilidade de ocorrência de nuvens, também disponibilizado para cada imagem da série, a partir do qual é feita uma identificação prévia dos tipos de nuvens e de determinadas classes de uso do solo. Apesar disso, notou-se que essa classificação apresenta certa confusão entre as classes. Por isso, as classes foram reavaliadas imagem por imagem, gerando máscaras de nuvens mais confiáveis. Após isso, as máscaras

foram aplicadas individualmente a cada imagem da série temporal, gerando novas imagens livres de nuvens e sombra de nuvens. A

Figura 4.6 apresenta uma sequência de imagens representativa das etapas até chegar nas máscaras de nuvens.

Figura 4.6 - Etapas do processo de geração de máscara de nuvens.



A figura apresenta como exemplo uma imagem da data de 07/03/2018. A: imagem em cor natural; B: mapa de probabilidade de ocorrência de nuvens; C: classificação das nuvens quanto à probabilidade de ocorrência; D: máscara de nuvens e sombra de nuvens; E: imagem em cor natural após a aplicação da máscara de nuvens/sombra; F: imagem em cor natural após a aplicação das máscaras de nuvens/sombra e do Mapbiomas.

Fonte: Produção do autor.

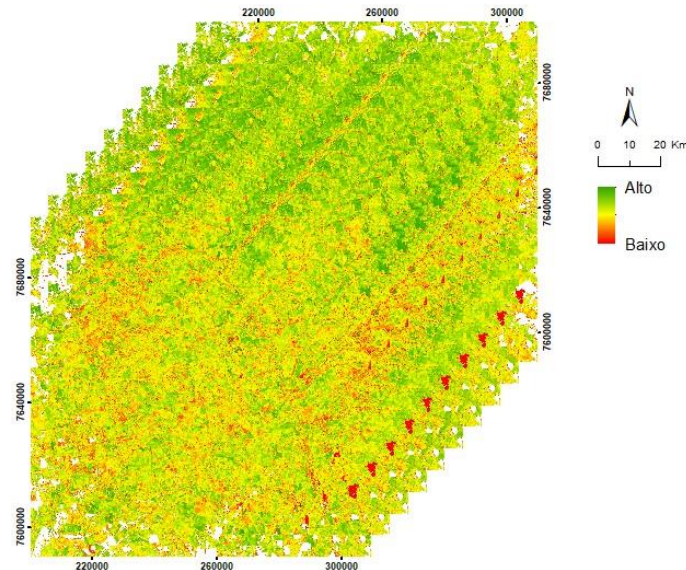
4.2. Análise por componentes principais

A Análise por Componentes Principais (ACP) corresponde a uma técnica estatística de transformação de variáveis correlacionadas em variáveis não

correlacionadas por meio de uma transformação linear (FONSECA et al., 2011). Neste processo, as principais componentes das variáveis originais são extraídas e reorganizadas em ordem decrescente de variância (JIAJU et al., 1988). O número de componentes principais extraídas é igual à quantidade de variáveis da análise (SINGH; HARRISON, 1985). As primeiras principais componentes concentram a maior variabilidade das variáveis originais, portanto, possuem a maior variância (SILJESTROM et al., 1997). Deste modo, as principais componentes podem ser utilizadas para reduzir a dimensionalidade dos dados, eliminando informações redundantes e otimizando os processos.

Ao invés de utilizar todas as bandas espectrais de cada observação da série temporal, extraiu-se as principais componentes de cada data. Observou-se que a primeira principal componente concentrava mais de 90% da informação presente em todas as bandas. Foi criado, então, um cubo de dados com a primeira principal componente extraída de cada data. Foram utilizadas neste trabalho um total de 390 bandas espectrais, que correspondem às seis bandas de cada data multiplicado pela quantidade de observações da série temporal. Utilizando a ACP foi possível, então, reduzir este total para 65 bandas, mantendo ainda as informações mais importantes de todas as bandas espectrais, e otimizando o tempo de processamento. A Figura 4.7 ilustra o cubo de dados gerado com a primeira componente principal de cada data da série temporal.

Figura 4.7 - Cubo de dados gerado para a primeira principal componente de cada data da série temporal.



Fonte: Produção do autor.

4.3. Índices de espectrais

Os índices espectrais correspondem a operações matemáticas em que as próprias bandas espectrais são os termos da equação (JUSTICE et al., 1998). Dessa forma, os valores de reflectância das bandas são convertidos para um conjunto de valores dentro do intervalo do índice. Seu uso apresenta como vantagem a redução da dimensionalidade dos dados, uma vez que concentra em apenas uma imagem, resultante do cálculo, a informação presente em várias bandas. Os índices espectrais são importantes por ressaltarem os alvos na superfície terrestre, facilitando sua identificação. Deste modo, existem centenas de índices espectrais, cada um voltado para a análise de um alvo específico, como a vegetação, o solo, e a água, por exemplo.

No caso da vegetação, um dos índices mais utilizados é o Índice de Vegetação por Diferença Normalizada, mais conhecido como NDVI, do inglês, *Normalized Difference Vegetation Index* (ROUSE et al., 1973). Este índice usa apenas as bandas do vermelho e do infra-vermelho, e embora bastante utilizado, apresenta

como problema o fato de atingir a saturação dependendo da quantidade de biomassa e do Índice de Área Foliar (IAF). Neste contexto, o Índice de Vegetação Melhorado, ou *Enhanced Vegetation Index* - EVI (HUETE et al., 1997), se apresenta como um índice mais estável, uma vez que considera outros fatores que auxiliam a corrigir a saturação, como os efeitos da superfície do solo abaixo dos dosséis e os efeitos atmosféricos. Além das bandas constantes no NDVI, o EVI considera também a banda do azul. A seguir, é apresentada sua fórmula:

$$EVI = G \left(\frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + C1 * \rho_{red} - C2 * \rho_{blue} + L} \right) \quad (4.1)$$

Onde,

ρ_{nir} é reflectância na banda do infra-vermelho próximo,

ρ_{red} é reflectância na banda do vermelho,

ρ_{blue} é reflectância na banda do azul,

G é o fator de ganho,

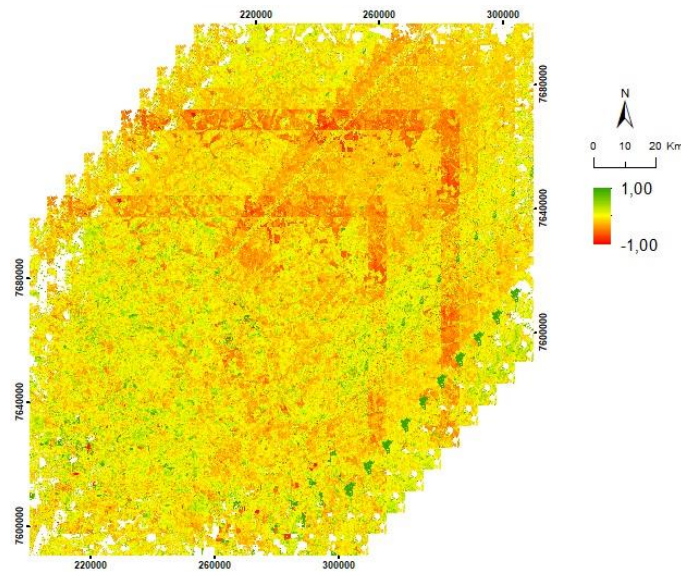
L é o fator de ajuste para o solo,

$C1$ e $C2$ são coeficientes de ajuste para efeito de aerossóis na atmosfera.

Os valores dos coeficientes adotados pelo algoritmo do EVI são: $L=1$, $C1=6$, $C2=7,5$ e $G= 2,5$ (HUETE et al., 1997; JUSTICE et al., 1998).

Como no pasto arbustivo existe uma diversidade maior de espécies vegetais, evidenciada pela presença de indivíduos arbustivos e arbóreos, pode-se afirmar que existe também uma diferença de biomassa em relação ao pasto herbáceo. Assim, neste trabalho tentou-se explorar essa diferença de biomassa na distinção entre esses dois tipos de cobertura. Foi feito, portanto, um EVI para cada data de observação da série temporal, gerando um novo cubo de dados, que foi posteriormente utilizado para classificação. A Figura 4.8 ilustra o cubo de dados de EVI.

Figura 4.8 - Cubo de EVI gerado para a série temporal.



Fonte: Produção do autor.

4.4. Modelo linear de mistura espectral

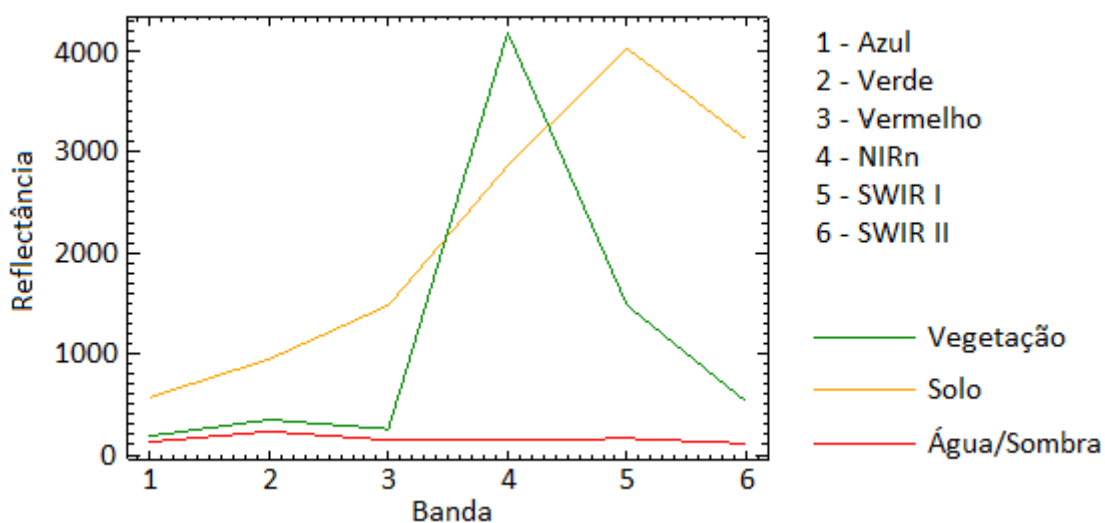
O sinal refletido da superfície que é registrado pelos sensores, dificilmente representa um único alvo (HORWITZ et al., 1971). Geralmente, os pixels em uma imagem de satélite apresentam a reflectância de um conjunto de alvos, os quais se localizam na projeção da área do pixel sobre a superfície, o que em sensoriamento remoto é conhecido como IFOV - *Instantaneous Field of View*, ou Campo de Visada Instantâneo. Isso se agrava nas imagens de baixa resolução espacial, em que os pixels têm dimensões maiores, e por isso a mistura de reflectâncias se torna ainda maior (DETMENDY; PACE, 1972). O Modelo Linear de Mistura Espectral - MLME, se apresenta, então, como uma ferramenta que possibilita a distinção dos tipos de alvo que ocorrem na área projetada de cada pixel (SHIMABUKURO; SMITH, 1991).

Por meio de uma equação baseada nos mínimos quadrados, o algoritmo do MLME calcula a proporção de cada componente da mistura nos pixels de uma imagem de satélite. Para isso, é necessária a identificação prévia de pixels puros de cada componente, isto é, exemplos de pixels em que não haveria mistura espectral, cuja reflectância seria correspondente a um tipo de alvo apenas

(RANSON, 1975). Após o fornecimento desses pixels puros, também conhecidos como *endmembers*, são geradas imagens-fração das componentes, além de uma imagem de erro, cujos pixels variam entre 0 e 1, uma vez que representam proporções.

Para este trabalho, foram consideradas as componentes de vegetação, solo e água/sombra para extração das imagens-fração. A Figura 4.9 apresenta um gráfico do perfil espectral de cada *endmember* escolhido para as componentes na imagem do dia 15/07/2018, que foi escolhida como exemplo devido à ausência de cobertura de nuvens. Os *endmembers* das demais imagens da série temporal apresentam perfis espectrais semelhantes.

Figura 4.9 - Perfis espectrais dos *endmembers* escolhidos para as componentes do MLME gerado para a data de 15/07/2018.

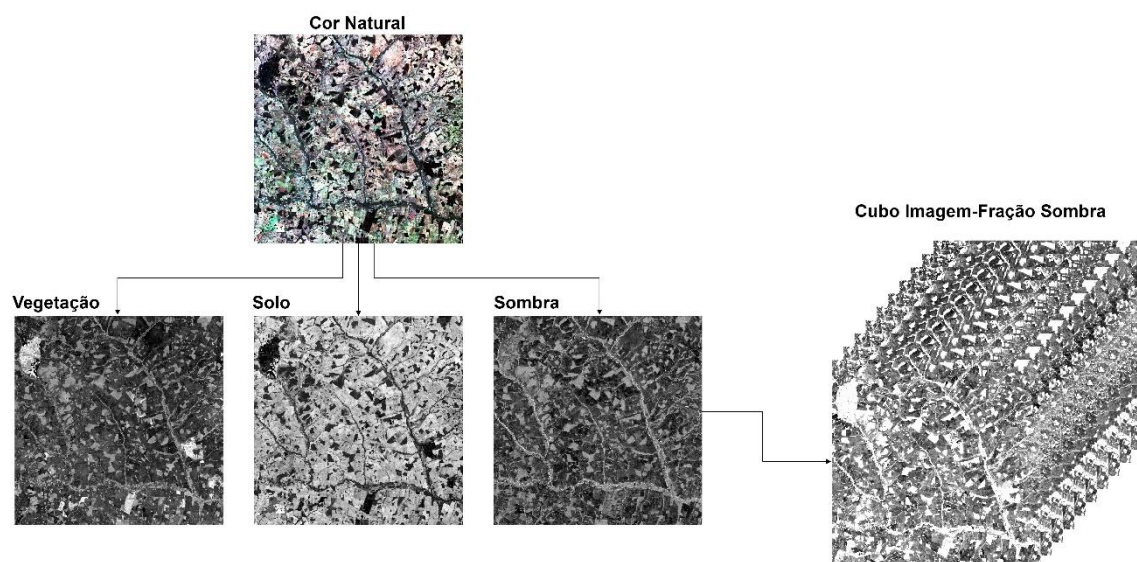


Fonte: Produção do autor.

Conforme já mencionado, o porte das espécies vegetais é o principal aspecto que diferencia pastagem arbustiva de pastagem herbácea. Como o pasto arbustivo é caracterizado pela maior incidência de indivíduos arbustivos e arbóreos, que apresentam porte mais elevado que a vegetação herbácea predominante no pasto herbáceo, espera-se que a incidência de sombras no pasto arbustivo também seja maior que no pasto herbáceo. Considerando isso,

a partir do MLME que foi gerado para cada data da série temporal, foram extraídas imagens-fração das três componentes anteriormente citadas, e as imagens-fração da componente sombra foram utilizadas para compor um dos cubos de dados que foi utilizado nas classificações, a fim de auxiliar na distinção entre as classes de pastagem arbustiva e pastagem herbácea. A Figura 4.10 apresenta como exemplo, as imagens-fração extraídas do MLME gerado para a imagem do dia 15/07/2018, e o cubo de dados gerado com a componente sombra.

Figura 4.10 - Frações espectrais das componentes utilizadas no MLME e cubo de dados da imagem-fração sombra.



Fonte: Produção do autor.

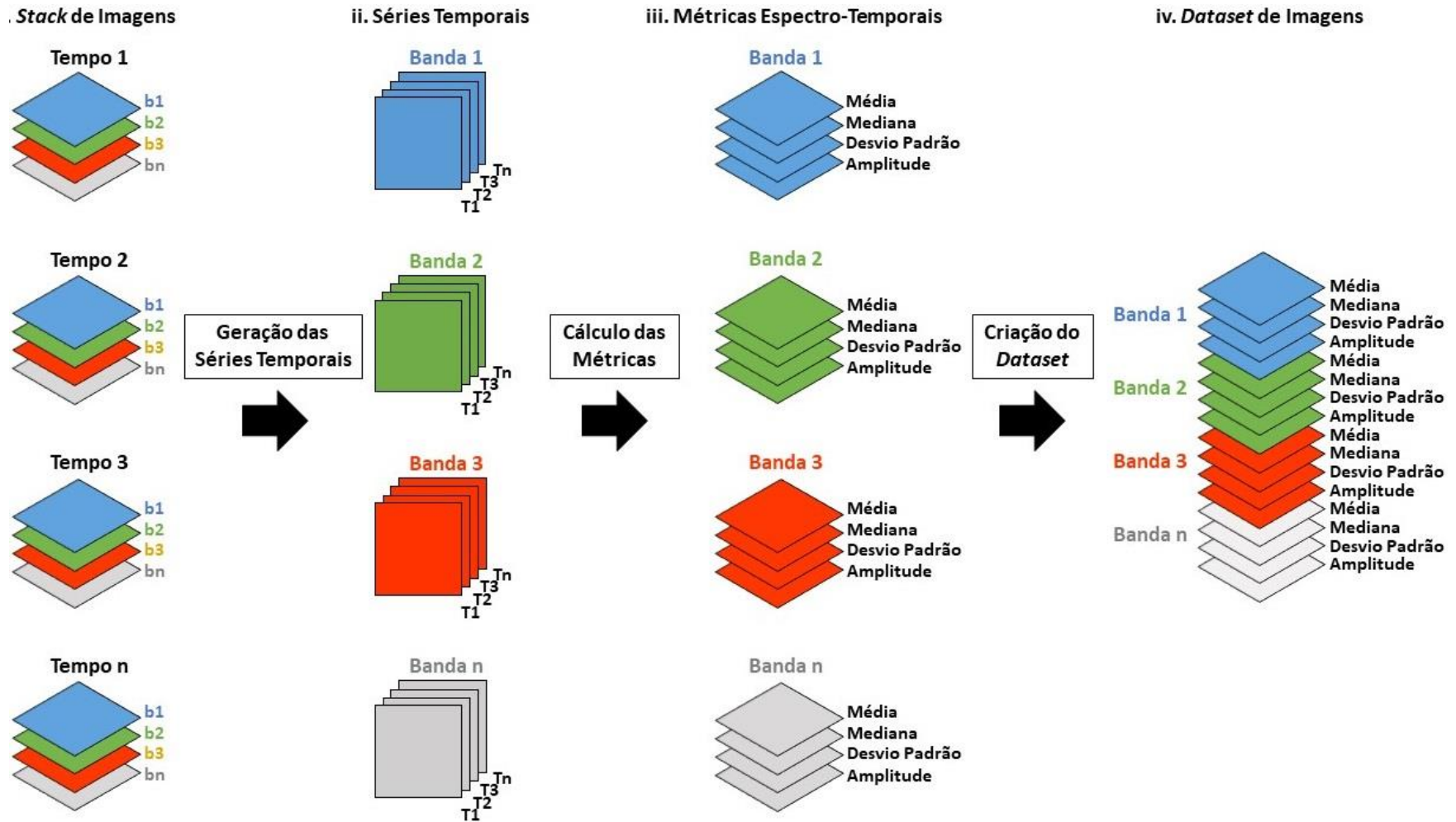
4.5. Métricas espectro-temporais

Conforme já mencionado, neste trabalho estão sendo usadas 6 bandas espectrais do Sentinel-2 A/B. Para cada uma dessas bandas, foram constituídas séries temporais utilizando todas as observações do ano de 2018 na área de estudo, com a exceção de 5 datas devido à total cobertura de nuvens. Dessa forma, cada banda possui séries temporais compostas por 65 observações. Entretanto, é importante lembrar que as séries são irregulares, ou seja, apesar

de terem sido consideradas 65 observações, nem todos os pixels foram observados em todas as datas devido à cobertura de nuvens.

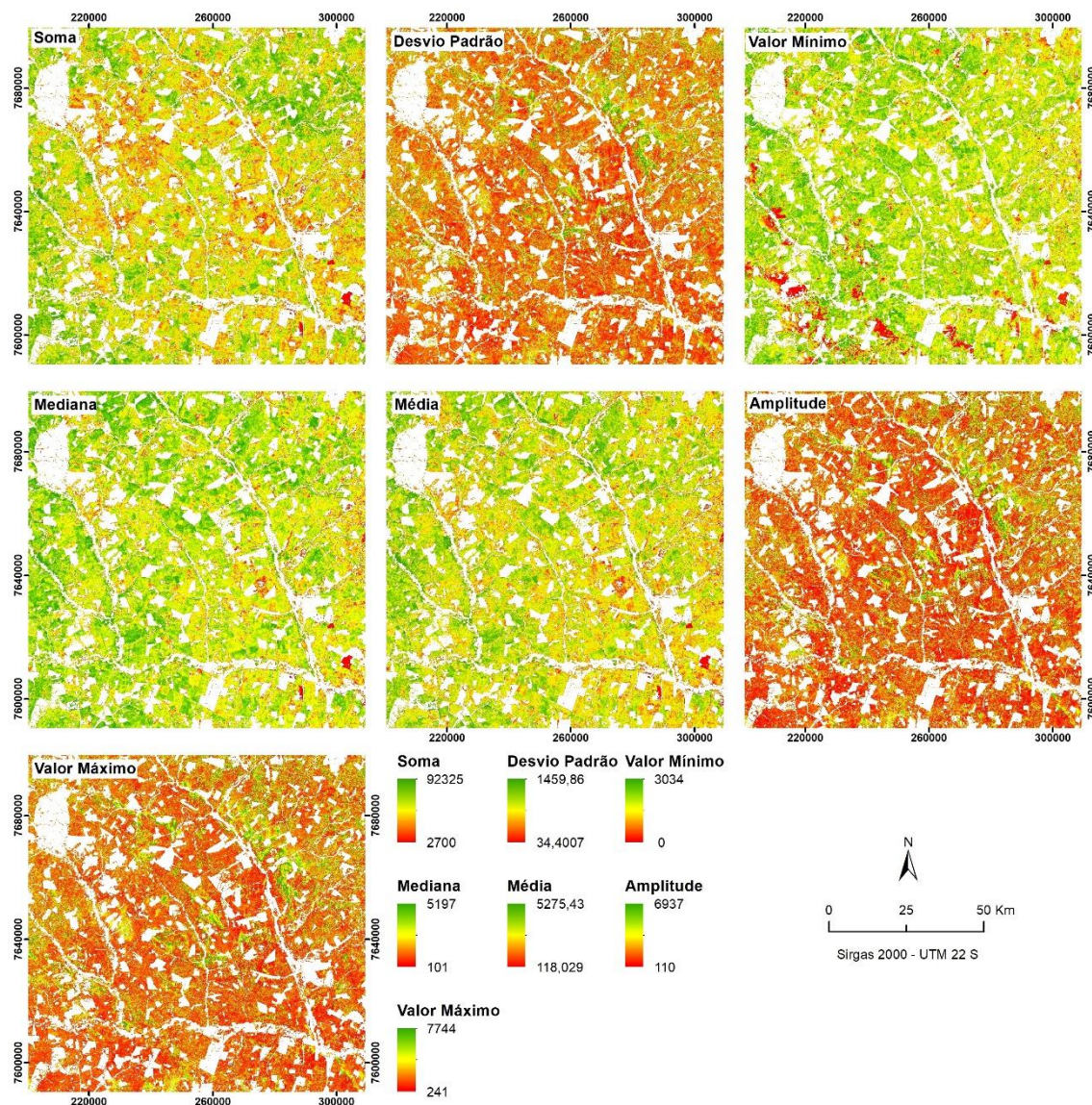
As métricas espectro-temporais correspondem a cálculos estatísticos aplicados às séries temporais, com o objetivo de concentrar em apenas uma imagem um valor que represente a série temporal em cada pixel (KÖRTING et al., 2013). Assim, o valor resultante de cada métrica representa o comportamento espectral de cada pixel ao longo de um tempo determinado (PARENTE et al., 2019). Neste trabalho, foram calculadas 7 métricas para cada uma das 6 bandas, a fim de representar o comportamento espectral dos pixels ao longo do ano de 2018. Foram consideradas as seguintes métricas: (i) média; (ii) mediana; (iii) desvio padrão; (iv) soma; (v) valor mínimo; (vi) valor máximo; e (vii) amplitude. Com o cálculo das métricas, foram obtidas no total 42 imagens, as quais foram reunidas, dando origem ao cubo de dados das Métricas Espectro-Temporais (MET). A Figura 4.11 apresenta um esquema de como as métricas e o cubo de dados resultante foram gerados. A Figura 4.12 apresenta como exemplo as imagens resultantes da aplicação de cada métrica sobre a série temporal da banda 2.

Figura 4.11 - Esquema da geração das métricas espectro-temporais e cubo de dados resultante.



Fonte: Produção do autor.

Figura 4.12 - Imagens das métricas espectro-temporais utilizadas para a geração do cubo de dados.



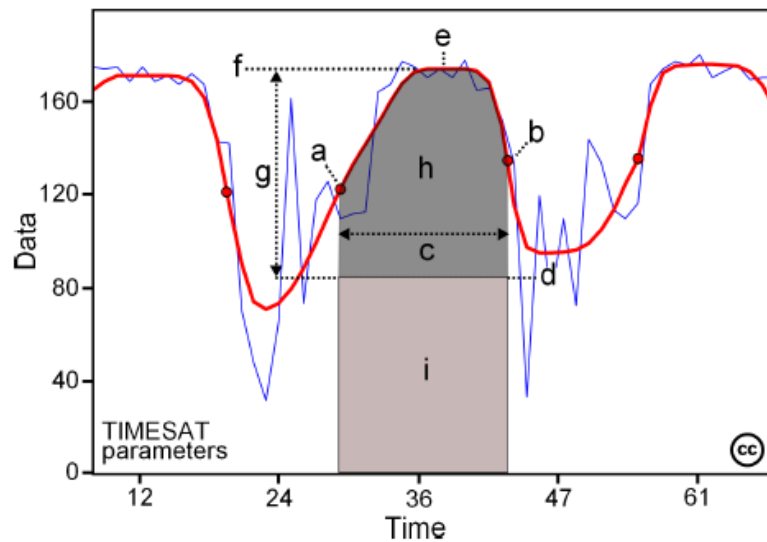
Fonte: Produção do autor.

4.6. Métricas fenológicas

A fenologia diz respeito ao padrão cíclico apresentado pela vegetação, que é influenciado pelas condições climáticas e sazonais, e que pode ser observado através do perfil espectral dos alvos de interesse (JÖNSSON; EKLUNDH, 2002). Dentro deste contexto, as métricas fenológicas correspondem a medições de parâmetros extraídos das curvas espectrais, através dos quais é possível fazer a interpretação da fenologia (JÖNSSON; EKLUNDH, 2004). É comum dentro do

sensoriamento remoto a utilização de índices de vegetação para gerar as curvas espectrais e extrair os parâmetros sazonais que representam o comportamento fenológico da vegetação. Neste trabalho foi utilizado o EVI para extração dos parâmetros de fenologia das classes de pasto arbustivo e pasto herbáceo. No total, foram extraídas 13 métricas utilizando o *software* TIMESAT versão 3.3. As métricas utilizadas são ilustradas na Figura 4.13 e os parâmetros extraídos apresentados na Tabela 4.1.

Figura 4.13 - Métricas fenológicas geradas pelo TIMESAT utilizando os ciclos sazonais da vegetação.



A descrição dos códigos da legenda é encontrada na Tabela 4.1.

Fonte: Eklundh and Jonsson (2015).

Tabela 4.1 - Métricas fenológicas extraídas do Timesat.

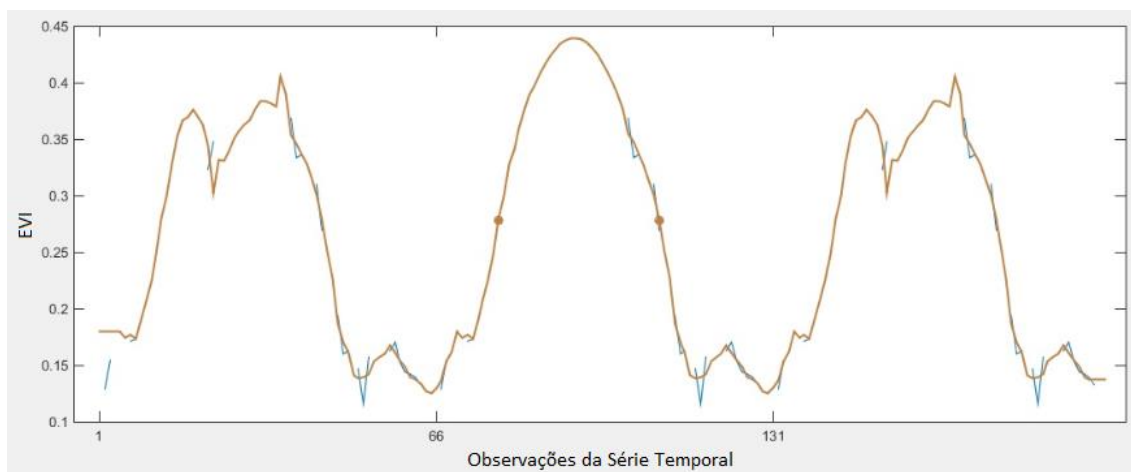
Legenda	Nome da Métrica Fenológica	Sigla
a	<i>Time for the start of the season</i>	<i>sos</i>
a	<i>Value for the start of the season</i>	<i>sosval</i>
b	<i>Time for the end of the season</i>	<i>eos</i>
b	<i>Value for the end of the season</i>	<i>eosval</i>
c	<i>Length of the season</i>	<i>los</i>
d	<i>Base level</i>	<i>base</i>
e	<i>Time for the mid of the season</i>	<i>mid</i>
f	<i>Largest data value for the fitted function during the season</i>	<i>peak</i>
g	<i>Seasonal amplitude</i>	<i>amp</i>
h	<i>Small seasonal integral</i>	<i>sint</i>
h + i	<i>Large seasonal integral</i>	<i>lint</i>
-	<i>Rate of increase at the beginning of the season</i>	<i>lder</i>
-	<i>Rate of decrease at the end of the season</i>	<i>rder</i>

Fonte: Eklundh and Jonsson (2015).

Uma peculiaridade deste *software* é que as curvas espectrais precisam ser ajustadas para que o pico vegetativo coincida com a posição central da série. As condições climáticas do Cerrado tornam o calendário agrícola não coincidente com o calendário gregoriano. Como requisito, então, para utilização do referido *software*, as séries temporais foram centradas no período de ocorrência do pico vegetativo, que no Cerrado, corresponde aos períodos de chuva.

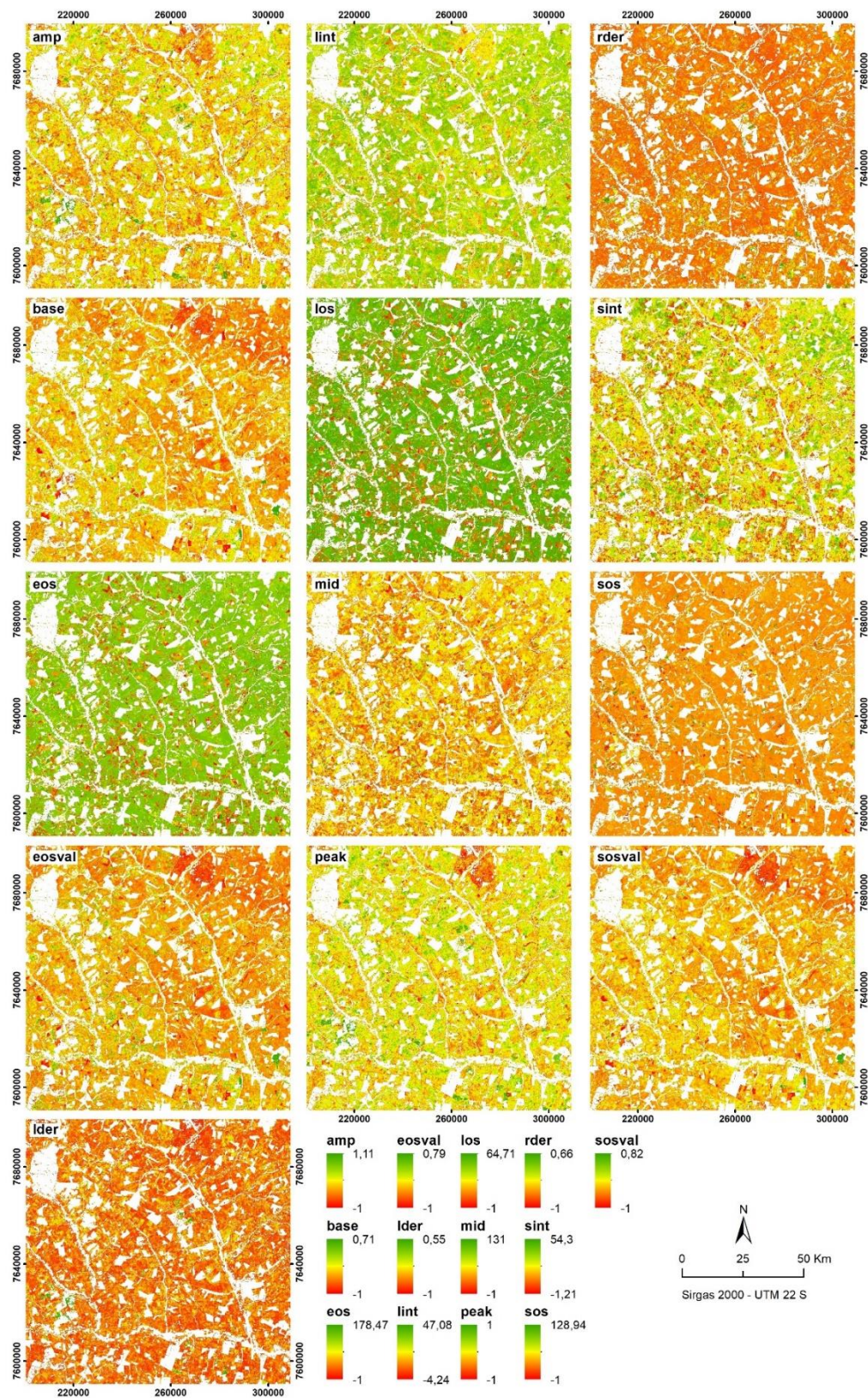
Como as séries temporais são irregulares devido à cobertura de nuvens, foi necessária a aplicação de um filtro para preencher as lacunas e suavizar os *outliers*. Para isso, utilizou-se o filtro Savitzky-Golay. A análise das séries temporais revelou a ocorrência de dois ciclos sazonais nas imagens de EVI, sendo um mais elevado e com maior duração, e o outro menos elevado e com menor duração, conforme apresenta a Figura 4.14. Os 13 parâmetros sazonais foram então extraídos para cada um dos dois ciclos. Deste modo, 26 imagens foram geradas no total, representando cada métrica fenológica do TIMESAT para cada um dos ciclos observados. Essas imagens foram também reunidas, formando o cubo de dados das Métricas Fenológicas (MF). A Figura 4.15 apresenta como exemplo as imagens das métricas fenológicas extraídas para o primeiro ciclo.

Figura 4.14 - Exemplos dos ciclos sazonais de pastagem encontrados na área e estudo, sendo um mais elevado e de maior duração e outro menos elevado e de menor duração.



Fonte: Produção do autor.

Figura 4.15 - Imagens das métricas fenológicas extraídas do Timesat para o primeiro ciclo.



Fonte: Produção do autor.

4.7. Segmentação de imagens

A segmentação consiste no particionamento de uma imagem em diversos polígonos, agrupando pixels semelhantes entre si e que diferem dos grupos adjacentes. A semelhança entre os pixels é dada por um limiar de similaridade, a partir do qual o algoritmo segmentador define os pixels que farão parte de cada grupo. Neste trabalho, foi utilizado o algoritmo *Multiresolution Segmentation*, traduzido como Segmentação Multirresolução (BAATZ; SCHÄPE, 2000), disponível no *software* eCognition Developer, versão 9.0. Este algoritmo realiza a segmentação por meio da técnica de crescimento de regiões, na qual os pixels são comparados par a par, a partir de pixels isolados que são identificados como sementes (TRIMBLE, 2014). Essa comparação é feita no espaço de atributos, tendo como base a distância euclidiana. Caso essa distância seja inferior ao valor estabelecido como o limiar de similaridade, então os pixels passam a compor o mesmo objeto, caso não, permanecem em objetos distintos. O cálculo da distância euclidiana é feito utilizando a seguinte equação:

$$D_{(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4.2)$$

Onde, $D_{(x,y)}$ é a distância euclidiana entre os pontos x e y .

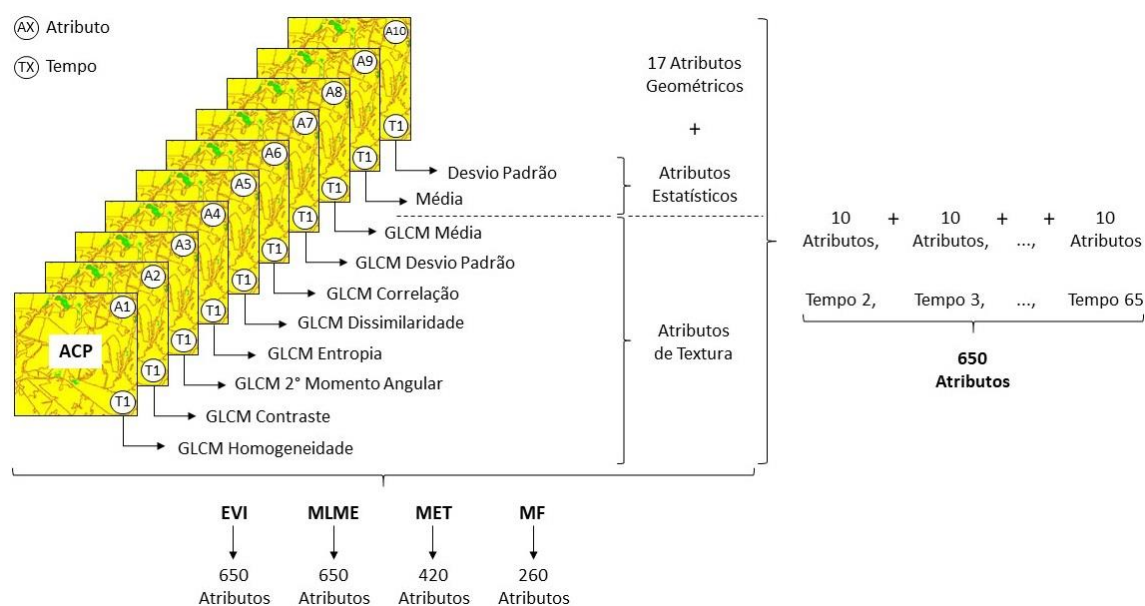
Para realizar a segmentação, foi utilizada uma única imagem que representasse todo o conjunto de imagens, a fim de contemplar todas as características da série temporal. Para isso, todas as bandas espectrais de todas as observações da série temporal foram agrupadas em uma única pilha, gerando-se em seguida uma imagem da média dos pixels. Como parâmetros de segmentação, foi utilizado para o fator de escala (*Scale parameter*) o valor de 200, para a forma (*Shape*) o valor de 0,1 e para a compactidade (*Compactness*) o valor de 0,5. O resultado da segmentação foi então utilizado para a extração de atributos de todos os cubos de dados mencionados anteriormente.

4.8. Extração de atributos

Conforme apresentado na Seção 2.6, a extração de atributos é parte integrante da mineração de dados, e é a etapa pela qual se inicia o processo de

reconhecimento de padrões. Esta etapa foi executada no *software* eCognition, versão 9. Três tipos de atributos foram considerados, sendo eles os atributos estatísticos, os atributos de textura e os atributos geométricos, os quais serão descritos adiante. Como este trabalho utiliza a abordagem baseada em GEOBIA, é importante dizer que os atributos foram extraídos por objeto, ou seja, o valor de cada atributo é relativo a cada polígono gerado através da segmentação. Vale ainda dizer que cada cubo de dados abordado anteriormente foi submetido à extração de atributos de forma individual, originando um banco de dados de atributos para cada cubo de dados, tendo como limites comuns a segmentação. A Figura 4.16 exemplifica o modo como ocorreu a extração de atributos.

Figura 4.16 - Esquema representando a extração de atributos.



Fonte: Produção do autor.

Juntos, os atributos texturais e espectrais compreendem 10 atributos extraídos de cada imagem em cada um dos cubos de dados. Esse valor deve ser, então, multiplicado pelo número de imagens que compõem cada cubo de dado, a fim de obter a quantidade total de atributos em cada cubo. Os atributos espaciais são agregados ao final deste processo, uma vez que são os mesmos para cada conjunto de dado. Nas classificações combinadas, os atributos espaciais são contabilizados apenas uma vez.

4.8.1. Atributos estatísticos

Estes atributos correspondem a medidas estatísticas extraídas do grupo de pixels dentro de cada objeto ao longo das imagens que formam os cubos de dados. Para esta classe de atributos, foram considerados dois tipos, sendo eles a média dos pixels e o desvio padrão dos pixels. A média é importante por constituir um valor representativo do conjunto de pixels, enquanto o desvio padrão expressa a dimensão da variabilidade entre os valores de cada conjunto. Utilizando esses parâmetros estatísticos é possível conhecer o comportamento dos conjuntos de pixels no que tange aos valores dos atributos que representam.

4.8.2. Atributos de textura

Os atributos de textura correspondem a medidas da rugosidade das imagens de satélite (HALL-BEYER, 2017a). A rugosidade pode ser definida pelo gradiente que ocorre entre os pixels, isto é, a diferença entre os valores dos níveis de cinza da imagem. Dessa forma, uma imagem é tão rugosa quanto maior for o seu gradiente, e de modo contrário, uma imagem é considerada mais lisa quanto menor for a diferença entre os valores dos números digitais que a compõem. A textura é analisada por meio da Matriz de Co-ocorrência de Níveis de Cinza, do inglês, *Gray Level Co-occurrence Matrix* (GLCM), a partir da qual são extraídos os atributos de textura (HALL-BEYER, 2017b). A diferença entre os pixels é, então, analisada por meio dessa matriz considerando diferentes direções (0°, 45°, 90°, 135° ou todas). Considerando que a pastagem é um tipo de uso do solo essencialmente heterogêneo, coberto por vegetação herbácea, arbustiva e arbórea, pode-se dizer que a análise da textura tem potencial para auxiliar na distinção entre as classes de pasto arbustivo e pasto herbáceo, dado que a rugosidade é um dos principais aspectos que as diferencia nas imagens de satélite.

Neste trabalho foram utilizados oito atributos de textura, conforme Haralick (1993), a fim de diferenciar a pastagem arbustiva da pastagem herbácea usando a rugosidade. Foram considerados, então, os seguintes atributos texturais: (i) GLCM Homogeneidade; (ii) GLCM Contraste; (iii) GLCM 2º Momento Angular; (iv) GLCM Entropia; (v) GLCM Dissimilaridade; (vi) GLCM Correlação; (vii) GLCM

Desvio Padrão; e (viii) GLCM Média. Estes atributos foram extraídos considerando todas as direções.

4.8.3. Atributos geométricos

A análise de atributos geométricos é possibilitada pela abordagem baseada em objetos, haja vista ser considerada nessa abordagem a dimensão espacial para a extração das características da cobertura do solo. Devido a isso, os atributos geométricos são também chamados de atributos espaciais. Conforme assinalado por Tobler (1970), todas as coisas estão correlacionadas, entretanto, coisas próximas estão mais correlacionadas do que coisas distantes. De acordo com este princípio, os atributos espaciais trazem consigo a vantagem de considerar a relação espacial entre os pixels vizinhos, etapa que é realizada durante a segmentação, e não apenas a relação temporal. Dentro deste contexto, Assis et al. (2017) discutem a relação do tempo e do espaço nas análises de séries temporais. De acordo com estes autores, é mais adequado considerar em primeiro lugar o tempo, e depois o espaço, pois do contrário, erros de classificação poderiam levar a inconsistências que tenderiam a aumentar continuamente conforme se incrementam variáveis à análise. Neste trabalho foi adotada esta prática. Foram considerados, então, 17 atributos geométricos do eCognition, conforme apresentado na Tabela 4.2. Para maior entendimento da definição de cada atributo, consultar Trimble (2014).

Tabela 4.2 - Atributos geométricos utilizados, por tipo de medida.

Número do Atributo	Nome do Atributo	Tipo de medida
1	<i>Border length (Pxl)</i>	Extensão
2	<i>Width (Pxl)</i>	Extensão
3	<i>Asymmetry</i>	Forma
4	<i>Rel. Border to Image Border</i>	Extensão
5	<i>Elliptic Fit</i>	Forma
6	<i>Density</i>	Forma
7	<i>Number of Pixels</i>	Extensão
8	<i>Radius of smallest Enclosing Ellipse</i>	Forma
9	<i>Rectangular Fit</i>	Forma
10	<i>Length (Pxl)</i>	Extensão
11	<i>Length/Width</i>	Extensão
12	<i>Radius of Largest Enclosed Ellipse</i>	Forma
13	<i>Main Direction</i>	Forma
14	<i>Shape Index</i>	Forma
15	<i>Compactness</i>	Forma
16	<i>Roundness</i>	Forma
17	<i>Border Index</i>	Forma

Fonte: Produção do autor.

4.9. Seleção de atributos

De acordo com Hall (1999), a seleção de atributos potencializa a predição dos algoritmos classificadores. Após a extração de atributos, foi utilizado como seletor o algoritmo *Correlation-based Feature Selection - CFS* (Seleção de Atributos baseada na Correlação), disponível no *software* Weka versão 3.8.3, para diminuir a dimensionalidade dos dados. Este algoritmo é baseado na correlação de Pearson e compara pares de atributos e cada atributo com as classes, criando uma ordem de mérito entre os atributos. Assim, os atributos considerados irrelevantes ou redundantes são eliminados do conjunto de amostras tanto de treinamento quanto de teste, mantendo somente aqueles atributos identificados com elevada capacidade de predição.

$$M_S = \frac{k\bar{r}_{cf}}{\sqrt{k+k(k-1)\bar{r}_{ff}}} \quad (4.3)$$

Onde,

M_S é o mérito de um conjunto S com k atributos,

$\overline{r_{cf}}$ é a média da correlação entre os atributos e as classes, e

$\overline{r_{ff}}$ é a média da correlação entre os atributos.

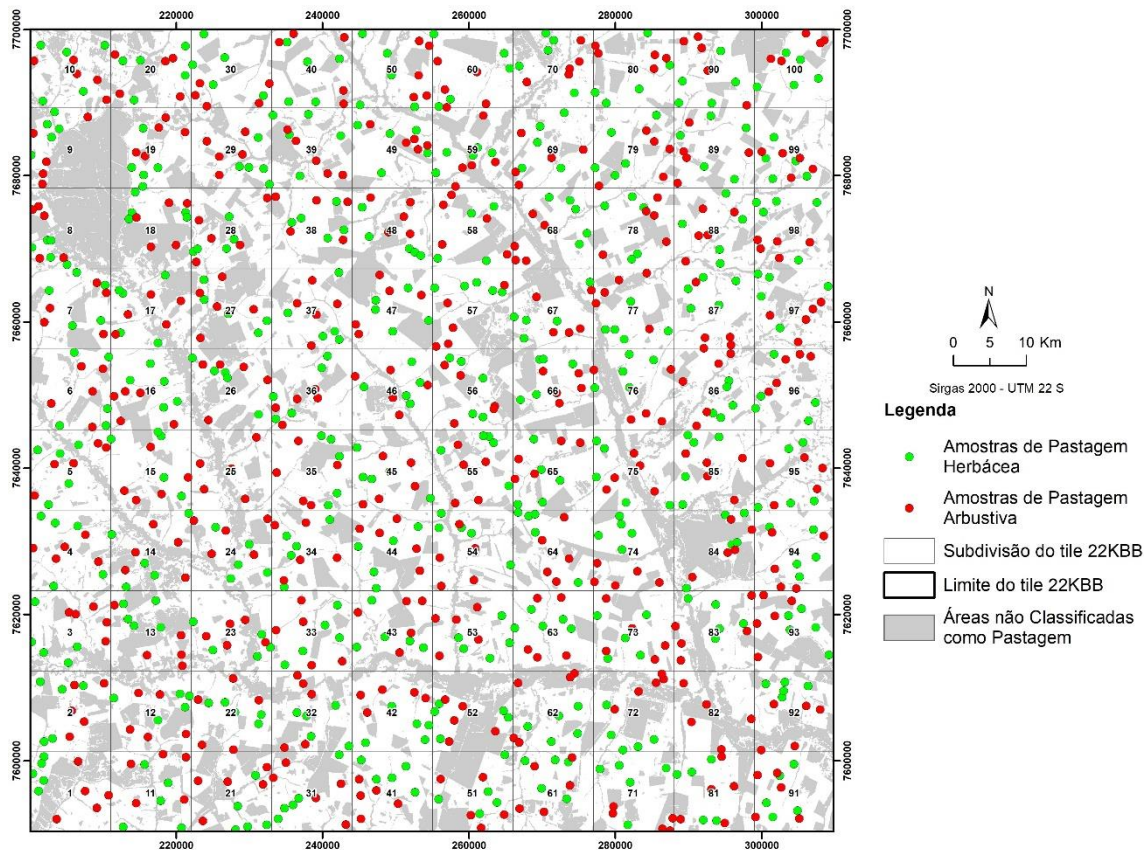
O numerador da equação é responsável por indicar o quão preditivo um conjunto de atributos é para uma determinada classe, e o denominador, o grau de redundância entre os atributos.

4.10. Processo de amostragem

A seleção de amostras foi feita utilizando as imagens de satélite do Google Earth (GE), devido à sua alta resolução espacial, associando à visualização do terreno através de fotos disponíveis no Google Maps. Como não se tem controle sobre a data das imagens disponibilizadas pela plataforma do Google, as amostras foram validadas utilizando imagens PlanetScope, de 4,77 m de resolução espacial, do mesmo ano das imagens da série temporal, ou seja, de 2018. Com isso, foi possível conciliar a alta resolução espacial das imagens do GE e a temporalidade das imagens da Planet, possibilitando a geração de amostras consistentes no espaço e no tempo. Essas imagens foram adquiridas pelo governo da Noruega através do NICFI (*Norway's International Climate and Forest Initiative*), e disponibilizadas gratuitamente durante um período em mosaicos temporais bianuais, ou seja, dois mosaicos por ano, para regiões entre os trópicos. Neste trabalho foi utilizado o mosaico temporal de junho de 2018.

Foi adotado um método de amostragem sistemática (OLOFSSON et al., 2014), no qual a área de estudo (*tile 22KBB* da grade do Sentinel-2) foi dividida regularmente em 10 quadrantes por linha e 10 quadrantes por coluna, totalizando 100 quadrantes (Figura 4.17). Para garantir que a área de estudo fosse amostrada de forma homogênea, em cada quadrante foram obtidas cinco amostras de pasto arbustivo e cinco amostras de pasto herbáceo, resultando em 10 amostras por quadrante e 1000 no total, sendo 500 amostras para cada uma das duas classes. Contudo, em dois quadrantes não se verificou a ocorrência da classe de pasto herbáceo. Deste modo, ao invés de 500 amostras para esta classe, utilizou-se 490, reduzindo também a quantidade total de 1000 para 990 amostras.

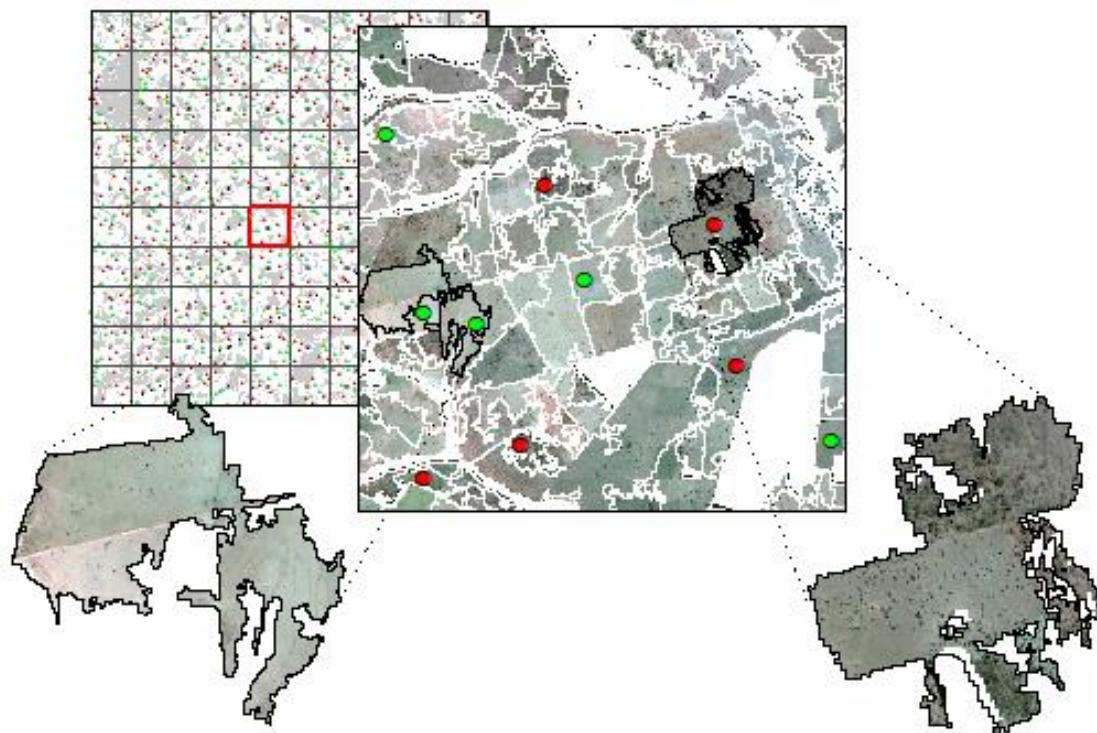
Figura 4.17 - Representação esquemática do processo de amostragem.



Fonte: Produção do autor.

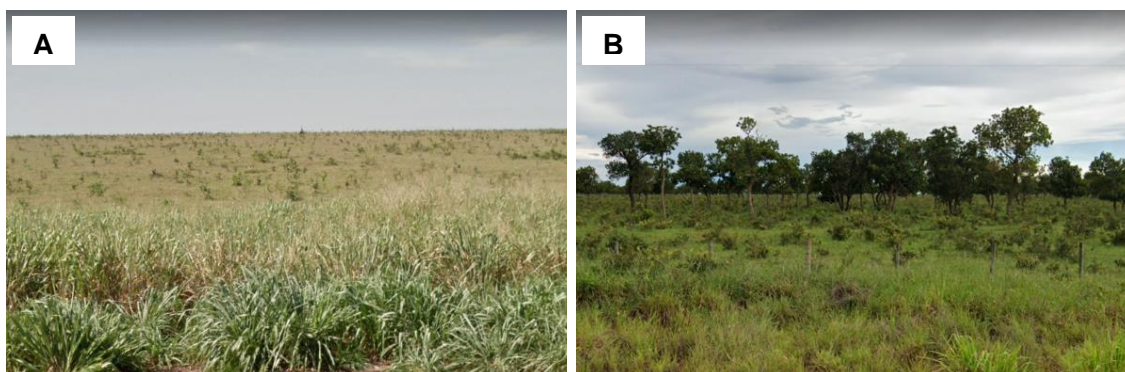
Vale também mencionar que como este trabalho se baseou na abordagem do GEOBIA, o que constitui as amostras são os polígonos provenientes da segmentação. A amostragem baseada em objetos se mostrou eficaz para os tipos de classe considerados neste trabalho, uma vez que os polígonos foram aderentes às áreas com maior e menor densidade de indivíduos arbustivos e arbóreos, não se atendo necessariamente aos limites dos talhões e dos imóveis rurais, se adequando, portanto, à conceituação das classes conforme proposto por Coutinho et al. (2020). A Figura 4.18 e a Figura 4.19 apresentam exemplos das classes conforme a imagem de satélite e observação de fotos no Google Maps, respectivamente.

Figura 4.18 - Detalhamento das amostras de pastagem arbustiva (em vermelho) e pastagem herbácea (em verde), e sua relação com a segmentação.



Fonte: Produção do autor.

Figura 4.19 - Classes de amostras conforme visão detalhada de campo: pastagem herbácea (A) e pastagem arbustiva (B).



Fonte: Google Maps (2021).

4.11. Classificação e avaliação

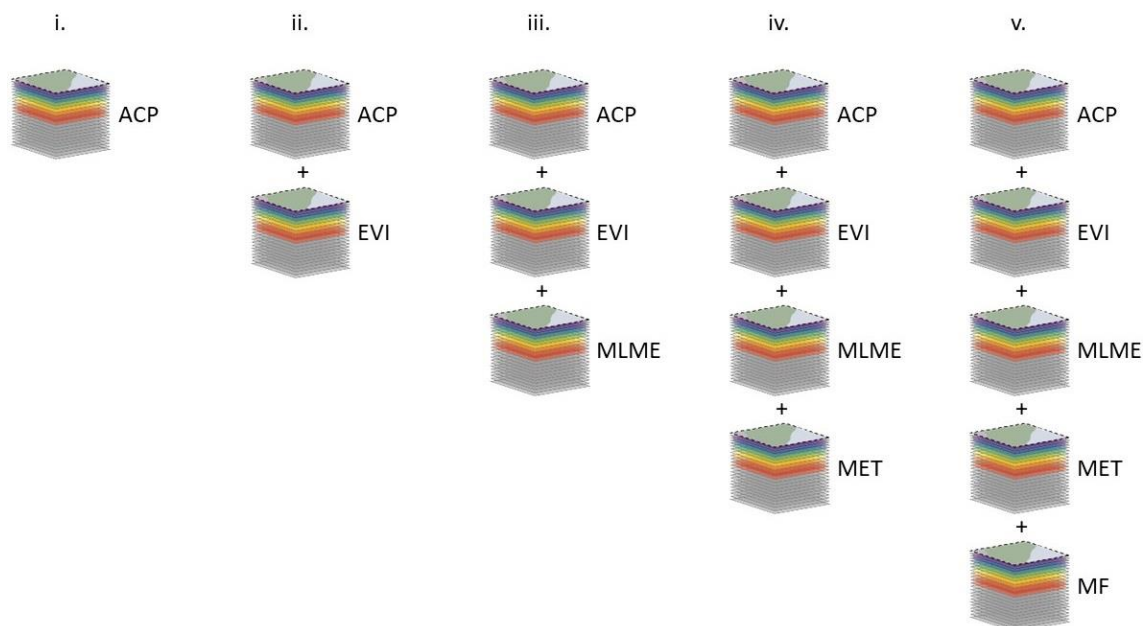
Para a classificação, as amostras foram divididas em 70% para treinamento e 30% para validação. Deste modo, para a classe de pasto arbustivo foram

consideradas 350 amostras de treinamento e 150 de validação, e para a classe de pasto herbáceo, 343 amostras para treinamento e 147 para validação. Foi utilizado como classificador o método *Random Forest*, disponível no *software* Weka versão 3.8.3, devido à capacidade deste algoritmo de lidar com grandes volumes de dados evitando o ajuste excessivo. As classificações foram realizadas utilizando 100 árvores (BREIMAN, 2001).

Foram realizadas classificações utilizando como insumo os cinco cubos de dados anteriormente citados. No total, foram feitas 18 classificações, as quais são agrupadas em dois grupos, classificações individuais e classificações combinadas. Foi feita essa divisão para avaliar a classificação com cada tipo de dado separadamente, e a contribuição de cada um utilizando-os em conjunto. Além disso, cada conjunto de dado, individual ou combinado, foi classificado duas vezes, uma considerando todos os atributos extraídos e outra apenas com os atributos selecionados. Isso foi feito para avaliar também a contribuição da seleção de atributos para a classificação.

As classificações combinadas foram feitas de acordo com a seguinte ordem: ACP, EVI, MLME, MET e MF, sendo cada cubo de dados agregado por vez, como ilustra a Figura 4.20. Portanto, 10 classificações compõem o grupo das classificações individuais e oito o grupo das classificações combinadas. Vale mencionar que em todas as classificações foram usadas as mesmas amostras, o mesmo classificador e os mesmos parâmetros de classificação, variando apenas os dados de entrada em cada modelo, já que o objetivo é analisar a contribuição dos cubos de dados para a classificação.

Figura 4.20 - Representação do esquema de agregação de dados utilizado nas classificações combinadas.



ACP: Análise por Componentes Principais; EVI: *Enhanced Vegetation Index* (Índice de Vegetação Melhorado); MLME: Modelo Linear de Mistura Espectral; MET: Métricas Espectro-Temporais; MF: Métricas Fenológicas.

Fonte: Produção do autor.

A avaliação das classificações foi feita utilizando matrizes de confusão, por meio das quais foi possível analisar os erros e a acurácia dos modelos. Essas matrizes foram geradas utilizando as amostras que foram destinadas a validação. Para manter a coerência no processo de avaliação, os dados de validação foram gerados utilizando um material com resolução espacial mais detalhada (imagens do GE e Planet) que a do material usado para classificação (imagens do Sentinel-2 A/B).

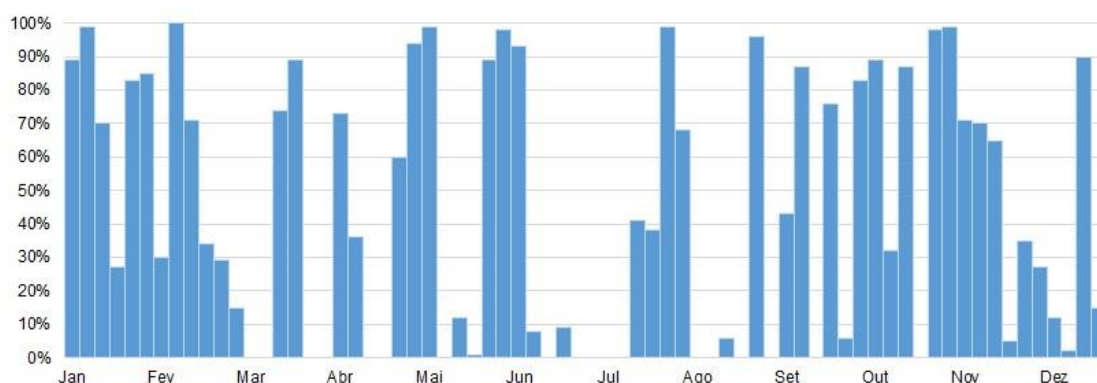
5. RESULTADOS E DISCUSSÃO

5.1. Análise dos resultados

5.1.1. Cobertura de nuvens

Mesmo utilizando dados de sensoriamento remoto com alta resolução temporal, ou seja, com frequência de observação de 5 dias, como a proporcionada pelos satélites Sentinel-2 A/B combinados, a cobertura de nuvens na área de estudo ainda representou um desafio para análise das séries temporais (PRUDENTE et al., 2020). O Apêndice A apresenta imagens sequenciais de toda a série temporal utilizada e a Figura 5.1 apresenta um gráfico com o percentual de cobertura de nuvens em cada data de observação dos satélites. Conforme pode ser observado, a maioria das imagens possui cobertura de nuvens, com algumas tendo apresentando cobertura completa, ou bem próximo a isso. Cerca de um quarto do total não apresentou cobertura de nuvens. Apesar disso, por menor que tenha sido a cobertura, isso ainda assim causou irregularidade nas séries, haja vista as nuvens terem impedido a observação de vários grupos de pixels na superfície.

Figura 5.1 - Percentual de cobertura de nuvens em cada data de observação da série temporal.



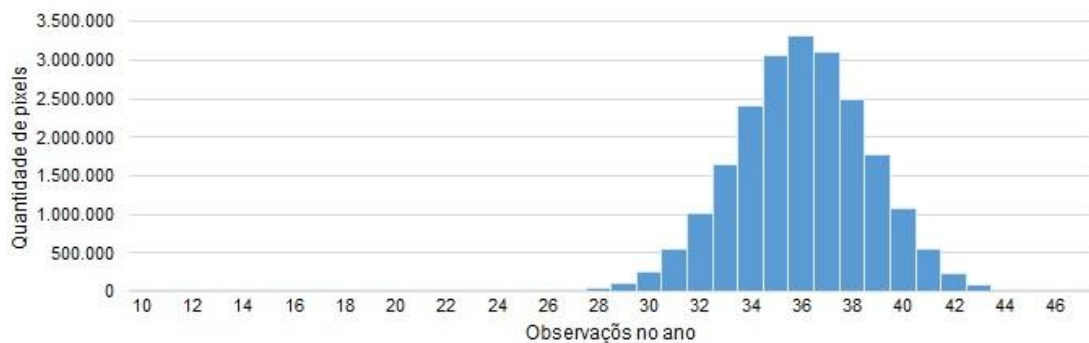
Fonte: Produção do autor.

Ainda de acordo com a Figura 5.1, pode-se observar que as datas com as menores taxas de cobertura de nuvens ocorrem no período seco, aproximadamente de março a setembro, apesar de alguns casos ainda com

percentuais elevados de cobertura. Por outro lado, a maior quantidade de dias consecutivos com cobertura de nuvens ocorre durante a estação chuvosa, aproximadamente de outubro a fevereiro.

A série temporal utilizada possui um total de 70 imagens, contudo, os pixels não foram observados de forma homogênea. A Figura 5.2 apresenta a relação entre a quantidade de observações livres de nuvens e a quantidade de pixels. A quantidade de pixels livres de nuvens apresenta uma distribuição normal, com pico no valor 36. A menor quantidade de observações da série corresponde a 10 observações, e a maior, a 47 observações. A maioria das séries possuem 36 observações.

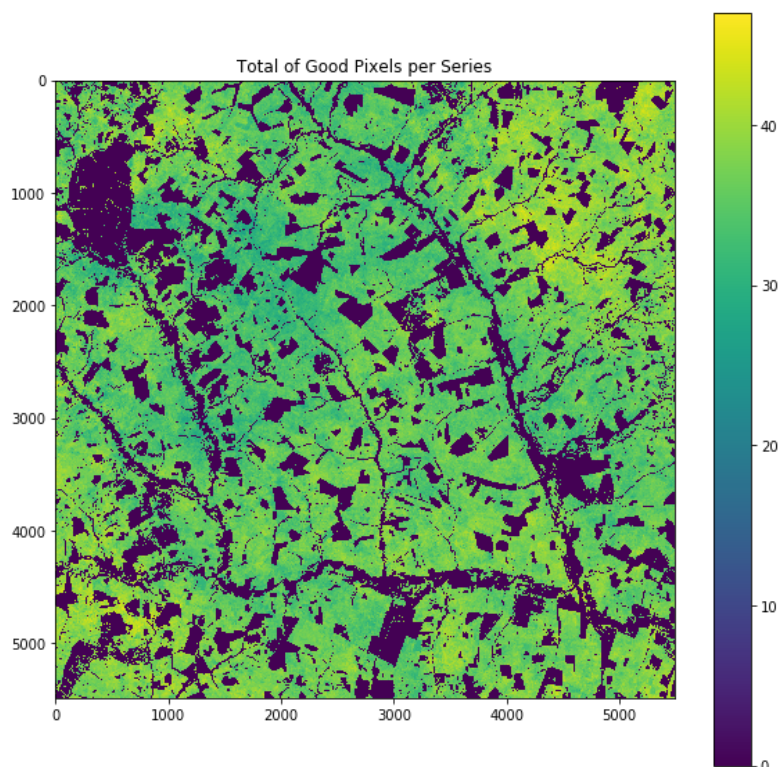
Figura 5.2 - Quantidade de observações livres de nuvens pela quantidade de pixels observados.



Fonte: Produção do autor.

A Figura 5.3 apresenta a distribuição espacial da quantidade de observações por pixel na área de estudo. É possível notar, portanto, que mesmo utilizando séries temporais densas, e em uma área do Cerrado com comportamento irregular de chuvas, com um longo período considerado seco, a cobertura de nuvens representa um desafio. Considerando isso, métodos de interpolação de séries temporais, que simulam séries regulares, podem ser um importante recurso para solucionar esse tipo de problema (RUFIN et al., 2019; SCHWIEDER et al., 2016).

Figura 5.3 - Distribuição espacial da quantidade de pixels livres de nuvens observados.



Os pixels com valor 0 correspondem à máscara utilizada para eliminar outras classes de uso do solo diferentes de pastagem.

Fonte: Produção do autor.

5.1.2. Seleção de atributos

Com a seleção de atributos a dimensionalidade dos dados foi bastante reduzida. A Tabela 5.1 apresenta a quantidade total de atributos antes e depois da etapa de seleção, além da representatividade dos atributos selecionados no total de atributos extraídos. De acordo com essa tabela, o percentual de atributos selecionados variou de 4,4% a 13,5% do total de atributos extraídos dos conjuntos de dados.

Tabela 5.1 - Quantidade total de atributos extraídos em valores absolutos e relativos, com e sem seleção de atributos, por tipo de conjunto de dados.

Conjuntos de Dados	Atributos Extraídos	Atributos Selecionado	%
[Individual] Análise por Componentes Principais (ACP)	667	73	10.9
[Individual] <i>Enhanced Vegetation Index</i> (EVI)	667	67	10.0
[Individual] Modelo Linear de Mistura Espectral (MLME)	667	63	9.4
[Individual] Métricas Espectro-Temporais (MET)	437	59	13.5
[Individual] Métricas Fenológicas (MF)	277	23	8.3
[Combinado] ACP + EVI	1317	102	7.7
[Combinado] ACP + EVI + MLME	1967	95	4.8
[Combinado] ACP + EVI + MLME + MET	2387	114	4.8
[Combinado] ACP + EVI + MLME + MET + MF	2647	117	4.4

Fonte: Produção do autor.

A fim de identificar os atributos que mais contribuíram para as classificações, foi avaliado o grau de importância de cada atributo selecionado por meio de uma análise de frequência. Para isso, a ocorrência dos atributos foi contabilizada em cada conjunto de dado, e os totais podem ser observados na Tabela 5.2. Vale dizer que um mesmo tipo de atributo pode ter sido selecionado mais de uma vez, desde que o algoritmo da seleção de atributos o tenha selecionado em mais de uma data ou em mais de uma métrica nos cubos de dados. Assim, um mesmo atributo pode ter várias ocorrências em cada um dos conjuntos de dados.

Embora haja uma pequena variação quanto aos atributos mais comuns nos conjuntos de dados, de maneira geral pode ser verificado que os atributos com as maiores ocorrências são o Desvio Padrão (espectral) e o GLCM Dissimilaridade, seguidos por GLCM Entropia, GLCM 2º Momento Angular e GLCM Correlação (Tabela 5.2). Os demais atributos tiveram menor importância para as classificações, uma vez que ocorreram em quantidade inferior à metade da ocorrência do atributo mais comum. Ainda considerando a Tabela 5.2, é possível observar que nenhum atributo geométrico esteve presente entre os atributos selecionados em todos os conjuntos de dados.

Tabela 5.2 - Quantidade total de atributos selecionados, por tipo de atributo e conjunto de dados. Os valores mais elevados indicam maior importância para a classificação.

Atributo	ACP	EVI	MLME	MET	MF	ACP + EVI	ACP + EVI + MLME	ACP + EVI + MLME + MET	ACP + EVI + MLME + MET + MF	Total
Desvio Padrão	17	10	13	13	-	22	22	25	23	145
GLCM Dissimilaridade	12	7	12	5	2	19	18	21	20	116
GLCM Entropia	12	9	11	2	2	14	13	15	15	93
GLCM 2° Momento Ang.	5	6	6	10	5	8	11	15	19	85
GLCM Correlação	9	7	9	8	3	11	11	12	13	83
Média	9	9	-	8	2	15	7	9	7	66
GLCM Homogeneidade	4	4	5	5	6	5	5	7	8	49
GLCM Média	1	7	3	4	-	4	3	5	5	32
GLCM Contraste	2	3	4	3	1	1	4	4	5	27
GLCM Desvio Padrão	2	5	-	1	2	3	1	1	2	17
Total	73	67	63	59	23	102	95	114	117	713

Fonte: Produção do autor.

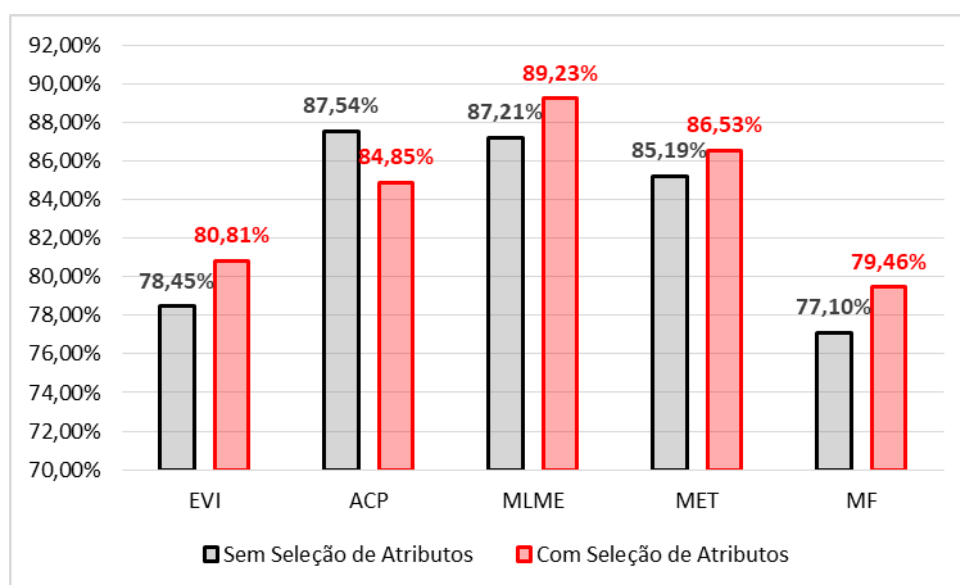
5.1.3. Assertividade dos modelos

A Figura 5.4 apresenta a acurácia das classificações dos conjuntos de dados individuais, enquanto na Figura 5.5 é apresentada a acurácia das classificações dos conjuntos de dados combinados. O Apêndice B apresenta as matrizes de confusão geradas a partir dos conjuntos amostrais de validação para cada uma das classificações. Considerando que a combinação dos dados se inicia pela ACP, este conjunto de dados está presente tanto na figura das classificações individuais quanto na figura das classificações combinadas. Conforme apresentado na

Figura 5.4, as classificações individuais, tanto com seleção de atributos quanto sem, obtiveram acurácias próximas e apresentaram comportamento semelhante, com a acurácia das classificações com seleção de atributos superior à acurácia das classificações sem seleção de atributos, exceto para a ACP, em que esse padrão é invertido. Ainda considerando as classificações individuais, a menor acurácia é atribuída ao conjunto de dados das métricas

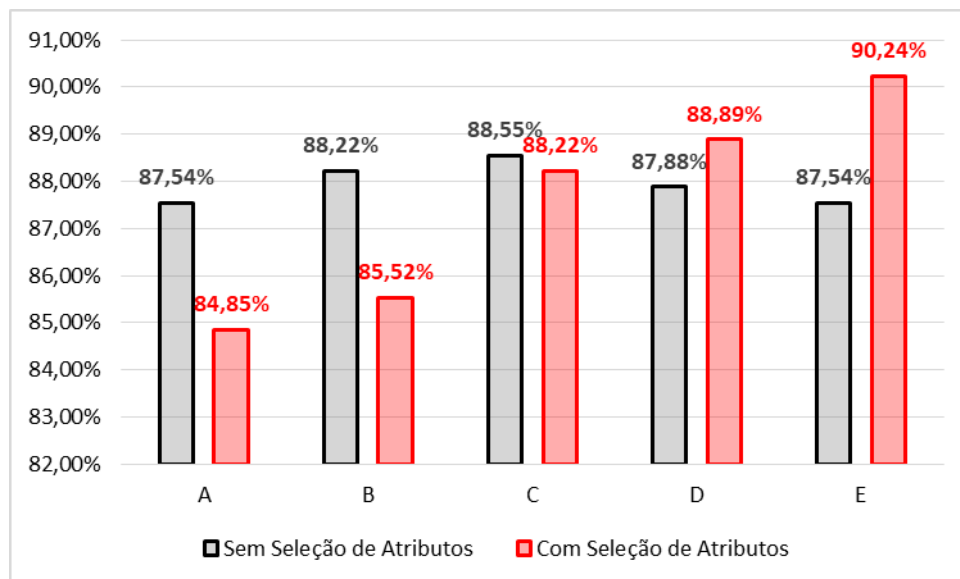
fenológicas, tanto na classificação com seleção de atributos quanto na classificação sem seleção. O valor de acurácia mais elevado está associado ao conjunto de dados do MLME (89,23%) considerando a classificação com seleção de atributos. Para as classificações sem seleção de atributos, o valor mais elevado de acurácia se associa ao conjunto de dados da ACP (87,54%), embora a acurácia do MLME tenha sido pouco inferior a esta (87,21%).

Figura 5.4 - Valores de acurácia global atingidos nas classificações individuais, com e sem seleção de atributos.



Fonte: Produção do autor.

Figura 5.5 - Valores de acurácia global atingidos nas classificações combinadas, com e sem seleção de atributos.



A: ACP / B: ACP + EVI / C: APC + EVI + MLME / D: ACP + EVI + MLME + MET / E: ACP + EVI + MLME + MET + MF.

Fonte: Produção do autor.

Diferentemente das classificações dos conjuntos de dados individuais, no que tange aos conjuntos de dados combinados, a acurácia das classificações com seleção de atributos e sem seleção seguem padrões diferentes. Enquanto a acurácia das classificações com seleção de atributos apresentou uma tendência contínua de aumento, variando de 84,85% a 90,24%, a acurácia das classificações sem seleção de atributos apresentou inicialmente uma tendência de crescimento, tornando-se decrescente, contudo, a partir do incremento do conjunto de dados das MET. A discrepância observada entre as acurácias das classificações com seleção de atributos é maior comparada à discrepância observada entre as classificações sem seleção de atributos, cuja variação foi de 87,54% a 88,55%. Considerando as classificações com seleção de atributos, a acurácia aumentou ao passo que cada conjunto de dados foi agregado aos modelos de classificação, e a acurácia mais elevada foi obtida com o incremento do conjunto de dados relativo às MF. Quanto às classificações sem seleção de

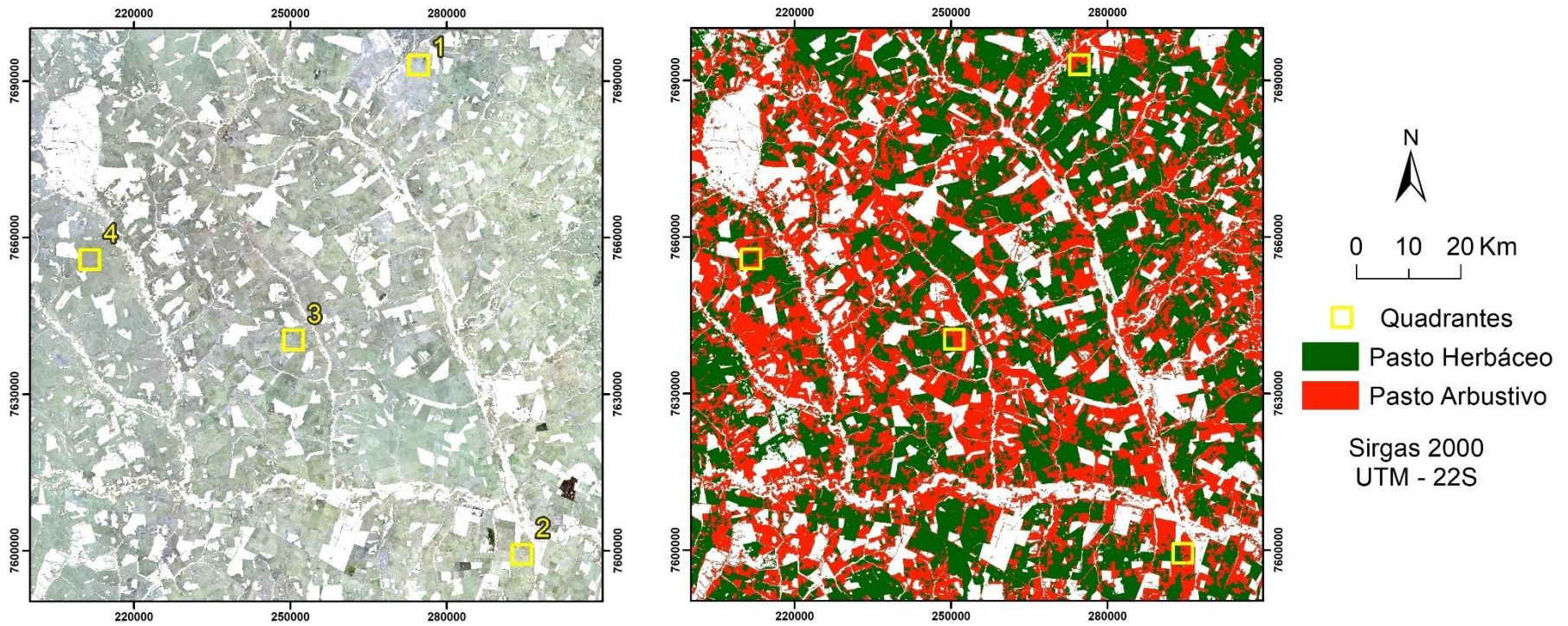
atributos, a acurácia mais elevada foi atingida com o incremento do conjunto de dados do MLME, e a menos elevada se associa igualmente ao primeiro e ao último conjunto, os quais compartilham o mesmo valor.

5.2. Discussão dos resultados

5.2.1. Classificações e dispersão amostral

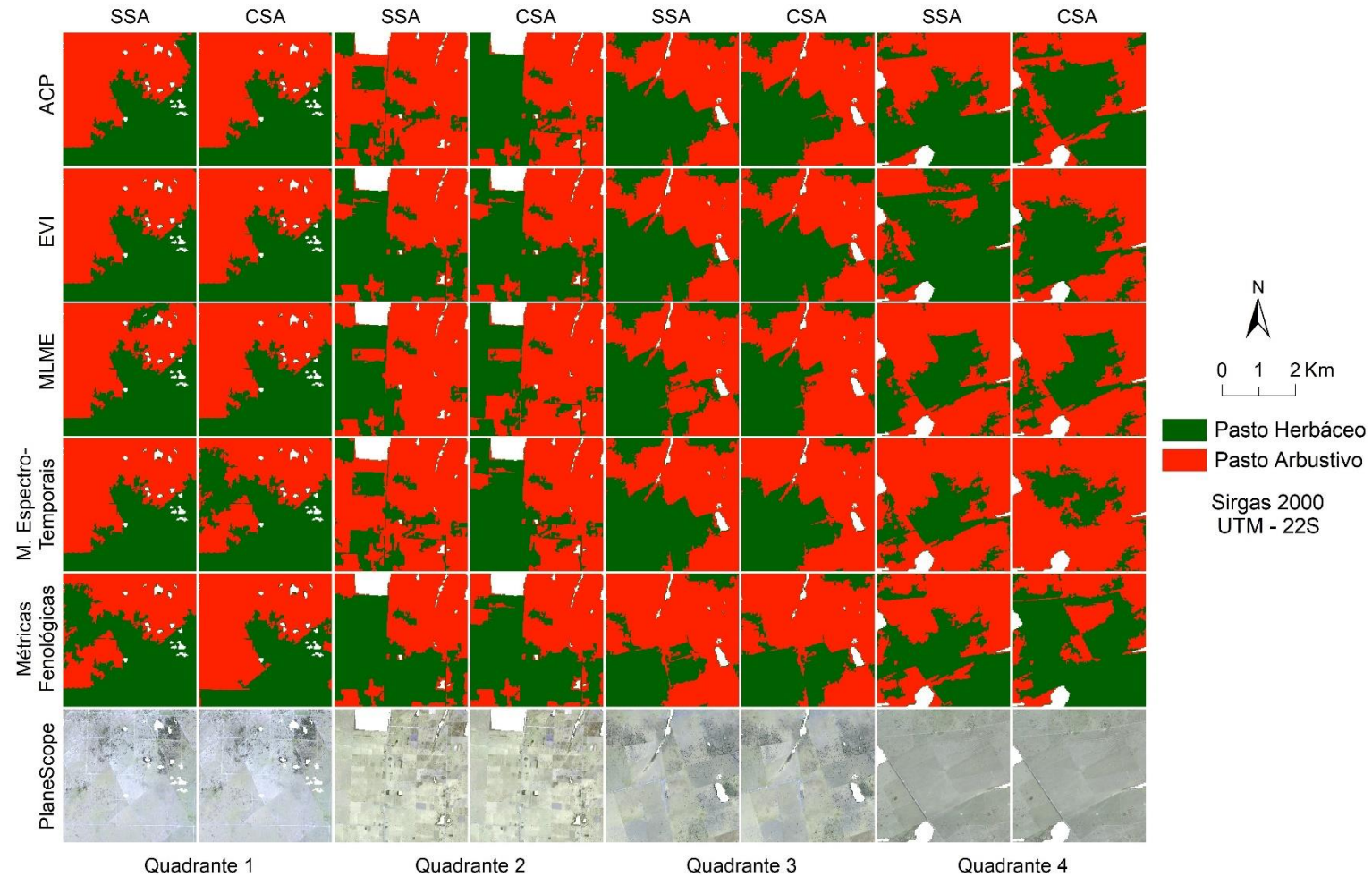
Na Figura 5.6 observa-se a classificação com o melhor resultado obtido para a separação entre pastagem arbustiva e pastagem herbácea, bem como a localização dos quadrantes de ampliação para detalhamento e comparação com as demais classificações. A Figura 5.7 e a Figura 5.8 apresentam respectivamente, em ampliações dos quadrantes da figura anterior, as classificações individuais e combinadas, com e sem seleção de atributos. Ambas as figuras apresentam ainda a imagem de satélite em cor natural para comparação das classes com as feições no terreno. Na Figura 5.9 se apresenta a análise da distribuição das amostras por classe e por tipo de atributo. É importante observar que os atributos espectrais e de textura são apresentados conjuntamente para cada conjunto de dado, enquanto os atributos geométricos são apresentados separadamente, haja vista se referirem aos objetos e não aos conjuntos de dados.

Figura 5.6 - Classificação com a maior acurácia e localização dos quadrantes de ampliação.



Fonte: Produção do autor.

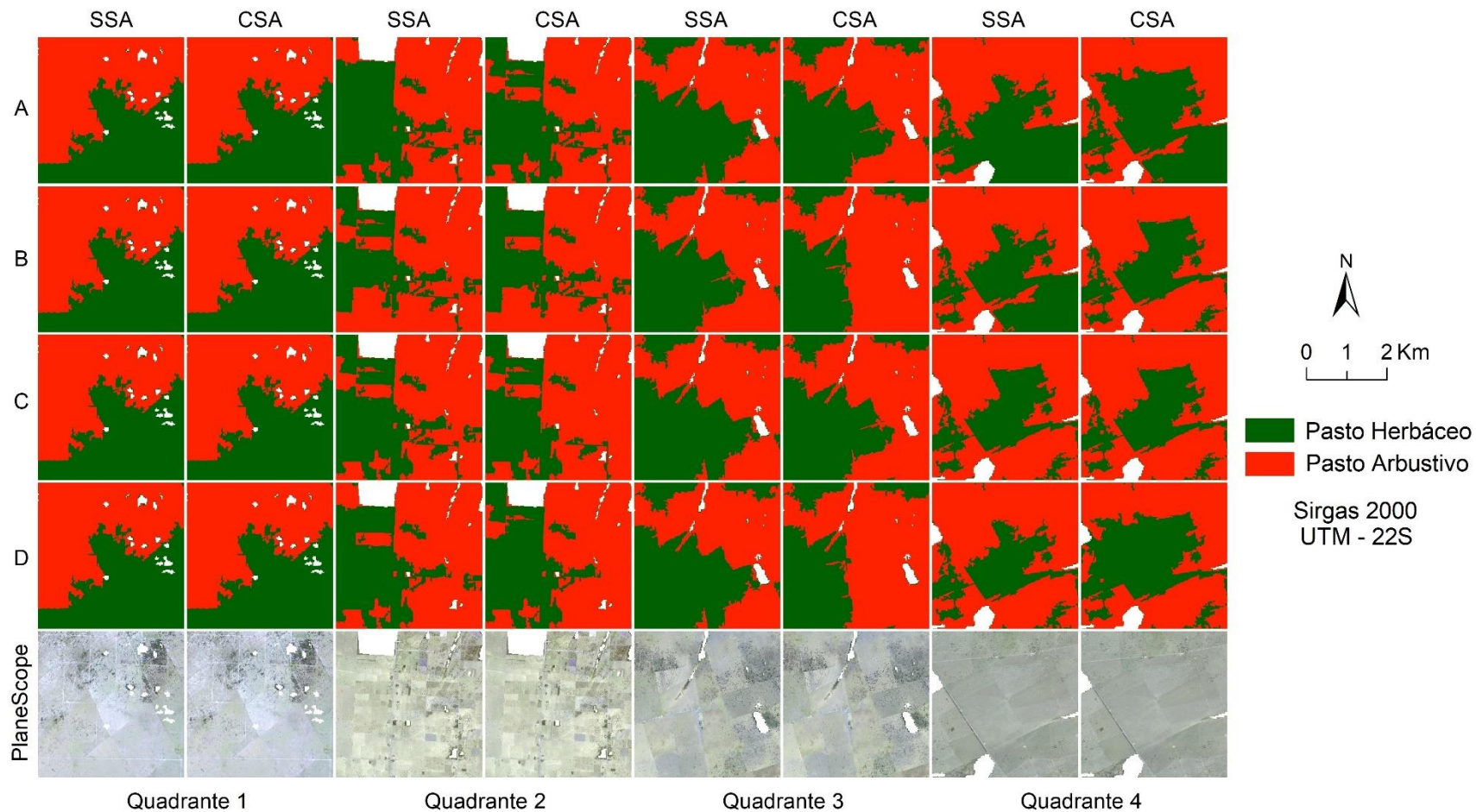
Figura 5.7 - Ampliações para detalhamento e comparação entre as classificações individuais, com e sem seleção de atributos.



SSA: sem seleção de atributos; CSA: com seleção de atributos.

Fonte: Produção do autor.

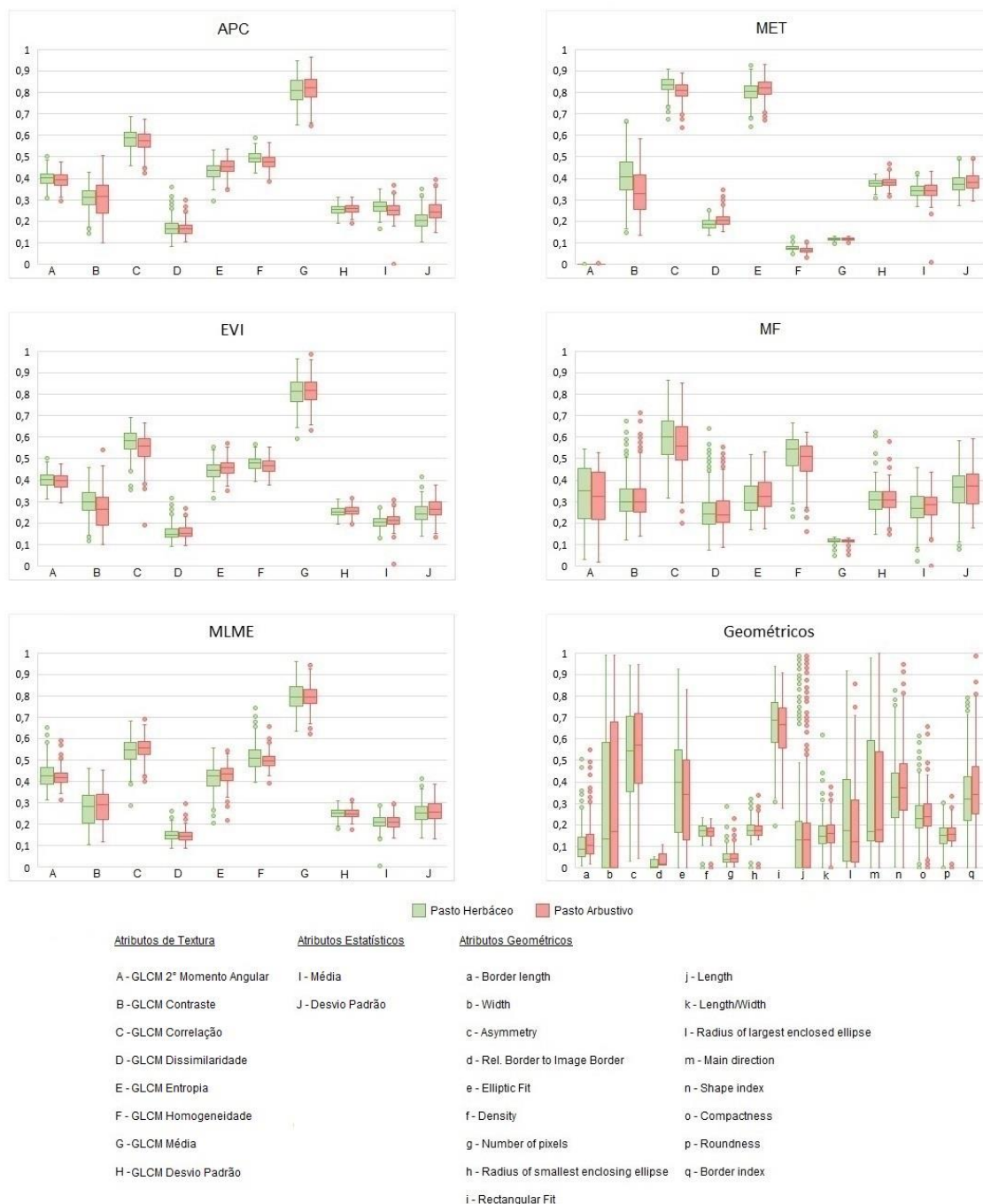
Figura 5.8 - Ampliações para detalhamento e comparação entre as classificações combinadas, com e sem seleção de atributos.



SSA: sem seleção de atributos; CSA: com seleção de atributos. A: ACP + EVI; B: ACP + EVI + MLME; C: ACP + EVI + MLME + MET; D: ACP + EVI + MLME + MET + MF.

Fonte: Produção do autor.

Figura 5.9 - Gráficos boxplot representando a dispersão de cada classe de amostra por tipo de atributo e conjunto de dado.



Fonte: Produção do autor.

Considerando a Figura 5.9, as amostras de pasto arbustivo e pasto herbáceo apresentaram-se muito próximas, isto é, com comportamento similar em todos os conjuntos de dados. Esse resultado, todavia, já era esperado, uma vez que

essas classes possuem aspecto também semelhante, resultando, por conseguinte, em amostras similares. Resultados preliminares mostraram situação semelhante ao tentar separar pastagem de agricultura no Cerrado (RODRIGUES et al., 2020b). Apesar disso, a classificação obtida atingiu 80% de acurácia com uma abordagem baseada em objetos e utilizando séries temporais.

Observa-se na Figura 5.9 que nos conjuntos de dados ACP, EVI e MLME o comportamento amostral foi similar nos três conjuntos. Essa possivelmente é a razão pela qual o acréscimo dos conjuntos de dados EVI e MLME não contribuiu consideravelmente para o aumento da acurácia nas classificações sem seleção de atributos (Figura 5.5). Apesar disso, é possível perceber que em todos os conjuntos de dados as amostras estão adequadamente distribuídas entre os atributos considerados.

Nos conjuntos de dados das MET e MF, por sua vez, as amostras apresentaram distribuição diferente tanto entre estes dois conjuntos quanto entre os demais conjuntos de dados. Embora fosse esperado que diferentes distribuições corroborassem para o aumento da acurácia nas classificações, isso não foi observado. Ao invés disso, a acurácia diminuiu a partir do acréscimo destes conjuntos (Figura 5.5). Vale lembrar aqui que o acréscimo dos conjuntos de dados aos modelos de classificação aumentou consideravelmente o volume de atributos (Figura 5.1), reduzindo o desempenho do algoritmo classificador (BELLMAN, 1961). Isso pode responder pelo decréscimo da acurácia nas classificações sem seleção de atributos com a incorporação dos conjuntos de dados das MET e MF. Essa observação reforça a importância da seleção de atributos para manter a qualidade dos modelos de classificação.

5.2.2. Acurácia global e conjuntos de dados

No que tange às classificações individuais, os resultados podem ser separados em dois grupos, um com acurácias elevadas e outro com acurácias menos elevadas, considerando de forma indistinta as classificações com e sem seleção de atributos. Do primeiro grupo fazem parte os conjuntos de dados da ACP, do MLME e das MET, enquanto do segundo fazem parte os conjuntos de dados do EVI e das MF (Figura 5.4). Os conjuntos que compõem o primeiro

grupo são voltados à exploração da variabilidade dos dados, ao que se deve, possivelmente, as acurácias mais elevadas, ao passo que nos componentes do segundo grupo, devido à similaridade entre as classes analisadas, encontrar variação entre os dados seria mais difícil.

Conforme anteriormente dito, a maior ou menor incidência de espécies arbustivas e arbóreas é a diferença basilar entre a pastagem arbustiva e a pastagem herbácea. Assim, pode-se afirmar que as gramíneas compõem a paisagem tanto de uma classe quanto da outra. Os valores de acurácia menos elevados encontrados para o conjunto de dados das MF podem ser, então, associados à conseqüente similaridade entre os ciclos sazonais das duas classes, o que pode ter reduzido a capacidade de distinção entre as classes usando este conjunto. Esse mesmo entendimento pode também ser utilizado para explicar os valores de acurácia menos elevados encontrados para o conjunto de dados do EVI. Apesar disso, Bendini et al. (2020) obtiveram acurácias superiores a 90% classificando diferentes tipos de cultura agrícola no Cerrado utilizando métricas fenológicas extraídas de séries temporais de EVI. Isso elucida que os valores de acurácia menos elevados obtidos com a classificação dos conjuntos de dados do EVI e das MF se devem à similaridade existente entre as classes, não sendo, portanto, um demérito destes conjuntos de dados.

O valor mais elevado de acurácia foi obtido com a incorporação do conjunto de dados das MF ao modelo de classificação, resultando numa acurácia global correspondente 90,24%. Apesar disso, não é razoável associar essa melhor performance exclusivamente às métricas fenológicas. Fosse este o caso, o conjunto de dados das MF teria atingido valores mais elevados de acurácia nas classificações individuais (Figura 5.4). A análise que foi feita leva a crer que este resultado se deve à classificação de múltiplos conjuntos de dados combinados, associada à seleção de atributos.

5.2.3. Aderência dos objetos às classes

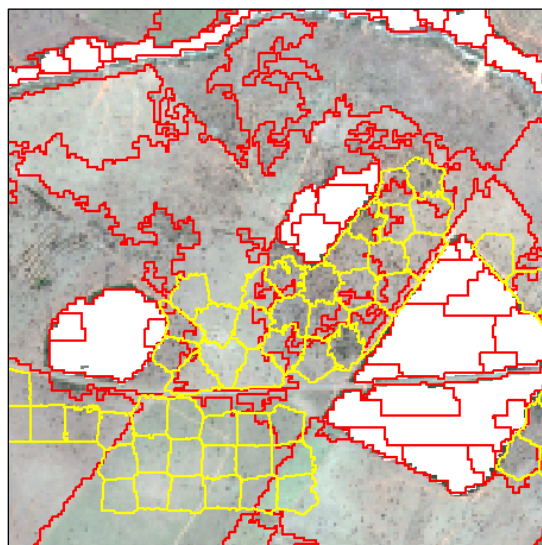
Girolamo-Neto et al. (2020) também se propuseram a separar pastagem arbustiva de pastagem herbácea no Cerrado utilizando dados do Sentinel-2. Em

seu trabalho, os autores também consideraram como área de estudo a célula 22KBB da articulação do Sentinel. Utilizando atributos de textura e a abordagem baseada em objetos, os autores analisaram imagens da estação seca e da estação chuvosa desta área por meio de índices de vegetação e imagens-fração do MLME. Diversos testes de classificação variando os dados de entrada foram realizados e o melhor resultado foi obtido com acurácia global de 87,03%. Como pôde ser visto, o resultado obtido neste trabalho superou o de Girolamo Neto et al. (2020), evidenciando o potencial da classificação de séries temporais, aliada às MET e MF, na distinção de alvos semelhantes.

Foi feita uma comparação entre a classificação de Girolamo-Neto *et al.* (2020) e a classificação com a acurácia mais elevada obtida neste trabalho por meio de uma análise de concordância. Como resultado, foi observado que os mapas são coincidentes em 68% de sua área. Era esperado que o grau de correspondência fosse mais elevado, haja vista o fato de que a metodologia utilizada nos dois trabalhos é semelhante em vários aspectos, como nos dados de sensoriamento remoto, nos tipos de atributos, na abordagem de classificação, além da área de estudo e das classes consideradas, que são as mesmas. Entretanto, existem alguns fatores que podem ter contribuído para a diferença entre os dois mapas.

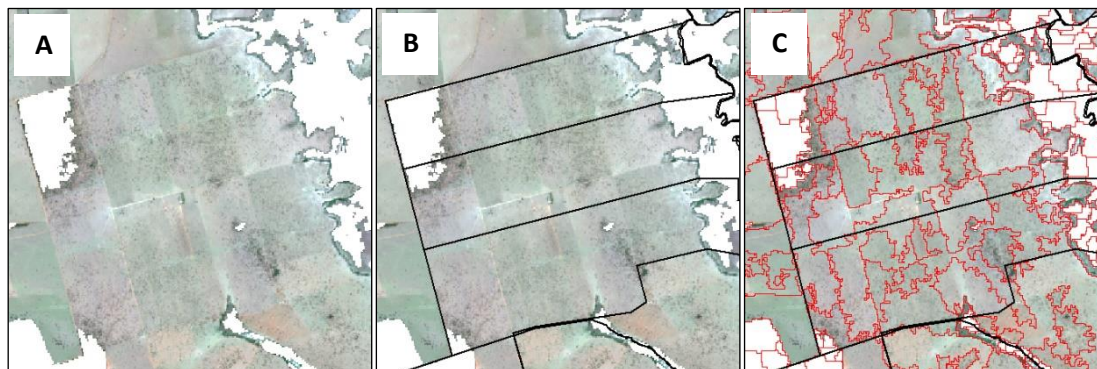
O tipo de segmentação utilizada nos dois trabalhos foi determinante para a definição do tamanho e da forma dos objetos. No trabalho de Girolamo Neto et al. (2020), os autores utilizaram a abordagem de superpixels (ACHANTA et al., 2012) para segmentar as imagens, o que conferiu polígonos menores e mais regulares ao mapa. Neste trabalho, a segmentação seguiu outra abordagem, que gerou polígonos maiores e irregulares (Figura 5.10). É importante dizer que as bordas dos polígonos, determinadas pela segmentação, não se atêm aos limites dos talhões ou das propriedades rurais (Figura 5.11), mostrando boa aderência às classes consideradas, de acordo com suas características no terreno. Vale ainda salientar que as classes de pasto arbustivo e pasto herbáceo não ocorrem em formas geométricas, se manifestando no espaço sem um formato definido. Esta pode ser a razão pela qual os atributos geométricos não corroboraram efetivamente para as classificações, uma vez que nenhum atributo geométrico esteve presente entre os atributos selecionados.

Figura 5.10 - Comparação entre a segmentação gerada com o algoritmo Multiresolution Segmentation (linhas vermelhas) e os superpixels gerados utilizando o algoritmo SLIC (linhas amarelas).



Fonte: Produção do autor.

Figura 5.11 - Relação entre os talhões das propriedades, o limite dos imóveis rurais e objetos provenientes da segmentação.



A: Talhões; B: Talhões e limites de propriedades rurais; C Talhões, limites de propriedades rurais e polígonos determinados pela segmentação (objetos).

Fonte: Produção do autor.

Outra possível razão para a discordância entre os mapas está ligada à estratégia de amostragem para treinamento e validação. Enquanto no trabalho de Girolamo Neto et al. (2020) foram utilizadas amostras provenientes de observações de campo, neste trabalho as amostras foram produzidas através de interpretação visual utilizando imagens de alta resolução espacial da constelação PlanetScope

(Seção 4.10). A criação de amostras utilizando imagens de alta resolução apresenta como vantagem a possibilidade de observar os objetos por inteiro, possibilitando uma amostragem consistente.

6. CONCLUSÕES E PERSPECTIVAS

Os atributos voltados à análise da heterogeneidade, sejam eles texturais ou espectrais, tiveram maior contribuição para as classificações, diferentemente dos atributos baseados na homogeneidade, que tiveram menor contribuição. Devido à natureza dos alvos objetos de classificação, os atributos geométricos foram desconsiderados na seleção de atributos, haja vista que as classes de pasto arbustivo e pasto herbáceo não se manifestam no espaço em formas geométricas. Soma-se a isso o fato dos polígonos da segmentação não seguirem necessariamente os limites das propriedades rurais e dos talhões internos às propriedades.

O aumento gradativo da quantidade de atributos reduziu a acurácia das classificações, revelando o importante papel que a seleção de atributos desempenha na classificação. A avaliação das classificações que foram realizadas mostrou que uma quantidade relativamente pequena de atributos, desde que não correlacionados, já é suficiente para produzir classificações com elevados níveis de confiabilidade. Os valores de acurácia mais elevados foram atingidos, portanto, quando se associou a seleção de atributos com a combinação de diferentes conjuntos de dados nas classificações.

As confusões que ocorreram nas classificações podem ser associadas à similaridade entre as amostras pertencentes às duas classes. Contudo, a análise de séries temporais, aliada ao uso de métricas espectro-temporais e fenológicas, apresentou perspectivas promissoras na separação de alvos semelhantes, ampliando as possibilidades da classificação de pastagens nos diferentes tipos em que ocorre. Como perspectiva de trabalhos futuros, se pretende utilizar dados previamente processados junto com imagens preenchidas temporalmente falando, a fim de aumentar a densidade das séries temporais e reproduzir a metodologia em áreas maiores.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACHANTA, R.; SHAJI, A.; SMITH, K.; LUCCHI, A.; FUA, P.; SÜSSTRUNK, S. SLIC superpixels compared to state-of-the-art superpixel methods. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 34, n. 11, p. 2274-2282, 2012.
- ADAMS, J. B.; SMITH, M. O. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site. **Journal of Geophysical Research**, v. 91, 1986.
- ALENCAR, A.; Z. SHIMBO, J.; LENTI, F.; BALZANI MARQUES, C.; ZIMBRES, B.; ROSA, M.; ARRUDA, V.; CASTRO, I.; FERNANDES MÁRCICO RIBEIRO, J.P.; VARELA, V.; ALENCAR, I.; PIONTEKOWSKI, V.; RIBEIRO, V.; M. C. BUSTAMANTE, M.; EYJI SANO, E.; BARROSO, M. Mapping three decades of changes in the Brazilian Savanna native vegetation using Landsat data processed in the Google Earth Engine platform. **Remote Sensing**, v.12, e924, 2020. Disponível em: <https://doi.org/10.3390/rs12060924>.
- 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.
- APPEL, M.; PEBESMA, E. On-demand processing of data cubes from satellite image collections with the gdalcubes Library. **Data**, v. 4, n. 92, 2019. Disponível em: <https://doi.org/10.3390/data4030092>.
- ASMARYAN, S.; MURADYAN, V.; TEPANOSYAN, G.; HOVSEPYAN, A.; SAGHATELYAN, A.; ASTSATRYAN, H.; GRIGORYAN, H.; ABRAHAMYAN, R.; GUIGOZ, Y.; GIULIANI, G. Paving the way towards an Armenian Data Cube. **Data**, v. 4, 117, 2019. Disponível em: <https://doi.org/10.3390/data4030117>. Acesso em: 12 jul. 2021.
- ASSIS, L. F. F. G.; QUEIROZ, G. R.; FERREIRA, K. R.; VINHAS, L.; LLAPA, E.; SANCHEZ, A.; MAUS, V.; CAMARA NETO, G. Big data streaming for remote sensing time series analytics using mapreduce. **Revista Brasileira de Cartografia**, v. 69, p. 991-1000, 2017.
- ASSIS, L. F. F. G.; FERREIRA, K. R.; VINHAS, L.; MAURANO, L.; ALMEIDA, C.; CARVALHO, A.; RODRIGUES, J.; MACIEL, A.; CAMARGO, C. TerraBrasilis: a spatial data analytics infrastructure for large-scale thematic mapping. **ISPRS International Journal of Geo-Information**, v. 8, n. 513, 2019.
- ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS EXPORTADORAS DE CARNE (ABIEC). **Beef report**: perfil da pecuária no Brasil. São Paulo – SP: ABIEC, 2018. 48p.
- AZEVEDO, T.; SOUZA-JUNIOR, C. M.; J., SHIMBO, J.; ALENCAR, A. MapBiomass initiative: mapping annual land cover and land use changes in

Brazil from 1985 to 2017. In: AGU FALL MEETING, 2018. **Abstracts...** 2018. p.B2.

BAATZ, M.; SCHÄPE, A. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In: STROBL, J.; BLASCHKE, T.; GRIESEBNER, W. (Ed.). **Angewandte geographische informationsverarbeitung** XII. Heidelberg: Wichmann-Verlag, 2000.

BELGIU, M.; DRAGUT, L. Random forest in remote sensing: a review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 114, p. 24-31, 2016.

BELLMAN, R. **Adaptive control processes**: a guided tour. New Jersey: Princeton University Press, 1961. p. 255.

BENDINI, H. N.; FONSECA, L. M. G.; SCHWIEDER, M.; KÖRTING, T. S.; RUFIN, P.; SANCHES, I. D.; LEITÃO, P. J.; HOSTERT, P. Detailed agricultural land classification in the Brazilian Cerrado based on phenological information from dense satellite image time series. **International Journal of Applied Earth Observation and Geoinformation**, v. 82, p. 1-10, 2019.

BENDINI, H. N.; FONSECA, L. M. G.; SOARES, A. R.; RUFIN, P.; SCHWIEDER, M.; RODRIGUES, M. A. A.; MARETTO, R. V.; KORTING, T. S.; LEITAO, P. J.; SANCHES, I. D.; HOSTERT, P. Applying a Phenological Object-Based Image Analysis (Phenobia) for agricultural land classification: a study case in the brazilian Cerrado. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS), 2020, Waikoloa-Hawaii. **Proceedings...**2020. p. 1078.

BLASCHKE, T. Object based image analysis for remote sensing. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 65, n. 1, p. 2–16, 2010.

BLASCHKE, T.; HAY, G. J.; KELLY, M.; LANG, S.; HOFMANN, P.; ADDINK, E.; QUEIROZ FEITOSA, R.; VAN DER MEER, F.; VAN DER WERFF, H.; VAN COILLIE, F.; TIEDE, D. Geographic Object-Based Image Analysis - towards a new paradigm. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 87, p. 180–191, 2014.

BLASCHKE, T.; LANG, S.; LORUP, E.; STROBL, J.; ZEIL, P. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In: CREMERS, A.; GREVE, K. (Ed.). **Environmental information for planning, politics and the public**. Marburg: Metropolis Verlag, 2000.

BRASIL. MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO (MAPA). **Projeções do agronegócio**: - Brasil 2017/2018 a 2027/28: projeções de longo prazo. Brasília-DF: MAPA, 2018. 114p.

BRASIL MINISTÉRIO DO MEIO AMBIENTE (MMA). **Lei n. 12.187, de 29 de dezembro de 2009**: institui a Política Nacional sobre Mudança do Clima-PNMC e dá outras providências. Brasília: MMA, 2009.

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). **Lei nº 12.651, de 25 de maio de 2012**: institui o novo código florestal brasileiro. Brasília: MMA, 2012.

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). **Mapeamento do uso e cobertura da terra do Cerrado**: Projeto TerraClass Cerrado 2013. Brasília: MMA, 2015a. 69p.

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). **Plano de ação para a prevenção e controle do desmatamento na Amazônia Legal**: 1ª fase. Brasília-DF: MMA, 2004. 156p.

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). **Plano de ação para a prevenção e controle do desmatamento e das queimadas no Cerrado**: 1ª fase. Brasília-DF: MMA, 2010. 202p.

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). Plano de ação para a prevenção e controle do desmatamento e das queimadas no Cerrado: 3ª fase: 2016 - 2020. Brasília-DF: MMA, 2018. 184p.

BRASIL. MINISTÉRIO DO MEIO AMBIENTE (MMA). **Portaria nº 365, de 27 de novembro de 2015**: institui o Programa de Monitoramento Ambiental dos Biomas Brasileiros. Brasília: DOU Diário Oficial da União, 2015b. Seção 1, n. 228. P. 114-115. Disponível em:

<https://www.mma.gov.br/images/arquivo/80049/Cerrado/Legislacao/PORTARIA%20N%20365%20DE%2027%20DE%20NOVEMBRO%20DE%202015.pdf>.

Acesso em: 07 abr. 2020.

BREIMAN, L. Random forests. **Machine Learning Journal**, v. 45, n. 1, p. 5 - 32, 2001.

CAMBRAIA NETO, A. J.; RODRIGUES, L. N. Impact of land use and occupation on potential groundwater recharge in a Brazilian savannah watershed. **Water International**, v. 46, n. 3, p. 348-364, 2021. DOI: 10.1080/02508060.2021.1898862.

CAMPOS, A. R.; FONSECA, D. F.; SAMPAIO, D. M.; FARIA, S. D.; ELMIRO, M. A. T. Classificação orientada a objeto para mapeamento do uso e cobertura do solo do município de Rio Acima - MG: por crescimento de regiões e por extração de objetos. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO (SBSR), 16., 2013, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2013.

CASTRO, A. I.; PEÑA, J. M.; TORRES-SÁNCHEZ, J.; JIMÉNEZ-BRENES, F. M.; VALENCIA-GREDILLA, F.; RECASENS, J.; LÓPEZ-GRANADOS, F. Mapping cynodon dactylon infesting cover crops with an automatic decision tree-OBIA procedure and UAV imagery for precision viticulture. **Remote Sensing**, v. 12, n. 1, p. 56, 2019.

CHIAVARI, J.; LOPES, C. L.; ARAÚJO, J. N. **Onde estamos na implementação do Código Florestal?** radiografia do CAR e do PRA nos estados brasileiros. Rio de Janeiro: Climate Policy Initiative, 2020.

COLDITZ, R. R. An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. **Remote Sensing**, v. 7, p. 9655-9681, 2015. Disponível em: <https://doi.org/10.3390/rs70809655>.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). **Calendário de plantio e colheita de grãos no Brasil**. Brasília-DF: CONAB, 2019. 75p.

COSTA, W. S.; FONSECA, L. M. G.; KORTING, T. S.; BENDINI, H. D. N.; SOUZA, R. C. M. Spatio-temporal segmentation applied to optical remote sensing image time series. **IEEE Geoscience and Remote Sensing Letters**, v. 15, n. 8, p. 1299–1303, 2018.

COUTINHO, A. C.; ALMEIDA, C. A.; SANCHES, I. D. A.; ANTUNES, J. F. G.; ESQUERDO, J. C. D. A.; ALENCAR, T. R. **Relatório de descrição das classes temáticas utilizadas no projeto gestão integrada da paisagem no bioma Cerrado**: componente 1: desenvolvimento institucional e capacitação para o gerenciamento da paisagem: mapeamento TerraClass. [S.l.]: GIZ, 2020.

COUTINHO, A. C.; ALMEIDA, C.; VENTURIERI, A.; ESQUERDO, J. C. D. M.; SILVA, M. **Projeto TerraClass**: uso e cobertura da terra nas áreas desflorestadas na Amazônia Legal. Brasília, DF: Embrapa; Belém: INPE, 2013.

DETCHEMENDY, D. M.; PACE, W. H. A model for spectral signature variability for mixtures. In: SHAHROKHI, F. (Ed.). **Remote sensing of Earth resources**. Tullahoma: University of Tennessee, 1972. v. 1, p. 596-620.

DIAS FILHO, M. B. **Diagnóstico das pastagens no Brasil**. Belém: Embrapa Amazônia Oriental, 2014.

DINIZ, C. G. et al. DETER-B: The new Amazon near real-time deforestation detection system. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 8, n. 7, 2015.

EKLUNDH, L.; JÖNSSON, P. **Timesat 3.2 software manual**. Sweden: Lund and Malmö University, 2015.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). **Sistema de análise temporal da vegetação - SATVeg**. Campinas, 2021. Disponível em: <https://www.satveg.cnptia.embrapa.br/satveg/login.html>. Acesso em: 12 jul. 2021.

EUROPEAN SPACE AGENCY (ESA). **Sentinel-2 user handbook**. [S.l.]: ESA, 2015. 64p.

EUROPEAN SPACE AGENCY (ESA). **Sen2Cor configuration and user manual**. [S.l.]: ESA, 2018. 54p.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AI Magazine**, v.17, n. 3, p. 37-54, 1996.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, R. The KDD process for extracting useful knowledge from volumes of data. **Communications of the ACM**, v. 39, n. 11, p. 27-34, 1996.

FERREIRA, K. R.; QUEIROZ, G. R.; VINHAS, L.; MARUJO, R. F. B.; SIMOES, R. E. O.; PICOLI, M. C. A.; CAMARA, G.; CARTAXO, R.; GOMES, V. C. F.; SANTOS, L. A.; SANCHEZ, A. H.; ARCANJO, J. S.; FRONZA, J. G.; NORONHA, C. A.; COSTA, R. W.; ZAGLIA, M. C.; ZIOTI, F.; KORTING, T. S.; SOARES, A. R.; CHAVES, M. E. D.; FONSECA, L. M. G. Earth observation data cubes for Brazil: requirements, methodology and products. **Remote Sensing**, v. 12, e 4033, 2020.

FONSECA, L.; NAMIKAWA, L.; CASTEJON, E.; CARVALHO, L.; PINHO, C.; PAGAMISSE, A. **Image fusion for remote sensing applications**. [S.I.]: InTech, 2011.

FRANK; E.; HALL, M. A.; WITTEN, I. H. **The WEKA workbench: online appendix for data mining: practical machine learning tools and techniques**. 4.ed. [S.I.]: Morgan Kaufmann, 2016.

FUNDAÇÃO DE CIÊNCIA, APLICAÇÃO E TECNOLOGIA ESPACIAIS (FUNCATE). **Monitoramento ambiental dos biomas brasileiros por satélites: Mata Atlântica, Caatinga, Pampa, Pantanal**. Fundo Amazônia/BNDES. Disponível em: <http://biomas.funcate.org.br/>. Acesso em: 12 jul. 2021.

GAO, X.; HUETE, A. R.; NI, W.; MIURA, T. Optical-biophysical relationships of vegetation spectra without background contamination. **Remote Sensing of Environment**, v. 74, p. 609-620, 2000.

GIROLAMO NETO, C. D.; SATO, L. Y.; SANCHES, I. D.; SILVA, I. C. O.; ROCHA, J. C. S.; ALMEIDA, C. A. Object based image analysis and texture features for pasture classification in Brazilian Savannah. **ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 3, p. 453-460, 2020.

GIROLAMO NETO, C.; FONSECA, L. M. G.; KORTING, T. S. Assessment of texture features for Brazilian savanna classification: a case study in Brasília National Park. RBC. **Revista Brasileira de Cartografia**, v. 5, p. 891 - 901, 2017.

GIULIANI, G.; CHATENOUX, B.; DE BONO, A.; RODILA, D.; RICHARD, J.P.; ALLENBACH, K.; DAO, H.; PEDUZZI, P. Building an Earth observations data cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). **Big Earth Data**, v. 1, p.100-117, 2017.

GOMES, L.; MIRANDA, H. S.; SOARES FILHO, B.; RODRIGUES, L.; OLIVEIRA, U.; BUSTAMANTE, M. M. C. Responses of plant biomass in the Brazilian Savanna to frequent fires. **Frontiers for Global Change**, v. 3, e 507710, 2020. DOI: 10.3389/ffgc.2020.507710.

GREGORY, J. M. **The CF metadata standard**. [S.I.]: CLIVAR, 2003. Technical report, 8.

GUAN, H.; LI, J.; CHAPMAN, M.; DENG, F.; JI, Z.; YANG, X. Integration of orthoimagery and lidar data for object-based urban thematic mapping using

Random Forests. **International Journal of Remote Sensing**, v. 34, n. 14, p. 5166-5186, 2013.

HALL, M. A. **Correlation-based feature selection for machine learning**. Dissertation (Master) - Waikato University, Hamilton, 1998.

HALL-BEYER, M. **GLCM Texture: a tutorial v.3.0**. University of Calgary, 2017a. Disponível em: <http://dx.doi.org/10.11575/PRISM/33280>.

HALL-BEYER, M. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. **International Journal of Remote Sensing**, v. 38, n. 5, p. 1312-1338, 2017b. DOI: 10.1080/01431161.2016.1278314.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man and Cybernetics**, v. 6, p. 610-621, 1973.

HAY, G.J.; CASTILLA, G. Geographic Object-Based Image Analysis (GEOBIA): a new name for a new discipline. In: BLASCHKE, T.; LANG, S.; HAY, G. (Ed.). **Object based image analysis**. Berlin: Springer, 2008. p.93112.

HORWITZ, H. M.; NALEPKA, R. F.; HYDE, P. D.; MORGENSTERN, J. P. Estimating the proportions of objects within a single resolution element of a multispectral scanner. In: INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, 7., 1971, Ann Arbor, MI. **Proceedings...** 1971. p. 1307-1320.

HUETE, A. R.; LIU, H. Q.; BATCHILY, K.; VAN LEEUWEN, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. **Remote Sensing of Environment**, v. 59, p. 440-451, 1997.

HUSSAIN, M.; CHEN, D.; CHENG, A.; WEI, H.; STANLEY, D. Change detection from remotely sensed images: from pixel-based to object-based approaches. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 80, p. 91-106, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo agropecuário 2017**: resultados definitivos. Rio de Janeiro-RJ: IBGE, 2019. 109p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Monitoramento da cobertura e uso da terra do Brasil**: 2016 - 2018. Rio de Janeiro-RJ: IBGE, 2020. 27p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Pesquisa da pecuária municipal 2018**. Rio de Janeiro-RJ: IBGE, 2019.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). COORDENAÇÃO GERAL DE OBSERVAÇÃO DA TERRA (OBT). PROGRAMA DE MONITORAMENTO DA AMAZÔNIA E DEMAIS BIOMAS. **Desmatamento**: bioma Cerrado. Disponível em: <http://terrabrasilis.dpi.inpe.br/downloads/>. Acesso em: 10 jun. 2020.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE).
COORDENAÇÃO GERAL DE OBSERVAÇÃO DA TERRA (OBT). **Mapa de uso e cobertura da terra: Cerrado - 2013**. Disponível em: http://www.dpi.inpe.br/tccerrado/dados/Mapa_TCcerrado_A1.jpg. Acesso em: 07 ago. 2020.

JIAJU, L. Development of principal component analysis applied to multitemporal Landsat TM data. **International Journal of Remote Sensing**, v.9, n.12, p.1895-1907, 1988.

JOHNSON, B. A.; MA, L. Image segmentation and object-based image analysis for environmental monitoring: recent areas of interest, researchers' views on the future priorities. **Remote Sensing**, v. 12, n. 1772, 2020. Disponível em: <https://doi.org/10.3390/rs12111772>.

JÖNSSON, P.; EKLUNDH, L. Seasonality extraction by function fitting to time-series of satellite sensor data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 40, n.8, p. 1824-1832, 2002.

JÖNSSON, P.; EKLUNDH, L. TIMESAT - a program for analysing time-series of satellite sensor data. **Computers and Geosciences**, v. 30, p. 833-845, 2004.

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, W.; HUETE, A.R.; VAN LEEUWEN, W.; WOLFE, R.E.; GIGLIO, L.; MULLER, J.; LEWIS, P.; BARNSLEY, M.J. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. **IEEE Transactions on Geoscience and Remote Sensing**, v.36, p.1228-1249, 1998.

KILLOUGH, B. The impact of analysis ready data in the Africa regional data cube. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS), 2019, Yokohama, Japan. **Proceedings...** Piscataway, NJ, USA: IEEE, 2019. p. 5646–5649.

KÖRTING, T. S.; CAMARA, G.; FONSECA, L. M. G. Land cover detection using temporal features based on polar representation. In: CONFERENCE ON GRAPHICS, PATTERNS, AND IMAGES, 2013, Arequipa. **Proceedings...** 2013.

LANEY, D. **3D data management**: controlling data volume, velocity, and variety. [S.l.: s.n.], 2001.

LEWIS, A.; OLIVER, S.; LYMBURNER, L.; EVANS, B.; WYBORN, L.; MUELLER, N.; RAEVKSI, G.; HOOKE, J.; WOODCOCK, R.; SIXSMITH, J.; WU, W.; TAN, P.; LI, F.; KILLOUGH, B.; MINCHIN, S.; ROBERTS, D.; AYERS, D.; BALA, B.; DWYER, J.; DEKKER, A.; DHU, T.; HICKS, A.; IP, A.; PURSS, M.; RICHARDS, C.; SAGAR, S.; TRENHAM, C.; WANG, P.; WANG, L. The Australian geoscience data cube-foundations and lessons learned. **Remote Sensing of Environment**, v. 202, p. 276–292, 2017.

LI, J.; ROY, D. P. A Global analysis of Sentinel-2A, Sentinel-2B and Landsat-8 data revisit intervals and implications for terrestrial monitoring. **Remote Sensing**, v. 9, e 902, 2017. Disponível em: <https://doi.org/10.3390/rs9090902>. Acesso em: 20 jun. 2020.

MASO, J.; ZABALA, A.; SERRAL, I.; PONS, X. A portal offering standard visualization and analysis on top of an open data cube for sub-national regions: the Catalan data cube example. **Data**, v. 4, n. 96, 2019.

MATO GROSSO DO SUL. SECRETARIA DE ESTADO DE MEIO AMBIENTE, DESENVOLVIMENTO ECONÔMICO, PRODUÇÃO E AGRICULTURA FAMILIAR (SEMAGRO). **Contas regionais: Produto Interno Bruto - Mato Grosso do Sul 2010/2018**. Campo Grande-MS: SEMAGRO, 2020. 30p.

MAUS, V.; CAMARA, G.; CARTAXO, R.; SANCHEZ, A.; RAMOS, F. M.; QUEIROZ, G. R. A Time-weighted dynamic time warping method for land-use and land-cover mapping. **IEEE J-STARS**, p. 1-11, 2016.

MONTEIRO, L. M.; BRUM, F. T.; PRESSEY, R. L.; MORELLATO, L. P. C.; SOARES-FILHO, B.; LIMA-RIBEIRO, M. S.; LOYOLA, R. Evaluating the impact of future actions in minimizing vegetation loss from land conversion in the Brazilian Cerrado under climate change. **Biodiversity and Conservation**, v. 29, p. 1701–1722, 2020. Disponível em: <https://doi.org/10.1007/s10531-018-1627-6>. Acesso em: 10 jun. 2020.

MÜLLER, H.; RUFIN, P.; GRIFFITHS, P.; SIQUEIRA, A. J. B.; HOSTERT, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. **Remote Sensing of Environment**, v. 156, p. 490-499, 2015.

NAGEL, G. W.; NOVO, E. M. L. M.; KAMPEL, M. Nanosatellites applied to optical Earth observation: a review. **Revista Ambiente e Agua**, v. 15, p. 1, 2020.

NEVES, A. K.; BENDINI, H. N.; KORTING, T. S.; FONSECA, L. M. G. Combining time series features and data mining to detect land cover patterns: a case study in northern Mato Grosso State, Brazil. **Revista Brasileira de Cartografia**, v. 68, p. 1133-1142, 2016.

NEVES, A. K.; KORTING, T. S.; GIROLAMO NETO, C.; FONSECA, L. M. G. Mineração de dados de sensoriamento remoto para detecção e classificação de áreas de pastagem na Amazônia Legal. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 18., 2017, Santos - SP. **Anais...** São José dos Campos: INPE, 2017. p. 2508-2515.

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.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD); FOOD AND AGRICULTURE ORGANIZATION (FAO). **Meat in OECD-**

FAO agricultural outlook 2019-2028. Paris: OECD Publishing, 2019.
Disponível em: https://doi.org/10.1787/agr_outlook-2019-en. Acesso em: 10 jun. 2020.

PARENTE, L.; FERREIRA, L. Assessing the spatial and occupation dynamics of the brazilian pasturelands based on the automated classification of MODIS images from 2000 to 2016. **Remote Sensing**, v. 10, n. 606, 2018. Disponível em: <https://doi.org/10.3390/rs10040606>.

PARENTE, L.; FERREIRA, L.; FARIA, A.; NOGUEIRA, S.; ARAÚJO, F.; TEIXEIRA, L.; HAGEN, S. Monitoring the brazilian pasturelands: a new mapping approach based on the landsat 8 spectral and temporal domains. **International Journal of Applied Earth Observation and Geoinformation**, v. 62, p. 135–143, 2017.

PARENTE, L.; MESQUITA, V.; MIZIARA, F.; BAUMANN, L.; FERREIRA, L. Assessing the pasturelands and livestock dynamics in Brazil, from 1985 to 2017: a novel approach based on high spatial resolution imagery and Google Earth Engine cloud computing. **Remote Sensing of Environment**, v. 232, p. 111301, 2019.

POTAPOV, P.; HANSEN, M.C.; KOMMAREDDY, I.; KOMMAREDDY, A.; TURUBANOVA, S.; PICKENS, A.; ADUSEI, B.; TYUKAVINA, A.; YING, Q. Landsat analysis ready data for global land cover and land cover change mapping. **Remote Sensing**, v. 12, e 426, 2020. Disponível em: <https://doi.org/10.3390/rs12030426>. Acesso em: 12 jul. 2021.

PRUDENTE, V. H. R.; MARTINS, V. S.; VIEIRA, D. C.; SILVA, N. R. F.; 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. 1, e100414, 2020.

RANSON, K. J. **Computer assisted classification of mixtures with simulated spectral signatures.** Thesis (PhD) - Colorado State University, 1975.

RIBEIRO, J. F.; WALTER, B. M. T. As principais fitofisionomias do bioma Cerrado. In: SANO, S. M.; ALMEIDA, S. P.; RIBEIRO, J. F. (Ed.). **Cerrado: ecologia e flora.** Brasília: EMBRAPA, 2008. p. 152 - 212.

RODRIGUES, M. A. A.; BENDINI, H. N.; SOARES, A. R.; KORTING, T. S.; FONSECA, L. M. G. Assessing differentiation between pasture and croplands using remote sensing image time series metrics. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS), 2020, Waikoloa-Hawaii. **Proceedings... IEEE**, 2020. p.168.

RODRIGUES, M. A. A.; BENDINI, H. N.; SOARES, A. R.; KORTING, T. S.; FONSECA, L. M. G. Remote sensing image time series metrics for distinction between pasture and croplands using the Random Forest classifier. In: LATIN AMERICAN GRSS & ISPRS REMOTE SENSING CONFERENCE (LAGIRS), 2020, Santiago. **Proceedings... IEEE**, 2020. p. 149.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the great plains with ERTS. In: EARTH RESOURCES TECHNOLOGY SATELLITE-1 SYMPOSIUM, 3., 1973. Washington.

Proceedings... Washington: NASA, 1974, v.1, p.309-317.

RUFIN, P.; MÜLLER, H.; PFLUGMACHER, D.; HOSTERT, P. Land use intensity trajectories on Amazonian pastures derived from Landsat time series.

International Journal of Applied Earth Observation and Geoinformation, v. 41, p. 1-10, 2015.

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. DOI: 10.1080/01431160600981517.

SANO, E. E.; ROSA, R.; BRITO, J. L.; FERREIRA, L. G. Land cover mapping of the tropical savanna region in Brazil. **Environmental Monitoring and Assessment**, v. 166, p. 113-124, 2010.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. **Analytical Chemistry**, v. 36, p. 1627–1639, 1964.

SEYLER, F.; CHAPLOT, V.; MULLER, F.; CERRI, C. E. P.; BERNOUX, M.; BALLESTER, V.; FELLER, C.; CERRI, C. C. C. Pasture mapping by classification of Landsat TM images: analysis of the spectral behaviour of the pasture class in a real medium-scale environment: the case of the Piracicaba Catchment (12 400 km², Brazil). **International Journal of Remote Sensing**, v. 23, n. 23, p. 4985 - 5004, 2002. DOI: 10.1080/01431160210146217.

SHIMABUKURO, Y. E.; SMITH, J. A. The least-squares mixing models to generate fraction images derived from Remote Sensing multispectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 29, n. 1, p. 16 - 20, 1991.

SILJESTROM, P. A.; MORENO, A.; VIKGREN, K.; CACERES, L. M. The application of selective principal components analysis (SPCA) to a Thematic Mapper (TM) image for the recognition of geomorphologic features configuration. **International Journal of Remote Sensing**, v. 18, n. 18, p. 3843-3852, 1997.

SIMÕES, R.; CAMARA, G.; QUEIROZ, G.; SOUZA, F.; ANDRADE, P.R.; SANTOS, L.; CARVALHO, A.; FERREIRA, K. Satellite image time series analysis for big Earth observation data. **Remote Sensing**, v. 13, 2021. Disponível em: <https://doi.org/10.3390/rs13132428>. Acesso em: 12 jul. 2021.

SINGH, A.; HARRISON, A. Standardized principal components. **International Journal of Remote Sensing**, v. 6, n. 6, 1985.

SIQUEIRA, A.; TADONO, T.; ROSENQVIST, A.; LACEY, J.; LEWIS, A.; THANKAPPAN, M.; SZANTOI, Z.; GORYL, P.; LABAHN, S.; ROSS, J.; HOSFORD, S. MECKLENBURG, S. TADONO, T. ROSENQVIST, A. LACEY, J.

CEOS analysis ready data for land: an overview on the current and future work. In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS), 2019, Yokohama, Japan. **Proceedings...** IEEE, 2019. p. 5536–5537.

SOUZA JUNIOR, C.; AZEVEDO, T. **MapBiomass general handbook**. [S.l.: s.n.], 2017.

TOBLER, W. R. Supplement: proceedings: International Geographical Union: commission on quantitative methods: a computer movie simulating urban growth in the Detroit region. **Economic Geography**, v. 46, p. 234–240, 1970. DOI:10.2307/143141.

TRIMBLE. **eCognition developer**: user guide. Munich, Germany: Trimble GmbH, 2014. 270p.

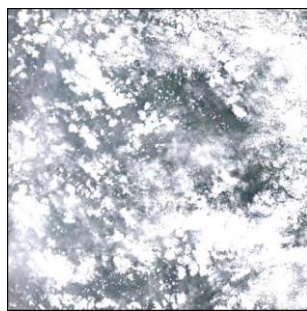
VALE, G. M.; VACOVAIS, M. V.; ZANIN, R. B. Metodologia de filtragem de estruturas rodoviárias em imagens digitais via índice de artificialidade e espaço de cor. In: SIMPÓSIO DE GEOTECNOLOGIAS DO PANTANAL (GEOPANTANAL), 5., 2014, Campo Grande - MS. **Anais...** 2014. p. 864-875.

VERIKAS, A.; GELZINIS, A.; BACAUSKIENE, M. Mining data with random forests: a survey and results of new tests. **Pattern Recognition**, v. 44, p. 330–349, 2011.

XU, L.; LI, J.; BRENNING, A.; A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. **Remote Sensing of Environment**, v. 141, 14-23, 2014.

YU, Q.; GONG, P.; CHINTON, N.; BIGING, G.; KELLY, M.; SCHIROKAUER, D. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. **Photogrammetric Engineering & Remote Sensing**, v. 72, n. 7, p.799 - 811, 2006.

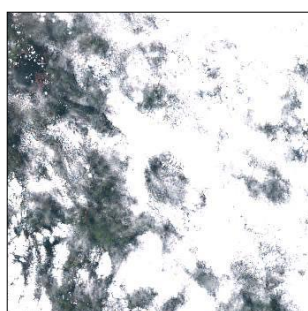
APÊNDICE A - COBERTURA DE NUVENS NA SÉRIE TEMPORAL



01/01/2018 - 89%



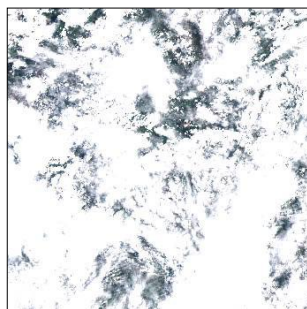
11/01/2018 - 99%



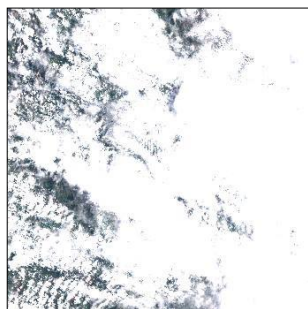
16/01/2018 - 70%



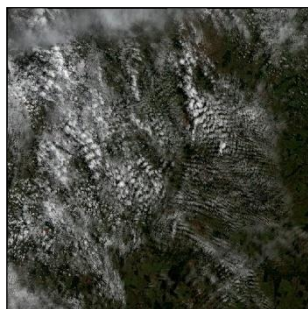
21/01/2018 - 27%



31/01/2018 - 83%



10/02/2018 - 85%



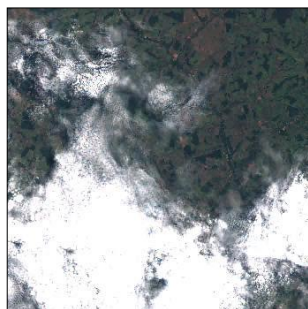
15/02/2018 - 30%



20/02/2018 - 100%



25/02/2018 - 71%



02/03/2018 - 34%



07/03/2018 - 29%



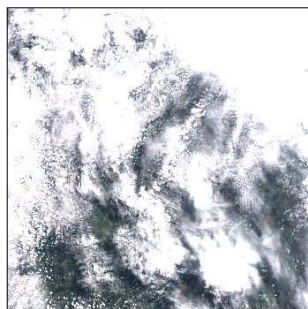
12/03/2018 - 15%



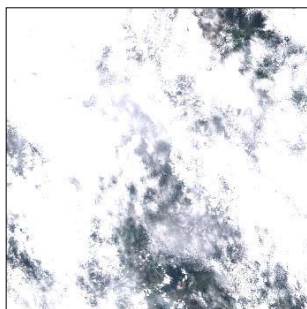
17/03/2018 - 0%



22/03/2018 - 0%



27/03/2018 - 74%



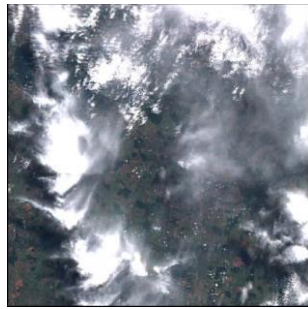
01/04/2018 - 89%



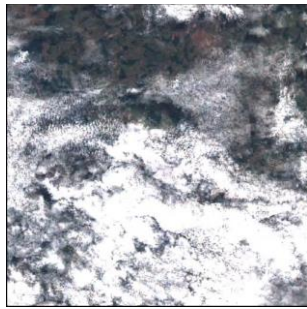
06/04/2018 - 0%



11/04/2018 - 0%



16/04/2018 - 73%



21/04/2018 - 36%



26/04/2018 - 0%



01/05/2018 - 0%



06/05/2018 - 60%



11/05/2018 - 94%



16/05/2018 - 99%



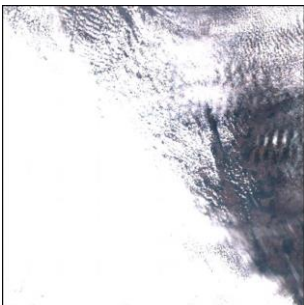
21/05/2018 - 0%



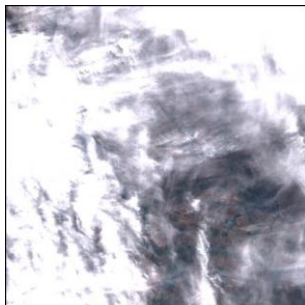
26/05/2018 - 12%



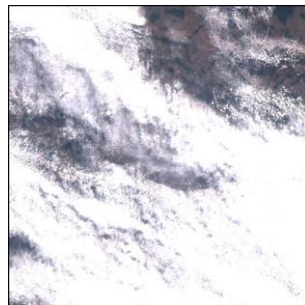
31/05/2018 - 1%



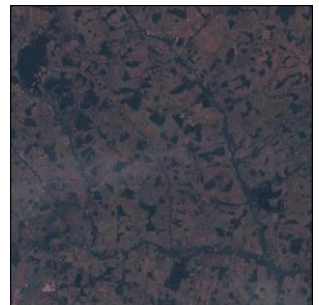
05/06/2018 - 89%



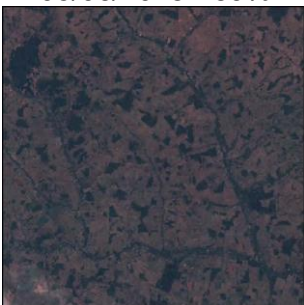
10/06/2018 - 98%



15/06/2018 - 93%



20/06/2018 - 8%



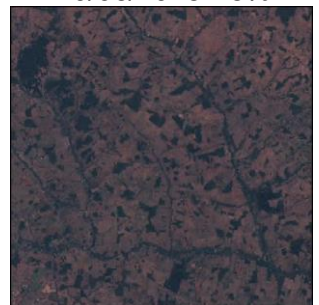
25/06/2018 - 0%



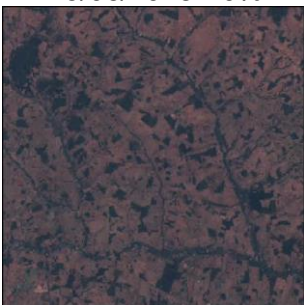
30/06/2018 - 9%



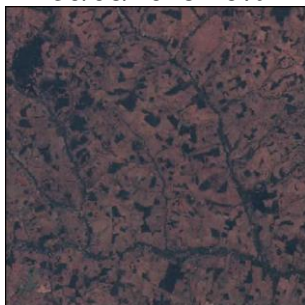
05/07/2018 - 0%



10/07/2018 - 0%



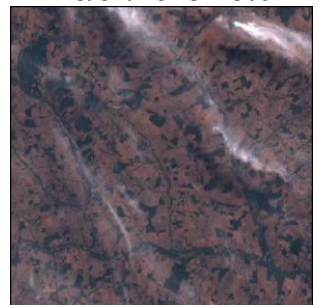
15/07/2018 - 0%



20/07/2018 - 0%

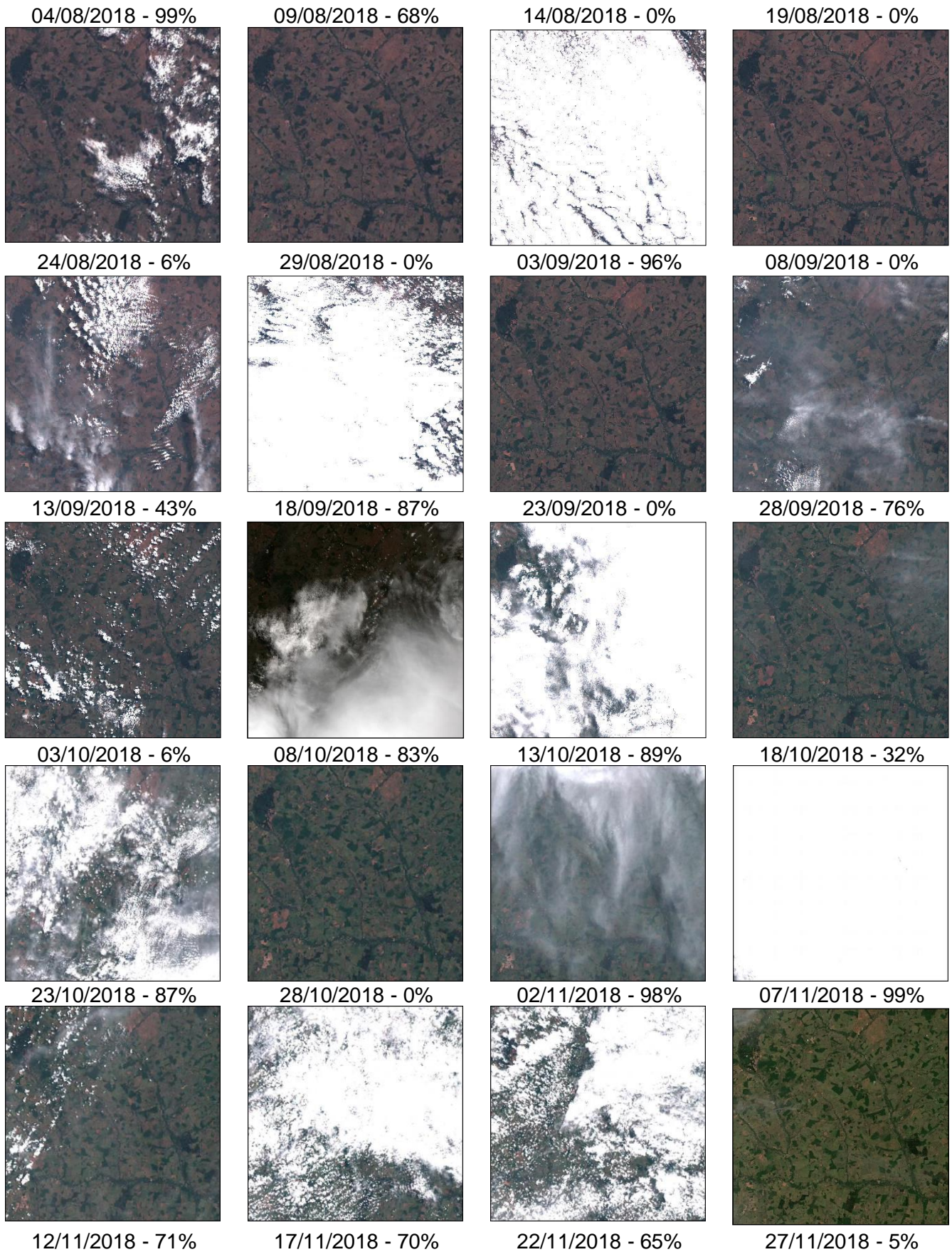


25/07/2018 - 41%



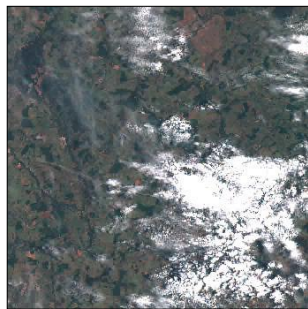
30/07/2018 - 38%



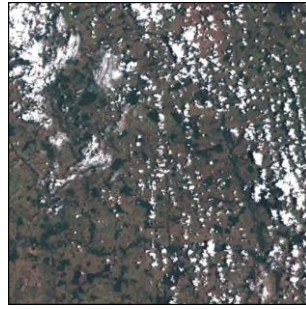




02/12/2018 - 35%



07/12/2018 - 27%



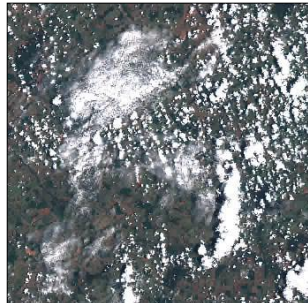
12/12/2018 - 12%



17/12/2018 - 2%



22/12/2018 - 90%



27/12/2018 - 15%

APÊNDICE B - MATRIZES DE CONFUSÃO

CLASSIFICAÇÕES INDIVIDUAIS

Tabela B.1 - Análise por Componentes Principais (ACP), sem seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	128	18	146	87,67%
	Pastagem Arbustiva	19	132	151	87,42%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	87,07%	88,00%	Índice Kappa	0,75 / 87,54%

Fonte: Produção do autor.

Tabela B.2 - Análise por Componentes Principais (ACP), com seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	117	15	132	88,64%
	Pastagem Arbustiva	30	135	165	81,82%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	79,59%	90,00%	Índice Kappa	0,70 / 84,85%

Fonte: Produção do autor.

Tabela B.3 - Índice de Vegetação Melhorado - *Enhanced Vegetation Index* (EVI), sem seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	119	36	155	76,77%
	Pastagem Arbustiva	28	114	142	80,28%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	80,95%	76,00%	Índice Kappa	78,45%
					0,57

Fonte: Produção do autor.

Tabela B.4 - Índice de Vegetação Melhorado - *Enhanced Vegetation Index* (EVI), com seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	119	29	148	80,41%
	Pastagem Arbustiva	28	121	149	81,21%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	80,95%	80,67%	Índice Kappa	80,81%
					0,62

Fonte: Produção do autor.

Tabela B.5 - Modelo Linear de Mistura Espectral (MLME), sem seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	131	22	153	85,62%
	Pastagem Arbustiva	16	128	144	88,89%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	89,12%	85,33%	Índice Kappa	87,21%
					0,74

Fonte: Produção do autor.

Tabela B.6 - Modelo Linear de Mistura Espectral (MLME), com seleção de atributos.

		Referência			
		Pastagem Herbácea	Pastagem Arbustiva	Total	Acurácia do Usuário
Classificação	Pastagem Herbácea	131	16	147	89,12%
	Pastagem Arbustiva	16	134	150	89,33%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	89,12%	89,33%	Índice Kappa	89,23% 0,78

Fonte: Produção do autor.

Tabela B.7 - Métricas Espectro-Temporais (MET), sem seleção de atributos.

		Referência			
		Pastagem Herbácea	Pastagem Arbustiva	Total	Acurácia do Usuário
Classificação	Pastagem Herbácea	128	25	153	83,66%
	Pastagem Arbustiva	19	125	144	86,81%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	87,07%	83,33%	Índice Kappa	85,19% 0,70

Fonte: Produção do autor.

Tabela B.8 - Métricas Espectro-Temporais (MET), com seleção de atributos.

		Referência			
		Pastagem Herbácea	Pastagem Arbustiva	Total	Acurácia do Usuário
Classificação	Pastagem Herbácea	126	19	145	86,90%
	Pastagem Arbustiva	21	131	152	86,18%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	85,71%	87,33%	Índice Kappa	86,53% 0,73

Fonte: Produção do autor.

Tabela B.9 - Métricas Fenológicas (MF), sem seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	113	34	147	76,87%
	Pastagem Arbustiva	34	116	150	77,33%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	76,87%	77,33%	Índice Kappa	77,10% 0,54

Fonte: Produção do autor.

Tabela B.10 - Métricas Fenológicas (MF), com seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	121	35	156	77,56%
	Pastagem Arbustiva	26	115	141	81,56%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	82,31%	76,67%	Índice Kappa	79,46% 0,59

Fonte: Produção do autor.

CLASSIFICAÇÕES COMBINADAS

Tabela B.11 - PCA + EVI, sem seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	131	19	150	87,33%
	Pastagem Arbustiva	16	131	147	89,12%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	89,12%	87,33%	Índice Kappa	88,22% 0,76

Fonte: Produção do autor.

Tabela B.12 - PCA + EVI, com seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	121	17	138	87,68%
	Pastagem Arbustiva	26	133	159	83,65%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	82,31%	88,67%	Índice Kappa	85,52%
				0,71	

Fonte: Produção do autor.

Tabela B.13 - PCA + EVI + MLME, sem seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	131	18	149	87,92%
	Pastagem Arbustiva	16	132	148	89,19%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	89,12%	88,00%	Índice Kappa	88,55%
				0,77	

Fonte: Produção do autor.

Tabela B.14 - PCA + EVI + MLME, com seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	129	17	146	88,36%
	Pastagem Arbustiva	18	133	151	88,08%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	87,76%	88,67%	Índice Kappa	88,22%
				0,76	

Fonte: Produção do autor.

Tabela B.15 - PCA + EVI + MLME + MET, sem seleção de atributos.

		Referência			
		Pastagem Herbácea	Pastagem Arbustiva	Total	Acurácia do Usuário
Classificação	Pastagem Herbácea	127	16	143	88,81%
	Pastagem Arbustiva	20	134	154	87,01%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	86,39%	89,33%	Índice Kappa	87,88%
					0,76

Fonte: Produção do autor.

Tabela B.16 - PCA + EVI + MLME + MET, com seleção de atributos.

		Referência			
		Pastagem Herbácea	Pastagem Arbustiva	Total	Acurácia do Usuário
Classificação	Pastagem Herbácea	128	14	142	90,14%
	Pastagem Arbustiva	19	136	155	87,74%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	87,07%	90,67%	Índice Kappa	88,89%
					0,78

Fonte: Produção do autor.

Tabela B.17 - PCA + EVI + MLME + MET + MF, sem seleção de atributos.

		Referência			
		Pastagem Herbácea	Pastagem Arbustiva	Total	Acurácia do Usuário
Classificação	Pastagem Herbácea	128	18	146	87,67%
	Pastagem Arbustiva	19	132	151	87,42%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	87,07%	88,00%	Índice Kappa	87,54%
					0,75

Fonte: Produção do autor.

Tabela B.18 - PCA + EVI + MLME + MET + MF, com seleção de atributos.

		Referência			Acurácia do Usuário
		Pastagem Herbácea	Pastagem Arbustiva	Total	
Classificação	Pastagem Herbácea	134	16	150	89,33%
	Pastagem Arbustiva	13	134	147	91,16%
	Total	147	150	297	Acurácia Global
	Acurácia do Produtor	91,16%	89,33%	Índice Kappa	90,24%
				0,80	

Fonte: Produção do autor.